Stochastics and Stochastics Reports, Vol. 51, pp. 241–268
Reprints available directly from the publisher
Photocopying permitted by license only

INDUCTIVE METHODS AND RATES OF R$^{\text{th}}$-MEAN CONVERGENCE IN ADAPTIVE FILTERING

M. A. KOURITZIN*

Department of Mathematics and Statistics, Carleton University, Ottawa, Ontario, Canada K1S 5B6

(Received 3 June 1993; in final form 28 April 1994)

An inductive method for establishing rates of r^{th}-mean convergence of linear, decreasing-gain adaptive algorithms is introduced. This inductive method does not require any sophisticated probability theory or mathematics and allows such weak forms of dependence as weak-multiplicativity, strong mixing and decaying covariance. In exposing our method, we establish rates of r^{th}-mean convergence for the process $\{h_k,\ k=1,2,3,\ldots\}$, generated by the basic linear adaptive filtering algorithm $h_{k+1}=h_k+\mu_k(b_k-A_kh_k)$ with $\{b_k,\ k=1,2,3,\ldots\}$ and $\{A_k,\ k=1,2,3,\ldots\}$ being weakly-stationary stochastic processes which drive the algorithm and $\{\mu_k\}_{k=0}^{\infty}$ being a sequence of the form $k^{-\alpha}$ for some $0<\alpha\le 1$. Specifically, we show in most cases that $E\|h_k-h_*\|^r=O(\mu_k^{r/2})$, where r is a real constant greater than or equal to one and $h_*\triangleq[EA_1]^{-1}Eb_1$. Finally, our method takes advantage of moment bounds for partial sums of non-stationary random sequences, some of which are established within this note.

KEY WORDS: Adaptive filtering, rates of r^{th}-mean convergence, moment bounds

1 INTRODUCTION

Adaptive algorithms have been alloted an important role in parameter estimation applications pertaining to system identification, adaptive control, transmission systems, signal processing and pattern recognition. Indeed, due to their inherent simplicity, linear (in h_k) decreasing-gain adaptive filtering algorithms of the form

$$h_{k+1}=h_k+\mu_k(b_k-A_kh_k)\qquad k=1,2,3,\ldots,\tag{1}$$

for some weakly-stationary, positive semi-definite matrix-valued stochastic process $\{A_k,\ k=1,2,3,\ldots\}$ and some weakly-stationary, vector-valued stochastic process $\{b_k,\ k=1,2,3,\ldots\}$, are frequently used to estimate the non-random vector

$$h_*\overset{\Delta}{=}[EA_1]^{-1}Eb_1\tag{2}$$

(assuming EA_1 is invertible), when the statistics of $\{A_k,\ k=1,2,3,\ldots\}$ and $\{b_k,\ k=1,2,3,\ldots\}$ are unknown. In equation (1), $\{\mu_k\}_{k=1}^{\infty}$ is a sequence of real

---

* This work was done while the author was on an NSERC sponsored visit to the Institut für Mathematische Stochastik at the Universität Freiburg. The author is grateful to both NSERC and the aforementioned Institut for their support.

numbers tending to zero as $k \to \infty$ and for some fixed positive integer M and constant $\mu > 0$, $\{A_k,\ k = 1, 2, 3, \ldots\}$ and $\{b_k,\ k = 1, 2, 3, \ldots\}$ are usually defined by

$$A_k \stackrel{\Delta}{=} \frac{\mu}{M} \sum_{l=0}^{M-1} X_{k-l} X_{k-l}^T, \quad b_k \stackrel{\Delta}{=} \frac{\mu}{M} \sum_{l=0}^{M-1} X_{k-l} \alpha_{k-l+1}, \quad k = 1, 2, 3, \ldots, \tag{3}$$

where $\{X_j, -\infty < j < \infty\}$ and $\{\alpha_j, -\infty < j < \infty\}$ are given (usually highly-dependent) weakly-stationary, second-order stochastic processes. Under these conditions, h_* minimizes (with respect to non-random vector g) the value of

$$E(\alpha_{k+1} - X_k^T g)^2 \quad \text{for each } k = 1, 2, 3, \ldots \tag{4}$$

It is often desirable that the sequence of stochastic estimates given by a (not necessarily linear) decreasing-gain adaptive algorithm converge in the r^{th}-mean sense to the best set of parameters at a reasonable rate. A recent paper by Gerencsér [5] establishes rates of r^{th}-mean convergence for a fairly wide class of algorithms and contains a summary of the previous work in this area. One characteristic of this work is the use of a "resetting" mechanism to ensure that the trajectory of parameter estimates stays in a compact domain. On the other hand, Watanabe [12] proved that rates of r^{th}-mean convergence can be established without the use of "resetting" for the linear algorithm described by (1) and (3) (with $M = 1$ and $\mu_k = 1/k$).

The methods used by Watanabe have some very nice characteristics. He does not require the use of a resetting mechanism, maximal inequalities, unverifiable conditions or sophisticated probability theory. The essence of his proof is: (i) Establish that $E^{1/r}\|h_k - h_*\|^r$ does not grow too fast in k (step 2 on page 137 of [12]); (ii) use i to determine that $E^{1/r}\|h_k - h_*\|^r$ in fact converges to zero (step 3 on page 137 of [12]); and (iii) use ii to deduce a rate of convergence for $E^{1/r}\|h_k - h_*\|^r$ (see (54) in the proof of Theorem 3 of [12]). Motivated to some degree by Watanabe's work, we introduce methods in this paper where we first establish that $E^{1/r}\|h_k - h_*\|^r$ is bounded or not diverging too fast in k and then we inductively refine the bound on $E^{1/r}\|h_k - h_*\|^r$ to eventually obtain the desired rate of convergence. Some key features of our method are: All refinements are done in exactly the same manner so a very clean induction can be used; no specific type of mixing or dependence assumption but rather only a general ergodic type theorem is required; appropriate subsequences are carefully selected to simplify calculations and to make proper use of the stability imposed by the positive definite assumption on EA_1; and an auxiliary process $\{u_k,\ k = 1, 2, 3, \ldots\}$ is defined to reduce the analysis of (1) to two inherently simpler algorithms.

Although the methods of Watanabe were motivational, his actual results have some drawbacks: He assumes rather specific dependence conditions (ϕ- and ψ-mixing); he attains rates only when $\mu_k = k^{-1}$; he proposes rates that are very conservative when r is large; he allows r only to be even integer rather than real valued; and he utilizes somewhat excessive moment conditions (in his ϕ-mixing case). In introducing our methods in this note, we extend Watanabe's work by addressing the above concerns. It is expected that our results will allow Heunis's strong invariance principle [6, Proposition 2.1 (i)] (which relies on Watanabe's results) to be restated under a more general set of conditions. However, it appears that to fully integrate our work into Heunis's strong invariance principle would require some work and has not been attempted.

An interesting aspect of Watanabe's work is that the rate of convergence he obtains is highly dependent on the minimum eigenvalue of EA_1. This is likely unavoidable since such reliance has been independently verified by simulations and unfortunate since adaptive filtering algorithms would normally only be used when EA_1 is unknown. As with Watanabe's results, the rates of convergence obtained in this note, when $\mu_k = k^{-1}$, are dependent on the minimum eigenvalue of EA_1. Indeed, in this case we show that there exists a constant $c_\eta > 0$ independent of k such that

$$E^{1/r}\|h_k - h_*\|^r \le c_\eta k^{-\eta} \qquad \text{for} \quad k = 1, 2, 3, \ldots, \tag{5}$$

where η is any constant such that $\eta \le \frac{1}{2}$ and η is strictly less than minimum eigenvalue of EA_1. On the other hand, our results yield rates of convergence, when $\mu_k = k^{-\alpha}$ for some $0 < \alpha < 1$, whose dependence on the statistics of $\{A_k,\ k = 1, 2, 3, \ldots\}$ can be only through multiplicative constants c_α^A, i.e. we obtain

$$E^{1/r}\|h_k^\alpha - h_*\|^r \le c_\alpha^A \mu_k^{1/2} \quad \text{for all } k = 1, 2, 3, \ldots, \tag{6}$$

regardless of the minimum eigenvalue of EA_1, where $\{h_k^\alpha, k = 1, 2, 3, \ldots\}$ is defined by (1) when $\mu_k = k^{-\alpha}$.

In Section 2 we state our assumptions, give our main results and illustrate our dependence condition through four examples. Contemplating typical problems in signal processing, we have striven for general dependence conditions at the expense of stringent moment conditions (almost as stringent as those of Watanabe). In Section 3 we use our inductive methods to prove the main results. Section 4 contains subsiduary results which are important to the proof of the main results. Finally, Section 5 contains the development of moment bounds which are critical to the examples in Section 2 and our claim that our results hold under fairly general dependence conditions. As pointed out by a reviewer, our strong mixing bounds, Corollary 11 of Section 5, might be of independent interest because they provide alternatives to Yokoyama [14, Theorems 1 and 2] when the random sequence is non-stationary or when explicit constants are required.

2 MAIN RESULT

Suppose $(\Omega, \mathscr{F}, P)$ is a probability space, d is a fixed positive integer and $0 < \alpha \le 1$ is a constant (which will later be used to define $\mu_k \triangleq k^{-\alpha}$ as in the introduction). Further, let $\|\cdot\|$ denote the norm of Euclidean distance on R^d and

$$|||A||| \overset{0}{=} \sup_{x \in R^d} \frac{\| Ax \|}{\| x \|}$$

the corresponding matrix norm for any d by d matrix A. Finally, assume $\{A_k, k = 1,2,3,\ldots,\}$ and $\{b_k, k = 1,2,3,\ldots,\}$ are respectively a $R^{d\times d}$-valued and a R^d-valued stochastic process on $(\Omega, \mathscr{F}, P)$ satisfying the following conditions (which are discussed in Remark 1 below):

(1) For each $k = 1,2,3,\ldots$, the distribution of A_k does not depend on k and $A_* \triangleq EA_1$ is well defined. The expectation of b_k, Eb_k, is well defined and equal to a constant vector b_* for each $k = 1,2,3,\ldots$.

(2) A_* and $A_k(\omega)$ are symmetric, A_* is positive definite and $A_k(\omega)$ is positive semi-definite for each k and ω.

In preparation for the statement of the final two conditions, suppose $0 < \sigma < 1$ and $\delta > 2$ are real constants satisfying

$$\frac{1}{\sigma} + \max\left\{\left\lceil\frac{2(1-\alpha)}{\alpha}\right\rceil, \left\lfloor\frac{4}{3\alpha}\right\rfloor\right\} < \delta, \tag{7}$$

where $\lceil x\rceil \triangleq \min\{i \in \mathbb{N}_0 : i \geq x\}$ and $\lfloor x\rfloor \triangleq \max\{i \in \mathbb{N}_0 : i \leq x\}$ for any $x \geq 0$. Admittedly, (7) is seemingly complicated. However, it is unavoidable (see the statement of Proposition 2 and Remark 3 which follow) and it reduces to the trivial expression $1/\sigma + 1 < \delta$ in the case of most practical interest, where $\alpha > \frac{2}{3}$. Regardless of the value of α, one finds that if σ is chosen close to one then (7) allows smaller values of δ and thereby, through Condition 3, a less stringent dependency condition. On the other hand, such a choice of σ implies a stronger moment assumption in Condition 4. A larger value of δ (for fixed σ and α) imposes more stringent dependency and moment conditions but provides a stronger result by allowing r to be larger in Proposition 2 below.

