axioms

MDPI

Article

Invariance Principle and Berry–Esseen Bound for Error Variance Estimator for Pth-Order Nonlinear Autoregressive Models

Kaiyu Liang [1,2], Yong Zhang [1] and Xue Ding [1,*]

[1] School of Mathematics, Jilin University, Changchun 130012, China
[2] Jilin Provincial Institute of Education, Changchun 130022, China
* Correspondence: dingxue83@jlu.edu.cn

Abstract: The invariance principle and Berry–Esseen bound for an error variance estimator based on the residuals are established by using a Taylor expansion and the classical invariance principle and Berry–Esseen bound for independent random variables. Some examples are given to illustrate their applications.

Keywords: nonlinear autoregressive models; invariance principle; Berry–Esseen bound; error variance estimator; residuals

MSC: 60F17; 60F05

1. Introduction and Main Results

Let $\{X_k, k \in \mathbb{Z}\}$ be a sequence of strictly stationary random variables that fulfills the following nonlinear autoregressive models of order p (NAR(p), for short):

$$X_k = g_{\boldsymbol{\theta}}\left(X_{k-1}, \ldots, X_{k-p}\right) + \varepsilon_k, \tag{1}$$

where $g_{\boldsymbol{\theta}} : \mathbb{R}^p \to \mathbb{R}$ is a collection of known measurable functions, $\boldsymbol{\theta} = \left(\theta_1, \ldots, \theta_q\right)' \in \Theta \subset \mathbb{R}^q$, $\{\varepsilon_k, k \in \mathbb{Z}\}$ is a collection of mean zero i.i.d. random variables with finite variance σ^2.

In recent years, an increasing number of scholars have been studying the error sequence in Model (1). One of the research directions is an error density estimator, and some classic results can be found in Liebscher [1], Cheng and Sun [2], Fu and Yang [3], Cheng [4], Hilgert and Portier [5], Kim et al. [6], Li [7], Cheng [8], Liu and Zhang [9], and Wu et al. [10]. Another research direction is an error variance estimator. There are fewer results for the error variance estimator compared to the error density estimator. Cheng [11] and Liang and Zhang [12] discussed asymptotic normality and almost sure central limit theorem for the error variance estimator from NAR(p) with i.i.d. errors, respectively.

The invariance principle can be seen mathematically as the functional central limit theorem, which is more effective than the classical central limit theorem. By combining the invariance principle with the continuity of convergence according to the distribution, one can obtain the asymptotic distribution of various statistics. The Berry–Esseen bound is commonly used to represent the speed at which the distribution function $F_n(x)$ of the regularized sum of the first n terms of a random variable sequence $\{X_n, n \geq 1\}$ converges to zero compared to its limit distribution $\Phi(x)$. They have a wide range of applications in disciplines such as mathematical statistics, econometrics, and financial mathematics.

The purpose of this article is to establish the invariance principle and Berry–Esseen bound of the estimator of the error variance σ^2 by using the observation value $\{X_1, X_2, \cdots, X_n\}$ from Model (1). The biggest inconvenience is that the error sequence and the residual sequence cannot be observed directly. Taylor's expansion and some technical methods are needed to solve it. Firstly, we provide an estimator $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_q)'$ of $\boldsymbol{\theta}$. Secondly, on the basis of $\hat{\boldsymbol{\theta}}$ and Model (1) above, the residuals below are established.

check for updates

Citation: Liang, K.; Zhang, Y.; Ding, X. Invariance Principle and Berry–Esseen Bound for Error Variance Estimator for Pth-Order Nonlinear Autoregressive Models. *Axioms* **2024**, *13*, 746. https://doi.org/10.3390/axioms13110746

Academic Editor: Hari Mohan Srivastava

Received: 9 September 2024
Revised: 22 October 2024
Accepted: 28 October 2024
Published: 30 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

$$\hat{\varepsilon}_k = X_k - g_{\hat{\boldsymbol{\theta}}}\left(X_{k-1}, \ldots, X_{k-p}\right), \ \ k = 1, 2, \ldots, n. \tag{2}$$

Finally, by (2), the estimator of σ^2 is given by

$$\hat{\sigma}^2_{[nt]} = \frac{1}{n}\sum_{k=1}^{[nt]} \hat{\varepsilon}_k^2, \ \ 0 \le t \le 1. \tag{3}$$

The underlying assumptions below for Model (1) are needed throughout this paper.

Assumption 1A. *For arbitrary $z \in \mathbb{R}^p$, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_q)' \in \mathbb{V} \subset \Theta \subset \mathbb{R}^q$, suppose that*

$$\left|\frac{\partial}{\partial \theta_j} g_{\boldsymbol{\theta}}(z)\right| \le N_1(z), \ \ \left|\frac{\partial^2}{\partial \theta_j \partial \theta_l} g_{\boldsymbol{\theta}}(z)\right| \le N_2(z),$$

where $E[N_1^2(X_{k-1}, \ldots, X_{k-p})] < \infty$ and $E[N_2^2(X_{k-1}, \ldots, X_{k-p})] < \infty$ for every $k \ge 1$, $1 \le j, l \le q$.

Assumption 1B.

$$|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}|^2 = \sum_{j=1}^{q} (\hat{\theta}_j - \theta_j)^2 \le C\frac{\log\log n}{n}, \ \ a.s. \tag{4}$$

where $\hat{\boldsymbol{\theta}}$ is a strong consistent estimator of $\boldsymbol{\theta}$ and $C > 0$ is a constant.

Remark 1. **Assumption 1B** *can be satisfied for some models. Please see the following corollaries for details.*

We will state the first result concerning the invariance principle for error variance estimator $\hat{\sigma}^2_{[nt]}$.

Theorem 1. *For Model (1), under* **Assumptions 1A** *and* **1B**, *if $E\varepsilon_1 = 0$, $E\varepsilon_1^2 = \sigma^2$ and $E\varepsilon_1^4 < \infty$, then*

$$\sup_{0 \le t \le 1} \frac{n(\hat{\sigma}^2_{[nt]} - \frac{[nt]}{n}\sigma^2)}{\sqrt{nVar(\varepsilon_1^2)}} \xrightarrow{\mathrm{d}} \sup_{0 \le t \le 1} W(t), \ \ n \to \infty. \tag{5}$$

particularly for $t = 1$.

$$\sqrt{n}(\hat{\sigma}_n^2 - \sigma^2) \xrightarrow{\mathrm{d}} \mathcal{N}(0, Var(\varepsilon_1^2)), \ \ n \to \infty.$$

The second result is the Berry–Esseen bound for error variance estimator $\hat{\sigma}_n^2$.

