SIAM J. CONTROL OPTIM.
Vol. 49, No. 1, pp. 106–124

STATE OBSERVERS WITH RANDOM SAMPLING TIMES AND CONVERGENCE ANALYSIS OF DOUBLE-INDEXED AND RANDOMLY WEIGHTED SUMS OF MIXING PROCESSES*

LE VAN THANH†, G. GEORGE YIN‡, AND LE YI WANG§

Abstract. Algorithms for system identification, estimation, and adaptive control in stochastic systems rely mostly on different types of signal averaging to achieve uncertainty reduction, convergence, stability, and performance enhancement. The core of such algorithms is various types of laws of large numbers that reduce the effect of noises when they are averaged. Many of the noise sequences encountered are often correlated and nonwhite. In the case of state estimation using quantized information such as in networked systems, convergence must be analyzed on double-indexed and randomly weighted sums of mixing-type stochastic processes, which are correlated with the remote past and distant future being asymptotically independent. This paper presents new results on convergence analysis of such processes. Strong laws of large numbers and convergence rates for such problems are established. These results resolve some fundamental issues in state observer designs with random sampling times, quantized information processing, and other applications.

Key words. randomly weighted sum, mixing sequence, double-indexed sum, triangular array, state observer, quantized information

AMS subject classifications. 93E12, 60F15, 93E24, 93E35, 93B30

DOI. 10.1137/10078671X

1. Introduction.

Algorithms for system identification, estimation, and adaptive control in stochastic systems rely mostly on different types of signal averaging to achieve uncertainty reduction, convergence, stability, and performance enhancement. The essence of such algorithms is based on various types of laws of large numbers that reduce the effect of noises when they are averaged. Many of the noise sequences encountered are often correlated and nonwhite. In the traditional setup of linear-time-invariant systems, observation structures often lead to single-indexed weighted sums of noises for which convergence results are abundant for either independent or mixing-type dependent noises; see [11, 17, 18, 19] for various results on identification and related issues and [5, 13, 23] on mixing processes and corresponding asymptotic properties.

However, new technology developments in sensors and communications have inevitably introduced many new scenarios of stochastic systems, in which convergence analysis commonly amounts to sums of random variables with double-indexed weights (also called triangular arrays). Moreover, the weights are possibly random, although they are usually independent of the noises. Furthermore, the random noise process

*Received by the editors February 22, 2010; accepted for publication (in revised form) October 21, 2010; published electronically January 13, 2011.
http://www.siam.org/journals/sicon/49-1/78671.html

†Department of Mathematics, Vinh University, Nghe An 42118, Vietnam (levt@vinhuni.edu.vn). This author's research was supported in part by the National Science Foundation under DMS-0603287.

‡Department of Mathematics, Wayne State University, Detroit, MI 48202 (gyin@math.wayne.edu). This author's research was supported in part by the National Science Foundation under DMS-0907753, and in part by the Air Force Office of Scientific Research under FA9550-10-1-0210.

§Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202 (lywang@wayne.edu). This author's research was supported in part by the National Science Foundation under DMS-0624849, and in part by the Air Force Office of Scientific Research under FA9550-10-1-0210.

itself is often mixing-type correlated random variables. Typical examples that introduce naturally mixing-type noises include communication channels with delays or memories. In reference to event-based sampling and optimal filtering [3, 22, 24], in [29], state observers are devised under binary-valued output measurements. The control input was used to generate switching-time sequences so that additional information could be obtained to estimate the internal states. It can be shown that sensor uncertainties and irregular switching time sequences will lead to double-indexed and randomly weighted sums of mixing-type variables. For recent advances in identification, state estimation, and fault detection using binary or quantized outputs, the reader is directed to [25, 30, 29, 31, 32].

In our recent work [28], an observer design problem was considered in which the switching-time sequence is deterministic. Here we are considering more realistic problems in which switching times are random. The new challenge is that one has to treat a weighted sum of random variables whose weights are random with double indices. Few results are available in the literature on convergence of double-indexed or triangular-array processes with mixing driving noises and random weights. In the existing literature, for double-indexed processes similar to ours, random weights were treated only for noises of martingale difference type [15], whereas constant weights and mixing noises were considered in [2]. Their results and methods on triangular arrays cannot be directly applied to the current setting. This paper presents new results on convergence analysis of such processes when the weights themselves are random and the driving noise is of mixing type. Strong laws of large numbers of Marcinkiewicz–Zygmund type and convergence rates for such problems are established. Under mild conditions, complete convergence and almost sure convergence are obtained, which contain as special cases the existing results in the literature. In addition, our conditions are weaker than the existing ones, and our results are more general and cover a wider range of scenarios. The asymptotic properties obtained in this paper are interesting in their own right. Their utility in observer design problems is demonstrated. Although the statements and proofs of our results are mostly probabilistic in nature and involve stochastic analysis, our asymptotic results resolve certain fundamental issues in quantized information processing and observer design and will be useful for adaptive systems, signal processing, and other applications.

The rest of the paper is arranged as follows. Section 2 presents some motivating applications for which the key technical results on convergence rely on double-indexed and randomly weighted sums of dependent noises. The central theme here is an algorithm of least mean-square type. Section 3 focuses on convergence analysis of randomly weighted sums of mixing sequences. For clarity and ease of presentation, our main results are first developed for strictly stationary random processes. Then we demonstrate that the strict stationarity on the driving noises can be relaxed to accommodate more practical situations in which noise characteristics change with time. Conditions and results of random sequences satisfying stochastic dominance are provided. In addition, we show by a counterexample that the conditions that we obtained are sharp. Applications of our results to state observation problems resolving the key convergence issues are presented in section 4. Finally, the paper is concluded with further remarks in section 5.

2. Motivating scenarios.

2.1. Main issues. Throughout this paper, let $(\Omega, \mathcal{F}, P)$ be the probability space. We focus on convergence of double-indexed summations of mixing random variables of the form

$$\frac{1}{n^r}\sum_{i=1}^{n} A_{ni} d_i \tag{2.1}$$

for some $1/2 < r < 1$. Properties of A_{ni} depend on applications and introduce different technical issues in the convergence analysis. Some of the technical issues will be resolved in this paper. As a standing condition, we use the following assumption throughout this paper.

Assumption 2.1. The noise $\{d_k\}$ is a stationary sequence of ρ^*-mixing random variables satisfying $Ed_k = 0$.

Loosely, a ρ^*-mixing process has decreasing correlation among its members that tends to 0 as the time difference between them goes to ∞. A more precise definition will be given in section 3. To proceed, we list some common ρ^*-mixing processes.

Example 2.2. Let $\{w_n\}$ be a sequence of independent and identically distributed (i.i.d.) random variables with zero mean and finite variance. Define

$$d_n = \sum_{j=0}^{l} c_j w_{n-j}$$

for some $l > 0$ and constants c_j, $j = 0, \ldots, l$. Then $\{d_n\}$ is known as a moving averaging process with order l. It can be easily verified that such a $\{d_n\}$ is a ρ^*-mixing process. A noisy communication channel with a finite memory can be modeled by such a system.

Example 2.3. Let $\{d_n : n \geq 1\}$ be a strictly stationary, finite-state, irreducible, and aperiodic Markov chain. Then it is a ρ^*-mixing sequence. Moreover, there exists $K > 0$ such that $\rho_n^* = o(e^{-Kn})$ as $n \to \infty$; see [7, Theorem 1.3]. Clearly, in this case, the correlation decays exponentially fast to 0.

In the rest of this section, we provide evidence that triangular-array processes with random weights and mixing driving noise are at the core of state observers with randomized sampling times and observations.