(3) There exists some constant $c_\delta > 0$ such that for any $m, n \in \{1,2,\ldots,d\}$, integers $k \geq j \geq 1$ and real constants $\{f_\ell\}_{\ell=j}^k$ we have

$$(i)\ E\left|\sum_{\ell=j}^{k}\left(b_\ell^{(m)} - b_*^{(m)}\right)f_\ell\right|^\delta \leq c_\delta\left(\sum_{\ell=j}^{k} f_\ell^2\right)^{\delta/2},$$

$$(ii)\ E\left|\sum_{\ell=j}^{k}\left(a_\ell^{(m,n)} - a_*^{(m,n)}\right)f_\ell\right|^\delta \leq c_\delta\left(\sum_{\ell=j}^{k} f_\ell^2\right)^{\delta/2},$$

where $b_\ell^{(m)}, b_*^{(m)}, a_\ell^{(m,n)}$ and $a_*^{(m,n)}$ denote particular components of b_ℓ, b_*, A_ℓ and A_* respectively.

(4) It is assumed that

$$\inf_{\epsilon>0} E\exp\left\{\frac{\delta\sigma}{1-\sigma}\epsilon^{-1}\|A_1\|^{(1+\epsilon)/\alpha}\right\} < \infty.$$

Remark 1 For certain applications it would be appealing to replace Condition 1 (and its counterpart 1' to follow) with some form of asymptotic stationarity.

Although the author believes that this may be possible, he wishes to keep the methods introduced in this note as clear as possible. Hence, no attempt to weaken Condition 1 or Condition 1' has been made. 2. Condition 2 imposes a degree of stability on Equation (1) and, thereby, helps ensure that the solution of (1) converges. Everything in this condition other than the strictly positive assumption on A_* can be trivially verified without knowledge of any distributions in all applications that the author is familiar with. However, for the proofs in the sequel, we will find it more convenient to utilize the bounds

$$\|I-\beta A_k\| \le \max\{1,\beta\|A_k\|\} \quad \text{for all } \beta>0,\ k=1,2,3,\dots \text{ and } \omega\in\Omega, \tag{8}$$

and

$$\|I-\beta A_*\| \le \max\{1-\beta\lambda_{\min},\beta\lambda_{\max}-1\} \quad \text{for all } \beta>0\,, \tag{9}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ represent the minimum and maximum eigenvalues of A_*. The bounds in (8) and (9) follow from Condition 2, the definition of $\|\cdot\|$ and standard theory for symmetric matrices (see e.g. [13, pp. 57-8]). 3. Condition 3 is our only dependence/independence condition and will be shown below to be a mild condition including several common dependence type conditions e.g. strong mixing, L-mixing and weak multiplicativity. This condition implies that there is a real constant c'_δ (which may depend on d as well as δ) such that for all integers $1\le j\le k$ and (non-random) d by d matrices $\{F_l\}_{l=j}^k$ one has that

$$(i)\quad E\left\{\left\|\sum_{l=j}^k F_l(b_l-b_*)\right\|^\delta\right\} \le c'_\delta\left(\sum_{l=j}^k \|F_l\|^2\right)^{\delta/2}, \tag{10}$$

$$(ii)\quad E\left\{\left\|\sum_{l=j}^k F_l(A_l-A_*)\right\|^\delta\right\} \le c'_\delta\left(\sum_{l=j}^k \|F_l\|^2\right)^{\delta/2}, \tag{11}$$

and, as an obvious consequence of (i), that

$$(iii)\quad \sup_k E\{\|b_k\|^\delta\} < \infty\,. \tag{12}$$

Condition 3 will only be utilized through (10–12). Still, the current form of Condition 3 was chosen because it appears to be easier to motivate through examples (see below). 4. Condition 4 is a fairly strong moment condition but is still weaker than the corresponding condition in Watanabe: Suppose corresponding to Assumption A2 a) in Watanabe [12] that we have $E\|A_1\|^{2k} \le \gamma^k k!$ for some $\gamma>0$ and all $k=1,2,3,\dots$ Then, for $\alpha>\frac{1}{2}$ and $\epsilon<2\alpha-1$ one finds

$$E\exp\left(\frac{\delta\sigma}{1-\sigma}\epsilon^{-1}\|A_1\|^{(1+\epsilon)/\alpha}\right) \le 1+\sum_{k=1}^\infty \frac{\left(\epsilon^{-1}\delta\sigma/(1-\sigma)\right)^k\left(\gamma^k k!\right)^{(1+\epsilon)/2\alpha}}{k!} < \infty \tag{13}$$

using Jensen's inequality and Stirling's Formula. Condition 4 follows for $\frac{1}{2} < \alpha \le 1$.

Now in preparation for the statement of the main results, let us define (labeling for easy reference throughout this note): **(A)** $\mu_k \stackrel{\varrho}{=} k^{-\alpha}$ for $k = 1, 2, \ldots$; (B) the random sequence $\{h_k, k = 1, 2, 3, \ldots\}$ is defined by $h_{k+1} = h_k + \mu_k(b_k - A_k h_k)$ for $k = 1, 2, 3, \ldots$ and h_1 is a non-random constant; **(C)** $h_* \stackrel{\varrho}{=} A_*^{-1} b_*$; **(D)** $\eta > 0$ is any constant such that $\eta \le \frac{1}{2}$ and $\eta < \lambda_{\min}$ (with $\lambda_{\min}$ as in (9)); (E) for each $k = 1, 2, 3, \ldots$, $\gamma_k \triangleq k^{-\eta}$ when $\alpha = 1$ and $\gamma_k \triangleq \mu_k^{1/2}$ when $0 < \alpha < 1$; **(F)** $|\xi|_\nu \stackrel{\varrho}{=} E^{1/\nu}|\xi|^\nu$, $\| z \|_\nu \stackrel{\varrho}{=} E^{1/\nu} \| z \|^\nu$ and $\| Y \|_\nu \stackrel{\varrho}{=} E^{1/\nu} \| Y \|^\nu$ for any $\nu \ge 1$, $\mathscr{R}$-valued random variable ξ, $\mathscr{R}^d$-valued random element z and any $\mathscr{R}^{d \times d}$-valued random element Y.

The following proposition represents our first main result and is proven in Section 3.

Proposition 2 *Suppose Conditions* (1–4) are satisfied with the given constants σ and δ. Moreover, suppose $r \ge 1$ is a real constant such that

$$\frac{1}{\sigma} + \max\left\{ \left\lceil \frac{2(1-\alpha)}{\alpha} \right\rceil, \left\lfloor \frac{4}{3\alpha} \right\rfloor \right\} < \frac{\delta}{r}. \tag{14}$$

Then, there exists a constant $c > 0$ such that

$$\| h_k - h_* \|_r \le c\gamma_k \quad \text{for } k = 1, 2, 3, \ldots$$

Remark 3 (14) is obviously true for r close enough to 1 by (7) and is necessary in order that we can find a real constant $\Theta > 0$ such that

$$\frac{1}{\sigma} = \frac{\delta}{r}\left(\frac{\Theta}{\Theta+1}\right)^{M_\alpha} - \sum_{j=1}^{M_\alpha}\left(\frac{\Theta}{\Theta+1}\right)^j, \tag{15}$$

where $M_\alpha \triangleq \max\{\lceil 2(1-\alpha)/\alpha \rceil, \lfloor 4/(3\alpha) \rfloor\}$. Observing that the expression on the right of (15) is zero when $\Theta = 0$, is continuous in Θ, and heads to $(\delta/r) - M_\alpha$ as $\Theta \to \infty$, we conclude by (14) that such a $\Theta > 0$ exists. Moreover, simple manipulation shows that such a Θ solves

$$\frac{1}{\sigma} = \left(\frac{\Theta}{\Theta+1}\right)^{M_\alpha}\left(\frac{\delta}{r} + \Theta\right) - \Theta, \tag{16}$$

which will be essential to the proof of Proposition 2.

Comparing Proposition 2 (with $\alpha = 1$) with Theorem 3 of Watanabe [12], we note that among other things, our result yields faster convergence (when r is not equal to two) and holds under more general dependence conditions and less stringent moment bounds. Still, our conditions can be characterized by strong moment and non-stringent dependency conditions (with some trade-off between the two through the choice of σ) which is indeed the situation for which the methods introduced in this note were designed to handle. With this in mind, we elect not to impose some more discriminating dependence assumption and try to reduce the moment conditions as Watanabe did but rather to consider next the practical situation where $\{A_k,\ k=1,2,3,...\}$ is essentially bounded but very little is known about the applicability of standard dependency conditions to $\{A_k,\ k=1,2,3,...\}$ and $\{b_k,\ k=1,2,3,...\}$. The compensation for insisting that $\{A_k,\ k=1,2,3,...\}$ is essentially bounded will turn out to be the applicability of our results under only decaying covariance assumptions, a weakening of other dependence assumptions and a relaxation of the stationarity assumption on $\{A_k,\ k=1,2,3,...\}$.

In preparation for our second and last main result we state alternative conditions to Conditions 1–4 which will be imposed for the proof of Proposition 4 in Section 3.

(1') Eb_k and EA_k exist with $Eb_k = b_*$ and $EA_k = A_*$ for all $k=1,2,3,...$

(2') We assume Condition 2.

For the statement of Conditions 3', we require a constant $\delta \geq 2$ such that

$$1+\max\left\{\left\lceil\frac{2(1-\alpha)}{\alpha}\right\rceil, \left\lfloor\frac{4}{3\alpha}\right\rfloor\right\} \leq \delta\,. \tag{17}$$

(3') We assume Condition 3 but now allow $\delta = 2$ and only insist that it and α satisfy (17).

(4') It is assumed that there exists a $c' > 0$ such that

$$\operatorname*{ess\,sup}_{\omega} \quad \|A_k(\omega)\| \leq c' \quad \text{for all } \ k=1,2,3,... \tag{18}$$

Clearly Conditions 1' and 3' are weaker than 1 and 3; while 4' is stronger than 4. We now state the second and final main result.