Theorem 2. *For Model (1), under Assumptions A and B, if $E\varepsilon_1 = 0$, $E\varepsilon_1^2 = \sigma^2$ and $E\varepsilon_1^{4+2\delta} < \infty$, for $0 < \delta \le 1$, we have*

$$\sup_{z \in \mathbb{R}} \left| P\left\{ \frac{\sqrt{n}}{\sqrt{Var(\varepsilon_1^2)}} \left(\hat{\sigma}_n^2 - \sigma^2\right) \le z \right\} - \Phi(z) \right| \le C\left(\frac{1}{n^{\delta/2}} + \frac{(\log\log n)^{2/3}}{n^{1/3}} \right), \tag{6}$$

where $\Phi(\cdot)$ stands for the distribution function of the standard normal distribution.

Remark 2. *Theorem 1 generalizes Theorem 2.2 of Cheng [11] from the central limit theorem to the invariance principle for error variance estimator* $\hat{\sigma}^2_{[nt]}$.

Some typical examples are shown to test the invariance principle and Berry–Esseen bound for the error variance estimator for some special NAR(p). The first one is a degenerate case, which concerns AR(1) models.

Example 1. *Consider the following AR(1) model:*

$$X_k = \theta X_{k-1} + \varepsilon_k, \ \ k \geq 1,$$

where $|\theta| < 1$.

Assumption 1A obviously holds for AR(1) model. **Assumption 1B** can be satisfied for the least squares estimator $\hat{\theta}$ from Theorem 1 of Wang et al. [13] under the condition $E\exp\{\alpha|\varepsilon_k\varepsilon_l|\} < \infty$. Thus, Theorems 1 and 2 hold for the AR(1) model. We summarize it as the following corollary:

Corollary 1. *For the AR(1) model, if* $E\exp\{\alpha|\varepsilon_k\varepsilon_l|\} < \infty$ *for some* $\alpha > 0$ *and arbitrary* $k, l \geq 1$, *one can obtain*

$$\sup_{0\leq t\leq 1} \frac{n(\hat{\sigma}^2_{[nt]} - \frac{[nt]}{n}\sigma^2)}{\sqrt{nVar(\varepsilon_1^2)}} \xrightarrow{\mathrm{d}} \sup_{0\leq t\leq 1} W(t), \ \ n \to \infty,$$

and

$$\sup_{z\in\mathbb{R}} \left| P\left\{ \frac{\sqrt{n}}{\sqrt{Var(\varepsilon_1^2)}} \left(\hat{\sigma}_n^2 - \sigma^2\right) \leq z \right\} - \Phi(z) \right| \leq \frac{C}{n^{\delta/2}} + \frac{C(\log\log n)^{2/3}}{n^{1/3}},$$

where $0 < \delta \leq 1$.

The second example encompasses the self-exciting threshold autoregressive (SETAR) models.

Example 2. *Let* $\{X_k, k \geq p\}$ *satisfy the continuous SETAR*(p, l, d) *models, as follows:*

$$X_k = \begin{cases} a_0 + \sum_{m=1}^{p} a_m X_{k-m} + \varepsilon_k & \text{if } \ X_{k-d} \in R_1, \\ a_0 + \sum_{m=1}^{p} a_m X_{k-m} + \sum_{n=2}^{j} b_n (X_{k-d} - r_{n-1}) + \varepsilon_k & \text{if } \ X_{k-d} \in R_j, j = 2, \ldots, l \end{cases}$$

where $\{R_n = (r_{n-1}, r_n], 1 \leq n \leq l\}$ *are the disjoint intervals, and* $-\infty = r_0 < r_1 < r_2 < \cdots < r_{l-1} < r_l = +\infty$ *are the thresholds.* $\boldsymbol{\theta}_0 = (a_0, \ldots, a_p, b_2, \ldots, b_l, r_1, \ldots, r_{l-1})' \in \Theta \subset \mathbb{R}^q$ *denote the true parameters of the models, and* $\boldsymbol{\theta} = (\bar{a}_0, \ldots, \bar{a}_p, \bar{b}_2, \ldots, \bar{b}_l, \bar{r}_1, \ldots, \bar{r}_{l-1})' \in \Theta$ *denote the unknown parameters,* $\tilde{\boldsymbol{X}}_k = (X_k, \ldots, X_{k-p+1})'$, $q = 2l + p - 1$. $\{X_{k-d}, k \geq p\}$ *is the threshold process (sometimes referred to as a switching variable), which controls the switching or jump between the regime* R_j. *The integer* $0 < d \leq p$ *is called a delay parameter. The integer* $0 < d \leq p$ *is called a delay parameter. The SETAR*(p, l, d) *is an AR*(p) *model in each of the several regimes. As such, it is a piecewise model which is linear in each regime* R_j, *but the overall time series model is nonlinear. The piecewise nature of the SETAR model is able to capture some important nonlinear phenomena, such as sudden jumps, sub and higher harmonics, asymmetric limit cycles and chaos, and amplitude-dependent frequency.*

Assumption 1C. *Assume that* $\{\varepsilon_i\}$ *has the density function* h *and that the continuous density function* f *of* X_i *has a support including the interval* $[r_{\min} - \eta, r_{\max} + \eta]$, *where* $r_{\min} = \min\{\bar{r}_1 : \boldsymbol{\theta} \in \Theta\}$, $r_{\max} = \max\{\bar{r}_{l-1} : \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$, $\eta > 0$. $\bar{r}_{n-1} \leq \bar{r}_n - \epsilon$ *for every* $\boldsymbol{\theta} \in \Theta$ *and* $n = 2, \ldots, l$ *and some* $\epsilon > 0$.

Assumption 1B can be satisfied by using Corollary 3.1 of Liebscher [1], under $E|\varepsilon_k|^\beta < \infty$, $E||\tilde{\boldsymbol{X}}_k||^\beta < \infty$ and **Assumption 1C**, $\beta > 4$.