2.2. State observers for systems with irregular sampling times. Consider a multi-input-single-output linear-time-invariant system

$$\begin{cases} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t), \end{cases} \tag{2.2}$$

where $A \in \mathbb{R}^{m_0 \times m_0}$, $B \in \mathbb{R}^{m_0 \times m_1}$, $C \in \mathbb{R}^{1 \times m_0}$ are known system matrices. However, the initial state $x(0)$ is unknown, and consequently $x(t)$ is unknown and must be estimated from observations on y. The output $y(t)$ is measured only at a sequence of irregular sampling time instants $\{t_i\}$ with measured values $\gamma(t_i)$ and the noise $\{d_i\}$ that may be correlated:

$$y(t_i) = \gamma(t_i) + d_i. \tag{2.3}$$

We would like to estimate the state $x(t)$ from information on the control input $u(t)$, $\{t_i\}$, and $\{\gamma(t_i)\}$. In practical systems, the irregular sampling sequences $\gamma(t_i)$ can be generated by different means such as randomized sampling, event-triggered sampling, signal quantization, etc.

The solution to system (2.2) can be expressed as

$$x(t) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau. \tag{2.4}$$

From the sampling time sequence $\{t_i, i = 1, \ldots, n\}$, we have

$$\gamma(t_i) + d_i = y(t_i) = Ce^{A(t_i-t_n)}x(t_n) + C\int_{t_n}^{t_i} e^{A(t_i-\tau)}Bu(\tau)d\tau. \tag{2.5}$$

Denote

$$v(t_i, t_n) = C\int_{t_n}^{t_i} e^{A(t_i-\tau)}Bu(\tau)d\tau.$$

Then

$$Ce^{A(t_i-t_n)}x(t_n) = \gamma(t_i) - v(t_i, t_n) + d_i, \quad i = 1, \ldots, n. \tag{2.6}$$

Define

$$\Phi_n = \begin{bmatrix} Ce^{A(t_1-t_n)} \\ \vdots \\ Ce^{A(t_{n-1}-t_n)} \\ C \end{bmatrix}, \ \Gamma_n = \begin{bmatrix} \gamma(t_1) \\ \vdots \\ \gamma(t_{n-1}) \\ \gamma(t_n) \end{bmatrix},$$
$$V_n = \begin{bmatrix} v(t_1, t_n) \\ \vdots \\ v(t_{n-1}, t_n) \\ 0 \end{bmatrix}, \ D_n = \begin{bmatrix} d_1 \\ \vdots \\ d_{n-1} \\ d_n \end{bmatrix}.$$

Then (2.6) can be written as

$$\Phi_n x(t_n) = \Gamma_n - V_n + D_n. \tag{2.7}$$

Suppose that Φ_n is full rank. Then the least-squares estimate of $x(t_n)$ is given by

$$z_n = (\Phi_n'\Phi_n)^{-1}\Phi_n'(\Gamma_n - V_n). \tag{2.8}$$

In the above and hereafter, G' denotes the transpose of G. The observer will have the state estimate updated at the sampling time t_n by $\widehat{x}(t_n) = z_n$ and will run as an open loop observer for $t_n \le t < t_{n+1}$ according to

$$\dot{\widehat{x}} = A\widehat{x} + Bu.$$

From (2.7) and (2.8), the estimation error for $x(t_n)$ at the sampling time t_n is

$$e(t_n) = \widehat{x}(t_n) - x(t_n) = (\Phi_n'\Phi_n)^{-1}\Phi_n'D_n = \left(\frac{1}{n^r}\Phi_n'\Phi_n\right)^{-1}\frac{1}{n^r}\Phi_n'D_n \tag{2.9}$$

for some $1/2 < r < 1$. For convergence analysis, one must consider a typical entry in $\frac{1}{n^r}\Phi_n'D_n$. By the Cayley–Hamilton theorem [20],

$$e^{At} = \alpha_1(t)I + \cdots + \alpha_{m_0}(t)A^{m_0-1}, \tag{2.10}$$

where the functions $\alpha_i(t)$ can be derived by the Sylvester interpolation method [20]. This implies

$$Ce^{A(t_i-t_n)} = [\alpha_1(t_i-t_n),\dots,\alpha_{m_0}(t_i-t_n)]\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{m_0-1}\end{bmatrix} = \varphi'(t_i-t_n)W_o, \tag{2.11}$$

where $\varphi'(t_i-t_n) = [\alpha_1(t_i-t_n),\dots,\alpha_{m_0}(t_i-t_n)]$, and W_o is the observability matrix. Denote

$$\Psi_n = \begin{bmatrix} \varphi'(t_1-t_n) \\ \vdots \\ \varphi'(0)\end{bmatrix}.$$

Then,

$$\Phi_n = \Psi_n W_o,$$

which implies

$$\frac{1}{n^r}\Phi'_n\Phi_n = W'_o\frac{1}{n^r}\Psi'_n\Psi_n W_o,$$

$$\frac{1}{n^r}\Phi'_n D_n = \frac{1}{n^r}W'_o\Psi'_n D_n, \tag{2.12}$$

and

$$e(t_n) = \left(\frac{1}{n^r}\Phi'_n\Phi_n\right)^{-1}\frac{1}{n^r}\Phi'_n D_n = W'_o\left(\frac{1}{n^r}\Psi'_n\Psi_n\right)^{-1}\frac{1}{n^r}\Psi'_n D_n. \tag{2.13}$$

It was shown in [28] that by appropriate scaling and normalization the rows $\varphi'(t_i-t_n)$ of Ψ_n are uniformly bounded. Under the assumptions of observability and the persistent excitation (PE)-type condition $\frac{1}{n^r}\Psi'_n\Psi_n \geq M > 0$ a.s. for some $M>0$, convergence analysis is reduced to that of $\Psi'_n D_n/n^r$. Since a typical entry of $\frac{1}{n^r}\Psi'_n D_n$ is

$$\frac{1}{n^r}\sum_{i=1}^{n}\alpha_j(t_i-t_n)d_i, \tag{2.14}$$

convergence analysis for $e(t_n)$ is a special case of (2.1). Note that when the sampling time sequence is a random process, so are $\alpha_j(t_i-t_n)$ in (2.14), rendering a randomly weighted triangular array of noise driven by mixing processes.

3. Weighted sums of random sequences with random weights and double indices. This section is divided into several parts. First, a short introduction to the ρ^*-mixing processes is given. Then certain preliminary results are obtained. Next, complete convergence is established. Finally, the main convergence result in the almost sure sense is obtained.

3.1. ρ^*-mixing processes. On the probability space $(\Omega, \mathcal{F}, P)$, let $\mathcal{A}$ and $\mathcal{B}$ be two sub-σ-algebras of $\mathcal{F}$. We denote by $\mathcal{L}_2(\mathcal{A})$ the space of all square integrable and $\mathcal{A}$-measurable random variables. The maximal coefficient of correlation is defined by

$$\rho(\mathcal{A}, \mathcal{B}) = \sup_{f \in \mathcal{L}_2(\mathcal{A}),\ g \in \mathcal{L}_2(\mathcal{B})} |\mathrm{corr}(f, g)|.$$

In this section, we will establish the almost sure convergence of double-indexed array processes with mixing signals and random weights. The results obtained are of interest in their own right and can be used not only for the observer design but also many other applications that encounter double-indexed random processes. For general results, we will use the symbol $\{X_n\}$ for noise, in place of $\{d_n\}$ in observer design problems.

