mathematics

Article

Almost Sure Central Limit Theorem for Error Variance Estimator in Pth-Order Nonlinear Autoregressive Processes

Kaiyu Liang and Yong Zhang *

School of Mathematic, Jilin University, Changchun 130012, China; liangky16@mails.jlu.edu.cn
* Correspondence: zyong2661@jlu.edu.cn

Abstract: In this paper, under some suitable assumptions, using the Taylor expansion, Borel–Cantelli lemma and the almost sure central limit theorem for independent random variables, the almost sure central limit theorem for error variance estimator in the pth-order nonlinear autoregressive processes with independent and identical distributed errors was established. Four examples, first-order autoregressive processes, self-exciting threshold autoregressive processes, threshold-exponential AR progresses and multilayer perceptrons progress, are given to verify the results.

Keywords: almost sure central limit theorem; nonlinear autoregressive processes; error variance estimator; residuals

MSC: 60F15; 60G50

1. Introduction

Over the past twenty years, there has been an increasing interest in the nonlinear time series literature, for example, the monograph by Tong [1] represents a good account of nonlinear time series models. Compared to linear models, studying the properties of estimators in nonlinear time series models is technically more complex and difficult. In this paper, we will investigate the properties of estimators in nonlinear autoregressive processes.

Throughout this paper, we always assume that $\{\varepsilon_i, i \in \mathbb{Z}\}$ is a sequence of independent and identically distributed random variables with mean zero, finite variance σ^2. $\{X_i, i \in \mathbb{Z}\}$ is a sequence of strictly stationary real random variables which satisfies nonlinear autoregressive processes of order p

$$X_i = r_{\boldsymbol{\theta}}(X_{i-1}, \dots, X_{i-p}) + \varepsilon_i, \tag{1}$$

for some $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)' \in \Theta \subset \mathbb{R}^q$, where $r_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} \in \Theta$, is a family of known measurable functions from $\mathbb{R}^p \to \mathbb{R}$. Obviously, $X_{i-1}, \dots, X_{i-s}$ are independent of $\{\varepsilon_j, j \geq i\}$.

In recent years, many authors have studied the properties of estimators for the error sequence. One research interest is the error density estimator, for example, Liebscher [2] proved the law of logarithm and the law of iterated logarithm of the M-estimator in the nonlinear autoregressive processes of order p with independent errors. Cheng and Sun [3] studied the goodness-of-fit test of the errors in the nonlinear autoregressive processes of order p with independent and identical distributed errors. Fu and Yang [4] obtained the asymptotic normality of error kernel density estimators in the pth-order nonlinear autoregressive processes with independent and identical distributed errors. Cheng [5] obtained the asymptotic distribution of the maximum of a suitably normalized deviation of the density estimator from the expectation of the kernel error density. Li [6] established the asymptotic normality of the L_p-norms of error density estimators in the pth-order nonlinear autoregressive processes with independent and identical distributed errors. Kim et al. [7] considered the goodness-of-fit test of the errors in the nonlinear autoregressive processes of order p with a stationary α-mixing error. Cheng [8] considered the uniform strong

check for updates

Citation: Liang, K.; Zhang, Y. Almost Sure Central Limit Theorem for Error Variance Estimator in Pth-Order Nonlinear Autoregressive Processes. *Mathematics* **2024**, *12*, 1482. https://doi.org/10.3390/math12101482

Academic Editor: Ionut Florescu

Received: 2 April 2024
Revised: 6 May 2024
Accepted: 9 May 2024
Published: 10 May 2024

consistency of the classical Glivenko–Cantelli Theorem for the residual-based empirical error in the pth-order nonlinear autoregressive processes with independent and identical distributed errors. Liu and Zhang [9] established the law of the iterated logarithm for error density estimators in the pth-order nonlinear autoregressive processes with independent and identical distributed errors.

The other research interest is the error variance estimator. Cheng [10] obtained the consistency and asymptotic normality of the variance estimator in the pth-order nonlinear autoregressive processes with independent and identical distributed errors. As we know, there are few results about the error variance estimators except for Cheng [10], and there are no results for the almost sure central limit theorem for the error variance estimator, and therefore, we will study the almost sure central limit theorem for the error variance estimator in this paper.

The almost sure central limit theorem (ASCLT, for short) has been first introduced independently by Brosamler [11] and Schatte [12]. Since then many interesting results have been discovered in this field. The classical ASCLT for a sequence $\{X, X_n; n \geq 1\}$ of i.i.d. random variables with zero means states that when $\mathrm{Var}(X) = \sigma^2$,

$$\lim_{n\to\infty} \frac{1}{D_n} \sum_{k=1}^{n} d_k I\left\{ \frac{S_k}{\sqrt{k}\sigma} \leq x \right\} = \Phi(x) \qquad a.s. \tag{2}$$

for all $x \in \mathbb{R}$ with the logarithmic averages $d_k = 1/k$ and $D_n = \sum_{k=1}^{n} d_k$, $S_k = \sum_{j=1}^{k} X_j$. However, logarithmic averaging is not the only one providing a.s. convergence for partial sums of i.i.d. random variables. Peligrad and Révész [13] showed that (2) holds with $d_k = (\log k)^\alpha/k$, $\alpha > -1$. Berkes and Csáki [14] showed that (2) holds also if $d_k = \exp\{(\log k)^\alpha\}/k$, $0 \leq \alpha < 1/2$. To compare these results, Hörmann [15], Tong et al. [16], Miao [17], Li [18], Zhang [19,20], Wu and Jiang [21], and Li and Zhang [22–24] showed that the a.s. limit (2) holds for any weight sequence $\{d_k\}$ satisfying a mild growth condition similar to Kolmogorov's condition on the law of iterated logarithm.

The paper is organized as follows. In Section 1, the significance and background of research is introduced. Some assumptions and main results are stated in Section 2. Several useful lemmas are listed in Section 3. The proofs are listed in Section 4. Examples are stated in Section 5. In the sequel, we denote with $C, C_1, C_2, \cdots$ generic constants that may be different in each of its appearances. $I\{A\}$ denotes the indicator function of the set A. $\Phi(\cdot)$ denotes the distribution function of the standard normal random variable $\mathcal{N}$.

2. Main Results

The goal of this paper is to study the properties of the estimator of the error variance σ^2 by means of the observations $\{X_1, X_2, \cdots, X_n\}$ in model (1). The main difficulty is that we do not observe the error $\{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\}$, the structure of estimation of parameters is complex, and the residuals are unknown. We need to use Taylor's expansion and many other techniques to deal with it. This is the greatest contribution of this paper. We will follow the following steps. Firstly, we compute an estimator $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_q)'$ of unknown parameter $\boldsymbol{\theta}$. Secondly, based on the estimator $\hat{\boldsymbol{\theta}}$ and model (1), we calculate the following residuals

$$\hat{\varepsilon}_i = X_i - r_{\hat{\boldsymbol{\theta}}}(X_{i-1}, \ldots, X_{i-p}), \;\; i = 1, 2, \ldots, n. \tag{3}$$

Finally, using the above residuals, we estimate the error variance σ^2 by using the following equation

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^{n} \hat{\varepsilon}_i^2. \tag{4}$$

Before giving the main results, we need the following basic assumptions for model (1) which will be used throughout the paper. For $1 \le i \le n$ and $1 \le j \le q$, let

$$Y_{ij} \triangleq \frac{\partial}{\partial \theta_j} r_{\boldsymbol{\theta}}(X_{i-1}, \ldots, X_{i-p}), \quad Z_{ijl} \triangleq \frac{\partial^2}{\partial \theta_j \partial \theta_l} r_{\boldsymbol{\theta}^*}(X_{i-1}, \ldots, X_{i-p}),$$

where $\boldsymbol{\theta}^* = \boldsymbol{\theta} + \lambda(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ for some $\lambda \in (0,1)$. By the fact that $X_{i-1}, \ldots, X_{i-s}$ are independent of $\{\varepsilon_j, j \ge i\}$, we conclude that ε_i is independent of Y_{ij}.

