Adaptive Control of Parametric Nonlinear Autoregressive Models Via a New Martingale Approach

Bernard Bercu and Bruno Portier

***Abstract*—The purpose of this note is to investigate the stability and the optimality of the adaptive tracking for a wide class of parametric nonlinear autoregressive models, via a new martingale approach. Several asymptotic results for the standard least squares estimator of the unknown model parameter, such as a central limit theorem, a law of iterated logarithm, and strong laws of large numbers are also provided.**

***Index Terms*—Adaptive control, central limit theorem (CLT), law of iterated logarithm (LIL), least squares (LS), martingales, nonlinear autoregressive models, strong laws.**

I. Introduction

Consider the nonlinear autoregressive model of order $d \geq 1$ given, for all $n \geq 0$, by

$$X_{n+1} = \theta f(X_n, X_{n-1}, \ldots, X_{n-d+1}) + U_n + \varepsilon_{n+1} \tag{1}$$

where X_n, U_n and ε_n are the scalar system output, input and driven noise, respectively. The nonlinear function f is assumed to be known while θ is the real unknown parameter of the model. The standard least squares (LS) estimator $\widehat{\theta}_n$ of θ is given by

$$\widehat{\theta}_n = s_{n-1}^{-1} \sum_{k=1}^{n} \phi_{k-1} (X_k - U_{k-1}), \qquad s_n = \sum_{k=0}^{n} \phi_k^2 + s \tag{2}$$

where $\phi_n = f(X_n, \ldots, X_{n-d+1})$. The positive constant s is added in order to avoid useless invertibility assumption.

The crucial role played by U_n is to regulate the dynamic of the process (X_n) by forcing X_n to track, as closed as possible, a bounded predictable reference trajectory (x_n). Via the certainty equivalence principle U_n, commonly called the adaptive control of the system, is given, for any $n \geq 0$, by

$$U_n = -\widehat{\theta}_n \phi_n + x_{n+1}. \tag{3}$$

Consequently, by substituting (3) into (1), we obtain the closed-loop system

$$X_{n+1} - x_{n+1} = \pi_n + \varepsilon_{n+1} \text{ where } \pi_n = (\theta - \widehat{\theta}_n)\, \phi_n. \tag{4}$$

A wide range of literature concerning the strong consistency and the tracking optimality is available in the linear framework where the function f in (1) is linear. We refer the reader to the most recent advances [1], [3]–[7], [16], and the references theirin. One may naturally expect that some results established in the linear case can be extended to the nonlinear framework where the function f in (1) grows faster than linear. However, in contrast with the linear situation, very few theoretical results are available except the important contribution of Guo [8]. In the special case $d = 1$ and $f(x) = x^a$ with $a \geq 4$, he proved the instability of the closed-loop system (4), even if the LS estimator $\widehat{\theta}_n$ converges a.s. to the true parameter θ. Moreover, in the general case $d \geq 1$ and under some restrictive assumptions on the noise (ε_n), he established the global stability of (4), as soon as the growth rate of the nonlinear function f does not exceed the one of a polynomial of degree $a < 4$. Finally, let us mention that related works in the nonparametric framework can also be found in [9], [14], and [15].

The purpose of this note is to improve the stability results of [8] thanks to a new strong law of large numbers for powers of martingales [2], really suitable in the analysis of the asymptotic behavior of nonlinear regression models.

This note is organized as follows. Section II is devoted to the stability and optimality results of the closed-loop system (4). Moreover, we also show the strong consistency of the LS estimator $\widehat{\theta}_n$ of θ with a sharp almost sure rate of convergence. In Section III, under some suitable assumptions on the nonlinear function f in (1), we establish a central limit theorem (CLT), a law of iterated logarithm (LIL) and several strong laws of large numbers associated with $\widehat{\theta}_n$. Appendix A contains the new strong law of large numbers for powers of martingales [2] whereas all technical proofs are postponed to Appendixes B–D.

II. Stability, Optimality and Consistency

First of all, let us introduce some definitions. The first one specifies a wide class of functions associated with (1) whereas the second one extends the usual definition of global stability and optimality.

Definition 1: We shall say that a function f belongs to the class $\mathcal{C}(a, b)$ where $a, b \in \mathbb{N}$ and $a \geq 1$, if there exist four nonnegative constants c_1, c_2, c_3, c_4 such that, for all $x \in \mathbb{R}^d$

$$c_1 + c_2 \parallel x \parallel^b \leq |f(x)| \leq c_3 + c_4 \parallel x \parallel^a \tag{5}$$

with $b \geq 1$ if $c_1 = 0$ and $b \geq 0$ otherwise.