Let $\{X_n, n \geq 1\}$ be a sequence of random variables. For a subset S of $\mathbb{N} = \{1, 2, \dots\}$, $\sigma(S)$ means the σ-field generated by $\{X_n, n \in S\}$. For $1 \leq J \leq L \leq \infty$, define the σ-field

$$\mathcal{F}_J^L = \sigma(X_k, J \leq k \leq L).$$

For $n \geq 1$, define

$$\rho_n = \sup_{j \geq 1} \rho(\mathcal{F}_1^j, \mathcal{F}_{j+n}^\infty) \tag{3.1}$$

and

$$\rho_n^* = \sup \rho(\sigma(S), \sigma(T)), \tag{3.2}$$

where in (3.2) the supremum is taken over all pairs of nonempty finite sets S, T of $\mathbb{N}$ such that $\mathrm{dist}(S, T) \geq n$. The sequence $\{X_n, n \geq 1\}$ is said to be ρ-mixing (resp., ρ^*-mixing) if $\rho_n \to 0$ as $n \to \infty$ (resp., $\rho_n^* \to 0$ as $n \to \infty$).

A weaker mixing condition can be introduced based on the following coefficient. For $n \geq 1$, let

$$r_n^* = \sup |\mathrm{corr}(V; W)|,$$

where the supremum is taken over all finite subsets S, T of $\{1, 2, \dots\}$ such that $\mathrm{dist}(S, T) \geq n$ and all the linear combinations $V = \sum_{i \in S} a_i X_i$ and $W = \sum_{i \in T} b_i X_i$. For the stationary Gaussian sequences, the coefficients ρ_n^* and r_n^* are identical; see [16].

Note that m-dependence (which can be thought of as a moving average sequence driven by an i.i.d. random noise; see Example 2.2) implies ρ^*-mixing; and ρ^*-mixing implies ρ-mixing. However, ρ-mixing does not imply ρ^*-mixing [8]. Various limit properties under the ρ^*-mixing condition were studied in the literature; see [9] and the references therein. As noted in [27], the condition

$$\lim_{n \to \infty} \rho_n^* < 1 \tag{3.3}$$

is important in estimating the moments of partial sums or maxima of partial sums. It should be noted that since

$$0 \leq \cdots \leq \rho_n^* \leq \rho_{n-1}^* \leq \cdots \leq \rho_1^* \leq 1, \tag{3.4}$$

inequality (3.3) is equivalent to

$$\rho_N^* < 1 \quad \text{for some} \quad N \geq 1. \tag{3.5}$$

Example 3.1. Let $\{d_n : n \geq 1\}$ be a strictly stationary Gaussian sequence which has a spectral density $f(x)$ satisfying $0 < m < f(x) < M$. Let $k \geq 0$ and $g : \mathbb{R}^{k+1} \to \mathbb{R}$ be a measurable function, and $Z_i = g(d_i, d_{i+1}, \ldots, d_{i+k})$. Then $\{Z_n, n \geq 1\}$ satisfies

$$\lim_{n\to\infty} \rho_n^* \leq 1 - \frac{m}{M} < 1. \tag{3.6}$$

Especially, the sequence $\{d_n\}$ satisfies (3.6); see [6, the proof of Theorem 2].

Consider the sequence given in Example 2.3; if $f(x)$ is continuous, then $\{d_n\}$ is a ρ^*-mixing sequence. If $f(x)$ is not continuous, then by Theorem 1 in [6] (see also Remark 3 in [10]) $\lim_{n\to\infty} \rho_n^* > 0$. It means that the class of processes satisfying (3.3) is indeed larger than that of ρ^*-mixing processes.

3.2. Preliminary results. Convergence rates of the Marcinkiewicz–Zygmund strong law of large numbers were established in [21] for stationary sequences of random variables satisfying (3.3). In their results, there is a restriction on certain moments of the random sequences; see Remark 3.9(ii) in this paper. If the restriction is violated, the results do not hold under their setup. Recently, their results were extended in [2] to the constant weighted sums with the same restriction on the moments together with some conditions on the weights. In this paper, we study the problem in which the weights are random, relaxing the conditions in [2]. Also, we do not need to impose the restriction on the moments, as in [21] and [2].

To proceed, we present a series of lemmas first. The first lemma is a part of [9, Theorem 5.2]. Here, for a given sequence of random variables $\{X_k, k \geq 1\}$, the dependence coefficients ρ_n (resp., ρ_n^*) will be denoted by $\rho(X, n)$ (resp., $\rho^*(X, n)$).

LEMMA 3.2. *Let $X^{(n)} = \{X_k^{(n)}, k \geq 1\}$, $n \geq 1$, be sequences of random variables. Suppose these sequences $X^{(n)}$ are independent of each other. Suppose that for each $k \geq 1$, $h_k : \mathbb{R} \times \mathbb{R} \times \cdots \to \mathbb{R}$ is a Borel function. Define the sequence $\{X_k, k \geq 1\}$ of random variables by $X_k = h_k(X_k^{(1)}, X_k^{(2)}, \ldots), k \geq 1$. Then, for each $m \geq 1$,*

$$\rho(X, m) \leq \sup_{n\geq 1} \rho(X^{(n)}, m), \quad \rho^*(X, m) \leq \sup_{n\geq 1} \rho^*(X^{(n)}, m).$$

The following lemma is an immediate consequence of Lemma 3.2.

LEMMA 3.3. *Let $0 \leq r < 1$, and let N be a positive integer. Let $X = \{X_k, k \geq 1\}$ and $Y = \{Y_k, k \geq 1\}$ be two sequences of random variables satisfying $\rho^*(X, N) \leq r$, $\rho^*(Y, N) \leq r$, and let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a Borel function. Suppose that X is independent of Y. Then the sequence $f(X, Y) = \{f(X_k, Y_k), k \geq 1\}$ satisfies $\rho^*(f(X, Y), N) \leq r$.*

Lemma 3.4 is a Rosenthal-type inequality for a sequence of random variables satisfying (3.3). For a proof, see [27].

LEMMA 3.4. *Let $X = \{X_k, k \geq 1\}$ be a sequence of mean 0 random variables satisfying (3.3), and let $q \geq 2$. Then, for all $n \geq 1$,*

$$E\Big(\max_{1\leq j\leq n} \Big|\sum_{i=1}^{j} X_i\Big|^q\Big) \leq K\Big(\sum_{i=1}^{n} E|X_i|^q + \Big(\sum_{i=1}^{n} EX_i^2\Big)^{q/2}\Big), \tag{3.7}$$

where K is a constant independent of n.

Lemma 3.5 is the key machinery in proving Theorem 3.8. For convenience, here and below, a random variable A_{ni} is also denoted by $A(n; i)$ interchangeably. Set $A(n; i) = 0$ when $i > n$.