Assumption 1. *Let $\mathbb{U} \subset \Theta \subset \mathbb{R}^q$ be an open neighborhood of $\boldsymbol{\theta}$. For any $y \in \mathbb{R}^p$, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_q) \in \mathbb{U}$, $j, l = 1, \ldots, q$, assume that*

$$\left| \frac{\partial}{\partial \theta_j} r_{\boldsymbol{\theta}}(y) \right| \le M_1(y), \quad \left| \frac{\partial^2}{\partial \theta_j \partial \theta_l} r_{\boldsymbol{\theta}}(y) \right| \le M_2(y),$$

where $E[M_1^4(X_{i-1}, \ldots, X_{i-p})] < \infty$ and $E[M_2^4(X_{i-1}, \ldots, X_{i-p})] < \infty$ for each $i \ge 1$.

Assumption 2. *Let $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_q)'$ be a strong consistent estimator for $\boldsymbol{\theta}$ satisfying the following law of iterated logarithm*

$$\limsup_{n \to \infty} \sqrt{\frac{n}{\log \log n}} |\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}| \le C, \quad a.s., \tag{5}$$

where $|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}| = \sqrt{\sum_{j=1}^{q} (\hat{\theta}_j - \theta_j)^2}$ and C is a positive constant.

Remark 1. *By Corollary 2.2 of Klimko and Nelson [25], we know that the least square estimator for the stochastic process under some suitable conditions satisfies (5). For the first-order autoregressive progresses, Wang et al. [26] proved that the least square estimator of the unknown parameters meets (5). For smooth threshold autoregressive progresses, Chan and Tong [27] obtained the conditional least square estimators of the unknown parameters that satisfy (5). For general nonlinear autoregressive progresses of order p, Liebscher [2] established M-estimators for the unknown parameters that satisfy (5), Yao [28] obtained (5) for least square estimators of nonlinear autoregressive progresses.*

Now, we will state the main result for the almost sure central limit theorem of the error variance estimator $\hat{\sigma}^2$.

Theorem 1. *Suppose that $\{d_k\}$ is a sequence of positive numbers satisfying the following conditions:*
(C1) $\limsup_{k \to \infty} k d_k (\log D_k)^\rho / D_k < \infty$ for some $\rho > 1$, where $D_n = \sum_{k=1}^n d_k$.
(C2) $D_n \to \infty$, $D_n = o(n^\epsilon)$, for any $\epsilon > 0$.

For model (1), under the Assumptions 1 and 2, if $E\varepsilon_1^4 < \infty$, for all $x \in \mathbb{R}$, we have

$$\lim_{n \to \infty} \frac{1}{D_n} \sum_{k=1}^{n} d_k I\left\{ \frac{\sqrt{k}}{\sqrt{Var(\varepsilon_1^2)}} \left(\hat{\sigma}_k^2 - \sigma^2 \right) \le x \right\} = \Phi(x) \qquad a.s. \tag{6}$$

Corollary 1. *Let $c_n > 0$ with $c_n \uparrow \infty$ and $\lim_{n \to \infty} \frac{c_{n+1}}{c_n} = 1$ and $\frac{k}{l} \le (\frac{c_k}{c_l})^\gamma$, $k < l$ for some constant $\gamma > 0$. Denote*

$$d_k = \log \frac{c_{k+1}}{c_k} \exp\left(\log^\beta c_k\right), \quad D_n = \sum_{k=1}^{n} d_k, \quad 0 \le \beta < 1/2.$$

Then, under the assumptions of Theorem 1, (6) also holds.

Remark 2. *If the conditions (C1) and (C2) of Theorem 1 is satisfied for some sequence* $\{D_n\}$, *then it is also satisfied for any other sequence* $D_n^* = \Psi(D_n)$, *provided that* $\Psi : \mathbb{R}^+ \to \mathbb{R}^+$ *is differentiable,* $\Psi'(x) = O(\Psi(x)/x)$ *and* $\log \Psi'(x)$ *is uniformly continuous on* (B, ∞) *for some* $B > 0$. *Typical examples are* $\Psi(x) = x^\gamma$, $\Psi(x) = (\log x)^\gamma$, $\Psi(x) = (\log \log x)^\gamma$ *with some suitable* $\gamma > 0$.

Remark 3. *It is easy to show that* $d_k = l(k)/k$, *where* $l(x)$ *is slowly varying at infinity and* $D_n \to \infty$, *satisfies the conditions* (C1) *and* (C2). *So typical examples including* $d_k = 1/k$; $d_k = \log^\theta k/k$, $\theta > -1$; $d_k = \exp(\log^\gamma k)/k$, $0 \le \gamma < 1/2$, $1 < \rho < (1-\gamma)/\gamma$.

3. Preliminary Lemmas

Some useful lemmas which are needed to prove the main result are given in the following section.

Lemma 1 (Hall and Heyde [29], Theorem 2.11, P.23)**.** *Let* $X_1 = S_1$ *and* $X_i = S_i - S_{i-1}$, $2 \le i \le n$ *denote the differences of the sequences* $\{S_i, 1 \le i \le n\}$. *If* $\{S_i, \mathcal{F}_i, 1 \le i \le n\}$ *is a martingale and* $p > 0$, *then there exists constant C depending only on p such that*

$$E\left(\max_{1\le i\le n}|S_i|^p\right) \le C\left\{E\left[\left(\sum_{i=1}^n E(X_i^2|\mathcal{F}_{i-1})\right)^{p/2}\right] + E\left(\max_{1\le i\le n}|X_i|^p\right)\right\}.$$

Lemma 2. *For* $1 \le i \le n$, $1 \le j, l \le q$, *then for any* $2 \le t \le 4$, *one can obtain*

$$E|Y_{ij}|^t \le EM_1^t(X_{i-1}, \ldots, X_{i-p}) \le \left(EM_1^4(X_{i-1}, \ldots, X_{i-p})\right)^{t/4} < \infty,$$
$$E|Z_{ijl}|^t \le EM_2^t(X_{i-1}, \ldots, X_{i-p}) \le \left(EM_2^4(X_{i-1}, \ldots, X_{i-p})\right)^{t/4} < \infty.$$

Proof. The proof of Lemma 2 is obvious by Assumption 1 and the Hölder inequality. □

Lemma 3. *Assume that* $\{G_n, n \ge 1\}$ *is a sequence of random variables satisfying the ASCLT with the weight* $\{d_k\}$ *defined as in Theorem 1, that is*

$$\forall\, x \in \mathbb{R}, \quad \lim_{n\to\infty} \frac{1}{D_n}\sum_{k=1}^n d_k I\{G_k \le x\} = \Phi(x) \qquad a.s.$$

Let $\{R_n, n \ge 1\}$ *be a sequence of random variables converging almost surely to zero. Then,* $\{G_n + R_n, n \ge 1\}$ *also satisfies the ASCLT. That is*

$$\forall\, x \in \mathbb{R}, \quad \lim_{n\to\infty} \frac{1}{D_n}\sum_{k=1}^n d_k I\{G_k + R_k \le x\} = \Phi(x) \qquad a.s.$$

Proof. For fixed $x \in \mathbb{R}$ and $\eta > 0$, recall that $\{G_n, n \ge 1\}$ satisfies the ASCLT, then we have

$$T_{n,\eta} := \left|\frac{1}{D_n}\sum_{k=1}^n d_k I\{G_k \le x + \eta\} - \Phi(x+\eta)\right| \to 0, \quad a.s.$$

and

$$W_{n,\eta} := \left|\frac{1}{D_n}\sum_{k=1}^n d_k I\{G_k \le x - \eta\} - \Phi(x-\eta)\right| \to 0, \quad a.s.$$

Remark that

$$\{G_n + R_n \le x\} \subset \{G_n \le x + \eta\} \cup \{|R_n| > \eta\},$$

$$\{G_n \le x-\eta\} \subset \{G_n+R_n \le x\} \cup \{|R_n| > \eta\}.$$

Then, we can conclude that

$$\begin{aligned}
&\frac{1}{D_n}\sum_{k=1}^{n} d_k I\{G_k+R_k \le x\} - \Phi(x) \\
\le &\frac{1}{D_n}\sum_{k=1}^{n} d_k I\{G_k \le x+\eta\} - \Phi(x+\eta) + \frac{1}{D_n}\sum_{k=1}^{n} d_k I\{|R_k| > \eta\} + |\Phi(x+\eta)-\Phi(x)| \\
\le &T_{n,\eta} + \frac{1}{D_n}\sum_{k=1}^{n} d_k I\{|R_k| > \eta\} + \int_x^{x+\eta}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt \\
\le &T_{n,\eta} + \frac{1}{D_n}\sum_{k=1}^{n} d_k I\{|R_k| > \eta\} + \frac{\eta}{\sqrt{2\pi}},
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{D_n}\sum_{k=1}^{n} d_k I\{G_k+R_k \le x\} - \Phi(x) \\
\ge &\frac{1}{D_n}\sum_{k=1}^{n} d_k I\{G_k \le x-\eta\} - \Phi(x-\eta) - \frac{1}{D_n}\sum_{k=1}^{n} d_k I\{|R_k| > \eta\} + \Phi(x-\eta)-\Phi(x) \\
\ge & -W_{n,\eta} - \frac{1}{D_n}\sum_{k=1}^{n} d_k I\{|R_k| > \eta\} - \frac{\eta}{\sqrt{2\pi}}.
\end{aligned}$$