Definition 2: We shall say that the tracking is globally stable of order $p \geq 1$ if

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} (X_k - x_k)^{2p} < \infty \quad \text{a.s.} \tag{6}$$

whereas it is optimal of order $p \geq 1$ if

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} (X_k - x_k)^{2p} = \sigma(2p) \quad \text{a.s.} \tag{7}$$

where, for the nondecreasing sequence $(\mathcal{F}_n)$ of σ-algebras events occurring up to time n

$$\sigma(2p) = \lim_{n \to \infty} \mathbb{E}[\varepsilon_{n+1}^{2p} \mid \mathcal{F}_n] \quad \text{a.s.} \tag{8}$$

Our first result concerns the global stability and the strong consistency.

Theorem 1: Consider the nonlinear autoregressive model (1) where f belongs to the class $\mathcal{C}(a, b)$ with $a < 4$. Assume that (ε_n) is a martingale difference sequence such that

$$\liminf_{n \to \infty} \mathbb{E}[\varepsilon_{n+1}^2 \mid \mathcal{F}_n] > 0 \quad \text{a.s.} \tag{9}$$

and satisfying, for some $\alpha > 2(2a - 1)$

$$\sup_{n \geq 0} \mathbb{E}[|\varepsilon_{n+1}|^\alpha \mid \mathcal{F}_n] < \infty \quad \text{a.s.} \tag{10}$$

Then, for any $1 \leq p \leq a$, the adaptive tracking is globally stable of order p. More precisely, for any $1 \leq p \leq a$, we have

$$\sum_{k=1}^{n} (X_k - x_k - \varepsilon_k)^{2p} = O(\log n) \quad \text{a.s.} \tag{11}$$

Manuscript received November 9, 2001; revised April 1, 2002. Recommended by Associate Editor A. Datta.

The authors are with the Laboratoire de Mathématiques, UMR C 8628, Equipe de Probabilités, Statistique et Modélisation, Bâtiment 425, Université de Paris-Sud, 91 405 Orsay Cedex, France.

Publisher Item Identifier 10.1109/TAC.2002.802756.

In addition, $\widehat{\theta}_n$ is a strongly consistent estimator of θ with

$$\left(\widehat{\theta}_n - \theta\right)^2 = O\left(\frac{\log\log n}{n}\right) \quad \text{a.s.} \tag{12}$$

Proof: The proof is given in Appendix B. □

Remark 1: The classical global stability and result (11) with $p = 1$ were previously established in [8, Th. 2.2]. Nevertheless, our proof is totally different from that of Guo as it mainly relies on the new strong law of large numbers for powers of martingales given in Appendix A. In addition, one can realize that we can avoid the useless condition [8, eq. (23)] on the noise (ε_n). Finally, let us mention that (12) is the first strong consistency result established in the nonlinear framework with adaptive tracking.

Remark 2: The minoration assumption on $|f|$ in (5) allows us to prove that $n = O(s_n)$ a.s. and to infer that $\widehat{\theta}_n$ converges to θ a.s. It is worth noting that no lower bound for $|f|$ is required in [8]. This is due to the fact that the stability analysis of [8] is concentrated on the asymptotic behavior of the tracking error $X_n - x_n$ rather than the estimation error $\widehat{\theta}_n - \theta$.

We shall now focus our attention on the tracking optimality. To this end, denote for $p \geq 1$

$$C_n(p) = \frac{1}{n}\sum_{k=1}^{n}(X_k - x_k)^{2p} \quad \text{and} \quad \Gamma_n(p) = \frac{1}{n}\sum_{k=0}^{n}\varepsilon_k^{2p}.$$

Corollary 2: Assume that f belongs to the class $\mathcal{C}(a, b)$ with $a < 4$. In addition, suppose that (ε_n) is a martingale difference sequence such that $\mathbb{E}[\varepsilon_{n+1}^2 \mid \mathcal{F}_n] = \sigma^2$ a.s. and satisfying the moment condition (10). If one can find $1 \leq p \leq a$ such that $\sigma(2p)$, given by (8), exists, then the adaptive tracking is optimal of order p and we have

$$\left(C_n(p) - \Gamma_n(p)\right)^2 = O\left(\frac{\log n}{n}\right) \quad \text{a.s.} \tag{13}$$

In addition, if $\mathbb{E}[\varepsilon_{n+1}^{2p-1} \mid \mathcal{F}_n] = 0$ and if $\sigma(2p-2)$ exists, then

$$\lim_{n\to\infty} \frac{n}{\log n}\left(C_n(p) - \Gamma_n(p)\right) = C_{2p}^2\,\sigma(2)\,\sigma(2p-2) \quad \text{a.s.} \tag{14}$$

where $\sigma(2) = \sigma^2$ and $\sigma(0) = 1$.

Proof: The proof is given in Appendix B. □

III. CLT AND STRONG LAWS

The purpose of this section is to establish a CLT and several strong laws of large numbers associated with the LS estimator $\widehat{\theta}_n$ of θ. We shall avoid the minoration assumption on $|f|$ in (5) by restricting ourselves to the polynomial algebra $\mathcal{P}(a)$ with d variables and total degree $\leq a$ where $a \geq 1$.