Proposition 4 *Suppose Conditions* (1'–4') *are satisfied with the given constant* δ. *Moreover, suppose* $r \geq 1$ *is the real constant such that*

$$1+\max\left\{\left\lceil\frac{2(1-\alpha)}{\alpha}\right\rceil, \left\lfloor\frac{4}{3\alpha}\right\rfloor\right\} = \frac{\delta}{r}\,. \tag{19}$$

Then there exists a constant $c > 0$ *such that*

$$\| h_k - h_* \|_r \leq c\gamma_k \quad \textit{for } k=1,2,3,\ldots$$

The following examples illustrate the generality of Condition 3 and 3'. Since Conditions 3 and 3' are stated in a component by component basis and Conditions 4 and 4' imply similar bounds for each component of $\{A_l,\ l=1,2,3,...\}$, there is no loss of generality in assuming $d=1$ in the following examples. Moreover, there is no loss of generality in assuming $\{A_l,\ l=1,2,3,...\}$ is zero mean. To avoid having two different definitions for $\{A_l,\ l=1,2,3,...\}$ (one inside the examples and one out), we let $\{\tilde{a}_l,\ l=1,2,3,...\}$ denote the one-dimensional, zero-mean version of $\{A_l,\ l=1,2,3,...\}$.

Example 1 As mentioned in Eweda and Macchi [3, p. 121], decaying-covariance assumptions suit data-transmission problems in which linear, adaptive filtering algorithms might be used. Suppose that $\{\tilde{a}_l,\ l = 1, 2, 3, \ldots\}$ satisfies the following decaying covariance assumption

$$\sum_{\nu=0}^{\infty} \sup_{l\in\{1,2,3,\ldots\}} |E\{\tilde{a}_{l+\nu}\tilde{a}_l\}| < \infty,$$

then it follows by Lemma 9 of Section 5 that there is a $c > 0$ such that

$$E\left|\sum_{l=j}^{k} \tilde{a}_l f_l\right|^2 \le c \sum_{l=j}^{k} f_l^2$$

for all integers $1 \le j \le k$ and constants $\{f_l\}_{l=j}^{k}$. Hence if a similar bound is established for $\{b_l,\ l = 1, 2, 3, \ldots\}$ (possibly through a similar decaying-convariance condition) and Conditions 1', 2' and 4' are also satisfied, then we get rates of L_1-convergence from Proposition 4 i.e.

$$E\|h_k - h_*\| \le c\,\gamma_k \qquad \text{for all } k = 1, 2, 3, \ldots,$$

provided $\alpha > \frac{2}{3}$.

Example 2 Strong mixing conditions are widely used and are satisfied by a fairly substantial class of processes including a variety of ARMA processes (see Mokkadem [9]). Suppose the process $\{\tilde{a}_l, l = 1, 2, 3, \ldots\}$ satisfies Conditions 1 and 4 as well as the following strong mixing condition: There exists a monotonically non-increasing, $\Re$-valued function $\alpha_a(\cdot)$ and a real constant $z > 1$, such that

$$\sup_{k\ge 1} \sup_{\substack{E\in\sigma\{\tilde{a}_j,\, j=k+l,k+l+1,\ldots\}\\ D\in\sigma\{\tilde{a}_j,\, j=1,2,\ldots,k\}}} |P(D\cap E) - P(D)P(E)| \le \alpha_a(l), \tag{20}$$

for $l = 0, 1, 2, \ldots$ and

$$\sum_{l=0}^{\infty} [\alpha_a(l)]^{2/\delta z} < \infty. \tag{21}$$

Then, it follows by Corollary 11 (i) of Section 5 that there exists a constant $c > 0$ such that for any sequence of real numbers $\{f_\ell\}_{\ell=0}^{\infty}$

$$E\left|\sum_{\ell=j}^{k} \tilde{a}_\ell f_\ell\right|^\delta \le c\left(\sum_{\ell=j}^{k} f_\ell^2\right)^{\delta/2} \tag{22}$$

for all $1 \le j \le k$. On the other hand, if Condition 4' is satisfied then one no longer requires Condition 1 to be satisfied and can relax (21) to $\sum_{l=0}^{\infty} [\alpha_a(l)]^{\frac{2}{\delta}} < \infty$ and still achieve (22) by using Corollary 11 (ii).

Example 3 In this example we consider the "stably-generated" processes adopted by Davis and Vinter [1, Definition 5.1.1] and the "L-mixing" processes of Gerencsér [4]. With this in mind, we assume that $\{\tilde{a}_l,\ l=1,2,3,\ldots\}$ satisfies Condition 4 and there exists a family of σ-algebra pairs $\{(\mathscr{F}_l,\mathscr{F}_l^+),\ l\geq 1\}$ such that $\mathscr{F}_n\subset\mathscr{F}_l\subset\mathscr{F}$ and $\mathscr{F}_l^+\subset\mathscr{F}_n^+\subset\mathscr{F}$ for all integers $1\leq n\leq l$, $\mathscr{F}_l$ is independent of $\mathscr{F}_l^+$ for all integers $l\geq 1$, $\tilde{a}_l$ is $\mathscr{F}_l$-measurable for all integers $l\geq 1$ and for δ as in Condition 3

$$\sum_{l=0}^{\infty}\sup_{n>l}E^{1/\delta}|\tilde{a}_n-E(\tilde{a}_n|\mathscr{F}_{n-l}^+)|^\delta<\infty\,. \tag{23}$$

Both types of processes described above are variations on the "exponentially stable" mixing processes introduced by Ljung (see S3 on page 772 of Ljung [7]) and have been found useful in studying inference in control theory. From Proposition 13 one sees that Condition 3 (ii) holds. Moreover, under Condition 4', one only requires (23) to hold for a δ as in Condition 3' in order for Condition 3'(ii) to hold.

Remark 5 To see that Davis and Vinter's stably-generated processes satisfy (23), suppose that $\{\tilde{a}_l,\ l=1,2,3,\ldots\}$ satisfies Condition 4 and that $\{u_l,\ -\infty<l<\infty\}$ is an independent sequence of random variables such that $\tilde{a}_l$ is $\sigma\{u_m,\ -\infty<m\leq l\}$-measurable for all $l=1,2,3,\ldots$. Moreover, suppose there exist constants $c>0$ and $\lambda\in(0,1)$ such that for all integers l, n with $l\leq n+1$, $l\geq 1$ there is a $\sigma\{u_{n-l+1},\ldots,u_n\}$-measurable random variable $a_n[l]$ satisfying

$$E|\tilde{a}_n-a_n[l]\,|^{2q}\leq c\,\lambda^l\,.$$

Then, it follows by Jensen's inequality for conditional expectations that

$$\begin{aligned}&E|\tilde{a}_n-E(\tilde{a}_n|\sigma\{u_m,\ n-l<m<\infty\})|^{2q}\\&\quad\leq E|\tilde{a}_n-E(\tilde{a}_n|\sigma\{u_{n-l+1},\ldots,u_n\})|^{2q}\leq 2^{2q}\,c\,\lambda^l\end{aligned}$$

for all integers $l\leq n+1$, $l\geq 1$ and $\{\tilde{a}_m,\ m=0,1,2,\ldots\}$ will also satisfy (23) with $\mathscr{F}_l\triangleq\sigma\{u_m,\ -\infty<m\leq l\}$ and $\mathscr{F}_l^+\triangleq\sigma\{u_m,\ l<m<\infty\}$.

Example 4 In [8], Longnecker and Serfling define a general class of dependent processes, the so-called "weak multiplicative" processes, which satisfy moderate moment bounds and quasi-orthogonality dependence restrictions. Rather than repeat the definitions here we refer the reader to Section 2 of [8] where some half-dozen precise formulations of this concept are given. From Section 4 of [8], it is seen that if $\{\tilde{a}_l,\ l=0,1,2,\ldots\}$ is an appropriate weak-multiplicative process then Condition 3 (ii) or Condition 3'(ii) will hold with δ taking the value of some positive even integer.

3 PROOF OF PROPOSITIONS 2 AND 4

The proofs of Propositions 2 and 4 require a substantial number of definitions and the author feels that it is beneficial to the reader to put the majority of these definitions in one place and then refer to them as required throughout the following proofs. We label them for easy reference: (G) Let $\{n_i\}_{i=0}^{\infty}$ be the sequence of integers defined for each $i = 0, 1, 2, \ldots$ by $n_i \triangleq \left\lfloor (i(1-\alpha))^{\frac{1}{1-\alpha}} \right\rfloor$ if $0 < \alpha < 1$ or $n_i \triangleq \lfloor \exp(i-1) \rfloor$ if $\alpha = 1$. (H) For each $i = 0, 1, 2, \ldots$, let $I_i \triangleq \{n_i + 1, n_i + 2, \ldots, n_{i+1}\}$, some of which may be empty when $\alpha < 1$. (I) s_k denotes the unique index such that $k - 1 \in I_{s_k}$ for $k = 2, 3, 4, \ldots$ (J) For each pair of integers $1 \le j < k$, $F_{j,k} \triangleq \mu_j I$ if $j = k-1$ or $F_{j,k} \triangleq (I - \mu_{k-1} A_*)(I - \mu_{k-2} A_*) \cdots (I - \mu_{j+1} A_*) \mu_j$ if $j < k-1$. (K) Suppose $\lambda_{\max}$ and $\lambda_{\min}$ are the largest and smallest eigenvalues of A_*. Then

$$t \triangleq \left\lceil \left(\frac{\lambda_{\max} + \lambda_{\min}}{2} \right)^{\frac{1}{\alpha}} \right\rceil, \quad c'' \triangleq \exp\left\{ \sum_{l=1}^{t-1} \mu_l \lambda_{\min} \right\} \prod_{l=1}^{\lfloor (\lambda_{\max}/2)^{1/\alpha} \rfloor} (\mu_l \lambda_{\max} - 1).$$

(L) For each $k = 1, 2, 3, \ldots$, we define $\bar{A}_k \triangleq A_k - A_*$ and $\bar{b}_k \triangleq b_k - b_*$.

Remark 6 It follows by (G) as well as (A) of Section 2 that if l, m are integers such that $0 \le l < m$ and $n_l \ge 1$ then

$$m - l - 1 \le \int_{n_l}^{n_m+1} t^{-\alpha}\, dt - \mu_{n_l} \le \sum_{j=n_l+1}^{n_m} \mu_j$$
$$\le \int_{n_l+1}^{n_m} t^{-\alpha}\, dt + \mu_{n_l+1} \le m - l + 1 \tag{24}$$

provided $n_l \ge 1$. Moreover, in the case $n_l = 0$, we can still make use of the two right-most inequalities so we turn our attention on obtaining a lower bound on $\sum_{j=n_l+1}^{n_m} \mu_j$ when $n_l = 0$. This last case is handled easily by defining $c_\alpha \triangleq 1$ if $\alpha = 1$ or $c_\alpha \triangleq \frac{1}{1-\alpha}$ if $0 < \alpha < 1$, and noting that

$$m - l - c_\alpha \le m - c_\alpha \le \int_1^{n_m+1} t^{-\alpha}\, dt \le \sum_{j=n_l+1}^{n_m} \mu_j \tag{25}$$

when $n_l = 0$. Finally, by (24), (25) and (H) we find that

$$\left| \sum_{i=l}^{m-1} \sum_{j \in I_i} \mu_j - m + l \right| \le c_\alpha \quad \text{for all } m > l \ge 0, \tag{26}$$

where c_α is defined in the previous sentence. (26) will be useful in the following proofs.