Noting that $\{R_n, n \ge 1\}$ is a sequence of random variables converging almost surely to zero and the arbitrariness of η, the desired conclusion follows from above discussion. □

Lemma 4 (Zhang [30], Lemma 2.10, P.391)**.** *Let $\{\zeta_n, n \ge 1\}$ be a sequence of uniformly bounded random variables and $\{d_n\}$, $\{D_n\}$ be defined as in Theorem 1. If there exist constants $C > 0$ and $\delta > 0$ and a sequence of positive numbers $\{a(k)\}$ such that $\sum_{n=1}^{\infty} a(2^n) < \infty$ and*

$$E|\zeta_k\zeta_j| \le C((k/j)^{\delta} + a(k)), \text{ for } j/k > b_n = (\log D_n)^{\rho/\delta},$$

then

$$\lim_{n\to\infty}\frac{1}{D_n}\sum_{k=1}^{n} d_k\zeta_k = 0 \quad a.s.$$

Lemma 5. *Let $\{\varepsilon_i, i \ge 1\}$ be a sequence of independent and identically distributed random variables with mean zero, finite variance σ^2 and $E\varepsilon_1^4 < \infty$. Let $\{d_n\}$, $\{D_n\}$ be defined as in Theorem 1. Then for all $x \in \mathbb{R}$,*

$$\lim_{n\to\infty}\frac{1}{D_n}\sum_{k=1}^{n} d_k I\left\{\frac{1}{\sqrt{kVar(\varepsilon_1^2)}}\sum_{i=1}^{k}\left(\varepsilon_i^2-\sigma^2\right) \le x\right\} = \Phi(x) \quad a.s. \tag{7}$$

Proof. Denote $T_k = \sum_{i=1}^{k}(\varepsilon_i^2-\sigma^2)$. Suppose that f is a bounded Lipschitz function. By classical central limit theorem, we have

$$Ef\left(\frac{T_k}{\sqrt{kVar(\varepsilon_1^2)}}\right) \to Ef(\mathcal{N}) \quad \text{as} \quad k\to\infty.$$

By the conclusions in Section 2 of Peligrad and Shao [31] and Theorem 7.1 of Billingsley [32], we know that (7) is equivalent to

$$\lim_{n\to\infty} \frac{1}{D_n} \sum_{k=1}^{n} d_k f\left(\frac{T_k}{\sqrt{kVar(\varepsilon_1^2)}}\right) = Ef(\mathcal{N}),\ a.s.$$

Hence, to prove (7), it suffices to show that

$$\frac{1}{D_n} \sum_{k=1}^{n} d_k \left[f\left(\frac{T_k}{\sqrt{kVar(\varepsilon_1^2)}}\right) - Ef\left(\frac{T_k}{\sqrt{kVar(\varepsilon_1^2)}}\right)\right] \to 0,\ \ a.s.\ n \to \infty. \tag{8}$$

For convenience, let $W_k = f(\frac{T_k}{\sqrt{kVar(\varepsilon_1^2)}}) - Ef(\frac{T_k}{\sqrt{kVar(\varepsilon_1^2)}})$. Notice that $\{\varepsilon_i, i \in \mathbb{Z}\}$ are independent, both f and f' are bounded, then we conclude that for $1 \le k < j \le n$,

$$\begin{aligned}
|EW_kW_j| &= \left|\mathrm{Cov}\left(f\left(\frac{T_k}{\sqrt{kVar(\varepsilon_1^2)}}\right), f\left(\frac{T_j}{\sqrt{jVar(\varepsilon_1^2)}}\right)\right)\right| \\
&= \left|\mathrm{Cov}\left(f\left(\frac{T_k}{\sqrt{kVar(\varepsilon_1^2)}}\right), f\left(\frac{T_j}{\sqrt{jVar(\varepsilon_1^2)}}\right) - f\left(\frac{T_j - T_k}{\sqrt{jVar(\varepsilon_1^2)}}\right)\right)\right| \\
&\le C_1 E\left| f\left(\frac{T_j}{\sqrt{jVar(\varepsilon_1^2)}}\right) - f\left(\frac{T_j - T_k}{\sqrt{jVar(\varepsilon_1^2)}}\right)\right| \\
&\le C_2 \frac{E|T_k|}{\sqrt{jVar(\varepsilon_1^2)}} \le C_3 \frac{(ET_k^2)^{1/2}}{\sqrt{jVar(\varepsilon_1^2)}} \\
&\le C_4 \frac{\sqrt{kVar(\varepsilon_1^2)}}{\sqrt{jVar(\varepsilon_1^2)}} \le C_5 (\frac{k}{j})^{1/2},
\end{aligned}$$

then by Lemma 4 with $\delta = 1/2$ and $a(k) \equiv 0$, (8) holds, and therefore, the proof of (7) is completed. □

Lemma 6. *Under the assumptions of Theorem 1, for any $1 \le j \le q$, we have*

$$\limsup_{n\to\infty} \frac{\log\log n}{n^{3/2}} \sum_{i=1}^{n} Y_{ij}^2 = 0\ \ a.s.$$

Proof. Let $n_k = [k^\alpha]$, $\alpha > 2$. By Lemma 2 and the Markov inequality, for any $\epsilon > 0$, it is easy to known that

$$\begin{aligned}
&\sum_{k=1}^{\infty} P\left(\frac{\log\log n_{k+1}}{n_k^{3/2}} \sum_{i=1}^{n_k} Y_{ij}^2 > \epsilon\right) \\
\le\ & C_1 \sum_{k=1}^{\infty} \frac{\log\log n_{k+1}}{n_k^{3/2}} \sum_{i=1}^{n_k} EY_{ij}^2 \\
\le\ & C_2 \sum_{k=1}^{\infty} \frac{\log\log n_{k+1}}{n_k^{3/2}} \cdot n_k \\
\le\ & C_3 \sum_{k=1}^{\infty} \frac{\log\log(k+1)}{k^{\alpha/2}} < \infty.
\end{aligned}$$

Then by the Borel–Cantelli lemma, we obtain

$$\frac{\log\log n_{k+1}}{n_k^{3/2}} \sum_{i=1}^{n_k} Y_{ij}^2 \to 0 \quad a.s. \;\; as \;\; k \to \infty. \tag{9}$$

Similarly, by Lemma 2 and the Markov inequality, for any $\epsilon > 0$, one can obtain

$$\begin{aligned}
& \sum_{k=1}^{\infty} P\left(\frac{\log\log n_{k+1}}{n_k^{3/2}} \max_{n_k < n \leq n_{k+1}} \sum_{i=n_k+1}^{n} Y_{ij}^2 > \epsilon \right) \\
\leq \;\; & C_1 \sum_{k=1}^{\infty} \frac{\log\log n_{k+1}}{n_k^{3/2}} E\left[\max_{n_k < n \leq n_{k+1}} \sum_{i=n_k+1}^{n} Y_{ij}^2 \right] \\
\leq \;\; & C_2 \sum_{k=1}^{\infty} \frac{\log\log n_{k+1}}{n_k^{3/2}} \sum_{i=n_k+1}^{n_{k+1}} EY_{ij}^2 \\
\leq \;\; & C_3 \sum_{k=1}^{\infty} \frac{\log\log n_{k+1}}{n_k^{3/2}} \cdot [n_{k+1} - n_k] \\
\leq \;\; & C_4 \sum_{k=1}^{\infty} \frac{\log\log(k+1)}{k^{3\alpha/2}} \cdot [(k+1)^{\alpha} - k^{\alpha}] \\
\leq \;\; & C_5 \sum_{k=1}^{\infty} \frac{\log\log(k+1)}{k^{\alpha/2+1}} < \infty.
\end{aligned}$$