Lemma 3: Assume that f^2 belongs to the polynomial algebra $\mathcal{P}(2a)$ with $a \geq 1$. In addition, suppose that (ε_n) is a martingale difference sequence satisfying (9) and (10) together with

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=d}^{n} f^2(\varepsilon_k + x_k, \ldots, \varepsilon_{k-d+1} + x_{k-d+1}) > 0 \quad \text{a.s.} \tag{15}$$

Then, we have $n = O(s_n)$ a.s. Moreover, if $a < 4$ and (15) is strengthened by the convergence

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=d}^{n} f^2(\varepsilon_k + x_k, \ldots, \varepsilon_{k-d+1} + x_{k-d+1}) = L \quad \text{a.s.} \tag{16}$$

where L is a positive constant, then we have

$$\lim_{n\to\infty} \frac{s_n}{n} = L \quad \text{a.s.} \tag{17}$$

Proof: The proof is given in Appendix C. □

Theorem 4: Consider the nonlinear autoregressive model (1) where f^2 belongs to the polynomial algebra $\mathcal{P}(2a)$ with $a < 4$. In addition, suppose that (ε_n) is a martingale difference sequence such that $\mathbb{E}[\varepsilon_{n+1}^2 \mid \mathcal{F}_n] = \sigma^2$ a.s. and satisfying the moment condition (10). If (16) holds, then we have the CLT

$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\sigma^2}{L}\right) \tag{18}$$

and the LIL

$$\limsup_{n\to\infty}\left(\frac{n}{2\log\log n}\right)^{1/2}(\widehat{\theta}_n - \theta) = -\liminf_{n\to\infty}\left(\frac{n}{2\log\log n}\right)^{1/2}(\widehat{\theta}_n - \theta) = \frac{\sigma}{\sqrt{L}} \quad \text{a.s.} \tag{19}$$

In particular

$$\limsup_{n\to\infty}\left(\frac{n}{2\log\log n}\right)(\widehat{\theta}_n - \theta)^2 = \frac{\sigma^2}{L} \quad \text{a.s.} \tag{20}$$

Finally, for any $1 \leq p \leq a$, we also have the strong law

$$\lim_{n\to\infty} \frac{1}{\log n}\sum_{k=1}^{n} k^{p-1}(\widehat{\theta}_k - \theta)^{2p} = \frac{\sigma^{2p}(2p)!}{L^p\,2^p\,p!} \quad \text{a.s.} \tag{21}$$

Proof: The proof is given in Appendix D. □

Remark 3: One can observe that the condition (16) is not really restrictive as it is fulfilled in several tracking situations. For example, assume for the sake of simplicity that $x_n \to \tau$ a.s. In addition, suppose that for any $0 \leq q \leq 2a$

$$\lim_{n\to\infty} \mathbb{E}\left[\varepsilon_{n+1}^q \mid \mathcal{F}_n\right] = \sigma(q) \quad \text{a.s.}$$

with $\sigma(0) = 1$, $\sigma(1) = 0$ and $\sigma(2) = \sigma^2$. $\forall x \in \mathbb{R}^d$, let $P(x) = f^2(x)$. In the case $d = 1$, we find via a Taylor expansion of P that

$$L = \sum_{i=0}^{2a} \frac{\partial^i P(0)}{i\,!\partial x^i} \sum_{j=0}^{i} C_i^j\,\sigma(j)\,\tau^{i-j}$$

and, in the case $d = 2$, we obtain

$$L = \sum_{i=0}^{2a}\sum_{j=0}^{2a-i} \frac{\partial^{i+j} P(0,0)}{i\,!\,j!\partial x^i\,\partial y^j}\sum_{k=0}^{i}\sum_{\ell=0}^{j} C_i^k\,C_j^\ell\,\sigma(k)\sigma(\ell)\,\tau^{i+j-k-\ell}.$$

IV. CONCLUSION

In this note, we carried out the complete stability analysis of the adaptive tracking for the single-input–single-output nonlinear autoregressive model, linearly parameterized by a one-dimensional unknown parameter θ. We have extended and refined [8], thanks to a new strong law of large numbers for powers of martingales [2], really suitable in the analysis of the asymptotic behavior of nonlinear regression models. It would be a very attractive challenge for the control community to generalize our approach to the multidimensional parameter case.