Being based on the same method, the proofs of Propositions 2 and 4 are very similar. Therefore, we will prove Proposition 2 first and then only indicate the changes required to establish Proposition 4.

Proof of Proposition 2 For each $k = 1, 2, 3, \ldots$ and $\omega \in \Omega$, we use (A), (B) and (C) of Section 2 to define

$$g_k \stackrel{\Delta}{=} h_k - h_*, \tag{27}$$

$$z_k \stackrel{\Delta}{=} b_k - A_k h_*, \tag{28}$$

$$u_k \stackrel{\Delta}{=} (I - \mu_{k-1} A_*) u_{k-1} + \mu_{k-1} z_{k-1} \quad \text{if } k \geq 2 \quad \text{or } u_k \stackrel{\Delta}{=} h_1 - h_* \quad \text{if } k = 1 \tag{29}$$

and $w_k \stackrel{\Delta}{=} g_k - u_k$. Hence, using (B), (27–29) and (L), we see that $\{w_k,\ k = 1, 2, 3, \ldots\}$ is defined by

$$w_{k+1} = (I - \mu_k A_*) w_k - \mu_k \bar{A}_k g_k \quad \text{for } k = 1, 2, 3, \ldots \quad \text{subject to } w_I = 0. \tag{30}$$

Now by Lemma 6 to follow in Section 4, we have a $c_0 > 0$ such that

$$\|u_k\|_\delta \leq c_0 \gamma_k \qquad \text{for all } k = 1, 2, 3, \ldots \tag{31}$$

with $\delta > r$ being the constant in (14) of the statement of Proposition 2 so it follows immediately by Minkowski's and Jensen's inequalities and the fact $g_k = u_k + w_k$ (with $w_1 = 0$) that we only have to show that there is some $c > 0$ such that

$$\|w_k\|_r \leq c \gamma_k \quad \text{for all } k = 2, 3, 4, \ldots \tag{32}$$

First, we consider the relatively simple case where $\alpha = 1$. From Lemma 7 (in Section 4), we have a $c_1 > 0$ such that

$$\|w_k\|_{\sigma\delta} \leq c_1 \qquad \text{for all } k = 2, 3, 4, \ldots \tag{33}$$

and by Jensen's inequality, (31) and (33) there is a $c_2 > 0$ such that

$$\|g_k\|_{\sigma\delta} \leq c_2 \quad \text{for all } k = 2, 3, 4, \ldots \tag{34}$$

Now fix an integer $k \geq 2$ and an $\omega \in \Omega$. Then, we have by (30), (J), (H), (I), (27), (B), (28) and Minkowski's inequality that

$$
\begin{aligned}
\|w_k\| &= \left\| \sum_{j=1}^{k-1} F_{j,k}\bar{A}_j g_j \right\| \\
&\le \sum_{i=0}^{s_k} \left\| \sum_{\substack{j\in I_i\\ j<k}} F_{j,k}\bar{A}_j \Big(g_{n_i+1} + \sum_{\substack{n\in I_i\\ n<j}} (g_{n+1}-g_n)\Big) \right\| \\
&\le \sum_{i=0}^{s_k} \left\{ \left\| \sum_{\substack{j\in I_i\\ j<k}} F_{j,k}\bar{A}_j \right\| \cdot \|g_{n_{i+1}}\| + \left\| \sum_{\substack{j,n\in I_i\\ n<j<k}} F_{j,k}\bar{A}_j\mu_n(b_n - A_n h_n) \right\| \right\} \\
&\le \sum_{i=0}^{s_k} \left\{ \left\| \sum_{\substack{j\in I_i\\ j<k}} F_{j,k}\bar{A}_j \right\| \cdot \|g_{n_i+1}\| + \sum_{\substack{n\in I_i\\ n<k}} \left[\left\| \sum_{\substack{j\in I_i\\ n<j<k}} F_{j,k}\bar{A}_j \right\| \cdot \mu_n(\|z_n\| + \|A_n\| \cdot \|g_n\|) \right] \right\}.
\end{aligned}
\tag{35}
$$

Moreover by (16) in Remark 3 with $\alpha = 1$, it is not difficult to show that

$$
\frac{\Theta+1}{\Theta}\,\frac{r}{\sigma\delta} + \frac{r}{\delta} = 1, \tag{36}
$$

where Θ is a constant as described in Remark 3. Therefore by (35), (36), (11), (28), (12) and Condition 4, (34), (26) with $m = l+1$ and Hölder's inequality there are c_3, $c_4 > 0$ independent of k such that

$$
\begin{aligned}
\|w_k\|_r &\le \sum_{i=0}^{s_k} \left\{ \left\| \sum_{\substack{j\in I_i\\ j<k}} F_{j,k}\bar{A}_j \right\|_\delta \cdot \|g_{n_i+1}\|_{\frac{\Theta\sigma\delta}{\Theta+1}} \right. \\
&\qquad \left. + \sum_{\substack{n\in I_i\\ n<k}} \left[\left\| \sum_{\substack{j\in I_i\\ n<j<k}} F_{j,k}\bar{A}_j \right\|_\delta \cdot \mu_n(\|z_n\|_{\frac{\Theta\sigma\delta}{\Theta+1}} + \|A_n\|_{\Theta\sigma\delta} \cdot \|g_n\|_{\sigma\delta}) \right] \right\} \\
&\le c_3 \sum_{i=0}^{s_k} \left\{ \Big(\sum_{\substack{j\in I_i\\ j<k}} \|F_{j,k}\|^2 \Big)^{\frac{1}{2}} \cdot \Big(\|g_{n_i+1}\|_{\sigma\delta} + \sum_{\substack{n\in I_i\\ n<k}} \mu_n + \sum_{\substack{n\in I_i\\ n<k}} \mu_n \|g_n\|_{\sigma\delta} \Big) \right\} \\
&\le c_4 \sum_{i=0}^{s_k} \Big(\sum_{\substack{j\in I_i\\ j<k}} \|F_{j,k}\|^2 \Big)^{\frac{1}{2}}.
\end{aligned}
\tag{37}
$$

Now, it follows from the definitions of $F_{j,k}$ and μ_j (see (J) and (A) when $\alpha = 1$) as well as (9) and (K) that

$$\begin{aligned}
\sum_{\substack{j\in I_i\\ j<k}} \|F_{j,k}\|^2 &\le \sum_{j\in I_i} \mu_j^2 \left[\prod_{l=j+1}^{k-1} \max\{1-\mu_l\lambda_{\min}, \mu_l\lambda_{\max}-1\}\right]^2 \\
&\le \sum_{j\in I_i} \mu_j^2 \left[\prod_{l=j+1}^{\min\{t-1,k-1\}} (\mu_l\lambda_{\max}-1) \cdot \prod_{l=\max\{j+1,t\}}^{k-1} (1-\mu_l\lambda_{\min})\right]^2 \\
&\le \sum_{j\in I_i} \mu_j^2 \left[\prod_{l\ge 1:\, \mu_l\lambda_{\max}>2} (\mu_l\lambda_{\max}-1) \cdot \exp\left\{-\lambda_{\min} \sum_{l=\max\{j+1,t\}}^{k-1} \mu_l\right\}\right]^2,
\end{aligned} \tag{38}$$

where we have used (and continue to use) the convention that a product over zero or a negative number of elements is 1. Therefore, by (38), (K), (I), (H), (26) and (G), there must be a $c_5 > 0$ such that

$$\begin{aligned}
\sum_{\substack{j\in I_i\\ j<k}} \|F_{j,k}\|^2 &\le c''^2 \sum_{j\in I_i} \mu_j^2 \exp\left\{-2\lambda_{\min}\left(\sum_{l=1}^{t-1}\mu_l + \sum_{l=\max\{j+1,t\}}^{k-1} \mu_l\right)\right\} \\
&\le c''^2 \sum_{j\in I_i} \mu_j^2 \exp\left\{2\lambda_{\min}\left(-\sum_{m=i}^{s_k-1}\sum_{l\in I_m}\mu_l + \sum_{l=n_i+1}^{j}\mu_l + \sum_{l=k}^{n_{s_k}}\mu_l\right)\right\} \qquad (39) \\
&\le c''^2 \mu_{n_i+1} \exp\{2\lambda_{\min}(i-s_k+3)\} \cdot \sum_{j\in I_i} \mu_j \\
&\le c_5 \exp\{-i+2\lambda_{\min}(i-s_k)\} \qquad \text{for all } i = 0,1,2,\ldots,s_k\,.
\end{aligned}$$

Finally by (39), (D), (G), (H) and (I) there are c_6, $c_7 > 0$ independent of k such that

$$\begin{aligned}
\sum_{i=0}^{s_k}\left(\sum_{\substack{j\in I_i\\ j<k}} \|F_{j,k}\|^2\right)^{\frac{1}{2}} &\le c_6 \exp\{-\eta s_k\} \sum_{i=0}^{s_k} \exp\{-s_k(\lambda_{\min}-\eta) + i(\lambda_{\min} - 1/2)\} \\
&\le c_6(\lfloor \exp\{s_k\}\rfloor)^{-\eta} \sum_{j=0}^{\infty} \exp\{-(\lambda_{\min}-\eta)j\} \\
&\le c_7(k-1)^{-\eta} \le c_7\, 2^{\eta} k^{-\eta}
\end{aligned} \tag{40}$$

since $\lambda_{\min} > \eta$. The result for $\alpha = 1$ follows from (37) and (40).

Now we establish the result in the case $0 < \alpha < 1$ by showing inductively that there exists constants $\{c'_m\}_{m=0}^{M_\alpha}$ independent of k such that

$$\|w_k\|_{v_m} \leq \begin{cases} c'_m\{[k^{1-\frac{3}{2}\alpha} + k^{\frac{2}{3}-\alpha}\log(k)]\,k^{-\frac{m\alpha}{2}} + 1\} & \text{for } m = 0,1,\dots,M_\alpha - 1, \quad k = 2,3,\dots \\ c'_m\, k^{-\frac{\alpha}{2}} & \text{for } m = M_\alpha, \quad k = 2,3,\dots, \end{cases} \tag{41}$$

where $M_\alpha \triangleq \max\left\{\left\lceil\frac{2(1-\alpha)}{\alpha}\right\rceil, \left\lfloor\frac{4}{3\alpha}\right\rfloor\right\}$, Θ is a constant as discussed in Remark 3 and

$$v_m \triangleq \frac{1}{\left(\frac{\Theta}{\Theta+1}\right)^{M_\alpha - m}\left(\frac{1}{r} + \frac{\Theta}{\delta}\right) - \frac{\Theta}{\delta}} \quad \text{for } m = 0,1,2,\dots,M_\alpha. \tag{42}$$