By the Borel–Cantelli lemma, it follows that

$$\frac{\log\log n_{k+1}}{n_k^{3/2}} \max_{n_k < n \leq n_{k+1}} \sum_{i=n_k+1}^{n} Y_{ij}^2 \to 0 \quad a.s. \;\; as \;\; k \to \infty. \tag{10}$$

Then combining (9) with (10), for $n_k < n \leq n_{k+1}$, one can obtain

$$\begin{aligned}
& \limsup_{n\to\infty} \frac{\log\log n}{n^{3/2}} \sum_{i=1}^{n} Y_{ij}^2 \\
\leq \;\; & \limsup_{k\to\infty} \frac{\log\log n_{k+1}}{n_k^{3/2}} \sum_{i=1}^{n_k} Y_{ij}^2 + \limsup_{k\to\infty} \frac{\log\log n_{k+1}}{n_k^{3/2}} \max_{n_k < n \leq n_{k+1}} \sum_{i=n_k+1}^{n} Y_{ij}^2 \\
\to \;\; & 0 \quad a.s. \;\; as \;\; k \to \infty.
\end{aligned}$$

Thus, the proof of Lemma 6 is completed. □

Lemma 7. *Under the assumptions of Theorem 1, for any* $1 \leq j, l \leq q$*, we have*

$$\limsup_{n\to\infty} \frac{(\log\log n)^2}{n^{5/2}} \sum_{i=1}^{n} Z_{ijl}^2 = 0 \;\; a.s.$$

Proof. By Lemma 2 and the Markov inequality, for any $\epsilon > 0$, it is easy to see that

$$\begin{aligned}
& \sum_{n=1}^{\infty} P\left(\frac{(\log\log n)^2}{n^{5/2}} \sum_{i=1}^{n} Z_{ijl}^2 > \epsilon \right) \\
\leq \;\; & C_1 \sum_{n=1}^{\infty} \frac{(\log\log n)^2}{n^{5/2}} \sum_{i=1}^{n} EZ_{ijl}^2 \\
\leq \;\; & C_2 \sum_{n=1}^{\infty} \frac{(\log\log n)^2}{n^{5/2}} n
\end{aligned}$$

$$\leq \quad C_3 \sum_{n=1}^{\infty} \frac{(\log \log n)^2}{n^{3/2}} < \infty.$$

By the Borel–Cantelli lemma, one can obtain

$$\frac{(\log \log n)^2}{n^{5/2}} \sum_{i=1}^{n} Z_{ijl}^2 \to 0 \quad a.s. \ \ as \ \ n \to \infty.$$

The proof of Lemma 7 is completed. □

Lemma 8. *Under the assumptions of Theorem 1, for any $1 \leq j \leq q$, we have*

$$\limsup_{n\to\infty} \frac{(\log \log n)^{1/2}}{n} \left| \sum_{i=1}^{n} Y_{ij} \varepsilon_i \right| = 0 \ \ a.s.$$

Proof. Let

$$Y_m = \sum_{i=1}^{m} Y_{ij} \varepsilon_i, \ \ 1 \leq m \leq n.$$

Let $\mathcal{F}_m$ be the σ-algebra generated by the random variables $\{\varepsilon_i, 1 \leq i \leq m\}$. By the fact that Y_{ij} and ε_i are independent, it is easy to compute that the process $\{Y_m, \mathcal{F}_m, 1 \leq m \leq n\}$ is a martingale. By Lemmas 1 and 2, for some $2 < t < 4$, we know

$$\begin{aligned}
E|Y_n|^t &= E\left|\sum_{i=1}^{n} Y_{ij}\varepsilon_i\right|^t \\
&\leq C_1 E\left[\left(\sum_{i=1}^{n} E\left(Y_{ij}^2 \varepsilon_i^2 | \mathcal{F}_{i-1}\right)\right)^{t/2}\right] + C_2 E\left(\max_{1\leq i\leq n} |Y_{ij}\varepsilon_i|^t\right) \\
&\leq C_3 E\left[\left(\sum_{i=1}^{n} EY_{ij}^2\right)^{t/2}\right] \cdot [E\varepsilon_1^2]^{t/2} + C_4 \sum_{i=1}^{n} E|Y_{ij}|^t \cdot E|\varepsilon_1|^t \\
&\leq C_5 n^{t/2} + C_6 n \leq C_7 n^{t/2}.
\end{aligned} \tag{11}$$

By the Markov inequality and (11), for any $\epsilon > 0$, it is easy to obtain

$$\begin{aligned}
&\sum_{n=1}^{\infty} P\left(\frac{(\log \log n)^{1/2}}{n} |\sum_{i=1}^{n} Y_{ij}\varepsilon_i| > \epsilon\right) \\
\leq \quad & C_1 \sum_{n=1}^{\infty} \frac{(\log \log n)^{t/2}}{n^t} E\left|\sum_{i=1}^{n} Y_{ij}\varepsilon_i\right|^t \\
\leq \quad & C_2 \sum_{n=1}^{\infty} \frac{(\log \log n)^{t/2}}{n^t} n^{t/2} \\
\leq \quad & C_3 \sum_{n=1}^{\infty} \frac{(\log \log n)^{t/2}}{n^{t/2}} < \infty.
\end{aligned}$$

By the Borel–Cantelli lemma, we can obtain

$$\frac{(\log \log n)^{1/2}}{n} \left|\sum_{i=1}^{n} Y_{ij}\varepsilon_i\right| \to 0 \quad a.s. \ \ as \ \ n \to \infty.$$

The proof of Lemma 8 is completed. □

Lemma 9. *Under the assumptions of Theorem 1, for any $1 \le j, l \le q$, we have*

$$\limsup_{n\to\infty} \frac{\log\log n}{n^{3/2}} \left| \sum_{i=1}^{n} Z_{ijl}\varepsilon_i \right| = 0 \ \ a.s.$$

Proof. Let $n_k = [k^\alpha]$, $\alpha > 2$. By Lemma 2, the Markov inequality, C_r inequality and Cauchy–Schwarz inequality, for any $\epsilon > 0$, it is easy to see that

$$\begin{aligned}
& \sum_{k=1}^{\infty} P\left(\frac{\log\log n_{k+1}}{n_k^{3/2}} \left| \sum_{i=1}^{n_k} Z_{ijl}\varepsilon_i \right| > \epsilon \right) \\
\le\ & C_1 \sum_{k=1}^{\infty} \frac{(\log\log n_{k+1})^2}{n_k^3} E\left(\sum_{i=1}^{n_k} Z_{ijl}\varepsilon_i \right)^2 \\
\le\ & C_2 \sum_{k=1}^{\infty} \frac{(\log\log n_{k+1})^2}{n_k^3} \cdot n_k \sum_{i=1}^{n_k} E Z_{ijl}^2 \varepsilon_i^2 \\
\le\ & C_3 \sum_{k=1}^{\infty} \frac{(\log\log n_{k+1})^2}{n_k^2} \sum_{i=1}^{n_k} (E Z_{ijl}^4)^{1/2} (E\varepsilon_i^4)^{1/2} \\
\le\ & C_4 \sum_{k=1}^{\infty} \frac{(\log\log n_{k+1})^2}{n_k} \le C_5 \sum_{k=1}^{\infty} \frac{(\log\log(k+1))^2}{k^\alpha} < \infty.
\end{aligned}$$

Then by the Borel–Cantelli lemma, we obtain

$$\frac{\log\log n_{k+1}}{n_k^{3/2}} \left| \sum_{i=1}^{n_k} Z_{ijl}\varepsilon_i \right| \to 0 \quad a.s. \ \ as \ \ k \to \infty. \tag{12}$$

