DISSERTATION

PARAMETER ESTIMATION FOR ALL-PASS TIME SERIES MODELS

Submitted by

Margaret Elizabeth Andrews

Department of Statistics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2003

COLORADO STATE UNIVERSITY

August 8, 2003

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY MARGARET ELIZABETH ANDREWS ENTITLED PARAMETER ESTIMATION FOR ALL-PASS TIME SERIES MODELS BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

F. J. Breidt

[illegible]

Jan Hannig

Richard A. Davis
Adviser

Richard A. Davis
Department Head

ABSTRACT OF DISSERTATION

PARAMETER ESTIMATION FOR ALL-PASS TIME SERIES MODELS

All-pass models are autoregressive-moving average models in which the roots of the autoregressive polynomial are reciprocals of roots of the moving average polynomial and vice versa. They generate uncorrelated (white noise) time series, but these series are not independent in the non-Gaussian case. All-pass models can be used to simplify the process of fitting noncausal autoregressive and noninvertible moving average models and, in this dissertation, these procedures are used in the deconvolution of a simulated water gun seismogram and to fit stock market trading volume data.

Because all-pass series are uncorrelated, estimation methods based on Gaussian likelihood, least-squares, or related second-order moment techniques cannot identify all-pass models. Consequently, least absolute deviations, maximum likelihood, and rank techniques are used to obtain parameter estimates. The rank estimator considered was first proposed by Louis Jaeckel for estimating linear regression parameters. Jaeckel's estimator minimizes the sum of model residuals weighted by a function of residual rank. The asymptotic properties of the three types of estimators are examined and their behavior is studied for finite samples via simulation.

For all-pass series with finite variance, maximum likelihood and rank estimation are studied in-depth, and some extensions of previous results for least absolute deviations estimation are given in the Appendix. In the infinite variance case, least absolute deviations estimation is considered when the noise distribution is in the domain of attraction of a non-Gaussian stable distribution, and maximum likelihood estimation is considered when the noise distribution is non-Gaussian stable. Under general conditions, it is shown that the estimators are asymptotically normal in the finite variance case, and, in the infinite variance case, the estimators converge in distribution to nondegenerate maxima of stochastic processes.

Margaret Elizabeth Andrews
Department of Statistics
Colorado State University
Fort Collins, Colorado 80523
Fall 2003

ACKNOWLEDGEMENTS

I am very grateful to my adviser Richard Davis for all his help with this dissertation and for the opportunity to work with him. I also want to thank my committee, Jay Breidt, Don Estep, and Jan Hannig, for their assistance, particularly Jay Breidt who has read much of this dissertation multiple times. For serving on the committee previously, I would like to thank Peter Brockwell and Hari Iyer. In addition, I want to thank my fellow graduate students and the department faculty for helping to make my time at Colorado State extremely rewarding, and I would like to thank my family and friends for their support and encouragement. Finally, I gratefully acknowledge the support of NSF Grants DMS9972015 and DMS0308109.

To my family

Contents

Chapter 1

Introduction

Autoregressive-moving average (ARMA) models are linear time series models that play a fundamental role in modern time series analysis. They have diverse applications and have been used, for example, to model the Dow-Jones utilities index, the level of Lake Huron, and the sunspot numbers (Brockwell and Davis, 1996). The series $\{X_t\}$ is an ARMA(p,q) process if it satisfies the difference equations

$$\phi(B)X_t = \theta(B)Z_t,$$

where B denotes the backshift operator ($B^k X_t = X_{t-k}$, $k = 0, \pm 1, \pm 2, \ldots$),

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$$

is an autoregressive polynomial of order p such that $\phi(z) \neq 0$ for $|z| = 1$,

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$$

is a moving average polynomial of order q, and $\{Z_t\}$ is a noise sequence.