Thus, Theorems 1 and 2 hold for the SETAR(p, l, d) model. We summarize it as the following corollary:

Corollary 2. *For SETAR*(p, l, d) *models, under* **Assumption 1A** *and* **Assumption 1C**, *if* $E||\tilde{\boldsymbol{X}}_k||^\beta < \infty$, $E|\varepsilon_k|^{(4+2\delta)\vee\beta} < \infty$, $\beta > 4$, $0 < \delta \leq 1$; *thus,*

$$\sup_{0\leq t\leq 1} \frac{n(\hat{\sigma}^2_{[nt]} - \frac{[nt]}{n}\sigma^2)}{\sqrt{nVar(\varepsilon_1^2)}} \xrightarrow{\mathrm{d}} \sup_{0\leq t\leq 1} W(t), \quad n \to \infty,$$

and

$$\sup_{z\in\mathbb{R}} \left| P\left\{ \frac{\sqrt{n}}{\sqrt{Var(\varepsilon_1^2)}} \left(\hat{\sigma}_n^2 - \sigma^2\right) \leq z \right\} - \Phi(z) \right| \leq \frac{C}{n^{\delta/2}} + \frac{C(\log\log n)^{2/3}}{n^{1/3}}.$$

The third example encompasses the threshold-exponential AR models.

Example 3. *The threshold-exponential AR models are defined below*

$$X_k = \sum_{l=1}^{K} (\alpha_l + \beta_l X_{k-1}) I\{X_{k-1} \in R_l\} + ce^{-\gamma X_{k-1}^2} X_{k-1} + \varepsilon_k,$$

with $X_0 = x_0$ *and* $\{R_l,\ 1 \leq l \leq K\}$ *being non-empty and non-overlapping intervals of* $\mathbb{R}$ *with* $\bigcup_l R_l = \mathbb{R}$. $\boldsymbol{\theta}_0 = (\alpha_1, \cdots, \alpha_K, \beta_1, \cdots, \beta_K, c, \gamma)' \in \Theta \subset \mathbb{R}^q$ *denote the true parameters,* $\boldsymbol{\theta} = (\overline{\alpha}_1, \cdots, \overline{\alpha}_K, \overline{\beta}_1, \cdots, \overline{\beta}_K, \overline{c}, \overline{\gamma})'$ *denote the unknown parameters, and* $\{X_{k-1}\}$ *is the threshold variable,* $q = 2K + 2$.

By Theorem 4 of Yao [14], **Assumption 1B** can be satisfied if $c \neq 0$, $\gamma > 0$, $|\beta_l| < 1$, $l = 1, \cdots, K$ and $E|\varepsilon_k|^{2+\eta} < \infty$ for some $\eta > 0$. Thus, Theorems 1 and 2 hold for threshold-exponential AR progresses. We summarize it as the following corollary:

Corollary 3. *For threshold-exponential AR models, if* $c \neq 0$, $\gamma > 0$, $|\beta_l| < 1$, $l = 1, \cdots, K$, *then under* **Assumption 1A**, *for* $E|\varepsilon_1|^{(2+\eta)\vee 4} < \infty$ *for some* $\eta > 0$, *one can obtain*

$$\sup_{0\leq t\leq 1} \frac{n(\hat{\sigma}^2_{[nt]} - \frac{[nt]}{n}\sigma^2)}{\sqrt{nVar(\varepsilon_1^2)}} \xrightarrow{\mathrm{d}} \sup_{0\leq t\leq 1} W(t), \quad n \to \infty,$$

for $E|\varepsilon_1|^{(2+\eta)\vee(4+2\delta)} < \infty$ *for some* $\eta > 0$ *and* $0 < \delta \leq 1$, *one can obtain*

$$\sup_{z\in\mathbb{R}} \left| P\left\{ \frac{\sqrt{n}}{\sqrt{Var(\varepsilon_1^2)}} \left(\hat{\sigma}_n^2 - \sigma^2\right) \leq z \right\} - \Phi(z) \right| \leq \frac{C}{n^{\delta/2}} + \frac{C(\log\log n)^{2/3}}{n^{1/3}}.$$

The last example is the multilayer perceptron model.

Example 4. *Consider the following multilayer perceptrons model:*

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

where $\boldsymbol{\theta}_0 = (\alpha_0, \cdots, \alpha_K, \beta_{lj}, 0 \le l \le p, 1 \le j \le K)' \in \Theta \subset \mathbb{R}^q$ *denote the the true parameters,* $\boldsymbol{\theta} = (\overline{\alpha}_0, \cdots, \overline{\alpha}_K, \overline{\beta}_{lj}, 0 \le l \le p, 1 \le j \le K)'$ *denote the unknown parameters, and* $\psi(\cdot)$ *is called an activation function,* $q = 1 + K(p+1)$*. The above models can describe that it has one output unit which provides the variable* X_k*,* p *input units being fed by variables* $X_{k-1}, \cdots, X_{k-p}$ *at time* k *and a hidden layer with* K *units. Due to its universal approximation ability, it has become popular in nonlinear models. We will take* $\psi(x) = tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ *in this example as in Yao [14].*

By Theorem 5 of Yao [14], **Assumption 1B** can be satisfied if $E|\varepsilon_k|^{6+\eta} < \infty$ for some $\eta > 0$, $r_{\boldsymbol{\theta}}(x) \neq r_{\boldsymbol{\theta}_0}(x)$ for every $\boldsymbol{\theta}$ different from $\boldsymbol{\theta}_0$ and some $x \in \mathbb{R}^p$, $\psi(x) = tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ and the matrix I_0 is regular, where

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

Thus Theorems 1 and 2 hold for multilayer perceptrons models. We summarize it as the following corollary:

Corollary 4. *For multilayer perceptrons models with* $E|\varepsilon_k|^{6+\eta} < \infty$ *for some* $\eta > 0$, $r_{\boldsymbol{\theta}}(x) \neq r_{\boldsymbol{\theta}_0}(x)$ *for every* $\boldsymbol{\theta}$ *different from* $\boldsymbol{\theta}_0$ *and some* $x \in \mathbb{R}^p$, $\psi(x) = tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$*, and with the matrix* I_0 *being regular, if* **Assumption 1A** *holds, then one can obtain*

$$\sup_{0 \le t \le 1} \frac{n(\hat{\sigma}^2_{[nt]} - \frac{[nt]}{n}\sigma^2)}{\sqrt{nVar(\varepsilon_1^2)}} \xrightarrow{\mathrm{d}} \sup_{0 \le t \le 1} W(t), \;\; n \to \infty.$$

and

$$\sup_{z \in \mathbb{R}} \left| P\left\{ \frac{\sqrt{n}}{\sqrt{Var(\varepsilon_1^2)}} \left(\hat{\sigma}_n^2 - \sigma^2\right) \le z \right\} - \Phi(z) \right| \le \frac{C}{n^{\delta/2}} + \frac{C(\log\log n)^{2/3}}{n^{1/3}},$$

where $0 < \delta \le 1$.