LEMMA 3.5. *Let $\alpha > 0$, $p \geq 1$, and $X = \{X_k, k \geq 1\}$ be a sequence of strictly stationary random variables with $E|X_1|^p < \infty$. $\{A_{ni}, 1 \leq i \leq n, n \geq 1\}$ is an array of random variables such that, for each $i \geq 1$, A_{ni} is independent of X_i for all $n \geq 1$. Assume that $\alpha p \geq 1$, $EX_1 = 0$ when $\alpha \leq 1$. Let $\{\nu_n, n \geq 1\}$ be a sequence of positive integers such that $\nu_n \leq n$, $n \geq 1$. If*

$$\sum_{i=1}^{\nu_n} E|A(\nu_n; i)| = O(n), \tag{3.8}$$

then

$$\lim_{n\to\infty} n^{-\alpha} \sum_{i=1}^{n} |E(A(\nu_n; i)X_{ni})| = 0, \tag{3.9}$$

where

$$X_{ni} = X_i I(|X_i| \leq n^\alpha).$$

Proof. Since $\nu_n \leq n$, it follows from (3.8) that

$$\sum_{i=1}^{n} E|A(\nu_n, i)| = \sum_{i=1}^{\nu_n} E|A(\nu_n; i)| = O(n). \tag{3.10}$$

For $n \geq 1$, since $A(\nu_n; i)$ is independent of X_{ni} and $|EA(\nu_n; i)| \leq E|A(\nu_n; i)|$,

$$\begin{aligned} &n^{-\alpha} \sum_{i=1}^{n} |E(A(\nu_n; i)X_{ni})| \\ &\leq n^{-\alpha} \sum_{i=1}^{n} E|A(\nu_n; i)||EX_{ni}| \\ &= n^{-\alpha} \Big(\sum_{i=1}^{n} E|A(\nu_n; i)|\Big) \big|E(X_1 I(|X_1| \leq n^\alpha))\big| \\ &\leq K n^{1-\alpha} \big|E(X_1 I(|X_1| \leq n^\alpha))\big| \quad \text{(by (3.10))}. \end{aligned} \tag{3.11}$$

If $\alpha > 1$, then (3.11) implies (3.9). If $0 < \alpha \leq 1$, then

$$\begin{aligned} &n^{1-\alpha}\big|E(X_1 I(|X_1| \leq n^\alpha))\big| \\ &\quad = n^{1-\alpha}\big|E(X_1 I(|X_1| > n^\alpha))\big| \text{ (since } EX_1 = 0) \\ &\quad \leq n^{1-\alpha} E(|X_1| I(|X_1| > n^\alpha)) \\ &\quad \leq E(|X_1|^{1/\alpha} I(|X_1| > n^\alpha)) \\ &\quad \leq \Big(E(|X_1|^p I(|X_1| > n^\alpha))\Big)^{1/(\alpha p)} \quad \text{(since } \alpha p \geq 1). \end{aligned} \tag{3.12}$$

Since $E|X_1|^p < \infty$,

$$\lim_{n\to\infty} E(|X_1|^p I(|X_1| > n^\alpha)) = 0. \tag{3.13}$$

Combining (3.11), (3.12), and (3.13), we obtain (3.9). ∎

3.3. Complete convergence. The concept of complete convergence was introduced in [14]. A sequence of real-valued random variables $\{X_n, n \geq 1\}$ is said to *converge completely* to 0 if

$$\sum_{n=1}^{\infty} P\big(|X_i| > \varepsilon\big) < \infty \quad \text{for all} \quad \varepsilon > 0. \tag{3.14}$$

Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d. random variables. The following result was proved in [14, 12].

PROPOSITION 3.6. *Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d. random variables. Then $EX_1 = 0$ and $E|X_1|^2 < \infty$ if and only if*

$$\sum_{n=1}^{\infty} P\Big(\Big|\sum_{i=1}^{n} X_i\Big| > \varepsilon n\Big) < \infty \quad \textit{for all} \quad \varepsilon > 0. \tag{3.15}$$

The above result was extended in [4]. For the Marcinkiewicz–Zygmund strong law of large numbers, the rate of convergence was given as follows.

PROPOSITION 3.7. *Let $\alpha > 1/2$, $p \geq 1$, and $\{X_n, n \geq 1\}$ be a sequence of i.i.d. random variables. Then $EX_1 = 0$ and $E|X_1|^p < \infty$ if and only if*

$$\sum_{n=1}^{\infty} n^{\alpha p-2} P\Big(\Big|\sum_{i=1}^{n} X_i\Big| > \varepsilon n^{\alpha}\Big) < \infty \quad \textit{for all} \quad \varepsilon > 0. \tag{3.16}$$

In [21], the above complete convergence was further extended to the ρ^*-mixing case. In the following theorem, we establish the complete convergence for randomly weighted sums of mixing random variables. It should be noted that many useful linear statistics based on a random sample are weighted sums of dependent random variables. Examples include least-squares estimators, nonparametric regression function estimators, and jackknife estimates, among others. In this respect, the studies of convergence for these weighted sums have made significant progress in probability theory and applications. Our theorem reduces to a result of [21] when $A_{ni} \equiv 1$ a.s., $\nu_n \equiv n$. In subsequent derivations, K denotes a generic positive constant whose value may be different for each appearance.

THEOREM 3.8. *Let $0 \leq r < 1$, and let N be a positive integer. Let $\alpha > 1/2$, $1 \leq p < 2$, and $X = \{X_n, n \geq 1\}$ be a sequence of strictly stationary random variables satisfying $\rho^*(X, N) \leq r$. $\{A_{ni}, n \geq 1, 1 \leq i \leq n\}$ is a triangular array of random variables, and $\{\nu_n, n \geq 1\}$ is a sequence of positive integers such that $\nu_n \leq n$, $n \geq 1$. Assume that $\alpha p \geq 1$, $EX_1 = 0$ when $\alpha \leq 1$. Suppose that $E|X_1|^p < \infty$ and, for all $n \geq 1$, the sequence $A_n = \{A_{ni}, 1 \leq i \leq n\}$ is independent of the sequence $\{X_i, i \geq 1\}$ and satisfies $\rho^*(A_n, N) \leq r$, and*

$$\sum_{i=1}^{\nu_n} E(A(\nu_n; i))^2 = O(n). \tag{3.17}$$

Then

$$\sum_{n=1}^{\infty} n^{\alpha p-2} P\Big(\max_{1\leq j\leq n}\Big|\sum_{i=1}^{j} A(\nu_n; i) X_i\Big| > \varepsilon n^{\alpha}\Big) < \infty \quad \textit{for all} \quad \varepsilon > 0. \tag{3.18}$$

Proof. For $n \ge 1$, set

$$X_{nj} = X_j I(|X_j| \le n^\alpha), \quad 1 \le j \le n,$$

and

$$S_{nj} = \sum_{i=1}^{j} \Big(A(\nu_n; i) X_{ni} - E(A(\nu_n; i) X_{ni}) \Big), \quad 1 \le j \le n.$$

For $\varepsilon > 0$,

(3.19)

$$\begin{aligned} P\Big(\max_{1\le j\le n} \Big|\sum_{i=1}^{j} A(\nu_n;i)X_i\Big| > \varepsilon n^\alpha\Big) &\le P\Big(\max_{1\le j\le n} |X_j| > n^\alpha\Big) + P\Big(\max_{1\le j\le n} \Big|\sum_{i=1}^{j} A(\nu_n;i)X_{ni}\Big| > \varepsilon n^\alpha\Big) \\ &\le P\Big(\max_{1\le j\le n} |X_j| > n^\alpha\Big) + P\Big(\max_{1\le j\le n} |S_{nj}| > \varepsilon n^\alpha - \max_{1\le j\le n}\Big|\sum_{i=1}^{j} E(A(\nu_n;i)X_{ni})\Big|\Big). \end{aligned}$$