(Clearly by (16) $\sigma\delta \geq v_m \geq r \geq 1$ so $\|\cdot\|_{v_m}$ is well-defined for $m = 0,1,2,\dots,M_\alpha$.) Now by (16), $v_0 = \sigma\delta$ and (41) follows with $m = 0$ by Lemma 7 so we will assume (41) holds for $m = m_1 - 1$, $1 \leq m_1 \leq M_\alpha$ and show that it holds for m_1. Fix an integer $k \geq 2$. It is not difficult (through substitution) to show that

$$\frac{v_{m_1}}{\delta} + \frac{(\Theta+1)v_{m_1}}{\Theta v_{m_1-1}} = v_{m_1}\left[\frac{1}{\delta} + \frac{\Theta+1}{\Theta v_{m_1-1}}\right] = 1 \tag{43}$$

and by (41) with $m = m_1 - 1$ as well as (31), the fact that $w_1 = 0$ and Jensen's inequality, we have a $c_8 > 0$ such that

$$\|g_j\|_{v_{m_1-1}} \leq c_8\{[j^{1-\frac{3\alpha}{2}} + j^{\frac{2}{3}-\alpha}\log(j)]\,j^{(1-m_1)\frac{\alpha}{2}} + 1\} \quad \text{for all } j = 1,2,3,\dots \tag{44}$$

Therefore, by (35), (43), (11), (12) and Condition 4, (44), and Hölder's inequality there is a $c_9 > 0$ independent of k such that

$$\begin{aligned} \|w_k\|_{v_{m_1}} &\leq \sum_{i=0}^{s_k}\Bigg\{\Bigg\|\sum_{\substack{j\in I_i\\ j<k}} F_{j,k}\bar{A}_j\Bigg\|_\delta \cdot \|g_{n_i+1}\|_{\frac{\Theta v_{m_1-1}}{\Theta+1}} \\ &\quad + \sum_{\substack{n\in I_i\\ n<k}}\Bigg[\Bigg\|\sum_{\substack{j\in I_i\\ n<j<k}} F_{j,k}\bar{A}_j\Bigg\|_\delta \cdot \mu_n\big(\|z_n\|_{\frac{\Theta v_{m_1-1}}{\Theta+1}} + \|A_n\|_{\Theta v_{m_1-1}} \cdot \|g_n\|_{v_{m_1-1}}\big)\Bigg]\Bigg\} \\ &\leq c_9\sum_{i=0}^{s_k}\Bigg\{\Bigg(\sum_{\substack{j\in I_i\\ j<k}}\|F_{j,k}\|^2\Bigg)^{\frac{1}{2}}\Bigg(\|g_{n_i+1}\|_{v_{m_1-1}} + \sum_{n\in I_i}\mu_n\Big\{1 + [n^{1-\frac{3\alpha}{2}} + n^{\frac{2}{3}-\alpha}\log(n)]\,n^{(1-m_1)\frac{\alpha}{2}}\Big\}\Bigg)\Bigg\} \end{aligned} \tag{45}$$

and, since (H) implies $n^{(1-m_1)\frac{\alpha}{2}} \le (n_i+1)^{(1-m_1)\frac{\alpha}{2}}$ and $n^{1-\frac{3\alpha}{2}} + n^{\frac{2}{3}-\alpha}\log(n) \le (n_i+1)^{1-\frac{3\alpha}{2}} + (n_i+1)^{\frac{2}{3}-\alpha}\log(n_i+1) + (n_{i+1})^{1-\frac{3\alpha}{2}} + (n_i+1)^{\frac{2}{3}-\alpha}\log(n_{i+1})$ for large enough i, it follows by (45), (44) and (26) that there is a $c_{10} > 0$ independent of k and i such that

$$\|w_k\|_{v_{m_1}} \le c_{10} \sum_{i=0}^{s_k} \left\{ \left(\sum_{\substack{j\in I_i \\ j<k}} \|F_{j,k}\|^2 \right)^{\frac{1}{2}} \left(1 + \left[(n_i+1)^{1-\frac{3\alpha}{2}} + (n_i+1)^{\frac{2}{3}-\alpha}\log(n_i+1) \right.\right.\right.$$
$$\left.\left.\left. + (n_{i+1})^{1-\frac{3\alpha}{2}} + (n_{i+1})^{\frac{2}{3}-\alpha}\log(n_{i+1}+1) \right] (n_i+1)^{\frac{(1-m_1)\alpha}{2}} \right) \right\} \tag{46}$$

(where we have added one in the last logarithm to avoid taking logarithms of n_{i+1} when it is zero). Moreover, by the fact that $\log(s_k+1) - \log(i) \le \log(s_k+2-i)$, i.e. $s_k\, i + 2i - i^2 - s_k - 1 = (s_k+1-i)(i-1) \ge 0$, for $i = 1, 2, \ldots, s_k$, when $s_k \ge 1$, and the fact that $\lambda_{\min} > 0$, (40) becomes in the case $0 < \alpha < 1$ (after repeating the work in (38) and (39))

$$\begin{aligned}
\sum_{i=0}^{s_k} \left(\sum_{\substack{j\in I_i \\ j<k}} \|F_{j,k}\|^2 \right)^{\frac{1}{2}} &\le c_{11} \sum_{i=0}^{s_k} \mu_{n_i+1}^{\frac{1}{2}} \exp\{\lambda_{\min}(i-s_k)\} \\
&\le c_{11} \left[\sum_{i=1}^{s_k} i^{\frac{-\alpha}{2(1-\alpha)}} \exp\{\lambda_{\min}(i-s_k)\} + \exp\{-\lambda_{\min} s_k\} \right] \\
&\le c_{11}(s_k+1)^{\frac{-\alpha}{2(1-\alpha)}} \left[\sum_{i=1}^{s_k} \exp\left\{ \lambda_{\min}(i-s_k) \right.\right. \\
&\qquad \left. + \frac{\alpha}{2(1-\alpha)}(\log(s_k+1) - \log(i)) \right\} \\
&\qquad \left. + \exp\left\{ -\lambda_{\min} s_k + \frac{\alpha}{2(1-\alpha)}\log(s_k+1) \right\} \right] \\
&\le c_{11} \left(\frac{n_{s_k+1}}{(1-\alpha)^{\frac{1}{1-\alpha}}} \right)^{-\frac{\alpha}{2}} \left[\sum_{j=0}^{s_k} \exp\left\{ -\lambda_{\min} j + \frac{\alpha}{2(1-\alpha)}\log(j+2) \right\} \right] \\
&\le c_{12}\, k^{-\frac{\alpha}{2}}
\end{aligned} \tag{47}$$

for some c_{11}, $c_{12} > 0$ independent of k. Next, using (G) for $\alpha < 1$, it is not difficult to show that there is an $i_0 = i_0(\alpha)$ such that $n_i + 1 \le n_{i+1} \le 2^{\frac{1}{1-\alpha}}(n_i+1)$

and $n_{i+1}+1 \le 3^{\frac{1}{1-\alpha}}(n_i+1)$ for all $i = i_0, i_0+1, i_0+2, \ldots$ so there exists a $c_{13} > 0$ independent of i and k such that

$$
\begin{aligned}
&1+\Big[(n_i+1)^{1-\frac{3\alpha}{2}}+(n_i+1)^{\frac{2}{3}-\alpha}\log(n_i+1)+(n_{i+1})^{1-\frac{3\alpha}{2}} \\
&+(n_{i+1})^{\frac{2}{3}-\alpha}\log(n_{i+1}+1)\Big](n_i+1)^{\frac{(1-m_1)\alpha}{2}} \\
&\qquad \le c_{13}\left\{1+\Big[(n_i+1)^{1-\frac{3\alpha}{2}}+(n_i+1)^{\frac{2}{3}-\alpha}\log(n_i+1)\Big](n_i+1)^{\frac{(1-m_1)\alpha}{2}}\right\}. \qquad (48)
\end{aligned}
$$

Finally, by the monotonicity of $i \to n_i$, the previous relation between n_i+1 and n_{i+1}, and (I), there must be constants $c_{14}^{m_1}, c_{15}^{m_1}, c_{16}^{m_1} > 0$ (which may depend on m_1 and α but are independent of i and k) such that

$$
\begin{aligned}
&(n_i+1)^{1-\frac{\alpha}{2}(2+m_1)}+(n_i+1)^{\frac{2}{3}-\frac{\alpha}{2}(1+m_1)}\log(n_i+1) \\
&\le c_{14}^{m_1}\left\{1+(n_{s_k}+1)^{1-\frac{\alpha}{2}(2+m_1)}+(n_{s_k}+1)^{\frac{2}{3}-\frac{\alpha}{2}(1+m_1)}\log(n_{s_k}+1)\right\} \\
&\le c_{15}^{m_1}\left\{1+(k-1)^{1-\frac{\alpha}{2}(2+m_1)}+(k-1)^{\frac{2}{3}-\frac{\alpha}{2}(1+m_1)}\log(k-1)\right\} \\
&\le c_{16}^{m_1}\left\{1+\Big[k^{1-\frac{3\alpha}{2}}+k^{\frac{2}{3}-\alpha}\log(k)\Big]k^{\frac{(1-m_1)\alpha}{2}}\right\} \qquad (49)
\end{aligned}
$$

for all $i = 0, 1, \ldots, s_k$. We conclude that (41) holds with $m = m_1$ by (46–49) and we have now shown the case where $0 < \alpha < 1$. $\square$

We now indicate how the proof of Proposition 2 can be modified to prove Proposition 4.

Proof of Proposition 4 We do not separate out the case where $\alpha = 1$ and we leave out most of the details because they are similar to the proof of Proposition 2. By Lemma 6, we only have to show inductively that there exists constants $\{c'_m\}_{m=0}^{M_\alpha}$ such that

$$
\|w_k\|_{v_m} \le \begin{cases} c'_m\{[\gamma_k k^{1-\alpha}+k^{\frac{2}{3}-\alpha}\log(k)]\,\gamma_k^m+1\} & \text{for } m=0,1,\ldots,M_\alpha-1, \quad k=2,3,\ldots \\ c'_m\,\gamma_k & \text{for } m=M_\alpha, \quad k=2,3,\ldots, \end{cases} \qquad (50)
$$

where $M_\alpha \stackrel{\Delta}{=} \max\left\{\left\lceil\frac{2(1-\alpha)}{\alpha}\right\rceil, \left\lfloor\frac{4}{3\alpha}\right\rfloor\right\}$ as before but now

$$
v_m \stackrel{\Delta}{=} \frac{1}{\frac{1}{r}-\frac{M_\alpha-m}{\delta}} \quad \text{for } m=0,1,\ldots,M_\alpha. \qquad (51)
$$

(Clearly, it follows that $\frac{1}{v_{m-1}}+\frac{1}{\delta}=\frac{1}{v_m}$ and, using (19), that $1 \le r = v_{M_\alpha} < v_{m-1} \le v_0 = \delta$ for $m = 1, 2, ..., M_\alpha$.) When $m = 0$, (50) follows from Lemma 8 of Section 4 and when $m = m_1$, we follow the same procedure used for Proposition 2. Specifically, by Lemma 6, Condition 4' and (50) with $m = m_1 - 1$, we have as in (45 – 49) of the proof of Proposition 2 that

$$
\begin{aligned}
\|w_k\|_{v_{m_1}} \le & \sum_{i=0}^{s_k}\left\{\left\|\sum_{\substack{j\in I_i\\ j<k}} F_{j,k}\bar{A}_j\right\|_\delta \cdot \|g_{n_i+1}\|_{v_{m_1-1}}\right.\\
& \left.+\sum_{\substack{n\in I_i\\ n<k}}\left[\left\|\sum_{\substack{j\in I_i\\ n<j<k}} F_{j,k}\bar{A}_j\right\|_\delta \cdot \mu_n(\|z_n\|_{v_{m_1-1}} + c'\|g_n\|_{v_{m_1-1}})\right]\right\}\\
& \le c_1 \max_{0\le i\le s_k}\left\{1+\left[\gamma_{n_i+1}(n_i+1)^{1-\alpha}+(n_i+1)^{\frac{2}{3}-\alpha}\log(n_i+1)\right]\gamma_{n_i+1}^{m_1-1}\right\}\\
& \times \sum_{i=0}^{s_k}\left(\sum_{\substack{j\in I_i\\ j<k}}\|F_{j,k}\|^2\right)^{\frac{1}{2}} \le c_2\left\{1+[\gamma_k k^{1-\alpha}+k^{\frac{2}{3}-\alpha}\log(k)]\ \gamma_k^{m_1-1}\right\}\cdot\gamma_k
\end{aligned}
\tag{52}
$$

for some $c_1,\ c_2 > 0$ and the result follows through induction. □

4 SUBSIDIARY RESULTS AND THEIR PROOFS

In this section we provide the technical results which were used in the proofs of Propositions 2 and 4.