The structure of this article is as follows. The background and main results of this paper are declared in Section 1. Auxiliary lemmas are stated in Section 2. The proofs are given in Section 3. Throughout this paper, C denotes a constant that may be distinctly subjected to its locations. Let $I\{A\}$ be the indicator function of the collection A.

2. Preliminary Lemmas

We will use the following lemmas:

Lemma 1 (Hall and Heyde [15])**.** *Let* $Y_1 = T_1$ *and* $Y_j = T_j - T_{j-1}$, $2 \le j \le n$*. If* $\{T_j, \mathscr{F}_j, 1 \le j \le n\}$ *is a martingale and* $q > 0$*, then*

$$E\left(\max_{1 \le j \le n} |T_j|^q\right) \le C_q \left\{ E\left[\left(\sum_{j=1}^{n} E(Y_j^2 | \mathscr{F}_{j-1})\right)^{q/2}\right] + E\left(\max_{1 \le j \le n} |Y_i|^q\right) \right\}.$$

where C_q *is a constant depending only on* q*.*

Lemma 2 (Liang and Zhang [12]). *Under the assumptions of Theorem 1 or 2, for* $1 \le i \le n$, $1 \le j, l \le q$; *then, for any* $2 \le s \le 4$, *one can obtain*

$$E|Y_{kj}|^s \le EN_1^s(X_{k-1}, \ldots, X_{k-p}) \le \left(EN_1^4(X_{k-1}, \ldots, X_{k-p})\right)^{s/4} < \infty,$$
$$E|Z_{kjl}|^s \le EN_2^s(X_{k-1}, \ldots, X_{k-p}) \le \left(EN_2^4(X_{k-1}, \ldots, X_{k-p})\right)^{s/4} < \infty,$$

where

$$Y_{kj} = \frac{\partial}{\partial \theta_j} g_{\boldsymbol{\theta}}(X_{k-1}, \ldots, X_{k-p}), \;\; Z_{kjl} = \frac{\partial^2}{\partial \theta_j \partial \theta_l} g_{\boldsymbol{\theta}^*}(X_{k-1}, \ldots, X_{k-p}),$$

$\boldsymbol{\theta}^* = \boldsymbol{\theta} + \mu(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ *for some* $\mu \in (0, 1)$.

Lemma 3 (Chao and Rao [16]). *Let* X_1, X_2 *be two random variables; then, for any* $z \in \mathbb{R}$ *and* $b > 0$,

$$\sup_{z \in \mathbb{R}} |P(X_1 + X_2 \le z) - \Phi(z)| \le \sup_{z \in \mathbb{R}} |P(X_1 \le z) - \Phi(z)| + \frac{b}{\sqrt{2\pi}} + P(|X_2| > b).$$

Lemma 4 (Petrov [17]). *Let* $X_1, \cdots, X_n$ *be independent random variables with zero means and* $E|X_k|^{2+\delta} < \infty$ *for* $0 < \delta \le 1$, $1 \le k \le n$. *Then, there exists a positive constant* A *such that*

$$\sup_{z \in \mathbb{R}} \left| P\left\{ \frac{1}{B_n^{1/2}} \sum_{k=1}^{n} X_k \le z \right\} - \Phi(z) \right| \le \frac{A}{B_n^{1+\delta/2}} \sum_{k=1}^{n} E|X_k|^{2+\delta},$$

where $B_n = \sum_{k=1}^{n} E|X_k|^2$.

3. Proofs

Proof of Theorem 1. In order to prove Theorem 1, it suffices to prove

$$\sup_{0 \le t \le 1} \left| \frac{n(\hat{\sigma}^2_{[nt]} - \frac{[nt]}{n}\sigma^2)}{B_n} - \frac{W(nt)}{\sqrt{n}} \right| \xrightarrow{\mathrm{P}} 0, \;\; n \to \infty, \tag{7}$$

where $B_n = \sqrt{nVar(\varepsilon_1^2)}$.

Noting (1) and (2), via Taylor's expansion, one can obtain

$$\begin{aligned} \hat{\varepsilon}_k &= \varepsilon_k - [g_{\hat{\boldsymbol{\theta}}}(X_{k-1}, \ldots, X_{k-p}) - g_{\boldsymbol{\theta}}(X_{k-1}, \ldots, X_{k-p})] \\ &= \varepsilon_j - \sum_{j=1}^{q} (\hat{\theta}_j - \theta_j) Y_{kj} - \frac{1}{2} \sum_{j=1}^{q} \sum_{l=1}^{q} (\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l) Z_{kjl}, \end{aligned} \tag{8}$$

where Y_{kj} and Z_{kjl} are defined as in Lemma 2. Using (8), one can conclude the following:

$$\begin{aligned} &\sup_{0 \le t \le 1} \left| \frac{n(\hat{\sigma}^2_{[nt]} - \frac{[nt]}{n}\sigma^2)}{B_n} - \frac{W(nt)}{\sqrt{n}} \right| \\ &= \sup_{0 \le t \le 1} \left| \frac{1}{B_n} \sum_{k=1}^{[nt]} (\hat{\varepsilon}_k^2 - \sigma^2) - \frac{W(nt)}{\sqrt{n}} \right| \\ &= \sup_{0 \le t \le 1} \left| \frac{1}{B_n} \sum_{k=1}^{[nt]} (\hat{\varepsilon}_k^2 - \varepsilon_k^2) - \frac{1}{B_n} \sum_{k=1}^{[nt]} (\varepsilon_k^2 - E\varepsilon_k^2) - \frac{W(nt)}{\sqrt{n}} \right| \\ &\le \sup_{0 \le t \le 1} \frac{1}{B_n} \sum_{k=1}^{[nt]} \left(\sum_{j=1}^{q} (\hat{\theta}_j - \theta_j) Y_{kj} \right)^2 \end{aligned}$$