Similarly, By Lemma 2 and the Markov inequality and C_r inequality, for any $\epsilon > 0$, one can obtain

$$\begin{aligned}
& \sum_{k=1}^{\infty} P\left(\frac{\log\log n_{k+1}}{n_k^{3/2}} \max_{n_k < n \le n_{k+1}} \left| \sum_{i=n_k+1}^{n} Z_{ijl}\varepsilon_i \right| > \epsilon \right) \\
\le\ & C_1 \sum_{k=1}^{\infty} \frac{(\log\log n_{k+1})^2}{n_k^3} E\left(\max_{n_k < n \le n_{k+1}} \left| \sum_{i=n_k+1}^{n} Z_{ijl}\varepsilon_i \right| \right)^2 \\
\le\ & C_2 \sum_{k=1}^{\infty} \frac{(\log\log n_{k+1})^2}{n_k^3} E\left(\sum_{i=n_k+1}^{n_{k+1}} |Z_{ijl}\varepsilon_i| \right)^2 \\
\le\ & C_3 \sum_{k=1}^{\infty} \frac{(\log\log n_{k+1})^2}{n_k^3} \cdot (n_{k+1} - n_k) \sum_{i=n_k+1}^{n_{k+1}} E\left[Z_{ijl}^2 \varepsilon_i^2 \right] \\
\le\ & C_4 \sum_{k=1}^{\infty} \frac{(\log\log n_{k+1})^2}{n_k^3} \cdot (n_{k+1} - n_k) \sum_{i=n_k+1}^{n_{k+1}} (E Z_{ijl}^4)^{1/2} (E\varepsilon_i^4)^{1/2} \\
\le\ & C_5 \sum_{k=1}^{\infty} \frac{(\log\log n_{k+1})^2}{n_k^3} \cdot (n_{k+1} - n_k)^2 \\
\le\ & C_6 \sum_{k=1}^{\infty} \frac{(\log\log(k+1))^2}{k^{3\alpha}} \cdot k^{2(\alpha-1)} \le C_7 \frac{(\log\log(k+1))^2}{k^{\alpha+2}} < \infty.
\end{aligned}$$

By the Borel–Cantelli lemma, it follows that

$$\frac{\log\log n_{k+1}}{n_k^{3/2}} \max_{n_k < n \le n_{k+1}} \left| \sum_{i=n_k+1}^{n} Z_{ijl}\varepsilon_i \right| \to 0 \quad a.s. \ \ as \ \ k \to \infty. \tag{13}$$

Then combining (12) with (13), for $n_k < n \le n_{k+1}$, one can obtain

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{\log\log n}{n^{3/2}} \left| \sum_{i=1}^{n} Z_{ijl}\varepsilon_i \right| \\
\le\ & \limsup_{k\to\infty} \frac{\log\log n_{k+1}}{n_k^{3/2}} \left| \sum_{i=1}^{n_k} Z_{ijl}\varepsilon_i \right| + \limsup_{k\to\infty} \frac{\log\log n_{k+1}}{n_k^{3/2}} \max_{n_k < n \le n_{k+1}} \left| \sum_{i=n_k+1}^{n} Z_{ijl}\varepsilon_i \right| \\
\to\ & 0 \quad a.s. \;\; as \;\; k \to \infty.
\end{aligned}$$

The proof of Lemma 9 is completed. □

Lemma 10. *Under the assumptions of Theorem 1, for any $1 \le j, l, k \le q$, we have*

$$\limsup_{n\to\infty} \frac{(\log\log n)^{3/2}}{n^2} \left| \sum_{i=1}^{n} Y_{ij} Z_{ilk} \right| = 0 \;\; a.s.$$

Proof. By Lemma 2, the Markov inequality, C_r inequality and Cauchy–Schwarz inequality, for any $\epsilon > 0$, it is easy to see that

$$\begin{aligned}
& \sum_{n=1}^{\infty} P\left(\frac{(\log\log n)^{3/2}}{n^2} \left| \sum_{i=1}^{n} Y_{ij} Z_{ilk} \right| > \epsilon \right) \\
\le\ & C_1 \sum_{n=1}^{\infty} \frac{(\log\log n)^{3t/4}}{n^t} E \left| \sum_{i=1}^{n} Y_{ij} Z_{ilk} \right|^{t/2} \\
\le\ & C_2 \sum_{n=1}^{\infty} \frac{(\log\log n)^{3t/4}}{n^t} n^{t/2-1} \sum_{i=1}^{n} E\left[|Y_{ij}|^{t/2} |Z_{ilk}|^{t/2} \right] \\
\le\ & C_3 \sum_{n=1}^{\infty} \frac{(\log\log n)^{3t/4}}{n^{t/2+1}} \sum_{i=1}^{n} \left(E|Y_{ij}|^t\right)^{1/2} \left(E|Z_{ilk}|^t\right)^{1/2} \\
\le\ & C_4 \sum_{n=1}^{\infty} \frac{(\log\log n)^{3t/4}}{n^{t/2}} < \infty.
\end{aligned}$$

where $2 < t \le 4$ is defined in Assumption 1.

By the Borel–Cantelli lemma, we obtain

$$\frac{(\log\log n)^{3/2}}{n^2} \left| \sum_{i=1}^{n} Y_{ij} Z_{ilk} \right| \to 0 \quad a.s. \;\; as \;\; n \to \infty.$$

The proof of Lemma 10 is completed. □

4. Proof

Recall (1) and (3), by Taylor's expansion expansion with the Lagrange remainder, there exists $\lambda \in (0,1)$, and $\boldsymbol{\theta}^* = \boldsymbol{\theta} + \lambda(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$

$$\begin{aligned}
\hat{\varepsilon}_i &= \varepsilon_i - [r_{\hat{\boldsymbol{\theta}}}(X_{i-1}, \ldots, X_{i-p}) - r_{\boldsymbol{\theta}}(X_{i-1}, \ldots, X_{i-p})] \\
&= \varepsilon_i - \sum_{j=1}^{q} (\hat{\theta}_j - \theta_j) Y_{ij} - \frac{1}{2} \sum_{j=1}^{q} \sum_{l=1}^{q} (\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l) Z_{ijl}.
\end{aligned} \tag{14}$$

Then by (14), we can obtain

$$\sqrt{\frac{n}{Var(\varepsilon_1^2)}} \left(\hat{\sigma}_n^2 - \sigma^2 \right)$$

$$
\begin{aligned}
&= \sqrt{\frac{n}{Var(\varepsilon_1^2)}}\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i^2 - \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 + \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 - \frac{1}{n}\sum_{i=1}^{n}E\varepsilon_i^2\right) \\
&= \sqrt{\frac{1}{nVar(\varepsilon_1^2)}}\left[\sum_{i=1}^{n}(\hat{\varepsilon}_i^2 - \varepsilon_i^2)\right] + \sqrt{\frac{1}{nVar(\varepsilon_1^2)}}\left[\sum_{i=1}^{n}(\varepsilon_i^2 - E\varepsilon_i^2)\right] \\
&= \sqrt{\frac{1}{nVar(\varepsilon_1^2)}}\left[\sum_{i=1}^{n}\left(\sum_{j=1}^{q}(\hat{\theta}_j - \theta_j)Y_{ij}\right)^2\right] \\
&\quad + \sqrt{\frac{1}{16nVar(\varepsilon_1^2)}}\left[\sum_{i=1}^{n}\left(\sum_{j=1}^{q}\sum_{l=1}^{q}(\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l)Z_{ijl}\right)^2\right] \\
&\quad - \sqrt{\frac{4}{nVar(\varepsilon_1^2)}}\left[\sum_{i=1}^{n}\sum_{j=1}^{q}(\hat{\theta}_j - \theta_j)Y_{ij}\varepsilon_i\right] \\
&\quad - \sqrt{\frac{1}{nVar(\varepsilon_1^2)}}\left[\sum_{i=1}^{n}\sum_{j=1}^{q}\sum_{l=1}^{q}(\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l)Z_{ijl}\varepsilon_i\right] \\
&\quad + \sqrt{\frac{1}{nVar(\varepsilon_1^2)}}\left[\sum_{i=1}^{n}\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{k=1}^{q}(\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l)(\hat{\theta}_k - \theta_k)Y_{ij}Z_{ilk}\right] \\
&\quad + \sqrt{\frac{1}{nVar(\varepsilon_1^2)}}\left[\sum_{i=1}^{n}(\varepsilon_i^2 - E\varepsilon_i^2)\right] \\
&=: I_{n1} + I_{n2} - I_{n3} - I_{n4} + I_{n5} + I_{n6}.
\end{aligned} \tag{15}
$$