LEMMA 6 *Under Conditions* 1–4 *or* 1'–4' *of Section 2, there exists a constant* $c > 0$ *independent of* k *such that*

$$\|u_k\|_\delta < c\,\gamma_k \qquad \text{for } k = 1, 2, 3, ...,$$

where u_k *and* γ_k *are defined respectively in (29) of Section 3 and (E) of Section 2 and* δ *is the constant described above Condition 3 or 3'.*

Proof Let k and j be positive integers such that $k \ge j$. Then, by (29) of Section 3 and (J) we have

$$\|u_k\|_\delta \le \left\|\!\left|(I-\mu_{k-1}A_*)(I-\mu_{k-2}A_*)\cdots(I-\mu_jA_*)\right|\!\right\| \cdot \|u_j\|_\delta + \left\|\sum_{m=j}^{k-1} F_{m,k}\, z_m\right\|_\delta, \tag{53}$$

where we interpret the matrix product as I when $k=j$. Moreover, $z_m = b_m - A_m h_* = \bar{b}_m - \bar{A}_m h_*$ by (C). Hence, (following an argument similar to that in (38) and the first inequality of (39) in Section 3) we have by (10), (11) and (K) that there is a $c_1 > 0$ independent of j and k such that

$$\begin{aligned}
\|u_k\|_\delta &\le c_1 \left\{ \prod_{l=j}^{k-1} \| I - \mu_l A_* \| \cdot \|u_j\|_\delta + \left(\sum_{m=j}^{k-1} \| F_{m,k} \|^2 \right)^{\frac{1}{2}} \right\} \\
&\le c_1 c'' \left\{ \exp\left(-\lambda_{\min} \sum_{l=j}^{k-1} \mu_l \right) \cdot \|u_j\|_\delta \right. \\
&\quad \left. + \left(\sum_{m=j}^{k-1} \exp\left[-2\lambda_{\min} \sum_{l=m+1}^{k-1} \mu_l \right] \mu_m^2 \right)^{\frac{1}{2}} \right\} \\
&\le c_1 c'' \left\{ \exp\left(-\lambda_{\min} \int_j^k t^{-\alpha}\, dt \right) \|u_j\|_\delta \right. \\
&\quad \left. + \left(\sum_{m=j}^{k-1} \exp\left[-2\lambda_{\min} \int_{m+1}^{k} t^{-\alpha}\, dt \right] \mu_m^2 \right)^{\frac{1}{2}} \right\}.
\end{aligned} \tag{54}$$

In the case $\alpha = 1$, we get from (54) with $j=1$ and (29) of Section 3 that there exists constants $c_2,\ c_3 > 0$ independent of k such that

$$\begin{aligned}
\|u_k\|_\delta &\le c_2 \left[k^{-\lambda_{\min}} + k^{-\lambda_{\min}} \left(\sum_{m=1}^{k-1} m^{-2} (m+1)^{2\lambda_{\min}} \right)^{\frac{1}{2}} \right] \\
&\le c_2 \left[k^{-\lambda_{\min}} + 2^{\lambda_{\min}} k^{-\lambda_{\min}} \left(\sum_{m=1}^{k-1} m^{2\lambda_{\min}-2} \right)^{\frac{1}{2}} \right] \\
&\le c_3 \left[k^{-\lambda_{\min}} + k^{-\eta} \right] \le 2 c_3 k^{-\eta},
\end{aligned} \tag{55}$$

where we have considered each of the cases $0 < \lambda_{\min} < 0.5$, $\lambda_{\min} = 0.5$, $0.5 < \lambda_{\min} < 1$ and $\lambda_{\min} \ge 1$ individually in the third inequality.

On the other hand, when $0 < \alpha < 1$ we define $f(x) \triangleq \exp\left(\frac{\lambda_{\min}}{1-\alpha} x \right)$ for all $x \in \mathscr{R}$ and, using (54), (A) and the fact $(t-2)^{-\alpha} \le 3^\alpha t^{-\alpha}$ for $t \ge 3$, get constants $c_4, c_5, c_6, c_7 > 0$ independent of j and k such that

$$
\begin{aligned}
\|u_k\|_\delta &\le c_4\Big[f(j^{1-\alpha}-k^{1-\alpha})\cdot\|u_j\|_\delta \\
&\quad+\Big\{\mu_j\sum_{m=j}^{k-1} m^{-\alpha} f\Big(2\big((m+1)^{1-\alpha}-k^{1-\alpha}\big)\Big)\Big\}^{\frac{1}{2}}\Big] \\
&\le c_5\Big[f(j^{1-\alpha}-k^{1-\alpha})\|u_j\|_\delta+\mu_j^{\frac{1}{2}} f(-k^{1-\alpha})\Big\{f\big(2(j+1)^{1-\alpha}\big) \\
&\quad+\int_{j+2}^{k+1}(t-2)^{-\alpha} f(2t^{1-\alpha})\,dt\Big\}^{\frac{1}{2}}\Big] \\
&\le c_6\Big[f(j^{1-\alpha}-k^{1-\alpha})\|u_j\|_\delta+\mu_j^{\frac{1}{2}} f(-k^{1-\alpha})\Big\{f\big(2(j+1)^{1-\alpha}\big) \\
&\quad+f\big(2(k+1)^{1-\alpha}\big)-f\big(2(j+2)^{1-\alpha}\big)\Big\}^{\frac{1}{2}}\Big] \\
&\le c_7\Big[f(j^{1-\alpha}-k^{1-\alpha})\|u_j\|_\delta+\mu_j^{\frac{1}{2}}\Big],
\end{aligned}
\tag{56}
$$

where the last inequality follows from the fact that $(x+1)^{1-\alpha}-x^{1-\alpha}\le 1$ for any $x\ge 0$ (see [8, Lemma 3.1]). Now if we let $j=1$ then we obtain from (56) and (29) of Section 3 that $\|u_k\|_\delta\le c_7\,[\,|h_1-h_*|+1]$ for all $k=1,2,3,\ldots$ Next, letting $j_k=\lceil k/2\rceil$ for any $k=1,2,3,\ldots$ and re-using the above simple consequence of [8, Lemma 3.1], we can find a $c_8>0$ such that

$$
\begin{aligned}
f(j_k^{1-\alpha}-k^{1-\alpha}) &\le f(1)f\left(\left(\frac{k}{2}\right)^{1-\alpha}-k^{1-\alpha}\right) \\
&= f(1)f\left(-\frac{(2^{1-\alpha}-1)}{2^{1-\alpha}}k^{1-\alpha}\right)\le c_8\,k^{-\frac{9}{2}}
\end{aligned}
\tag{57}
$$

for $k=1,2,3,\ldots$ (since b^{-k^c} with $b>1$ and $c>0$ decreases to zero faster than $k^{-\frac{9}{2}}$). Hence, the lemma follows by applying (56) a second time with $j=\lceil k/2\rceil$ this time.

Lemma 7 *Under Conditions* 1–4 *of Section 2, there exists a constant $c>0$ such that*

$$
\|w_k\|_{\sigma\delta}\le c\left[\gamma_k k^{1-\alpha}+k^{\frac{2}{3}-\alpha}\log(k)+1\right]\quad \text{for } k=2,3,4,\ldots,
$$

where w_k is defined in (30) of Section 3, γ_k is defined in (E) of Section 2 and δ is the constant described above Condition 3.

Proof From (27–30) of Section 3 and (B) of Section 2, we have for $k = 1, 2, 3, \ldots$ that

$$w_{k+1} = (I - \mu_k A_k) w_k - \mu_k \bar{A}_k u_k \quad \text{subject to } w_1 = 0 . \tag{58}$$

Moreover, by Condition 4, we can fix an integer $k \geq 2$ and an $\epsilon > 0$ such that

$$E \exp\left\{ \frac{\delta\sigma}{I - \sigma} \epsilon^{-I} \|A_I\|^{\frac{1+\epsilon}{\alpha}} \right\} < \infty . \tag{59}$$

Then, we have by (58), Minkowski's inequality and (8) that

$$\begin{aligned}
\|w_k\| &\leq \sum_{j=1}^{k-1} \|(I - \mu_{k-1}A_{k-1})(I - \mu_{k-2}A_{k-2}) \cdots (I - \mu_{j+1}A_{j+1})\| \cdot \mu_j \|\bar{A}_j u_j\| \\
&\leq \sum_{j=1}^{k-1} \left[\prod_{l=j+1}^{k-1} \max\{1, \mu_l \|A_l\|\} \cdot \mu_j \|\bar{A}_j u_j\| \right] \\
&\leq \prod_{l=2}^{k-1} \left(1 + (\mu_l \|A_l\|)^{\frac{1+\epsilon}{\alpha}}\right)^{\frac{\alpha}{1+\epsilon}} \cdot \sum_{j=1}^{k-1} \mu_j \|\bar{A}_j u_j\| \\
&\leq \exp\left\{ \frac{\alpha}{1+\epsilon} \sum_{l=2}^{k-1} (\mu_l \|A_l\|)^{\frac{1+\epsilon}{\alpha}} \right\} \cdot \sum_{j=1}^{k-1} \mu_j \|\bar{A}_j u_j\| \quad \text{for all } \omega \in \Omega.
\end{aligned} \tag{60}$$