$$
\begin{aligned}
&+ \sup_{0\le t\le 1} \frac{1}{B_n} \sum_{k=1}^{[nt]} \left(\sum_{j=1}^{q} \sum_{l=1}^{q} (\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l) Z_{kjl} \right)^2 \\
&+ \sup_{0\le t\le 1} \frac{2}{B_n} \left| \sum_{k=1}^{[nt]} \sum_{j=1}^{q} (\hat{\theta}_j - \theta_j) Y_{kj} \varepsilon_k \right| \\
&+ \sup_{0\le t\le 1} \frac{1}{B_n} \left| \sum_{k=1}^{[nt]} \sum_{j=1}^{q} \sum_{l=1}^{q} (\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l) Z_{kjl} \varepsilon_k \right| \\
&+ \sup_{0\le t\le 1} \frac{1}{B_n} \left| \sum_{k=1}^{[nt]} \sum_{j=1}^{q} \sum_{l=1}^{q} \sum_{m=1}^{q} (\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l)(\hat{\theta}_m - \theta_m) Y_{kj} Z_{klm} \right| \\
&+ \sup_{0\le t\le 1} \frac{1}{B_n} \left| \sum_{k=1}^{[nt]} (\varepsilon_k^2 - E\varepsilon_k^2) - \frac{W(nt)}{\sqrt{n}} \right| \\
&=: \sum_{j=1}^{6} G_{nj}. \qquad (9)
\end{aligned}
$$

By (9), in order to obtain (7), it is enough to show

$$G_{nj} \xrightarrow{\mathrm{P}} 0, \quad n \to \infty, \quad 1 \le j \le 6.$$

Noticing the fundamental inequality

$$\left(\sum_{j=1}^{q} c_j d_j \right)^2 \le \left(\sum_{j=1}^{q} c_j^2 \right) \left(\sum_{j=1}^{q} d_j^2 \right). \qquad (10)$$

For G_{n1}, via the Markov inequality, (4) and (10), Lemma 2, and $B_n = \sqrt{nVar(\varepsilon_1^2)}$, we obtain

$$
\begin{aligned}
P(|G_{n1}| > \varepsilon) &= P\left(\sup_{0\le t\le 1} \frac{1}{B_n} \sum_{k=1}^{[nt]} \left(\sum_{j=1}^{q} (\hat{\theta}_j - \theta_j) Y_{kj} \right)^2 > \varepsilon \right) \\
&\le \frac{1}{\varepsilon} \frac{1}{B_n} E \sup_{0\le t\le 1} \left[\sum_{k=1}^{[nt]} \left(\sum_{j=1}^{q} (\hat{\theta}_j - \theta_j) Y_{kj} \right)^2 \right] \\
&\le \frac{1}{\varepsilon} \frac{1}{B_n} E \left[\sum_{j=1}^{q} (\hat{\theta}_j - \theta_j)^2 \cdot \sum_{j=1}^{q} \sum_{k=1}^{n} Y_{kj}^2 \right] \\
&\le \frac{C}{\varepsilon} \frac{1}{B_n} \cdot \frac{\log\log n}{n} \cdot \sum_{j=1}^{q} \sum_{k=1}^{n} E Y_{kj}^2 \\
&\le \frac{C}{\varepsilon} \frac{1}{B_n} \cdot \frac{\log\log n}{n} \cdot n \\
&\le \frac{C}{\varepsilon} \frac{1}{\sqrt{nVar(\varepsilon_1^2)}} \cdot \log\log n \\
&\to 0, \quad n \to \infty. \qquad (11)
\end{aligned}
$$

For G_{n2}, via the Markov inequality, (4) and (10), Lemma 2, and $B_n = \sqrt{nVar(\varepsilon_1^2)}$, we have

$$P(|G_{n2}| > \varepsilon) = P\left(\sup_{0\le t\le 1} \frac{1}{B_n} \sum_{k=1}^{[nt]} \left(\sum_{j=1}^{q} \sum_{l=1}^{q} (\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l) Z_{kjl} \right)^2 > \varepsilon \right)$$

$$
\begin{aligned}
&\leq \frac{1}{\varepsilon}\frac{1}{B_n}E\sup_{0\leq t\leq 1}\left[\sum_{k=1}^{[nt]}\left(\sum_{j=1}^{q}\sum_{l=1}^{q}(\hat{\theta}_j-\theta_j)(\hat{\theta}_l-\theta_l)Z_{kjl}\right)^2\right]\\
&\leq \frac{1}{\varepsilon}\frac{1}{B_n}E\left[\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2\cdot\sum_{l=1}^{q}(\hat{\theta}_l-\theta_l)^2\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{k=1}^{n}Z_{kjl}^2\right]\\
&\leq \frac{C}{\varepsilon}\frac{1}{B_n}\cdot\frac{(\log\log n)^2}{n^2}\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{k=1}^{n}EZ_{kjl}^2\\
&\leq \frac{C}{\varepsilon}\frac{1}{B_n}\cdot\frac{(\log\log n)^2}{n^2}\cdot n\\
&\leq \frac{C}{\varepsilon}\frac{1}{\sqrt{nVar(\varepsilon_1^2)}}\cdot\frac{(\log\log n)^2}{n}\\
&\to 0,\ \ n\to\infty.
\end{aligned}
\tag{12}
$$

For G_{n3}, let

$$
Y_m=\sum_{k=1}^{m}Y_{kj}\varepsilon_k,\quad \mathcal{F}_m=\sigma(\{\varepsilon_k,1\leq k\leq m\}),\ \ 1\leq m\leq n.
$$

By (1) and the independence of $\{\varepsilon_k\}$, one can conclude that Y_{kj} and ε_k are independent; thus, $\{Y_m,\mathcal{F}_m,1\leq m\leq n\}$ is a martingale. Using Lemmas 1 and 2, we know

$$
\begin{aligned}
E\left(\max_{1\leq m\leq n}|Y_m|^2\right)&=E\sup_{0\leq t\leq 1}\left|\sum_{k=1}^{[nt]}Y_{kj}\varepsilon_k\right|^2\\
&\leq CE\left[\left(\sum_{k=1}^{n}E\left(Y_{kj}^2\varepsilon_k^2|\mathcal{F}_{k-1}\right)\right)\right]+CE\left(\max_{1\leq k\leq n}|Y_{kj}\varepsilon_k|^2\right)\\
&\leq CE\left[\left(\sum_{k=1}^{n}EY_{kj}^2\right)\right]\cdot[E\varepsilon_1^2]+C\sum_{k=1}^{n}E|Y_{kj}|^2\cdot E|\varepsilon_1|^2\\
&\leq Cn.
\end{aligned}
\tag{13}
$$