All-pass models are an interesting subclass of ARMA models in which the roots of $\phi(z)$ are reciprocals of the roots of $\theta(z)$ and vice versa. They generate uncorrelated (white noise) time series, but these series are not independent in the non-Gaussian case. Surprisingly, all-pass models can often be used to describe financial time series that display "nonlinear" behavior. In particular, financial time series, such as exchange rate log returns, frequently appear uncorrelated, have heavy-tailed marginal distributions, and exhibit volatility clustering. If a distribution assigns substantial probability to values relatively large in absolute value, it is called *heavy-tailed,* and, therefore, data from heavy-tailed distributions look "spiky." More specifically, if a distribution is called heavy-tailed, it is generally understood that there exists an $\alpha > 0$ such that the probability of seeing an observation larger in absolute value than x, $x > 0$, is roughly proportional to $x^{-\alpha}$. Volatility clustering is the tendency of observations relatively small in absolute value to be followed by other small observations and the tendency of observations relatively large in absolute value to be followed by other large observations. Therefore, nonlinear models with time-dependent conditional variances, such as generalized autoregressive conditionally heteroskedastic (GARCH) models and stochastic volatility models, are often chosen to describe financial data. However, linear all-pass models with heavy-tailed noise can also exhibit the three characteristics, and so they can sometimes be used as an alternative model for financial data. For example, in Chapter 3, an all-pass model is used to describe the daily log returns for the Canada/United States exchange rate.

An all-pass series can be obtained by fitting a causal, invertible ARMA model (all the roots of $\phi(z)$ and $\theta(z)$ are outside the unit circle in the complex plane) to a series generated by a causal, noninvertible ARMA model (all the roots of $\phi(z)$ are outside the unit circle and at least one root of $\theta(z)$ is inside the unit circle). The residuals follow an all-pass model of order r, where r is the number of roots of the true moving average polynomial inside the unit circle. Noninvertible ARMA models have been used, for example, in vocal tract filters (Chi and Kung, 1995; Chien, Yang, and Chi, 1997), in the analysis of unemployment rates (Huang and Pawitan, 2000), and in seismogram deconvolution (Lii and Rosenblatt, 1988). Similarly, an all-pass series can be obtained by fitting a causal autoregressive model to a series generated by a noncausal autoregressive model (Breidt, Davis, and Trindade, 2001). Applications for noncausal models include the analyis of stock market trading volume data (Calder, 1998), the deconvolution of absorption spectra (Blass and Halsey, 1981), the design of communications systems (Benveniste, Goursat, and Roget, 1980), the processing of blurry images (Donoho, 1981; Chien, Yang, and Chi, 1997), the deconvolution of seismic signals (Wiggins, 1978; Ooe and Ulrych, 1979; Donoho, 1981; Godfrey and Rocca, 1981; Hsueh and Mendel, 1985), and the analysis of astronomical data (Scargle, 1981).

Estimation methods based on Gaussian likelihood, least-squares, or related second-order moment techniques cannot identify all-pass models because Gaussian all-pass series are independent. Therefore, cumulant-based estimators, using cumulants of order greater than two, are often used to estimate the parameters of an all-pass process with finite second moments (Giannakis and Swami, 1990; Chi and Kung, 1995; Chien, Yang, and Chi, 1997). Also, Breidt, Davis, and Trindade (2001) consider a least absolute deviations (LAD) approach which is motivated by approximating the likelihood of an all-pass model with Laplace (two-sided exponential) noise. In Breidt, Davis, and Trindade (2001), LAD estimates are compared with estimates obtained by maximizing absolute residual kurtosis, a fourth-order moment approach. Based on these results, it appears that cumulant-based estimation is not the most efficient technique for estimating all-pass model parameters. In addition, because all-pass noise may have a distribution quite different from Laplace, LAD estimation cannot always be the most efficient either. Hence, in this dissertation, we consider general maximum likelihood (ML) and rank (R) estimation techniques. We also examine LAD estimation in the infinite variance case, and, in the Appendix, we correct some of the results in Breidt, Davis, and Trindade (2001).

In Chapter 2, we consider using ML estimation to estimate the parameters of an all-pass series with finite variance. Because they are based on specific information about the noise distribution, ML estimates tend to be less disperse than LAD and R-estimates. We give an approximate likelihood for all-pass model parameters and establish asymptotic normality and strong consistency for the ML estimators under general conditions. The behavior of the estimators for finite samples is studied via simulation, and order selection is considered. We then develop a two-step procedure for fitting noninvertible ARMA models using all-pass models and apply it to the deconvolution of a simulated water gun seismogram.