APPENDIX A

The theory of martingales is at the core of the greater part of investigations concerning stochastic regression models: a strong law of large numbers for martingales [4], [11], [13], [20] is extensively used to prove strong consistency results and the fluctuations are established via a CLT for martingales [10]. The strong law is really suitable for linear regression models and almost all theoretical advances has been made in the linear situation. Unfortunately, this strong law is not adapted to the nonlinear framework. The purpose of this appendix is to present a new strong law of large numbers for martingales [2], really suitable in the analysis of nonlinear regression models. More precisely, let (ε_n) be a martingale difference sequence adapted to an appropriate filtration $(\mathcal{F}_n)$ and let (ϕ_n) be an adapted sequence of random variables. For all $n \geq 0$, set

$$M_n = M_0 + \sum_{k=1}^{n} \phi_{k-1}\varepsilon_k \quad \text{and} \quad s_n = \sum_{k=0}^{n} \phi_k^2 + s. \tag{A.1}$$

Theorem A.1: Assume that (ε_n) is a martingale difference sequence such that, for some $p \geq 1$ and for some $\alpha > 2(2p-1)$

$$\sup_{n\geq 0} \mathbb{E}\left[\,|\varepsilon_{n+1}|^{\alpha}|\mathcal{F}_n\right] < \infty \quad \text{a.s.} \tag{A.2}$$

Assume also that s_n increases a.s. to infinity. Then

$$\left(\frac{M_n^2}{s_{n-1}}\right)^p = O(\log s_n) \quad \text{a.s.} \tag{A.3}$$

$$\sum_{k=1}^{n}\left(\frac{s_k^p - s_{k-1}^p}{s_k^p}\right)\left(\frac{M_k^2}{s_{k-1}}\right)^p = O(\log s_n) \quad \text{a.s.} \tag{A.4}$$

Theorem A.2: Assume that (ε_n) is a martingale difference sequence satisfying $\mathbb{E}\left[\,\varepsilon_{n+1}^2|\mathcal{F}_n\right] = \sigma^2$ a.s. and for some $p \geq 1$, the moment condition (A.2). Assume also that s_n increases a.s. to infinity with $\phi_n^2 = o(s_n)$ a.s. Then

$$\lim_{n\to\infty} \frac{1}{\log s_n}\sum_{k=1}^{n}\frac{\phi_k^2}{s_k}\left(\frac{M_k^2}{s_{k-1}}\right)^p = \frac{\sigma^{2p}\,(2p)!}{2^p\,p!} \quad \text{a.s.} \tag{A.5}$$

Remark 4: The first strong law of large numbers for martingales is due to [13]. It was refined in [11] and later in [19] and [20]. In the particular case $p = 1$, Theorem A.1 corresponds to [4, Th. 1.3.24], which is extensively used in the analysis of linear regression models. Theorem A.2 establishes the convergence of the moment of order $2p$ in the almost sure central limit theorem for martingales [2].

APPENDIX B

This appendix is devoted to the proofs of Section II. We first start with a very useful Lemma which is at the core of all investigations concerning the excitation and the stability of the sequence (ϕ_n) where $\phi_n = f(X_n, X_{n-1}, \ldots, X_{n-d+1})$.

Excitation and Stability

Lemma B.1: Assume that f belongs to the class $\mathcal{C}(a,b)$ with $a \geq 1$. In addition, suppose that (ε_n) is a martingale difference sequence satisfying (9) and (10). Then, we have

$$n = O(s_n) \quad \text{a.s.} \tag{B.1}$$

$$\log s_{n+1} = O(\log s_n) \quad \text{a.s.} \tag{B.2}$$

Moreover, if $a < 4$, we also have

$$s_n = O(n) \quad \text{a.s.} \tag{B.3}$$

Proof: For any $n \geq 0$, set $\varphi_n = (X_n, \ldots, X_{n-d+1})$. As the function f belongs to $\mathcal{C}(a,b)$ and $\phi_n = f(\varphi_n)$, it clearly follows from (5) that $\phi_n^2 \geq c_1^2 + c_2^2 \parallel \varphi_n \parallel^{2b}$. Consequently

$$s_n \geq s + nc_1^2 + c_2^2 \sum_{k=0}^{n} \parallel \varphi_k \parallel^{2b}. \tag{B.4}$$

On the one hand, the excitation result (B.1) obviously holds if $c_1 > 0$. On the other hand, assume that $c_1 = 0$ so that $b \geq 1$. We immediately deduce from (B.4) that

$$s_n \geq c_2^2 \sum_{l=1}^{d}\sum_{k=d}^{n} X_{k-l+1}^{2b}. \tag{B.5}$$

Moreover, we obtain from the closed-loop (4), together with the variance condition (9), that

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k^2 > 0 \quad \text{a.s.} \tag{B.6}$$