Then, via the Markov inequality, (4), (10), and (13), and $B_n=\sqrt{nVar(\varepsilon_1^2)}$, we know that

$$
\begin{aligned}
&P(|G_{n3}|>\varepsilon)=P\left(\sup_{0\leq t\leq 1}\frac{2}{B_n}\left|\sum_{k=1}^{[nt]}\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)Y_{kj}\varepsilon_k\right|>\varepsilon\right)\\
\leq\ &\frac{1}{\varepsilon^2}\frac{4}{B_n^2}E\sup_{0\leq t\leq 1}\left[\sum_{k=1}^{[nt]}\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)Y_{kj}\varepsilon_k\right]^2\\
\leq\ &\frac{1}{\varepsilon^2}\frac{4}{B_n^2}E\left[\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2\cdot\sum_{j=1}^{q}\sup_{0\leq t\leq 1}\left(\sum_{k=1}^{[nt]}Y_{kj}\varepsilon_k\right)^2\right]\\
\leq\ &\frac{C}{\varepsilon^2}\frac{1}{B_n^2}\cdot\frac{\log\log n}{n}\cdot\sum_{j=1}^{q}E\sup_{0\leq t\leq 1}\left(\sum_{k=1}^{[nt]}Y_{kj}\varepsilon_k\right)^2\\
\leq\ &\frac{C}{\varepsilon^2}\frac{1}{B_n^2}\cdot\frac{\log\log n}{n}\cdot n\\
\leq\ &\frac{C}{\varepsilon^2}\frac{1}{nVar(\varepsilon_1^2)}\cdot\log\log n\\
\to\ &0,\ \ n\to\infty.
\end{aligned}
\tag{14}
$$

For G_{n4}, via the Markov inequality, (4) and (10), C_r inequality, Cauchy–Schwarz inequality, Lemma 2, and $B_n = \sqrt{nVar(\varepsilon_1^2)}$, one can obtain

$$\begin{aligned}
P(|G_{n4}| > \varepsilon) &= P\left(\sup_{0\le t\le 1}\frac{1}{B_n}\left|\sum_{k=1}^{[nt]}\sum_{j=1}^{q}\sum_{l=1}^{q}(\hat{\theta}_j-\theta_j)(\hat{\theta}_l-\theta_l)Z_{kjl}\varepsilon_k\right| > \varepsilon\right)\\
&\le \frac{1}{\varepsilon^2}\frac{1}{B_n^2}E\sup_{0\le t\le 1}\left[\sum_{k=1}^{[nt]}\sum_{j=1}^{q}\sum_{l=1}^{q}(\hat{\theta}_j-\theta_j)(\hat{\theta}_l-\theta_l)Z_{kjl}\varepsilon_k\right]^2\\
&\le \frac{1}{\varepsilon^2}\frac{1}{B_n^2}E\sup_{0\le t\le 1}\left[\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2\cdot\sum_{l=1}^{q}(\hat{\theta}_l-\theta_l)^2\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}\left(\sum_{k=1}^{[nt]}Z_{kjl}\varepsilon_k\right)^2\right]\\
&\le \frac{C}{\varepsilon^2}\frac{1}{B_n^2}\frac{(\log\log n)^2}{n^2}\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}E\sup_{0\le t\le 1}\left(\sum_{k=1}^{[nt]}Z_{kjl}\varepsilon_k\right)^2\\
&\le \frac{C}{\varepsilon^2}\frac{1}{B_n^2}\frac{(\log\log n)^2}{n^2}\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}E\sup_{0\le t\le 1}[nt]\cdot\sum_{k=1}^{[nt]}Z_{kjl}^2\varepsilon_k^2\\
&\le \frac{C}{\varepsilon^2}\frac{1}{B_n^2}\frac{(\log\log n)^2}{n^2}\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}n\cdot\sum_{k=1}^{n}EZ_{kjl}^2\varepsilon_k^2\\
&\le \frac{C}{\varepsilon^2}\frac{1}{B_n^2}\frac{(\log\log n)^2}{n}\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{k=1}^{n}(EZ_{kjl}^4)^{1/2}(E\varepsilon_k^4)^{1/2}\\
&\le \frac{C}{\varepsilon^2}\frac{1}{B_n^2}\frac{(\log\log n)^2}{n}\cdot n\\
&\le \frac{C}{\varepsilon^2}\frac{1}{nVar(\varepsilon_1^2)}\cdot(\log\log n)^2\\
&\to 0,\quad n\to\infty.
\end{aligned}\tag{15}$$

For G_{n5}, via the Markov inequality, (4) and (10), C_r inequality, Cauchy–Schwarz inequality, Lemma 2, and $B_n = \sqrt{nVar(\varepsilon_1^2)}$, it follows that

$$\begin{aligned}
P(|G_{n5}| > \varepsilon) &= P\left(\sup_{0\le t\le 1}\frac{1}{B_n}\left|\sum_{k=1}^{[nt]}\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{m=1}^{q}(\hat{\theta}_j-\theta_j)(\hat{\theta}_l-\theta_l)(\hat{\theta}_m-\theta_m)Y_{kj}Z_{klm}\right| > \varepsilon\right)\\
&\le \frac{1}{\varepsilon^2}\frac{1}{B_n^2}E\sup_{0\le t\le 1}\left[\sum_{k=1}^{[nt]}\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{m=1}^{q}(\hat{\theta}_j-\theta_j)(\hat{\theta}_l-\theta_l)(\hat{\theta}_m-\theta_m)Y_{kj}Z_{klm}\right]^2\\
&\le \frac{1}{\varepsilon^2}\frac{1}{B_n^2}E\sup_{0\le t\le 1}\left[\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2\cdot\sum_{l=1}^{q}(\hat{\theta}_l-\theta_l)^2\cdot\sum_{m=1}^{q}(\hat{\theta}_m-\theta_m)^2\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{m=1}^{q}\left(\sum_{k=1}^{[nt]}Y_{kj}Z_{klm}\right)^2\right]\\
&\le \frac{C}{\varepsilon^2}\frac{1}{B_n^2}\frac{(\log\log n)^3}{n^3}\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{m=1}^{q}E\sup_{0\le t\le 1}\left(\sum_{k=1}^{[nt]}Y_{kj}Z_{klm}\right)^2\\
&\le \frac{C}{\varepsilon^2}\frac{1}{B_n^2}\frac{(\log\log n)^3}{n^3}\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{m=1}^{q}E\sup_{0\le t\le 1}[nt]\cdot\sum_{k=1}^{[nt]}Y_{kj}^2Z_{klm}^2\\
&\le \frac{C}{\varepsilon^2}\frac{1}{B_n^2}\frac{(\log\log n)^3}{n^3}\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{m=1}^{q}n\cdot\sum_{k=1}^{n}EY_{kj}^2Z_{klm}^2\\
&\le \frac{C}{\varepsilon^2}\frac{1}{B_n^2}\frac{(\log\log n)^3}{n^2}\cdot\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{m=1}^{q}\sum_{k=1}^{n}(EY_{kj}^4)^{1/2}(EZ_{kjm}^4)^{1/2}\\
&\le \frac{C}{\varepsilon^2}\frac{1}{B_n^2}\frac{(\log\log n)^3}{n^2}\cdot n
\end{aligned}$$