Recall the elementary inequality

$$
\left(\sum_{i=1}^{q}a_ib_i\right)^2 \le \left(\sum_{i=1}^{q}a_i^2\right)\left(\sum_{i=1}^{q}b_i^2\right). \tag{16}
$$

For I_{n1}, by (5) and (16) and Lemma 6, it is easy to know that

$$
\begin{aligned}
I_{n1} &= \sqrt{\frac{1}{nVar(\varepsilon_1^2)}}\left[\sum_{i=1}^{n}\left(\sum_{j=1}^{q}(\hat{\theta}_j - \theta_j)Y_{ij}\right)^2\right] \\
&\le \sqrt{\frac{1}{nVar(\varepsilon_1^2)}}\sum_{j=1}^{q}(\hat{\theta}_j - \theta_j)^2 \cdot \sum_{j=1}^{q}\sum_{i=1}^{n}Y_{ij}^2 \\
&= \frac{1}{\sqrt{Var(\varepsilon_1^2)}}\frac{n}{\log\log n}\sum_{j=1}^{q}(\hat{\theta}_j - \theta_j)^2 \cdot \sum_{j=1}^{q}\frac{\log\log n}{n^{3/2}}\sum_{i=1}^{n}Y_{ij}^2 \\
&\to 0 \quad a.s.\ \ as\ \ n \to \infty.
\end{aligned} \tag{17}
$$

For I_{n2}, by (5) and (16) and Lemma 7, one can obtain

$$
\begin{aligned}
I_{n2} &= \sqrt{\frac{1}{16nVar(\varepsilon_1^2)}}\left[\sum_{i=1}^{n}\left(\sum_{j=1}^{q}\sum_{l=1}^{q}(\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l)Z_{ijl}\right)^2\right] \\
&\le \sqrt{\frac{1}{16nVar(\varepsilon_1^2)}}\sum_{j=1}^{q}(\hat{\theta}_j - \theta_j)^2 \cdot \sum_{l=1}^{q}(\hat{\theta}_l - \theta_l)^2 \cdot \sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{i=1}^{n}Z_{ijl}^2 \\
&= \frac{1}{4\sqrt{Var(\varepsilon_1^2)}}\left(\frac{n}{\log\log n}\sum_{j=1}^{q}(\hat{\theta}_j - \theta_j)^2\right)^2 \cdot \sum_{j=1}^{q}\sum_{l=1}^{q}\frac{(\log\log n)^2}{n^{5/2}}\sum_{i=1}^{n}Z_{ijl}^2 \\
&\to 0 \quad a.s.\ \ as\ \ n \to \infty.
\end{aligned} \tag{18}
$$

For I_{n3}, by (5) and (16) and Lemma 8, one can obtain

$$\begin{aligned}I_{n3} =&\sqrt{\frac{4}{nVar(\varepsilon_1^2)}}\left[\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)\cdot\sum_{i=1}^{n}Y_{ij}\varepsilon_i\right]\\ \le&\sqrt{\frac{4}{nVar(\varepsilon_1^2)}}\left[\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2\cdot\sum_{j=1}^{q}\left(\sum_{i=1}^{n}Y_{ij}\varepsilon_i\right)^2\right]^{1/2}\\ =&\frac{2}{\sqrt{Var(\varepsilon_1^2)}}\left[\frac{n}{\log\log n}\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2\cdot\sum_{j=1}^{q}\left(\frac{(\log\log n)^{1/2}}{n}\sum_{i=1}^{n}Y_{ij}\varepsilon_i\right)^2\right]^{1/2}\\ \to&0\quad a.s.\ \ as\ \ n\to\infty.\end{aligned} \tag{19}$$

For I_{n4}, by (5) and (16) and Lemma 9, we have

$$\begin{aligned}I_{n4} =&\sqrt{\frac{1}{nVar(\varepsilon_1^2)}}\left[\sum_{j=1}^{q}\sum_{l=1}^{q}(\hat{\theta}_j-\theta_j)(\hat{\theta}_l-\theta_l)\sum_{i=1}^{n}Z_{ijl}\varepsilon_i\right]\\ \le&\sqrt{\frac{1}{nVar(\varepsilon_1^2)}}\left[\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2\cdot\sum_{l=1}^{q}(\hat{\theta}_l-\theta_l)^2\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}\left(\sum_{i=1}^{n}Z_{ijl}\varepsilon_i\right)^2\right]^{1/2}\\ =&\frac{1}{\sqrt{Var(\varepsilon_1^2)}}\left[\left(\frac{n}{\log\log n}\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2\right)^2\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}\left(\frac{\log\log n}{n^{3/2}}\sum_{i=1}^{n}Z_{ijl}\varepsilon_i\right)^2\right]^{1/2}\\ \to&0\quad a.s.\ \ as\ \ n\to\infty.\end{aligned} \tag{20}$$

For I_{n5}, by (5) and (16) and Lemma 10, we know

$$\begin{aligned}I_{n5} =&\sqrt{\frac{1}{nVar(\varepsilon_1^2)}}\left[\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{k=1}^{q}(\hat{\theta}_j-\theta_j)(\hat{\theta}_l-\theta_l)(\hat{\theta}_k-\theta_k)\sum_{i=1}^{n}Y_{ij}Z_{ilk}\right]\\ \le&\sqrt{\frac{1}{nVar(\varepsilon_1^2)}}\left[\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2\cdot\sum_{l=1}^{q}(\hat{\theta}_l-\theta_l)^2\cdot\sum_{k=1}^{q}(\hat{\theta}_k-\theta_k)^2\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{k=1}^{q}\left(\sum_{i=1}^{n}Y_{ij}Z_{ilk}\right)^2\right]^{1/2}\\ =&\frac{1}{\sqrt{Var(\varepsilon_1^2)}}\left[\left(\frac{n}{\log\log n}\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2\right)^3\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{k=1}^{q}\left(\frac{(\log\log n)^{3/2}}{n^2}\sum_{i=1}^{n}Y_{ij}Z_{ilk}\right)^2\right]^{1/2}\\ \to&0\quad a.s.\ \ as\ \ n\to\infty.\end{aligned} \tag{21}$$

Combing (17)–(21), one can obtain

$$I_{n1}+I_{n2}-I_{n3}-I_{n4}+I_{n5}\to 0\quad a.s.\ \ as\ \ n\to\infty. \tag{22}$$

For I_{n6}, by Lemma 5, it is obviously that

$$\lim_{n\to\infty}\frac{1}{D_n}\sum_{k=1}^{n}d_kI\left\{\frac{1}{\sqrt{kVar(\varepsilon_1^2)}}\sum_{i=1}^{k}(\varepsilon_i^2-E\varepsilon_i^2)\le x\right\}=\Phi(x)\qquad a.s. \tag{23}$$

Finally, (6) follows by combining (15), (22) with (23) and Lemma 3, thus the proof of Theorem 1 is completed.

5. Examples

Some examples are given in this section to verify the almost sure central limit theorem for the error variance estimator for some special nonlinear autoregressive models. The first example is a degenerate model, that is, AR(1) progresses.

Example 1. *An AR(1) model is a family of $\{X_i\}$ of random variables such that for every $i \geq 1$*

$$X_i = \theta X_{i-1} + \varepsilon_i,$$

where $\{\varepsilon_i, i \geq 1\}$ is a collection of i.i.d. random variables with zero mean and finite variance σ^2. We also assume that $E\exp\{\gamma|\varepsilon_i\varepsilon_j|\} < \infty$ for some $\gamma > 0$ and any $i, j \geq 1$. It is obviously that $\{X_i\}$ is a stationary model under the condition $|\theta| < 1$.

It is easy to check that the Assumption 1 holds naturally. For Assumption 2, by Theorem 1 of Wang et al. [26] and $E\exp\{\gamma|\varepsilon_i\varepsilon_j|\} < \infty$, (5) holds for the least squares estimator $\hat{\theta}$. Therefore, we have the following statement for AR(1) progression due to Theorem 1.

Theorem 2. *Suppose $\{d_k\}$ is a sequence of positive numbers satisfying conditions (C1) and (C2). For the above AR(1) model, if $E\exp\{\gamma|\varepsilon_i\varepsilon_j|\} < \infty$ for some $\gamma > 0$ and any $i, j \geq 1$, for any $x \in \mathbb{R}$, one can obtain*

$$\lim_{n\to\infty} \frac{1}{D_n} \sum_{k=1}^{n} d_k I\left\{ \frac{\sqrt{k}}{\sqrt{Var(\varepsilon_1^2)}} \left(\hat{\sigma}_k^2 - \sigma^2\right) \leq x \right\} = \Phi(x) \qquad a.s.$$

The next example concerns the self-exciting threshold autoregressive (SETAR) progresses.