Hence, (B.1) follows from (B.5), (B.6), and the Hölder inequality. Furthermore, we derive from (5) that $\phi_n^2 \leq 2c_3^2 + 2c_4^2 \parallel \varphi_n \parallel^{2a}$. Moreover, we obtain from (4) that $X_{n+1}^2 \leq 3(\pi_n^2 + x_{n+1}^2 + \varepsilon_{n+1}^2)$. Consequently, as the reference trajectory (x_n) is a.s. bounded, we have for all $n \geq d$

$$\phi_n^2 = O\left(\sum_{k=1}^{d} \pi_{n-k}^{2a} + \sum_{k=1}^{d} \varepsilon_{n-k+1}^{2a} + 1\right) \quad \text{a.s.} \tag{B.7}$$

which clearly leads to

$$s_{n+1} = O\left(\sum_{k=d}^{n}\left(\sum_{l=1}^{d} \pi_{k-l+1}^{2a} + \sum_{l=1}^{d} \varepsilon_{k-l+2}^{2a} + 1\right)\right) \quad \text{a.s.} \tag{B.8}$$

In addition, from the moment condition (10), it follows that $\sum_{k=1}^{n} \varepsilon_k^{2a} = O(n)$ a.s. However, we already saw that $n = O(s_n)$ a.s. Consequently, we obtain from (B.8) that one can find a positive random variable ξ such that, for n large enough

$$\begin{aligned} s_{n+1} &\leq \xi\left(\sum_{k=d}^{n}\sum_{l=1}^{d} \pi_{k-l+1}^{2a} + s_n\right) \quad \text{a.s.} \\ &\leq d\,\xi\, s_n^a \left(\sum_{k=1}^{n}\left(\frac{\pi_k^2}{s_k}\right)^a + 1\right) \quad \text{a.s.} \end{aligned} \tag{B.9}$$

Next, let us recall that $\pi_n = (\theta - \widehat{\theta}_n)\phi_n$. We can easily deduce from (1) and (2) that $\widehat{\theta}_n - \theta = s_{n-1}^{-1} M_n$ with M_n given by (A.1) where $M_0 = \theta s$. Hence, we immediately infer from (B.9) that

$$s_{n+1} \leq d\,\xi\, s_n^a \left(\sum_{k=1}^{n} \frac{M_k^{2a}\phi_k^{2a}}{s_k^a\, s_{k-1}^{2a}} + 1\right) \quad \text{a.s.} \tag{B.10}$$

Furthermore, by the elementary fact that $\phi_n^{2a} \leq s_n^a - s_{n-1}^a$ and $s_n \geq s$, we obtain from (B.10) that

$$s_{n+1} \leq d\,\xi\, s_n^a \left(\frac{1}{s^a}\sum_{k=1}^{n} \frac{s_k^a - s_{k-1}^a}{s_k^a}\left(\frac{M_k^2}{s_{k-1}}\right)^a + 1\right) \quad \text{a.s.} \tag{B.11}$$

Finally, (A.4) and (B.11) ensure that $s_{n+1} = O(s_n^a \log s_n)$ which implies (B.2). Now, the stability result (B.3) remains to be proved. It is well known that (B.3) holds for $a = 1$ via the classical stability analysis for linear regression models [1], [3]–[5]. Hereafter, we shall assume that $a \geq 2$. It follows from the moment condition (10) that $\sup_{k\leq n} |\varepsilon_k| = o(n^c)$ a.s. where $1/\alpha < c < 1/2(2a-1)$. Hence, (B.7) implies that

$$\phi_{n+1}^2 = O\left(\sum_{k=1}^{d} \pi_{n-k+1}^{2a}\right) + o(n^{2ac}) \quad \text{a.s.} \tag{B.12}$$

For any $n \geq 1$, set $r_n = \phi_n^2 \log s_{n-1} \,/\, s_{n-1}^{a-1}$. Similarly to (B.10), we deduce from (B.12) that

$$r_{n+1} = O\left(\frac{\log s_n}{s_n^{a-1}} \sum_{k=1}^{d} \frac{M_{n-k+1}^{2a}\phi_{n-k+1}^{2a}}{s_{n-k}^{2a}}\right) + o\left(\frac{n^{2ac}\log s_n}{s_n^{a-1}}\right) \quad \text{a.s.} \tag{B.13}$$

On the one hand, we already saw that $n = O(s_n)$ a.s. On the other hand, we know from (A.3) that $M_n^{2a} = O(s_{n-1}^a \log s_n)$ a.s. Thus, we find via (B.13) that

$$r_{n+1} = O\left(\log s_n \sum_{k=1}^{d} \frac{\phi_{n-k+1}^2 \log s_{n-k+1}}{s_{n-k}^a}\right) + o(1) \quad \text{a.s.} \tag{B.14}$$