Now we let $\zeta = \alpha(1-\sigma)/\sigma\delta(1+\epsilon)$ and note that, since $\epsilon\sigma < \epsilon$ and $1 + \alpha\sigma - \alpha - \sigma = (1-\alpha)(1-\sigma) \geq 0$, we have that $(1 + \epsilon - \alpha + \alpha\sigma)/(1+\epsilon)\sigma > 1$. Hence by (60), three applications of Hölder's inequality (with $p = (1+\epsilon)/\alpha(1-\sigma)$, $p = (1 + \epsilon - \alpha + \alpha\sigma)/(1+\epsilon)\sigma$ and $p = n$), Lemma 6 and Condition 1 there exists $c_1 > 0$ and $c_2 > 0$ independent of k such that

$$\begin{aligned}
\|w_k\|_{\sigma\delta} &\leq \left[E \exp\left\{ \frac{\delta\sigma}{1-\sigma} \sum_{l=2}^{k-1} \mu_l^{\frac{1+\epsilon}{\alpha}} \|A_l\|^{\frac{1+\epsilon}{\alpha}} \right\} \right]^{\zeta} \cdot \sum_{j=1}^{k-1} \mu_j \|\bar{A}_j u_j\|_{\frac{(1+\epsilon)\sigma\delta}{1+\epsilon-\alpha+\alpha\sigma}} \\
&\leq \left[1 + \sum_{n=1}^{\infty} \frac{\left(\frac{\delta\sigma}{1-\sigma}\right)^n E\left(\sum_{l=2}^{k-1} \mu_l^{\frac{(1+\epsilon)(n-1)}{\alpha n}} \mu_l^{\frac{1+\epsilon}{\alpha n}} \|A_l\|^{\frac{1+\epsilon}{\alpha}} \right)^n}{n!} \right]^{\zeta} \sum_{j=1}^{k-1} \mu_j \|\bar{A}_j\|_{\phi} \cdot \|u_j\|_{\delta} \\
&\leq c_1 \left[1 + \sum_{n=1}^{\infty} \frac{\left(\frac{\delta\sigma}{1-\sigma}\right)^n \left(\sum_{l=2}^{k-1} \mu_l^{\frac{1+\epsilon}{\alpha}} \right)^{n-1} \sum_{l=2}^{k-1} \mu_l^{\frac{1+\epsilon}{\alpha}} E\|A_1\|^{\frac{n(1+\epsilon)}{\alpha}}}{n!} \right]^{\zeta} \cdot \sum_{j=1}^{k-1} \mu_j \gamma_j \\
&\leq c_2 \left[E \exp\left(\frac{\delta\sigma}{1-\sigma} \epsilon^{-1} \|A_1\|^{\frac{1+\epsilon}{\alpha}} \right) \right]^{\zeta} \left\{ \gamma_k k^{1-\alpha} + k^{\frac{2}{3}-\alpha} \log(k) + 1 \right\} \text{ for } k = 2, 3, 4, \ldots,
\end{aligned} \tag{61}$$

where in the last inequality we have used the fact that $\sum_{l=2}^{\infty} l^{-(1+\epsilon)} \leq \epsilon^{-1}$ and have considered the cases $\alpha = 1$, $\frac{2}{3} < \alpha < 1$, $\alpha = \frac{2}{3}$ and $0 < \alpha < \frac{2}{3}$ separately. Moreover, for the sake of completeness $\phi = \frac{\sigma\delta(1+\epsilon)}{(1-\sigma)(1+\epsilon-\alpha)}$. Hence, by (61) and (59) we have a $c_3 > 0$ such that

$$\|w_k\|_{\sigma\delta} \leq c_3 \left\{ \gamma_k k^{1-\alpha} + k^{\frac{2}{3}-\alpha} \log(k) + 1 \right\} \quad \text{for } k = 2, 3, 4, \ldots \tag{62}$$

Now we establish an analog to Lemma 7 when we have essentially bounded random variables.

Lemma 8 *Under Conditions* 1'–4' *of Section* 2, *there exists a constant* $c > 0$ *such that*

$$\|w_k\|_\delta \leq c \left[\gamma_k k^{1-\alpha} + k^{\frac{2}{3}-\alpha} \log(k) + 1 \right] \quad \textit{for } k = 2, 3, 4, \ldots,$$

where w_k *is defined in* (30) *of Section* 3, γ_k *is defined in* (E) *of Section* 2 *and* δ *is the constant described above Condition* 3'.

Proof Let $k \geq 2$ be a fixed positive integer. Then, by the second inequality in (60) of Lemma 7 and Condition 4', we have a $c_1 > 0$ such that

$$\begin{aligned} \|w_k\| &\leq \sum_{j=1}^{k-1} \left[\prod_{l=j+1}^{k-1} \max\{1, \mu_l\, c'\} \cdot \mu_j\, 2\, c' \|u_j\| \right] \\ &\leq c_1 \sum_{j=1}^{k-1} \mu_j \|u_j\| \quad \text{for almost all } \omega \in \Omega, \end{aligned} \tag{63}$$

where c' is the constant of Condition 4'. Hence, utilizing Minkowski's inequality and Lemma 6 and recalling the argument used in the last inequality of (61), we can find a $c_2 > 0$ such that

$$\|w_k\|_\delta \leq c_1 \sum_{j=1}^{k-1} \mu_j \|u_j\|_\delta \leq c_2 (\gamma_k k^{1-\alpha} + k^{\frac{2}{3}-\alpha} \log(k) + 1) \quad \text{for } k = 2, 3, 4, \ldots, \tag{64}$$

□

5 MOMENT BOUNDS AS IN CONDITIONS 3 AND 3'

In this section, we establish (in order) the moment bounds required for Examples 1, 2 and 3 in Section 2. Hence, our first result, Lemma 9, is a bound for processes satisfying our decaying covariance assumption.

Lemma 9 *Suppose that* $\{u_k,\ k = 1, 2, 3, \ldots\}$ *is a* $\mathscr{R}$*-valued, zero-mean stochastic process on* $(\Omega, \mathscr{F}, P)$ *such that*

$$\beta(v) \triangleq \sup_{l=1,2,3,\ldots} |E\{u_{l+v} u_l\}| < \infty \quad \text{for all } v = 0, 1, 2, \ldots$$

Then, we have for any integers $1 \le j \le k$ *and sequence of real constants* $\{f_l\}_{l=j}^k$ *that*

$$E\left|\sum_{l=j}^k u_l f_l\right|^2 \le 2\sum_{v=0}^{k-j} \beta(v) \cdot \sum_{l=j}^k f_l^2 .$$

Proof To simplify the proof we define $f_l \triangleq 0$ for $l = k+1, k+2, \ldots$ Then, it follows by symmetry, a change of variables and Cauchy-Schwarz that

$$\begin{aligned} E\left|\sum_{l=j}^k u_l f_l\right|^2 &\le 2\sum_{l=j}^k \sum_{m=l}^k f_l f_m E\{u_l u_m\} \\ &\le 2\sum_{l=j}^k \sum_{v=0}^{k-j} |f_l| \cdot |f_{l+v}| \beta(v) \qquad (65) \\ &\le 2\sum_{v=0}^{k-j} \beta(v) \sqrt{\sum_{l=j}^k f_l{}^2} \cdot \sqrt{\sum_{l=j}^k f_{l+v}^2} \end{aligned}$$

and the lemma follows. □

Now we establish a moment bound for strong mixing processes which was partially motivated by Gerencśer (see the first remark on page 168 of [4]) and the following Davydov inequality (see Deo [2, Lemma 1]):

Lemma (Davydov) *Let* $r, s, t \ge 1$ *be constants such that* $r^{-1} + s^{-1} + t^{-1} = 1$ *and* $\mathcal{G}$, $\mathcal{H} \subset \mathcal{F}$ be *σ-algebras. Suppose* ξ, η *are random variables with* $\xi \in L_s(\Omega, \mathcal{G}, P)$ *(the Banach space of random variables* X *such that* $E|X|^s < \infty$*) and* $\eta \in L_t(\Omega, \mathcal{H}, P)$*. Then*

$$|E\xi\eta - E\xi\, E\eta| \le 12(\alpha(\mathcal{G}, \mathcal{H}))^{1/r} |\xi|_s |\eta|_t ,$$

where

$$\alpha(\mathcal{G}, \mathcal{H}) \triangleq \sup_{A \in \mathcal{G}, B \in \mathcal{H}} |P(A \cap B) - P(A)P(B)| .$$

Since the method of proof of our strong mixing bounds involves differentiation, the author feels that it is sensible to first establish the result in the continuous-time setting and then derive the desired discrete-time version as a Corollary. Secondly, it may not be obvious that the following proposition and its corollary are actually strong mixing bounds. However, it is clear that a function $\alpha_a(\cdot)$ satisfying (20) of Section 2 (with ξ_n replacing $\tilde{a}_j$) also satisfies (75) (to follow) for any $1 \le j \le k$ so strong mixing conditions imply that our bounds hold over any interval. (66) and (75) were used because they appear more consistent with the statments and proofs of these results.

Proposition 10 *Suppose* $T > 0$, $z > 1$ *and* $m \ge 2$ *are real constants and* $\{u_t,\ 0 \le t \le T\}$ *is a measurable,* $\mathcal{R}$*-valued, zero-mean stochastic process on* $(\Omega, \mathcal{F}, P)$ *satisfying the following moment bound*

$$M \triangleq \sup_{0\le t\le T} [E|u_t|^{\frac{mz}{z-1}}]^{\frac{z-1}{mz}} < \infty .$$

Moreover, suppose $\alpha(\cdot)$ *is a function such that*

$$\sup_{\substack{D\in\sigma\{u_t,\ 0\le t\le s\}\\ E\in\sigma\{u_{s+\tau}\}}} |P(D\cap E)-P(D)P(E)| \le \alpha(\tau) \tag{66}$$

for all $s,\tau\ge 0$ *such that* $s,\ s+\tau\in[0,T]$. *Then, for any function* $f\in L_2[0,T]$, $\int_0^T f_v u_v\,dv$ *is a random variable satisfying*

$$E\left|\int_0^T f_v u_v\,dv\right|^m \le \left((24m-24)\cdot\int_0^T [\alpha(v)]^{\frac{2}{mz}}\,dv\cdot M^2\cdot\int_0^T f_v^2\,dv\right)^{\frac{m}{2}} .$$