$$\begin{aligned} &\leq \frac{C}{\varepsilon^2}\frac{1}{nVar(\varepsilon_1^2)}\cdot\frac{(\log\log n)^3}{n} \\ &\to 0, \quad n\to\infty. \end{aligned} \tag{16}$$

By the invariance principle for i.i.d. random variable (see Lu [18]), we know that

$$P(|G_{n6}|>\varepsilon)\to 0, \quad n\to\infty. \tag{17}$$

In view of (9)–(17), we obtain that (9) holds. Hence, the proof of Theorem 1 is completed. □

Proof of Theorem 2. By (8), we can obtain

$$\begin{aligned} &\sqrt{\frac{n}{Var(\varepsilon_1^2)}}\left(\hat{\sigma}^2-\sigma^2\right) \\ =&\sqrt{\frac{n}{Var(\varepsilon_1^2)}}\left(\frac{1}{n}\sum_{ki=1}^{n}\hat{\varepsilon}_k^2-\frac{1}{n}\sum_{k=1}^{n}\varepsilon_k^2+\frac{1}{n}\sum_{k=1}^{n}\varepsilon_k^2-\frac{1}{n}\sum_{k=1}^{n}E\varepsilon_k^2\right) \\ =&\frac{1}{B_n}\left[\sum_{k=1}^{n}(\hat{\varepsilon}_k^2-\varepsilon_k^2)\right]+\frac{1}{B_n}\left[\sum_{k=1}^{n}(\varepsilon_k^2-E\varepsilon_k^2)\right] \\ =&\frac{1}{B_n}\left[\sum_{k=1}^{n}\left(\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)Y_{kj}\right)^2\right]+\frac{1}{B_n}\left[\sum_{k=1}^{n}\left(\sum_{j=1}^{q}\sum_{l=1}^{q}(\hat{\theta}_j-\theta_j)(\hat{\theta}_l-\theta_l)Z_{kjl}\right)^2\right] \\ &-\frac{1}{B_n}\left[\sum_{k=1}^{n}\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)Y_{kj}\varepsilon_k\right]-\frac{1}{B_n}\left[\sum_{k=1}^{n}\sum_{j=1}^{q}\sum_{l=1}^{q}(\hat{\theta}_j-\theta_j)(\hat{\theta}_l-\theta_l)Z_{kjl}\varepsilon_k\right] \\ &-\frac{1}{B_n}\left[\sum_{k=1}^{n}\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{m=1}^{q}(\hat{\theta}_j-\theta_j)(\hat{\theta}_l-\theta_l)(\hat{\theta}_m-\theta_m)Y_{kj}Z_{klm}\right] \\ &+\frac{1}{B_n}\left[\sum_{k=1}^{n}(\varepsilon_k^2-E\varepsilon_k^2)\right] \\ =:&L_{n1}+L_{n2}-L_{n3}-L_{n4}-L_{n5}+L_{n6}, \end{aligned} \tag{18}$$

where $B_n=\sqrt{nVar(\varepsilon_1^2)}$.

Then, by Lemma 3 and (18), one can obtain

$$\begin{aligned} &\sup_{z\in\mathbb{R}}\left|P\left\{\frac{\sqrt{n}}{\sqrt{Var(\varepsilon_1^2)}}\left(\hat{\sigma}_n^2-\sigma^2\right)\leq z\right\}-\Phi(z)\right| \\ \leq &\sup_{z\in\mathbb{R}}|P\{|L_{n6}|\leq z\}-\Phi(z)|+\frac{a}{\sqrt{2\pi}}+P(|L_{n1}+L_{n2}-L_{n3}-L_{n4}-L_{n5}|>a) \\ \leq &\sup_{z\in\mathbb{R}}|P\{|L_{n6}|\leq z\}-\Phi(z)|+\frac{a}{\sqrt{2\pi}}+\sum_{j=1}^{5}P(|L_{nj}|>a/5). \end{aligned} \tag{19}$$

Using a similar proof to that of (11) with $t=1$ and the Markov inequality, we obtain

$$\begin{aligned} &P(|L_{n1}|>a/5)\leq\frac{5}{a}E|L_{n1}| \\ = &\frac{5}{a}\sqrt{\frac{1}{nVar(\varepsilon_1^2)}}E\left[\sum_{k=1}^{n}\left(\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)Y_{kj}\right)^2\right] \\ \leq &\frac{C}{a}\cdot\frac{\log\log n}{n^{1/2}}. \end{aligned} \tag{20}$$

Using a similar method to that of (12) with $t = 1$ and the Markov inequality, we know that

$$\begin{aligned} & P(|L_{n2}| > a/5) \leq \frac{5}{a} E|L_{n2}| \\ = & \sqrt{\frac{1}{4nVar(\varepsilon_1^2)}} E\left[\sum_{k=1}^{n}\left(\sum_{j=1}^{q}\sum_{l=1}^{q}(\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l)Z_{kjl}\right)^2\right] \\ \leq & \frac{C}{a} \cdot \frac{(\log\log n)^2}{n^{3/2}}. \end{aligned} \tag{21}$$

Using a similar argument to that of (14) with $t = 1$ and the Markov inequality, we have

$$\begin{aligned} & P(|L_{n3}| > a/5) \leq \frac{25}{a^2} E|L_{n3}|^2 \\ = & \frac{25}{a^2} \frac{1}{nVar(\varepsilon_1^2)} E\left[\sum_{k=1}^{n}\sum_{j=1}^{q}\sum_{l=1}^{q}(\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l)Z_{kjl}\varepsilon_i\right]^2 \\ \leq & \frac{C}{a^2} \cdot \frac{\log\log n}{n}. \end{aligned} \tag{22}$$