We have

$$\begin{aligned} &\sum_{n=1}^{\infty} n^{\alpha p-2} P\Big(\max_{1\le j\le n} |X_j| > n^\alpha\Big) \\ &\quad\le \sum_{n=1}^{\infty} n^{\alpha p-2} \sum_{j=1}^{n} P\Big(|X_j| > n^\alpha\Big) \\ &\quad= \sum_{n=1}^{\infty} n^{\alpha p-1} P(|X_1| > n^\alpha) \\ &\quad= \sum_{n=1}^{\infty} n^{\alpha p-1} \sum_{i=n}^{\infty} P(i^\alpha < |X_1| \le (i+1)^\alpha) \\ &\quad= \sum_{i=1}^{\infty} \sum_{n=1}^{i} n^{\alpha p-1} P(i^\alpha < |X_1| \le (i+1)^\alpha) \\ &\quad\le K \sum_{i=1}^{\infty} i^{\alpha p} P(i^\alpha < |X_1| \le (i+1)^\alpha) \\ &\quad\le K E|X_1|^p < \infty. \end{aligned} \tag{3.20}$$

For $n \ge 1$, by Jensen's inequality,

$$\begin{aligned} \Big(\sum_{i=1}^{\nu_n} E|A(\nu_n;i)|\Big)^2 &\le \Big(\sum_{i=1}^{\nu_n} (EA^2(\nu_n;i))^{1/2}\Big)^2 \\ &\le n \sum_{i=1}^{\nu_n} EA^2(\nu_n;i). \end{aligned}$$

By (3.17), (3.8) holds. By Lemma 3.5,

$$
\begin{aligned}
0 &\le \lim_{n\to\infty} n^{-\alpha} \max_{1\le j\le n} \Big|\sum_{i=1}^{j} E(A(\nu_n;i)X_{ni})\Big| \\
&\le \lim_{n\to\infty} n^{-\alpha} \max_{1\le j\le n} \sum_{i=1}^{j} \Big|E(A(\nu_n;i)X_{ni})\Big| \\
&= \lim_{n\to\infty} n^{-\alpha} \sum_{i=1}^{n} \Big|E(A(\nu_n;i)X_{ni})\Big| = 0.
\end{aligned} \tag{3.21}
$$

From (3.19), (3.20), and (3.21), to obtain (3.18), it remains to show that

$$
\sum_{n=1}^{\infty} n^{\alpha p-2} P\Big(\max_{1\le j\le n} |S_{nj}| > n^{\alpha}\varepsilon/2\Big) < \infty. \tag{3.22}
$$

By Lemma 3.3, for all $n \ge 1$, the random variables $AX(\nu_n) = \{A(\nu_n;i)X_{ni}, 1 \le i \le n\}$ also satisfy $\rho^*(AX(\nu_n), N) \le r$. The conclusion (3.22) follows from the estimate

$$
\begin{aligned}
&\sum_{n=1}^{\infty} n^{\alpha p-2} P\Big(\max_{1\le j\le n} |S_{nj}| > n^{\alpha}\varepsilon/2\Big) \\
&\le \frac{4}{\varepsilon^2} \sum_{n=1}^{\infty} n^{\alpha p-2-2\alpha} E\Big(\max_{1\le j\le n} |S_{nj|}\Big)^2 \quad \text{(by Chebyshev's inequality)} \\
&\le K \sum_{n=1}^{\infty} n^{\alpha p-2-2\alpha} \sum_{i=1}^{n} E\Big(A(\nu_n;i)X_{ni} - E(A(\nu_n;i)X_{ni})\Big)^2 \quad \text{(by Lemma 3.4)} \\
&\le K \sum_{n=1}^{\infty} n^{\alpha p-2-2\alpha} \sum_{i=1}^{n} E(A(\nu_n;i)X_{ni})^2 \\
&= K \sum_{n=1}^{\infty} n^{\alpha p-2-2\alpha} \sum_{i=1}^{n} E(A(\nu_n;i))^2 E(X_{ni})^2 \\
&\qquad\qquad \text{(since } A(\nu_n;i) \text{ is independent of } X_{ni}\text{)} \\
&= K \sum_{n=1}^{\infty} n^{\alpha p-2-2\alpha} \Big(\sum_{i=1}^{n} E(A(\nu_n;i))^2\Big) E(X_1 I(|X_1| \le n^{\alpha}))^2 \\
&\le K \sum_{n=1}^{\infty} n^{\alpha p-1-2\alpha} E(X_1 I(|X_1| \le n^{\alpha}))^2 \\
&= K \sum_{n=1}^{\infty} n^{\alpha p-1-2\alpha} \sum_{i=1}^{n} E\big(X_1^2 I((i-1)^{\alpha} < |X_1| \le i^{\alpha})\big) \\
&= K \sum_{i=1}^{\infty} \sum_{n=i}^{\infty} n^{\alpha p-1-2\alpha} E\big(X_1^2 I((i-1)^{\alpha} < |X_1| \le i^{\alpha})\big) \\
&\le K \sum_{i=1}^{\infty} i^{\alpha p-2\alpha} E\big(X_1^2 I((i-1)^{\alpha} < |X_1| \le i^{\alpha})\big) \\
&\le K \sum_{i=1}^{\infty} i^{\alpha p} P((i-1)^{\alpha} < |X_1| \le i^{\alpha}) \\
&\le KE|X_1|^p < \infty.
\end{aligned}
$$

The proof of the theorem is completed. □

Remark 3.9. (i) If $p \geq 2$, then Theorem 3.8 still holds if (3.1) is replaced by

$$\sum_{i=1}^{\nu_n} E(|A(\nu_n;i)|^q) = O(n) \quad \text{for some} \quad q > 2(\alpha p - 1)/(2\alpha - 1). \tag{3.23}$$

(ii) The method in [21] requires that $\alpha p > 1$ even in the case $1 \leq p < 2$. So we cannot take $\alpha = 1/p$, which is more interesting in some cases.

(iii) A complete convergence was also established in [2] for constant weighted random sums, which is similar to Theorem 3.8. The authors used a different truncation method and a different approach from ours. One of the crucial conditions is $\alpha p > 1$ and

$$\sum_{i=1}^{n} |a_{ni}|^q = O(n^\delta) \quad \text{for some} \quad 0 < \delta < 1.$$

This condition is not satisfied if $a_{ni} \equiv 1$, so their result does not reduce to that of [21]. The Marcinkiewicz–Zygmund strong law of large numbers was also deduced for weighted sums from their complete convergence results. In general, for the weighted case, we cannot obtain

$$\lim_{n\to\infty} \frac{1}{n^{1/p}} \sum_{i=1}^{n} A_{ni}X_i = 0 \quad \text{a.s.}$$

from

$$\sum_{n=1}^{\infty} n^{-1} P\Big(\max_{1\leq j\leq n} \Big| \sum_{i=1}^{j} A_{ni}X_i \Big| > \varepsilon n^{1/p} \Big) < \infty$$

because the sequence $\{\max_{1\leq j\leq n} \big| \sum_{i=1}^{j} A_{ni}X_i \big|, n \geq 1\}$ is not increasing (see Example 3.14).

In the following theorem, the conditions for $\{A_{ni}, n \geq 1, 1 \leq i \leq n\}$ are stronger than (3.17). We obtain the conclusion (3.25), which is stronger than that of (3.18). Although the truncation method in the proof of Theorem 3.10 is the same as that of Theorem 3.8, the estimations are different.