Example 2. *Let $\{X_i, i \geq p\}$ be a sequence of stationary and geometrically ergodic random variable satisfying the following continuous SETAR(p, l, d) progresses.*

$$X_i = \begin{cases} a_0 + \sum_{m=1}^{p} a_m X_{i-m} + \varepsilon_i, & \text{if } X_{i-d} \in R_1, \\ a_0 + \sum_{m=1}^{p} a_j X_{i-m} + \sum_{k=2}^{j} b_k (X_{i-d} - r_{k-1}) + \varepsilon_i, & \text{if } X_{i-d} \in R_j, j = 2, \ldots, l \end{cases}$$

where $\{\varepsilon_i\}$ is a collection of i.i.d. random variables with zero mean and finite variance σ^2, $R_1, \ldots, R_l$ are the different regions with $R_s = (r_{s-1}, r_s]$ for $1 \leq s \leq l$, and $-\infty = r_0 < r_1 < r_2 < \cdots < r_{l-1} < r_l = +\infty$ are the thresholds. Let $\boldsymbol{\theta}_0 = (a_0, \ldots, a_p, b_2, \ldots, b_l, r_1, \ldots, r_{l-1})^\top \in \Theta \subset \mathbb{R}^q$ be the true parameters of the progresses and $\boldsymbol{\theta} = (\bar{a}_0, \ldots, \bar{a}_p, \bar{b}_2, \ldots, \bar{b}_l, \bar{r}_1, \ldots, \bar{r}_{l-1})^\top$, $\tilde{\mathbf{X}}_i = (X_i, \ldots, X_{i-p+1})^\top$, $q = p + 2l - 1$.

Condition $\mathfrak{C}$ Suppose that $\{\varepsilon_i\}$ has the density h and the density f of X_i is continuous and has a support including the interval $[r_{\min} - \eta, r_{\max} + \eta], \eta > 0$ where $r_{\min} = \min\{\bar{r}_1 : \boldsymbol{\theta} \in \Theta\}$, $r_{\max} = \max\{\bar{r}_{l-1} : \boldsymbol{\theta} \in \Theta\}$. There is some $\varepsilon > 0$ such that $\bar{r}_{k-1} \leq \bar{r}_k - \varepsilon$ for all $\boldsymbol{\theta} \in \Theta$ and $k = 2, \ldots, l$.

By Corollary 3.1 of Liebscher [2], under Condition $\mathfrak{C}$ and $E|\varepsilon_i|^\gamma < \infty$, $E||\tilde{\mathbf{X}}_i||^\gamma < \infty$, $\gamma > 4$, the Assumption 2 holds. Therefore, we have the following result for SETAR progresses due to Theorem 1.

Theorem 3. *Suppose $\{d_k\}$ is a sequence of positive numbers satisfying conditions (C1) and (C2). For the above SETAR(p, l, d) progresses, under Assumption 1 and Condition $\mathfrak{C}$ and $E|\varepsilon_i|^\gamma < \infty$, $E||\tilde{\mathbf{X}}_i||^\gamma < \infty$, $\gamma > 4$, for any $x \in \mathbb{R}$, one can obtain*

$$\lim_{n\to\infty} \frac{1}{D_n} \sum_{k=1}^{n} d_k I\left\{ \frac{\sqrt{k}}{\sqrt{Var(\varepsilon_1^2)}} \left(\hat{\sigma}_k^2 - \sigma^2 \right) \le x \right\} = \Phi(x) \qquad a.s.$$

Next, we will consider the threshold-exponential AR progresses.

Example 3. *Let R_j, $j = 1, \cdots, K$ be non-overlapping and non-empty intervals of $\mathbb{R}$ such that $\bigcup_j R_j = \mathbb{R}$. A combined threshold-exponential AR progresses is defined by*

$$X_i = \sum_{j=1}^{K} (\alpha_j + \beta_j X_{i-1}) I\{X_{i-1} \in R_j\} + c e^{-\gamma X_{i-1}^2} X_{i-1} + \varepsilon_i,$$

with $X_0 = x_0$ and $\{\varepsilon_i\}$ is a collection of i.i.d. random variables with zero mean. Let the true parameters $\boldsymbol{\theta}_0 = (\alpha_1, \cdots, \alpha_K, \beta_1, \cdots, \beta_K, c, \gamma)^\top \in \Theta \subset \mathbb{R}^q$ and the parameters $\boldsymbol{\theta} = (\overline{\alpha}_1, \cdots, \overline{\alpha}_K, \overline{\beta}_1, \cdots, \overline{\beta}_K, \overline{c}, \overline{\gamma})^\top$ with $q = 2K + 2$.

For Assumption 2, if $c \neq 0$, $\gamma > 0$, $|\beta_j| < 1$, $j = 1, \cdots, K$ and $E|\varepsilon_i|^{2+\delta} < \infty$ for some $\delta > 0$, by Theorem 4 of Yao [28], (5) holds for the least squares estimator $\hat{\boldsymbol{\theta}}$. Therefore, we have the following statement for threshold-exponential AR progresses due to Theorem 1.

Theorem 4. *Suppose $\{d_k\}$ is a sequence of positive numbers satisfying conditions (C1) and (C2). For the above threshold-exponential AR progresses, if $c \neq 0$, $\gamma > 0$, $|\beta_j| < 1$, $j = 1, \cdots, K$ and $E|\varepsilon_i|^{2+\delta} < \infty$ for some $\delta > 0$ and $E\varepsilon_1^4 < \infty$, then under Assumption 1, for any $x \in \mathbb{R}$, one can obtain*

$$\lim_{n\to\infty} \frac{1}{D_n} \sum_{k=1}^{n} d_k I\left\{ \frac{\sqrt{k}}{\sqrt{Var(\varepsilon_1^2)}} \left(\hat{\sigma}_k^2 - \sigma^2 \right) \le x \right\} = \Phi(x) \qquad a.s.$$

Next, we will consider the multilayer perceptrons progress.

Example 4. *Multilayer perceptrons progressesw have become popular in nonlinear modeling due to its universal approximation ability. Such an example is the model described below which has p input units feeding by variables $X_{i-1}, \cdots, X_{i-p}$ at time i, a hidden layer with K units and one output unit which provides the variable X_i*

$$X_i = \sum_{j=1}^{K} \alpha_j \psi\left(\sum_{l=1}^{p} \beta_{lj} X_{i-l} + \beta_{0j} \right) + \alpha_0 + \varepsilon_i,$$

where $\{\varepsilon_i\}$ is a collection of i.i.d. random variables with zero mean. Let the true parameters $\boldsymbol{\theta}_0 = (\alpha_0, \cdots, \alpha_K, \beta_{lj}, 0 \le l \le p, 1 \le j \le K)^\top \in \Theta \subset \mathbb{R}^q$ and the parameters $\boldsymbol{\theta} = (\overline{\alpha}_0, \cdots, \overline{\alpha}_K, \overline{\beta}_{lj}, 0 \le l \le p, 1 \le j \le K)^\top$ with $q = 1 + K(p+1)$.

For Assumption 2 and sigmoid map $\psi(x) = tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$, if for all $\boldsymbol{\theta}$ are different from $\boldsymbol{\theta}_0$, there exists $x \in \mathbb{R}^p$ such that $r_{\boldsymbol{\theta}}(x) \neq r_{\boldsymbol{\theta}_0}(x)$, $E|\varepsilon_i|^{6+\delta} < \infty$ for some $\delta > 0$ and the matrix I_0 is regular, where

$$I_0 = 2 \int_{\mathbb{R}^p} M_{\boldsymbol{\theta}_0}(x) \mu_{\boldsymbol{\theta}_0}(dx), \;\; M_{\boldsymbol{\theta}}(x) = \left(\frac{\partial r_{\boldsymbol{\theta}}(x)}{\partial \theta_i} \cdot \frac{\partial r_{\boldsymbol{\theta}}(x)}{\partial \theta_j}\right)_{1 \le i,j \le q},$$

then by Theorem 5 of Yao [28], (5) holds for the least squares estimator $\hat{\boldsymbol{\theta}}$. Therefore, we have the following statement for multilayer perceptrons due to Theorem 1.