Proof It is not hard to show from the given measurability condition that for any $t\in[0,T]$, $(s,\ \omega)\to u_s(\omega):[0,t]\times\Omega\to\mathscr{R}$ is $\mathscr{B}([0,t])\otimes\mathscr{F}$-measurable. Hence, all measurability requirements follow easily from Fubini's theorem and will not be discussed in the body of the proof. Moreover, it follows immediately from Cauchy-Schwarz, Jensen's inequality, the moment condition and the fact $f\in L_2[0,T]$ that

$$E\left(\int_0^T |f_s u_s|\,ds\right)^m \le \left(\int_0^T f_s^2\,ds\right)^{\frac{m}{2}} T^{\frac{m}{2}-1}\int_0^T E|u_s|^m ds < \infty \tag{67}$$

and hence $\int_0^\tau f_s u_s\,ds$, $0\le\tau\le T$ is well defined for *a.a.* ω. Moreover, since $m\ge 2$, the functions $x\to|x|^m$ and $x\to|x|^{m-1}\mathrm{sgn}(x)$ are continuously differentiable on $\mathscr{R}$,

$$\frac{d}{dx}|x|^m = m|x|^{m-1}\sin(x) \tag{68}$$

and

$$\frac{d}{dx}\{|x|^{m-1}\mathrm{sgn}(x)\} = (m-1)|x|^{m-2} . \tag{69}$$

We now fix an ω such that

$$x_t \triangleq \int_0^t f_s u_s\,ds \qquad \text{for all } 0\le t\le T \tag{70}$$

is well defined and note by (67–70) and Natason [10, Theorem 2, p. 245] that $t\to|x_t|^m$ and $t\to|x_t|^{m-1}\mathrm{sgn}(x_t)$ are absolutely continuous on $[0,T]$. Hence, using (68–70), the Fundamental Theorem of Calculus and the chain rule (see [11, Theorem 5.5, p. 105]), we obtain

$$|x_T|^m = \int_0^T m\{|x_s|^{m-1}\sin(x_s)\}\frac{dx_s}{ds}\,ds = m(m-1)\int_0^T\int_0^s f_s u_s|x_t|^{m-2}u_t f_t\,dt\,ds. \tag{71}$$

Taking expectations on both sides of (71), bringing the expectation inside the integrals (which is justified by (67)), applying Davydov's and Hölder's inequalities and then interchanging the order of the two integrals, we obtain

$$
\begin{aligned}
E|x_T|^m &= m(m-1)\int_0^T\int_0^s f_s f_t \cdot E\{u_s|x_t|^{m-2}u_t\}\,dt\,ds \\
&\le 12m(m-1)\int_0^T\int_0^s |f_s|\cdot|f_t| \\
&\quad \cdot[\alpha(s-t)]^{\frac{2}{mz}}\cdot \||x_t|^{m-2}u_t\|_{\frac{mz}{mz-z-1}}\,M\,dt\,ds \\
&\le 12m(m-1)\int_0^T\int_0^s |f_s|\cdot|f_t| \\
&\quad \cdot[\alpha(s-t)]^{\frac{2}{mz}}[E|x_t|^m]^{1-2/m}\cdot M^2\,dt\,ds \\
&\le \int_0^T [E|x_t|^m]^{1-2/m}K_t\,dt,
\end{aligned}
\tag{72}
$$

where the above manipulations hold in the case $m=2$ and in the case $m>2$, and

$$
K_t \stackrel{\Delta}{=} 12m(m-1)M^2\cdot|f_t|\int_t^T |f_s|\cdot[\alpha(s-t)]^{\frac{2}{mz}}\,ds \quad \text{for all } 0\le t\le T\,. \tag{73}
$$

Now we have by (72), (73), (67) and Lemma 12 (to follow) that

$$
E|x_T|^m \le (24m-24)^{\frac{m}{2}}\cdot M^m\left(\int_0^T |f_t|\cdot\int_t^T |f_s|\cdot[\alpha(s-t)]^{\frac{2}{mz}}\,ds\,dt\right)^{\frac{m}{2}}. \tag{74}
$$

The result follows by defining $f_s=0$ for $s>T$ and using the arguments in the last two inequalities of (69) in the proof of Lemma 9.

COROLLARY 11 *Suppose $z>1$ and $m\ge 2$ are real constants, j and k are positive integers such that $k\ge j$ and $\{\xi_l,\ l=j,j+1,\dots,k\}$ is a $\mathscr{R}$-valued, zero-mean stochastic process on $(\Omega,\mathscr{F},P)$ such that*

$$
M_{\frac{mz}{z-1}} \stackrel{\Delta}{=} \max_{j\le l\le k}[E|\xi_l|^{\frac{mz}{z-1}}]^{\frac{z-1}{mz}} < \infty\,.
$$

Moreover, suppose $\{\tilde\alpha(l)\}_{l=0}^{k-j}$ is a non-increasing sequence such that

$$
\sup_{\substack{D\in\sigma\{\xi_n,\ j\le n\le r\}\\ E\in\sigma\{\xi_{r+l}\}}} |P(D\cap E)-P(D)P(E)| \le \tilde\alpha(l) \tag{75}
$$

for all integers r, l such that $r, r+l \in \{j, j+1, \dots, k\}$. *Then, for any constants* $\{g_l\}_{l=j}^{k}$ *we have*

$$\text{(i)}\quad E\left|\sum_{l=j}^{k} g_l \xi_l\right|^m \leq \left[(24m-24)\cdot \sum_{l=0}^{k-j} [\tilde{\alpha}(l)]^{\frac{2}{mz}} \cdot M^2_{\frac{mz}{z-1}} \cdot \sum_{l=j}^{k} g_l^2\right]^{\frac{m}{2}} .$$

Moreover, if

$$M_\infty \triangleq \max_{j\leq l\leq k} \operatorname{ess\,sup}_{\omega} |\xi_l| < \infty$$

then we can replace (i) *with*

$$\text{(ii)}\quad E\left|\sum_{l=j}^{k} g_l \xi_l\right|^m \leq \left[(24m-24)\cdot \sum_{l=0}^{k-j} [\tilde{\alpha}(l)]^{\frac{2}{m}} \cdot M^2_{\infty} \cdot \sum_{l=j}^{k} g_l^2\right]^{\frac{m}{2}} .$$

Proof We define $u_v = \xi_l$ and $f_v = g_l$ for all $v \in [l-j, l-j+1)$, $l = j, j+1, \dots, k$ and $u_{k-j+1} = 0$. Moreover, we define $\alpha(v) \triangleq \tilde{\alpha}(l)$ for all $v \in [l, l+1)$, $l = 0, 1, \dots, k-j$. Then (i) follows by applying Proposition 10 to $\{u_t,\ 0 \leq t \leq T\}$, where $T \triangleq k-j+1$. (ii) follows immediately by taking limits as $z \to 1$ on both sides of (i).

The following Gronwall-type inequality, required for the proof of Proposition 10, is itself proven on pages 174 and 175 of Gerencsér [4]. □

LEMMA 12 *Let* x_t, K_t, $0 \leq t \leq T$ *be non-negative functions such that* $\sup_{0\leq t\leq T} x_t < \infty$ *and*

$$x_t \leq \int_0^t x_s^{1-1/\rho} K_s\, ds \qquad \textit{for all } 0 \leq t \leq T$$

and some $\rho \geq 1$. *Then*

$$x_T \leq \left(\frac{1}{\rho}\int_0^T K_s\, ds\right)^{\rho} .$$

Finally, we state the stably-generated/L-mixing bound which we require in Example 3.

PROPOSITION 13 *Suppose* $m \geq 2$ *is a real constant, j and k are positive integers such that* $k \geq j$ *and* $\{\xi_l,\ l = j, j+1, \dots, k\}$ *is a* $\mathscr{R}$*-valued, zero-mean stochastic process on* $(\Omega, \mathscr{F}, P)$ *such that*

$$M_m \triangleq \max_{j\leq l\leq k} E^{\frac{1}{m}} |\xi_l|^m < \infty .$$

Moreover, suppose there exists a family of σ*-algebra pairs* $\{(\mathscr{F}_l, \mathscr{F}_l^+),\ l = j, j+1, \dots, k\}$ *such that* $\mathscr{F}_n \subset \mathscr{F}_l \subset \mathscr{F}$ *and* $\mathscr{F}_l^+ \subset \mathscr{F}_n^+ \subset \mathscr{F}$ *for all integers* $j \leq n \leq l \leq k$, $\mathscr{F}_l$ *is independent of* $\mathscr{F}_l^+$ *for all integers* $j \leq l \leq k$ *and* ξ_l *is*

$\mathscr{F}_l$-measurable for all $j \le l \le k$. Then, for any constants $\{g_l\}_{l=j}^{k}$ we have

$$E\left|\sum_{l=j}^{k} g_l \xi_l\right|^m \le \left[(2m-2) \cdot \sum_{l=0}^{k-j} \gamma_m(l) \cdot M_m \cdot \sum_{l=j}^{k} g_l^2\right]^{m/2},$$

where

$$\gamma_m(l) \triangleq \max_n E^{\frac{1}{m}}\left|\xi_n - E\{\xi_n | \mathcal{F}_{n-l}^{+}\}\right|^m, \quad l = 0, 1, \dots, k-j$$

and the above maximum is over the values of n such that n and $n - l \in \{j, j+1, \dots, k\}$.

Proof A continuous-time version of the above result can be proven following the proof of Theorem 1.1 in [4] when one makes the following observations:

(i) Although Theorem 1.1 of [4] is stated in terms of an L-mixing process, its proof clearly carries over to our more general setting.

(ii) Since Lemma 2.3 in [4] is clearly true without the 2 on the right hand side of (2.3), the constant C_m of Theorem 1.1 in [4] can be reduced.

Then, our desired discrete-time version follows in a similar manner to Corollary 11. □

Acknowledgement

The author thanks Lászĺo Gerencsér for inspecting a preliminary version of the inductive methods in this note as well as the reviewers for several suggestions which improved the clarity of this work.

References

[1] M. H. A. Davis and R. B. Vinter. *Stochastic Modelling and Control.* Chapman and Hall, London New York, 1985.
[2] C. M. Deo. A note on empirical processes of strong-mixing sequences. *Ann. Probability*, **1(5)** (1973), 870–875.
[3] E. Eweda and O. Macchi.Convergence of an adaptive linear estimation algorithm. *IEEE Trans. Auto. Control*, **AC-29(2)** (1984), 119–127.
[4] L. Gerencsér. On a class of mixing processes. *Stochastic*, **26** (1989), 165–191.
[5] L. Gerencsér. Rates of convergence of recursive estimators. *SIAM J. Control and Optimization*, **30(5)** (1992), 1200–1227.
[6] A. J. Heunis. Rates of convergence for an adaptive filtering algorithm driven by stationary dependent data. *SIAM Control and Optimization*, **32** (1994), 116–139.
[7] L. Ljung. Convergence analysis of parametric identification methods. *IEEE Trans. Auto. Control* **AC-23(5)** (1978), 770–782.
[8] M. Longnecker and R. J. Serfling. Moment inequalities for S_n under general dependence restrictions, with applications. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, **43** (1978), 1–21.
[9] A. Mokkadem. Mixing properties of ARMA processes. *Stochastic Processes Appl.*, **29** (1988), 309–315.
[10] I. P. Natason. *Theory of Functions of a Real Variable, volume 1.* Ungar, New York, 1955.
[11] W. Rudin. *Priciples of Mathematical Analysis*, third edition. McGraw Hill, Tokyo, 1976.
[12] M. Watanabe.The 2rth mean convergence of adaptive filters with stationary dependent random variables. *IEEE Trans. Inform. Theory*, **IT-30(2)** (1984), 134–140.
[13] J.H. Wilkinson. *The Algebraic Eigenvalue Problem.*Clarendon Press, Oxford, 1965.
[14] R. Yokoyama. Moment bounds for stationary mixing sequences.*Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, **52** (1980), 45–57.