THEOREM 3.10. *Let $0 \leq r < 1$, and let N be a positive integer. Let $\alpha > 0$, $1 < p < 2$, and $X = \{X_n, n \geq 1\}$ be a sequence of mean 0 strictly stationary random variables satisfying $\rho^*(X, N) \leq r$. Let $\{A_{ni}, n \geq 1, 1 \leq i \leq n\}$ be an array of random variables such that $|A_{ni}| \leq K$ a.s., $n \geq 1, 1 \leq i \leq n$, and let $\{\nu_n, n \geq 1\}$ be a sequence of positive integers such that $\nu_n \leq n$, $n \geq 1$. If $E|X_1|^p < \infty$ and, for all $n \geq 1$, the sequence $A_n = \{A_{ni}, 1 \leq i \leq n\}$ is independent of the sequence $\{X_i, i \geq 1\}$ and satisfies $\rho^*(A_n, N) \leq r$, and if*

$$\sup_{n\geq 1} \sum_{i=1}^{\nu_n} E|A(\nu_n;i)| \leq K, \tag{3.24}$$

then

$$\sum_{n=1}^{\infty} n^{\alpha p - 1} P\Big(\max_{1\leq j\leq n} \Big| \sum_{i=1}^{j} A(\nu_n;i)X_i \Big| > \varepsilon n^\alpha \Big) < \infty \quad \textit{for all} \quad \varepsilon > 0. \tag{3.25}$$

Proof. Let X_{nj} and S_{nj} be as in the proof of Theorem 3.8, and let

$$Y_{nj} = X_j - X_{nj}, T_{nj} = \sum_{i=1}^{j} \Big(A(\nu_n; i)Y_{ni} - E(A(\nu_n; i)Y_{ni})\Big), \quad n \geq 1, \quad 1 \leq j \leq n.$$

For $\varepsilon > 0$ arbitrary,

$$\begin{aligned} &P\Big(\max_{1\leq j\leq n}\Big|\sum_{i=1}^{j} A(\nu_n; i)X_i\Big| > \varepsilon n^\alpha\Big) \\ &\quad \leq P\Big(\max_{1\leq j\leq n}|S_{nj}| > n^\alpha\varepsilon/2\Big) + P\Big(\max_{1\leq j\leq n}|T_{nj}| > n^\alpha\varepsilon/2\Big). \end{aligned} \tag{3.26}$$

Similar to the proof of (3.10), we also have from (3.24) that

$$\sup_{n\geq 1}\sum_{i=1}^{n} E|A(\nu_n; i)| \leq K.$$

Now, we have

$$\begin{aligned} &\sum_{n=1}^{\infty} n^{\alpha p-1} P\Big(\max_{1\leq j\leq n}|T_{nj}| > n^\alpha\varepsilon/2\Big) \\ &\quad \leq \frac{2}{\varepsilon}\sum_{n=1}^{\infty} n^{\alpha p-1-\alpha} E\Big(\max_{1\leq j\leq n}|T_{nj}|\Big) \\ &\quad \leq \frac{4}{\varepsilon}\sum_{n=1}^{\infty} n^{\alpha p-1-\alpha}\sum_{i=1}^{n} E|A(\nu_n; i)|E|Y_{ni}| \\ &\quad = \frac{4}{\varepsilon}\sum_{n=1}^{\infty} n^{\alpha p-1-\alpha}\Big(\sum_{i=1}^{n} E|A(\nu_n; i)|\Big)E(|X_1|I(|X_1| > n^\alpha)) \\ &\quad \leq K\sum_{n=1}^{\infty} n^{\alpha p-1-\alpha}\sum_{i=n}^{\infty} E\Big(|X_1|I((i-1)^\alpha \leq |X_1| < i^\alpha)\Big) \\ &\quad \leq K\sum_{i=1}^{\infty}\sum_{n=1}^{i} n^{\alpha p-1-\alpha} E\Big(|X_1|I((i-1)^\alpha \leq |X_1| < i^\alpha)\Big) \\ &\quad \leq K\sum_{i=1}^{\infty} i^{\alpha p} P((i-1)^\alpha \leq |X_1| < i^\alpha) \\ &\quad \leq KE|X_1|^p < \infty. \end{aligned} \tag{3.27}$$

Note that

$$\sup_{n\geq 1, 1\leq i\leq n}|A_{ni}| \leq K \quad \text{a.s.},$$

so (3.24) implies

$$\sum_{i=1}^{n} E(A^2(\nu_n; i)) \leq K.$$

So, it follows from Lemma 3.4 that

$$\begin{aligned}
&\sum_{n=1}^{\infty} n^{\alpha p-1} P\Big(\max_{1\le j\le n} |S_{nj}| > n^{\alpha}\varepsilon/2\Big)\\
&\quad\le \frac{4}{\varepsilon^2}\sum_{n=1}^{\infty} n^{\alpha p-1-2\alpha} E\Big(\max_{1\le j\le n}|S_{nj}|\Big)^2\\
&\quad\le K\sum_{n=1}^{\infty} n^{\alpha p-1-2\alpha}\sum_{i=1}^{n} E(A(\nu_n;i))^2 E(X_{ni}^2)\\
&\quad\le K\sum_{n=1}^{\infty} n^{\alpha p-1-2\alpha} E(X_1^2 I(|X_1|\le n^{\alpha}))\\
&\quad= K\sum_{n=1}^{\infty} n^{\alpha p-1-2\alpha}\sum_{i=1}^{n} E\Big(X_1^2 I((i-1)^{\alpha}<|X_1|\le i^{\alpha})\Big)\\
&\quad\le K\sum_{i=1}^{\infty}\sum_{n=i}^{\infty} n^{\alpha p-1-2\alpha} E\Big(X_1^2 I((i-1)^{\alpha}<|X_1|\le i^{\alpha})\Big)\\
&\quad\le K\sum_{i=1}^{\infty} i^{\alpha p} P(i^{\alpha}<|X_1|\le (i+1)^{\alpha})\\
&\quad\le KE|X_1|^p<\infty.
\end{aligned}\tag{3.28}$$

Combining (3.26)–(3.28), we obtain (3.25). ∎

3.4. Almost sure convergence. The following theorem is a consequence of Remark 3.9.

THEOREM 3.11. *Let* $0 \le r < 1$, *and let* N *be a positive integer. Let* $1 \le \kappa < 2$, *and let* $\{X_n, n \ge 1\}$ *be a sequence of mean* 0 *strictly stationary random variables satisfying* $\rho^*(X, N) \le r$. *Suppose that* $\{A_{ni}, n \ge 1, 1 \le i \le n\}$ *is an array of random variables such that, for each* $n \ge 1$, *the sequence* $A_n = \{A_{ni}, 1 \le i \le n\}$ *satisfies* $\rho^*(A_n, N) \le r$, *and*

$$\sum_{i=1}^{n} E(|A_{ni}|^q) = O(n) \quad \textit{for some} \quad q > 2\kappa/(2-\kappa). \tag{3.29}$$

If $E|X_1|^{2\kappa} < \infty$, *and* $\{A_{ni}, n \ge 1, 1 \le i \le n\}$ *is independent of* $\{X_i, i \ge 1\}$, *then*

$$\lim_{n\to\infty} \frac{1}{n^{1/\kappa}} \sum_{i=1}^{n} A_{ni}X_i = 0 \quad \textit{a.s.} \tag{3.30}$$

Proof. Let $p = 2\kappa$, $\alpha = 2/p$, and $\nu_n \equiv n$. Then $\alpha > 1/2$, and (3.29) coincides with (3.23). So, by Remark 3.9(i), we get

$$\sum_{n=1}^{\infty} P\Big(\max_{1\le j\le n}\Big|\sum_{i=1}^{j} A_{ni}X_i\Big| > \varepsilon n^{1/\kappa}\Big) < \infty \quad \text{for all} \quad \varepsilon > 0. \tag{3.31}$$

By the Borel–Cantelli lemma, it follows from (3.31) that

$$\lim_{n\to\infty} \frac{1}{n^{1/\kappa}} \sum_{i=1}^{n} A_{ni}X_i = 0 \quad \text{a.s.} \tag{3.32}$$

The proof of the theorem is concluded. ∎

3.5. Remark on stochastic dominance. In the above, we have assumed that the driving noise is strictly stationary. This is adequate for our observer design problems. Nevertheless, more general nonstrictly stationary sequences can be treated, although technical derivations will become more complex. We note that the strict stationarity in Theorems 3.8, 3.10, and 3.11 can be much relaxed. It can be replaced by stochastic dominance. Recall that the sequence $\{X_n, n \geq 1\}$ is stochastically dominated by the random variable X if

$$\sup_{n\geq 1} P\big(|X_n| > t\big) \leq KP\big(|X| > t\big) \text{ for all } t \geq 0. \tag{3.33}$$

The proof of the following result can be carried out similar to that of the previous cases. We state the following result and omit the verbatim argument.