In addition, we obtain from (B.2) that, for any $1 \leq k \leq d$, $\log s_n = O(\log s_{n-k})$ a.s. Moreover, $\log s_n = o(s_n)$ a.s. Consequently, (B.14) reduces to

$$r_{n+1} = o\left(\sum_{k=1}^{d} r_{n-k+1}\right) + o(1) \quad \text{a.s.}$$

which leads to $r_n = o(1)$ a.s. Next, we claim that $\phi_n^2 = o(s_{n-1})$ a.s. As a matter of fact, if $t_n = \phi_n^2/s_{n-1}$, it follows once again from (A.3) and (B.12) that

$$t_{n+1} = O\left(\sum_{k=1}^{d} \frac{\phi_{n-k+1}^{2a}\log s_{n-k+1}}{s_n s_{n-k}^a}\right) + o\left(\frac{1}{s_n^{1-2ac}}\right) \quad \text{a.s.}$$

Consequently, we obtain that

$$t_{n+1} = O\left(\sum_{k=1}^{d} \frac{r_{n-k+1}t_{n-k+1}}{s_n^{3-a}}\right) + o(1) \quad \text{a.s.} \tag{B.15}$$

and, as soon as $a < 4$

$$t_{n+1} = O\left(\sum_{k=1}^{d} r_{n-k+1}t_{n-k+1}\right) + o(1) \quad \text{a.s.} \tag{B.16}$$

One can observe that it is only at the implication from (B.15) to (B.16) that we have to require $a < 4$. Next, as $r_n = o(1)$, we deduce from (B.16) that

$$t_{n+1} = o\left(\sum_{k=1}^{d} t_{n-k+1}\right) + o(1) \quad \text{a.s.}$$

so that $t_n = o(1)$ a.s. which means that $\phi_n^2 = o(s_{n-1})$ a.s. We are now in position to prove (B.3). We already saw that, for all $n \geq 1$, $\phi_n^{2a} \leq s_n^a - s_{n-1}^a$. Hence, we deduce from (A.4) that

$$\sum_{k=1}^{n}\left(\frac{s_{k-1}}{s_k}\right)^a \pi_k^{2a} = O(\log s_n) \quad \text{a.s.}$$

Consequently, as s_n is a.s. equivalent to s_{n-1}, we infer that

$$\sum_{k=1}^{n} \pi_k^{2a} = O(\log s_n) \quad \text{a.s.} \tag{B.17}$$

Therefore, it follows from (B.8), together with (B.17), that $s_n = O(\log s_n) + O(n)$ a.s. It clearly implies that $s_n = O(n)$ a.s. which completes the proof of Lemma B.1. □

Proof of Theorem 1

It follows from (A.4), together with the fact that s_n is a.s. equivalent to s_{n-1}, that, for any $1 \leq p \leq a$

$$\sum_{k=1}^{n} \pi_k^{2p} = O(\log s_n) \quad \text{a.s.} \tag{B.18}$$

Hence, as $s_n = O(n)$ a.s., (B.18) implies (11) and the global stability of order p. It remains to prove the consistency result (12). On the one hand, we deduce via (B.1), (B.12), and (B.18) that $\phi_n^2 = O(\log s_n) + o(s_n^{2ac}) = o(s_n^\delta)$ a.s. with $0 < \delta < 1$. Consequently, it follows from (A.1) and [19, Lemma 2] that $M_n^2 = O(s_n \log\log s_n)$ a.s. On the other hand, we already saw in the proof of Lemma B.1 that $\widehat{\theta}_n - \theta = s_{n-1}^{-1} M_n$. Thus, it clearly implies (12), completing the proof of Theorem 1. □

Proof of Corollary 2

One can easily see from (4) that, for any $n \geq 1$

$$n\Big(C_n(p) - \Gamma_n(p)\Big) = P_{n-1} + Q_n$$
$$P_n = \sum_{k=0}^{n} \pi_k^{2p}, \quad Q_n = \sum_{\ell=1}^{2p-1}\sum_{k=0}^{n-1} C_{2p}^{\ell}\pi_k^{2p-\ell}\varepsilon_{k+1}^{\ell}. \tag{B.19}$$

Then, we obtain (13) and (14) via the the same arguments as in the proof [2, Cor. 7]. □

APPENDIX C

This appendix is concerned with the proof of Lemma 3. We shall only carry out the proof for $d = 2$ inasmuch as it already contains all the features for the general case. For all $x, y \in \mathbb{R}$, we set $P(x, y) = f^2(x, y)$. An easy calculation shows that for any $x, y, u, v \in \mathbb{R}$

$$P(x+u, y+v) = P(x, y) + \sum_{i=0}^{2a-1}\sum_{j=0}^{2a-i-1} \Delta_{ij}(u, v)\, x^i\, y^j$$

where

$$\Delta_{ij}(u, v) = \frac{1}{i\,!\,j!}\left[\frac{\partial^{i+j}P(u, v)}{\partial x^i\,\partial y^j} - \frac{\partial^{i+j}P(0, 0)}{\partial x^i\,\partial y^j}\right].$$