By the similar proof to that of (15) with $t = 1$ and Markov inequality, it follows that

$$\begin{aligned} & P(|L_{n4}| > a/5) \leq \frac{25}{a^2} E|L_{n4}|^2 \\ = & \frac{25}{a^2} \frac{1}{nVar(\varepsilon_1^2)} E\left[\sum_{k=1}^{n}\sum_{j=1}^{q}\sum_{l=1}^{q}(\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l)Z_{kjl}\varepsilon_k\right]^2 \\ \leq & \frac{C}{a^2} \cdot \frac{(\log\log n)^2}{n}. \end{aligned} \tag{23}$$

Using a similar method to that of (16) with $t = 1$ and the Markov inequality, one can conclude that

$$\begin{aligned} & P(|L_{n5}| > a/5) \leq \frac{25}{a^2} E|L_{n5}| \\ = & \frac{25}{a^2} \frac{1}{nVar(\varepsilon_1^2)} E\left[\sum_{k=1}^{n}\sum_{j=1}^{q}\sum_{l=1}^{q}\sum_{m=1}^{q}(\hat{\theta}_j - \theta_j)(\hat{\theta}_l - \theta_l)(\hat{\theta}_m - \theta_m)Y_{kj}Z_{klm}\right]^2 \\ \leq & \frac{C}{a^2} \cdot \frac{(\log\log n)^3}{n^2}. \end{aligned} \tag{24}$$

Using Lemma 4 with $X_k = \varepsilon_k^2 - E\varepsilon_k^2$, we can easily see that

$$\begin{aligned} & \sup_{z\in\mathbb{R}} |P\{|L_{n6}| \leq z\} - \Phi(z)| \\ \leq & \frac{A}{(nVar(\varepsilon_1^2))^{1+\delta/2}} \sum_{k=1}^{n} E|\varepsilon_k^2 - E\varepsilon_k^2|^{2+\delta} \\ \leq & \frac{C}{n^{\delta/2}}. \end{aligned} \tag{25}$$

Then, by combining (19)–(25), it follows that

$$\sup_{z\in\mathbb{R}}\left|P\left\{\frac{\sqrt{n}}{\sqrt{Var(\varepsilon_1^2)}}\left(\hat{\sigma}_n^2 - \sigma^2\right) \leq z\right\} - \Phi(z)\right|$$

$$\begin{aligned}
&\leq C\left(\frac{1}{n^{\delta/2}}+a+\frac{1}{a}\cdot\frac{\log\log n}{n^{1/2}}+\frac{1}{a}\cdot\frac{(\log\log n)^2}{n^{3/2}}\right)\\
&\quad +C\left(\frac{1}{a^2}\cdot\frac{\log\log n}{n}+\frac{1}{a^2}\cdot\frac{(\log\log n)^2}{n}+\frac{1}{a^2}\cdot\frac{(\log\log n)^3}{n^2}\right)\\
&\leq C\left(\frac{1}{n^{\delta/2}}+a+\frac{1}{a}\cdot\frac{\log\log n}{n^{1/2}}+\frac{1}{a^2}\cdot\frac{(\log\log n)^2}{n}\right)\\
&\leq C\left(\frac{1}{n^{\delta/2}}+\frac{(\log\log n)^{2/3}}{n^{1/3}}\right),
\end{aligned}$$

and thus the last inequality is obtained by taking $a=\frac{(\log\log n)^{2/3}}{n^{1/3}}$. Therefore, (6) is proved. □

4. Conclusions

In this article, via Taylor's expansion, the classical invariance principle, and Berry–Esseen bound for independent random variables, the authors established the invariance principle and Berry–Esseen bound for an error variance estimator for NAR with i.i.d. errors. The central limit theorem for the error variance estimator was extended to the invariance principle. Four corollaries were given to test the results. In the future, the moderate deviation principle and large deviation principle for the error variance estimator for NAR with independent errors will be considered.

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

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

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

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

References

1. Liebscher, E. Strong convergence of estimators in nonlinear autoregressive models. *J. Multivar. Anal.* **2003**, *84*, 247–261. [CrossRef]
2. 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]
3. Fu, K.; Yang, X. Asymptotics of kernel density estimators in nonlinear autoregressive models. *J. Math. Chem.* **2008**, *44*, 831–838. [CrossRef]
4. Cheng, F. Global property of error density estimation in nonlinear autoregressive time series models. *Stat. Inference Stoch. Process.* **2010**, *13*, 43–53. [CrossRef]
5. Hilgert, N.; Portier, B. Strong uniform consistency and asymptotic normality of a kernel based error density estimator in functional autoregressive models. *J. Stat. Plan. Inference* **2012**, *15*, 105–125. [CrossRef]
6. 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.
7. 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]
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. Wu, S.; Yang, W.; Gao, M.; Fang, H. Asymptotic results of error density estimator in nonlinear autoregressive models. *J. Korean Stat. Soc.* **2024**, *53*, 563–582. [CrossRef]
11. Cheng, F. Variance estimation in nonlinear autoregressive time series models. *J. Stat. Plann. Inference* **2011**, *141*, 1588–1592. [CrossRef]

12. Liang, K.; Zhang, Y. Kaiyu Liang, Yong Zhang. Almost sure central limit theorem for error variance estimator in pth-order nonlinear autoregressive processes. *Mathematics* **2024**, *12*, 1482. [CrossRef]
13. Wang, Y.; Mao, M.; Hu, X.; He, T. The law of iterated logarithm for autoregressive processes. *Math. Probl. Eng.* **2014** , *2014*, 972712. [CrossRef]
14. Yao, J. On least squares estimation for stable nonlinear AR processes. *Ann. Inst. Stat. Math.* **2000**, *52*, 316–331. [CrossRef]
15. Hall, P.; Heyde, C. *Martingale Limit Theory and Its Application*; Academic Press: New York, NY, USA, 1980.
16. Chang, M.; Rao, P. Berry-Esseen bound for the Kaplan-Meier estimator. *Comm. Stat. Theory Methods* **1989**, *18*, 4647–4664. [CrossRef]
17. Petrov, V. *Limit Theorems of Probability Theory*; Oxford Univ. Press Inc.: New York, NY, USA, 1995.
18. Lu, C. The invariance principle for linear processes generated by a negatively associated sequence and its applications. *Acta Math. Appl. Sin. Engl. Ser.* **2003**, *19*, 641–646.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.