THEOREM 3.12. *In Theorems* 3.8, 3.10, *and* 3.11, *if* $\{X_n, n \geq 1\}$ *being strictly stationarity satisfying* $E|X_1|^p < \infty$ ($E|X_1|^{2\kappa} < \infty$ *in Theorem* 3.11) *is replaced by the assumption that there exists a random variable* X *with* $E|X|^p < \infty$ ($E|X|^{2\kappa} < \infty$ *in Theorem* 3.11) *such that* (3.33) *holds, then the conclusions of Theorems* 3.8, 3.10, *and* 3.11 *continue to hold.*

Remark 3.13. We note the following points.

(i) It follows from Taylor [26, Lemma 5.2.2, p. 123] that stochastic dominance can be accomplished by the sequence of random variables having a bounded absolute rth moment ($r > 0$). Specifically, if

$$\sup_{n\geq 1} E|X_n|^r < \infty \text{ for some } r > 0,$$

then there exists a random variable X with $E|X|^p < \infty$ for all $0 < p < r$ such that (3.33) holds with $K = 1$. (The proviso that $r > 1$ in [26, Lemma 5.2.2] is not needed, as was pointed out in [1]).

(ii) In Theorem 3.8, if we take $\alpha = 1/p$, $\nu_n \equiv n$, then we get

$$\sum_{n=1}^{\infty} n^{-1} P\Big(\max_{1\leq j\leq n}\Big|\sum_{i=1}^{j} A_{ni}X_i\Big| > \varepsilon n^{1/p}\Big) < \infty. \tag{3.34}$$

If $A_{ni} \equiv 1$ (nonweighted), then from (3.34) one can easily get

$$\lim_{n\to\infty} \frac{1}{n^{1/p}} \sum_{i=1}^{n} A_{ni}X_i = 0 \text{ a.s.} \tag{3.35}$$

However, the following example shows that in the weighted case we cannot obtain (3.35) from (3.34).

Example 3.14. Let $1 \leq p < 2$, $\alpha = 1/p$, and $\nu_n \equiv n$, and consider a sequence $\{X_n, n \geq 1\}$ of independent mean 0 random variables such that for all $n \geq 1$

$$P\{X_n = 0\} = 1 - 1/n \quad \text{and} \quad P\{X_n = -n^{1/2}\} = P\{X_n = n^{1/2}\} = 1/(2n). \tag{3.36}$$

Then $\sup_{n\geq 1} E|X_n|^2 = 1 < \infty$. By [26, Lemma 5.2.2], there exists a random variable X with $E|X|^p < \infty$ such that the sequence $\{X_n, n \geq 1\}$ is stochastically dominated by X. Let $\{A_{ni}, n \geq 1, 1 \leq i \leq n\}$ be an array of random variables such that for all $n \geq 1$

$$P\{A_{ni} = 0\} = 1 \quad \text{for all} \quad 1 \leq i < n \quad \text{and} \quad P\{A_{nn} = n^{1/2}\} = 1. \tag{3.37}$$

Then for all $n \geq 1$

$$\sum_{i=1}^{n} E|A_{ni}|^2 = n; \tag{3.38}$$

i.e., (3.17) holds. By Theorem 3.12, (3.34) holds. Now, let $Y_n = X_n/n^{1/2}$, $n \geq 1$. Then $\{Y_n, n \geq 1\}$ is a sequence of independent random variables and

$$\sum_{n=1}^{\infty} P\{|Y_n| \geq 1\} = \sum_{n=1}^{\infty} \frac{1}{n} = \infty. \tag{3.39}$$

By the Borel–Cantelli lemma, it follows from (3.39) that

$$\limsup_{n\to\infty} |Y_n| \geq 1 \text{ a.s.} \tag{3.40}$$

Now, for all $n \geq 1$,

$$n^{-1/p}\Big|\sum_{i=1}^{n} A_{ni}X_i\Big| = n^{-1/p}\Big|A_{nn}X_n\Big| = n^{1-1/p}|Y_n| \text{ a.s.} \tag{3.41}$$

It follows from $1 \leq p < 2$, (3.40), and (3.41) that (3.35) fails.

Remark 3.15. One question of interest is the following: In Theorem 3.8, if we impose further restrictions on the weights, will (3.35) still hold? The answer is positive. By requiring

$$\sup_{2^k \leq n < 2^{k+1},\ 1\leq i\leq n} |A(n;i) - A(2^k;i)| = o(2^{k/p}) \quad \text{a.s.} \tag{3.42}$$

and

$$\sup_{k\geq 0,\ 2^k \leq n < 2^{k+1}} \sum_{i=1}^{n} E|A(n;i) - A(2^k;i)|^2 \leq K, \tag{3.43}$$

we can deduce (3.35). Nevertheless, such a condition is not easily verifiable for our observer problems. To obtain the almost sure convergence in the form of (3.35), we require higher finite absolute moments as in Theorem 3.11, which are verifiable for the observer design problems.

4. Observers and state estimation.

We now return to the scenario of state estimation under random sampling time sequences discussed in section 2. It has been shown in [28] that by applying signal normalization using the modes of the A matrix it is always possible to reduce the state estimation problem to convergence analysis of the expression (2.12), in which the rows of Ψ_n are bounded a.s.

THEOREM 4.1. *Assume the following:*

(a) *The system is observable; i.e., the observability matrix*

$$W_o' = [C', (CA)' \ldots (CA^{m_0-1})']$$

has full rank.

(b) $\frac{1}{n^r}\Psi_n'\Psi_n \geq M > 0$ *a.s. for some* $M > 0$.

In addition, assume that the conditions of Theorem 3.11 *are satisfied with* $A_{ni} = \alpha_j(t_i - t_n)$ *in* (2.14) *and* $X_n = d_n$. *Then*

$$e(t_n) = \left(\frac{1}{n^r}\Psi_n'\Psi_n\right)^{-1}\frac{1}{n^r}\Psi_n' D_n \to 0 \quad a.s. \tag{4.1}$$

Proof. The proof is a direct application of Theorem 3.11. ∎

Remark 4.2. The condition $\frac{1}{n^r}\Psi_n'\Psi_n \geq M > 0$ a.s. is a PE-type condition. However, unlike typical identification problems in which PE conditions are satisfied by using input design, Ψ_n here can be affected only by sampling time sequences. Designing the time sequence to satisfy this condition is covered somewhere else.

One may ask the following question: Under what conditions will the hypotheses in Theorem 3.11 be fulfilled? The following theorem provides a sufficient condition for the conditions in Theorem 3.11 to hold. It states that if $\{d_n\}$ has finite fourth moments and the matrix Ψ_n satisfies certain boundedness conditions, then for all $q \geq 0$ the conditions in Theorem 3.11 are verified.