For any $n \geq 1$, set $\xi_n = \varepsilon_n + x_n$. We clearly have from (4)

$$s_{n+1} - s_1 = W_{n+1} + Q_{n+1} \tag{C.1}$$

where

$$W_{n+1} = \sum_{k=1}^{n} P(\xi_{k+1}, \xi_k). \quad Q_{n+1} = \sum_{i=0}^{2a-1}\sum_{j=0}^{2a-i-1} Q_{n+1}(i, j)$$

with $Q_{n+1}(i, j) = \sum_{k=1}^{n} \Delta_{ij}(\pi_k, \pi_{k-1})\, \xi_{k+1}^i\, \xi_k^j$. For any integer $1 \leq p \leq a$, denote $\nu_n(p) = \sum_{k=1}^{n} \pi_k^{2p}$ and $\nu_n = \nu_n(1)$. First of all, we deduce from the elementary fact that the polynomial $\Delta_{ij}(u, v)$ is of total degree $2a - (i + j)$ that

$$|Q_{n+1}(i, j)| = O\left(\sum_{k=1}^{n}\sum_{\ell=1}^{2a-i-j} |\xi_{k+1}|^i\, |\xi_k|^j \left(|\pi_k|^\ell + |\pi_{k-1}|^\ell\right)\right).$$

Hence, we obtain from the Cauchy–Schwarz inequality that

$$|Q_{n+1}(i,j)| = O(\sqrt{n\,\nu_n}) + O\left(\left(\sup_{k\le n}|\xi_{k+1}|\right)^{i+j}\sum_{\ell=2}^{2a-i-j}\sum_{k=1}^{n}|\pi_k|^{\ell}\right)$$

$$|Q_{n+1}(i,j)| = O(\sqrt{n\,\nu_n}) + o\left(n^{c(i+j)}\sum_{\ell=2}^{2a-(i+j)}\sum_{k=1}^{n}|\pi_k|^{\ell}\right) \quad \text{a.s.}$$

where $1/\alpha < c < 1/2(2a-1)$ and $i+j \le 2a-1$. Consequently, as $c(i+j) < 1/2$, we obtain that

$$|Q_{n+1}(i,j)| = O(\sqrt{n\,\nu_n}) + o\left(\sqrt{n}\sum_{p=1}^{a}\nu_n(p)\right) \quad \text{a.s.} \tag{C.2}$$

We are now in position to prove Lemma 3. First, assume that s_n converges a.s. Then, we clearly have $\phi_n = o(1)$ a.s. In addition, the classical strong law of large numbers for martingales ensures that M_n, given by (A.1), converges a.s. Next, let us recall that $\pi_n = -s_{n-1}^{-1}M_n\,\phi_n$. Consequently, we obtain that $\pi_n = o(1)$ a.s. so that, for any $p \ge 1$, $\nu_n(p) = o(n)$ a.s. Therefore, we derive that $Q_{n+1} = o(n)$ a.s. Finally, (C.1) immediately implies that $W_{n+1} = o(n)$ which is not possible by (15). Hence, s_n increases a.s. to infinity. However, at this stage, we do not know how fast grows s_n to infinity. On the one hand, assume that

$$\sup_{n\ge 1}\frac{s_{n+1}}{s_n} = \infty \quad \text{a.s.}$$

Then, for n large enough, $s_{n+1} \ge Cs_n$ with $C > 1$, which clearly leads to $n = O(s_n)$ a.s. On the other hand, assume that

$$\sup_{n\ge 1}\frac{s_{n+1}}{s_n} < \infty \quad \text{a.s.}$$

Hence, we deduce from (A.4) that for any $1 \le p \le a$, $\nu_n(p) = O(\log s_n)$ a.s. Therefore, (C.2) immediately implies that $|Q_{n+1}| = O(\sqrt{n}\,\log s_n)$ a.s. so that $Q_{n+1}^2 = o(ns_{n+1})$ a.s. Consequently, we infer from (C.1) that

$$\left(\frac{s_{n+1}}{n} - \frac{W_{n+1}}{n}\right)^2 = o\left(\frac{s_{n+1}}{n}\right) \quad \text{a.s.} \tag{C.3}$$

Hereafter, suppose that

$$\liminf_{n\to\infty}\frac{s_{n+1}}{n} = 0 \quad \text{a.s.}$$

Then, we deduce via (C.3) that $W_{n+1} = o(n)$ a.s. which is one more time not possible by (15). Hence

$$\liminf_{n\to\infty}\frac{s_{n+1}}{n} > 0 \quad \text{a.s.}$$

which obviously leads to $n = O(s_n)$ a.s. and achieves the proof of the first part of Lemma 3. Finally, the convergence result (17) remains to be proved. As f^2 belongs to the polynomial algebra $\mathcal{P}(2a)$ with $a < 4$ and $n = O(s_n)$, proceeding exactly as in the proof of Lemma B.1, we find that $s_n = O(n)$ and $\phi_n^2 = o(s_{n-1})$ a.s. Hence, s_n is a.s. equivalent to s_{n-1}. Consequently, we deduce from (C.2) that $Q_{n+1} = o(n)$ a.s. Therefore, (17) immediately follows from (C.1) together with (16), which completes the proof of Lemma 3. □