Theorem 5. *Suppose $\{d_k\}$ is a sequence of positive numbers satisfying conditions (C1) and (C2). For the univariate multilayer perceptrons progress with $\psi(x) = tanh(x)$, if for all $\boldsymbol{\theta}$ different from $\boldsymbol{\theta}_0$, there exists $x \in \mathbb{R}^p$ such that $r_{\boldsymbol{\theta}}(x) \neq r_{\boldsymbol{\theta}_0}(x)$, $E|\varepsilon_i|^{6+\delta} < \infty$ for some $\delta > 0$ and the matrix I_0 is regular, then under Assumption 1, for any $x \in \mathbb{R}$, one can obtain*

$$\lim_{n\to\infty} \frac{1}{D_n} \sum_{k=1}^{n} d_k I \left\{ \frac{\sqrt{k}}{\sqrt{Var(\varepsilon_1^2)}} \left(\hat{\sigma}_k^2 - \sigma^2 \right) \leq x \right\} = \Phi(x) \qquad a.s.$$

6. Conclusions

In this paper, using Taylor's expansion, the Borel–Cantelli lemma and the classical almost sure central limit theorem for independent random variables, the authors establish the almost sure central limit theorem for the error variance estimator for nonlinear autoregressive progresses with independent and identical distributed errors. The results extend the almost sure central limit theorem for the error variance estimator to the nonlinear autoregressive progresses. Four examples, first-order autoregressive processes, self-exciting threshold autoregressive processes, threshold-exponential AR progresses and multilayer perceptrons progress, are given to verify the results. In the future, we will try to investigate the almost sure central limit theorem for the error variance estimator for nonlinear autoregressive progresses with dependent errors and the moderate deviation principle for the error variance estimator for nonlinear autoregressive progresses with independent errors.

Author Contributions: Conceptualization, K.L.; formal analysis, K.L.; methodology, Y.Z.; validation, K.L. and Y.Z.; visualization, K.L. and Y.Z.; writing—original draft preparation, K.L.; writing—review and editing, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant No. 12171198, 11771178); the Science and Technology Development Program of Jilin Province (Grant No. 20210101467JC); and the Science and Technology Program of Jilin Educational Department during the "14th Five-Year" Plan Period (Grant No. JJKH20241239KJ).

Data Availability Statement: Data sharing is not applicable to this article as no datasets were generated or analyzed during the current (theoretical) study.

Acknowledgments: We would like to thank the anonymous referees for a very careful reading of the article.

Conflicts of Interest: The authors declare no conflicts of interest in this paper.

References

1. Tong, H. *Non-Linear Time Series: A Dynamical Approach*; Oxford University Press: New York, NY, USA, 1990.
2. Liebscher, E. Strong convergence of estimators in nonlinear autoregressive models. *J. Multivar. Anal.* **2003**, *84*, 247–261. [CrossRef]
3. Cheng, F.; Sun, S. A goodness-of-fit test of the errors in nonlinear autoregressive time series models. *Stat. Probab. Lett.* **2008**, *78*, 50–59. [CrossRef]
4. Fu, K.; Yang, X. Asymptotics of kernel density estimators in nonlinear autoregressive models. *J. Math. Chem.* **2008**, *44*, 831–838. [CrossRef]
5. Cheng, F. Global property of error density estimation in nonlinear autoregressive time series models. *Stat. Inference Stoch. Process.* **2010**, *13*, 43–53. [CrossRef]
6. Li, J. Asymptotics of the L_p-norms of density estimators in the nonlinear autoregressive models. *Commun. Stat. Theory Methods* **2014**, *43*, 4845–4855. [CrossRef]
7. Kim, K.; Sin, M.; Kim, O. A goodness-of-fit test of the errors in nonlinear autoregressive time series models with stationary α-mixing error terms. *ROMAI J.* **2014**, *10*, 63–70.
8. Cheng, F. Strong consistency of the distribution estimator in the nonlinear autoregressive time series. *J. Multivar. Anal.* **2015**, *142*, 41–47. [CrossRef]
9. Liu, T.; Zhang, Y. Law of the iterated logarithm for error density estimators in nonlinear autoregressive models. *Commun. Stat. Theory Methods* **2020**, *49*, 1082–1098. [CrossRef]

10. Cheng, F. Variance estimation in nonlinear autoregressive time series models. *J. Stat. Plann. Inference* **2011**, *141*, 1588–1592. [CrossRef]
11. Brosamler, G. An almost everywhere central limit theorem. *Math. Proc. Camb. Philos. Soc.* **1988**, *104*, 561–574. [CrossRef]
12. Schatte, P. On strong versions of the central limit theorem. *Math. Nachrichten* **1988**, *137*, 249–256. [CrossRef]
13. Peligrad, M.; Reévész, P. On the almost sure central limit theorem. In *Almost Everywhere Convergence*; Academic Press: Boston, MA, USA, 1989; Volume II, pp. 209–225.
14. Berkes, I.; Csáki, E. A universal result in almost sure central limit theory. *Stoch. Process. Appl.* **2001**, *94*, 105–134. [CrossRef]
15. Hörmann, S. An extension of almost surecentral limit theory. *Stat. Probab. Lett.* **2006**, *76*, 191–202.
16. Tong, B.; Peng, Z.; Saralees, N. An extension of almost surecentral limit theorem for order statistics. *Extremes* **2009**, *12*, 201–209.
17. Miao, Y. An extension of almost surecentral limit theory for the product of partial sums. *J. Dyn. Syst. Geom. Theor.* **2009**, *7*, 49–60.
18. Li, Y. An extension of the almost sure central limit theorem for products of sums under association. *Commun. Stat. Theory Methods* **2013**, *42*, 478–490. [CrossRef]
19. Zhang, Y. A universal result in almost sure central limit theorem for products of sums of partial sums under mixing sequence. *Stochastics* **2016**, *88*, 803–812. [CrossRef]
20. Zhang, Y. An extension of almost sure central limit theorem for self-normalized products of sums for mixing sequences. *Commun. Stat. Theory Methods* **2016**, *45*, 6625–6640. [CrossRef]
21. Wu, Q.; Jiang, Y. Almost sure central limit theorem for self-normalized partial sums and maxima. *Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat. RACSAM* **2016**, *110*, 699–710. [CrossRef]
22. Li, J.Y.; Zhang, Y. An almost sure central limit theorem for the stochastic heat equation. *Stat. Probab. Lett.* **2021**, *177*, 109149. [CrossRef]
23. Li, J.; Zhang, Y. An almost sure central limit theorem for the parabolic Anderson model with delta initial condition. *Stochastics* **2023**, *95*, 483–500. [CrossRef]
24. Li, J.; Zhang, Y. Almost sure central limit theorems for stochastic wave equations. *Electron. Commun. Probab.* **2023**, *28*, 9. [CrossRef]
25. Klimko, L.; Nelson, P. On conditional least squares estimation for stochastic processes. *Ann. Stat.* **1978**, *6*, 629–642. [CrossRef]
26. Wang, Y.; Mao, M.; Hu, X.; He, T. The law of iterated logarithm for autoregressive processes. *Math. Probl. Eng.* **2014**, *2014*, 972712. [CrossRef]
27. Chan, K.; Tong, H. On estimating thresholds in autoregressive models. *J. Time Ser. Anal.* **1986**, *7*, 179–190. [CrossRef]
28. Yao, J. On least squares estimation for stable nonlinear AR processes. *Ann. Inst. Stat. Math.* **2000**, *52*, 316–331. [CrossRef]
29. Hall, P.; Heyde, C. *Martingale Limit Theory and its Application*; Academic Press: New York, NY, USA, 1980.
30. Zhang, Y. Further research on limit theorems for self-normalized sums. *Commun. Stat. Theory Methods* **2020**, *49*, 385–402. [CrossRef]
31. Peligrad, M.; Shao, Q. A note on the almost sure central limit theorem. *Stat. Probab. Lett.* **1995**, *22*, 131–136. [CrossRef]
32. Billingsley, P. *Convergence of Probability Measures*; Wiley: New York, NY, USA, 1968.