THEOREM 4.3. *Suppose that conditions* (a) *and* (b) *in Theorem* 4.1 *are satisfied and that, for each* $1 \leq i \leq n$, $\{\varphi(t_i - t_n)\}$ *is uniformly bounded a.s. If* $E|d_1|^4 < \infty$ *and* $\{\varphi(t_1 - t_n), 1 \leq i \leq n\}$ *is independent of* $\{d_i, i \geq 1\}$, *then* $e(t_n)$ *in* (2.9) *satisfies* (4.1).

Proof. Let $\kappa = 1/r$. $1/2 < r < 1$ implies $1 < \kappa < 2$, and $E|d_1|^4 < \infty$ implies $E|d_1|^{2\kappa} < \infty$. Since the rows of Ψ_n are uniformly bounded a.s., the condition (3.29) is satisfied for any q. By Theorem 3.11,

$$\frac{1}{n^r}\Psi_n' D_n \to 0 \quad \text{a.s.},$$

which, together with the condition $\frac{1}{n^r}\Psi_n'\Psi_n \geq M > 0$ a.s., implies that $e(t_n) \to 0$ a.s. ∎

Remark 4.4. If we state the condition in Theorem 4.3 using the matrix Ψ_n, the boundedness can be stated as follows: The rows of Ψ_n are bounded uniformly a.s.

5. Further remarks. This work has been devoted to convergence analysis of state observers with random sampling times. The noise sequence is assumed to be of ρ^*-mixing type, and the sampling times are random. A central theme here is to show that the state estimation error sequence goes to 0 a.s. As a consequence, the convergence in observer design problems under randomized sampling times is obtained.

Our approach is based on establishing a novel result of Marcinkiewicz–Zygmund-type strong laws of large numbers. The essence is to treat double-indexed and randomly weighted sums of mixing variables. We established the probability one convergence of the aforementioned sequences together with results on rates of convergence. The results obtained are of interest in their own right and can be applied to many other applications that involve convergence analysis of such sequences.

In this work, we assume that the random weights and noises are independent. When the observers are used in feedback control or adaptation, the weights will become correlated with the noises, rendering a more challenging problem. This remains an open problem.

For the observer design problem, the original system in continuous time is a purely deterministic one. A related problem, important but much more difficult, is when the system is a diffusion with a Brownian motion included. Binary-valued observations introduce further challenges. How to solve such problems deserves more thoughts and further considerations.

REFERENCES

[1] A. Adler, A. Rosalsky, and R. L. Taylor, *Some strong laws of large numbers for sums of random elements*, Bull. Inst. Math. Acad. Sinica, 20 (1992), pp. 335–357.

[2] J. An and D. Yuan, *Complete convergence of weighted sums for ρ^*-mixing sequence of random variables*, Probab. Statist. Lett., 78 (2008), pp. 1466–1472.

[3] K. J. Åström and B. Bernhardsson, *Comparison of periodic and event based sampling for first-order stochastic systems*, in Proceedings of the 14th IFAC World Congress, Beijing, 1999, pp. 301–306.

[4] L. E. Baum and M. Katz, *Convergence rates in the law of large numbers*, Trans. Amer. Math. Soc., 120 (1965), pp. 108–123.

[5] P. Billingsley, *Convergence of Probability Measures*, John Wiley, New York, 1968.

[6] R. C. Bradley, *On the spectral density and asymptotic normality of weakly dependent random fields*, J. Theoret. Probab., 5 (1992), pp. 355–373.

[7] R. C. Bradley, *Every lower psi-mixing Markov chain is "interlaced rho-mixing,"* Stochastic Process Appl., 72 (1997), pp. 221–239.

[8] R. C. Bradley, *A stationary rho-mixing Markov chain which is not "interlaced" rho-mixing*, J. Theoret. Probab., 14 (2001), pp. 717–727.

[9] R. C. Bradley, *Basic properties of strong mixing conditions. A survey and some open questions*, Probab. Surv., 2 (2005), pp. 107–144.

[10] W. Bryc and W. Smolenski, *Moment conditions for almost sure convergence of weakly correlated random variables*, Proc. Amer. Math. Soc., 119 (2001), pp. 629–635.

[11] H. F. Chen and L. Guo, *Identification and Stochastic Adaptive Control*, Birkhäuser, Boston, 1991.

[12] P. Erdös, *On a theorem of Hsu and Robbins*, Ann. Math. Statistics, 20 (1949), pp. 286–291.

[13] P. Hall and C. C. Heyde, *Martingale Limit Theory and Its Application*, Academic Press, New York, 1980.

[14] P. L. Hsu and H. Robbins, *Complete convergence and the law of large numbers*, Proc. Nat. Acad. Sci. U.S.A., 33 (1947), pp. 25–31.

[15] D. W. Huang and L. Guo, *Estimation of nonstationary ARMAX models based on Hannan-Rissanen method*, Ann. Statist., 18 (1990), pp. 1729–1756.

[16] A. N. Kolmogorov and Y. A. Rozanov, *On strong mixing conditions for stationary Gaussian processes*, Theory Probab. Appl., 5 (1960), pp. 204–208.

[17] H. J. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, 2nd ed., Springer-Verlag, New York, 2003.

[18] L. Ljung, *System Identification: Theory for the User*, Prentice–Hall, Englewood Cliffs, NJ, 1987.

[19] M. Milanese and A. Vicino, *Information-based complexity and nonparametric worst-case system identification*, J. Complexity, 9 (1993), pp. 427–446.

[20] K. Ogata, *Modern Control Engineering*, 4th ed., Prentice–Hall, Englewood Cliffs, NJ, 2002.

[21] M. Peligrad and A. Gut, *Almost sure results for a class of dependent random variables*, J. Theoret. Probab., 12 (1999), pp. 87–104.

[22] N. Persson and F. Gustafsson, *Event based sampling with application to vibration analysis in pneumatic tire*, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2001, pp. 3885–3888.

[23] W. F. Stout, *Almost Sure Convergence*, Academic Press, New York, 1974.

[24] R. L. Stratonovich, *Application of the Markov processes theory to optimal filtering*, Radio Engrg. Electron. Phys., 5 (1960), pp. 1–19.

[25] J. Sur and B. E. Paden, *State observer for linear time-invariant systems with quantized output*, Trans. ASME Ser. G J. Dynam. Systems Measurement and Control, 120 (1998), pp. 423–426.

[26] R. L. Taylor, *Stochastic Convergence of Weighted Sums of Random Elements in Linear Spaces*, Lecture Notes in Math. 672, Springer-Verlag, Berlin, 1978.

[27] S. Utev and M. Peligrad, *Maximal inequalities and an invariance principle for a class of weakly dependent random variables*, J. Theoret. Probab., 16 (2003), pp. 101–115.

[28] L. Y. Wang, C. Li, G. Yin, L. Guo, and C. Z. Xu, *State Observers of Linear-Time-Invariant Systems under Irregular-Sampling and Sensor Limitations*, preprint, 2009.

[29] L. Y. Wang, G. H. Xu, and G. Yin, *State reconstruction for linear time-invariant systems*

with binary-valued output observations, Systems Control Lett., 57 (2008), pp. 958–963.
[30] L. Y. Wang and G. Yin, *Asymptotically efficient parameter estimation using quantized output observations*, Automatica, 43 (2007), pp. 1178–1191.
[31] L. Y. Wang, G. Yin, J. F. Zhang, and Y. L. Zhao, *System Identification with Quantized Observations*, Birkhäuser, Boston, 2010.
[32] L. Y. Wang, J. F. Zhang, and G. Yin, *System identification using binary sensors*, IEEE Trans. Automat. Control, 48 (2003), pp. 1892–1907.