APPENDIX D

We finally prove the asymptotic results of Theorem 4. Their proofs rely mainly on the almost sure convergence (17). On the one hand, since (ε_n) has finite conditional moment of order $\alpha > 2$, (18) immediately follows from (17) together with the martingale CLT (see, e.g., [10, Cor. 3.1]). On the other hand, we deduce from the moment condition (10) together with Chow's lemma that $\sum_{k=1}^{n}|\varepsilon_k|^{\beta} = O(n)$ a.s. where β is such that $2(2a-1) < \beta < \alpha$. In addition, we also obtain from (B.18) that $\sum_{k=1}^{n}|\pi_k|^{\beta} = o(n)$ a.s. Furthermore, as f^2 belongs to $\mathcal{P}(2a)$ with $a \ge 1$, it is not difficult to see that, similarly to (B.7)

$$\phi_n^2 = O\left(\sum_{k=1}^{d}\pi_{n-k}^{2a} + \sum_{k=1}^{d}\varepsilon_{n-k+1}^{2a} + 1\right) \quad \text{a.s.} \tag{D.1}$$

Consequently, if $\gamma = \beta/2$, we deduce that $\sum_{k=1}^{n}|\phi_k|^{\gamma} = O(n)$ a.s. Since $\gamma > 2$, it clearly implies that

$$\sum_{n=1}^{\infty}\left(\frac{|\phi_n|}{\sqrt{n}}\right)^{\gamma} < \infty \quad \text{a.s.} \tag{D.2}$$

Hence, (17) together with (D.2) ensure that the conditions of the martingale LIL are satisfied (see. e.g., [17, Th. 3]) which achieves the proof of (19) and (20). Finally the strong laws given by (21) immediately follow from (17) and [2, Cor. 9], completing the proof of Theorem 4. □

REFERENCES

[1] B. Bercu, "Central limit theorem and law of iterated logarithm for the least squares algorithms in adaptive tracking," *SIAM J. Control Optim.*, vol. 36, pp. 910–928, 1998.

[2] ——, "On the convergence of moments in the almost sure central limit theorem for martingales with statistical applications," , 2002, submitted for publication.

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

[4] M. Duflo, *Random Iterative Models*. Berlin, Germany: Springer-Verlag, 1997.

[5] L. Guo and H. F. Chen, "The Aström-Wittenmark self-tuning regulator revisited and ELS-based adaptive trackers," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 802–812, July 1991.

[6] L. Guo, "Further results on least squares based adaptive minimum variance control," *SIAM J. Control Optim.*, vol. 32, pp. 187–212, 1994.

[7] ——, "Convergence and logarithm laws of self-tuning regulators," *Automatica*, vol. 31, pp. 435–450, 1995.

[8] ——, "On critical stability of discrete-time adaptive nonlinear control," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 1488–1499, Nov. 1997.

[9] ——, "How much uncertainty can be dealt with by feedback?," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 2203–2217, Dec. 2000.

[10] P. Hall and C. C. Heyde, *Martingale Limit Theory and its Applications*. New York: Academic, 1980.

[11] T. L. Lai and C. Z. Wei, "Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems," *Ann. Math. Stat.*, vol. 10, pp. 154–166, 1982.

[12] ——, "Extended least squares and their applications to adaptive control and prediction in linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 898–906, 1986.

[13] J. Neveu, *Discrete Parameter Martingales*. Amsterdam, The Netherlands: North-Holland, 1975.

[14] B. Portier and A. Oulidi, "Nonparametric estimation and adaptive control of functional autoregressive models," *SIAM J. Control Optim.*, vol. 39, no. 2, pp. 411–432, 2000.

[15] B. Portier, "Nonparametric and parametric estimators for adaptive control of discrete-time nonlinear systems," *Commun. Inform. Syst.*, to be published.

[16] W. Ren and P. R. Kumar, "Stochastic adaptive prediction and model reference control," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 2047–2060, Oct. 1994.

[17] W. F. Stout, *Almost Sure Convergence*. New York: Academic, 1974.

[18] ——, "A martingale analogue of Kolmogorov's law of the iterated logarithm," *Z. Wahrscheinlichkeitstheorie*, vol. 15, pp. 279–290, 1970.

[19] C. Z. Wei, "Asymptotic properties of least squares estimates in stochastic regression models," *Ann. Stat.*, vol. 13, pp. 1498–1508, 1985.

[20] ——, "Adaptive prediction by least squares predictors in stochastic regression models with applications to time series," *Ann. Stat.*, vol. 15, pp. 1667–1682, 1987.