Because the objective functions for LAD and ML estimation are complicated functions of the model parameters, they tend to be quite bumpy and can, therefore, be hard to minimize or maximize. In addition, ML estimation is difficult to implement when the noise distribution is complicated or unknown. R-estimation can often overcome these limitations of LAD and ML estimation. Thus, in Chapter 3, we consider a R-estimator first proposed by Jaeckel (1972) for estimating linear regression parameters. Jaeckel's estimator minimizes the sum of model residuals weighted by a function of residual rank. In Chapter 3, we use this R-estimation technique to estimate the parameters of an all-pass process with finite variance. Under general conditions, the R-estimators are asymptotically normal. If the weight function is properly chosen, R-estimators can be nearly as asymptotically efficient as ML estimators, and the objective function for R-estimation can be fairly smooth and hence easy to minimize. We examine the behavior of the R-estimators for finite samples via simulation and again discuss all-pass order selection. R-estimation is used to fit an all-pass model to daily log returns for the Canada/United States exchange rate.

In Chapter 4, we consider LAD and ML estimation for all-pass series with infinite variance. In particular, we assume the noise distribution belongs to the domain of attraction of a non-Gaussian, stable distribution. These all-pass series tend to be very "spiky" because the noise distribution is heavy-tailed. In Davis (1996), LAD estimation is compared to least squares estimation for causal, invertible ARMA processes with noise distribution in the domain of attraction of a non-Gaussian, stable distribution. It is discovered that LAD estimation is more efficient than least squares estimation in this case. Hence, LAD estimation appears to be useful for ARMA processes with infinite variance, and so we study this type of estimator for all-pass series with infinite variance in Chapter 4. For ML estimation, stronger restrictions are required for the noise distribution. We assume that it is non-Gaussian stable, and not simply in the domain of attraction of a stable distribution. Under general conditions, these LAD and ML estimators have nondegenerate limiting distributions. The behavior of the LAD estimators for finite samples is studied using simulation and the estimation procedure is used to fit a noncausal autoregressive model to Microsoft trading volume data.

To summarize, this dissertation presents a variety of useful approaches for estimating all-pass model parameters and gives a number of applications for all-pass models–fitting noncausal autoregressive and noninvertible ARMA models, and modeling uncorrelated financial data. Interesting extensions of this dissertation work include finding other applications for all-pass models and developing procedures to assess the accuracy of fitted all-pass models using sample correlations of the squares and absolute values. Also, rank estimation appears to be very useful for estimating all-pass model parameters in Chapter 3 and thus this estimation technique merits further examination. It would be worthwhile to consider rank estimation for infinite variance all-pass series, ARMA series, and possibly GARCH series. A procedure for finding an optimal weight function for the rank objective function is also a subject for future study.

Chapter 2

Maximum Likelihood Estimation for All-Pass Time Series Models

2.1 Introduction

All-pass models are autoregressive-moving average (ARMA) models in which the roots of the autoregressive polynomial are reciprocals of roots of the moving average polynomial and vice versa. They generate uncorrelated (white noise) time series, but these series are not independent in the non-Gaussian case. An all-pass series can be obtained by fitting a causal, invertible ARMA model (all the roots of the autoregressive and moving average polynomials are outside the unit circle) to a series generated by a causal, noninvertible ARMA model (all the roots of the autoregressive polynomial are outside the unit circle and at least one root of the moving average polynomial is inside the unit circle). The residuals follow an all-pass model of order r, where r is the number of roots of the true moving average polynomial inside the unit circle. Noninvertible ARMA models have appeared, for example, in vocal tract filters (Chi and Kung, 1995; Chien, Yang, and Chi, 1997) and in the analysis of unemployment rates (Huang and Pawitan, 2000). We use them in this chapter in the deconvolution of a simulated water gun seismogram. Other deconvolution approaches are discussed in Wiggins (1978), Ooe and Ulrych (1979), Blass and Halsey (1981), Donoho (1981), Godfrey and Rocca (1981), Hsueh and Mendel (1985), and Lii and Rosenblatt (1988).

Estimation methods based on Gaussian likelihood, least-squares, or related second-order moment techniques cannot identify all-pass models. Therefore, cumulant-based estimators, using cumulants of order greater than two, are often used to estimate such models (Giannakis and Swami, 1990; Chi and Kung, 1995; Chien, Yang, and Chi, 1997). Breidt, Davis, and Trindade (2001) consider a least absolute deviations (LAD) approach which is motivated by approximating the likelihood of an all-pass model with Laplace (two-sided exponential) noise. Under general conditions, the LAD estimators are asymptotically normal.

In this chapter, we use a maximum likelihood (ML) approach to estimate all-pass model parameters. Related likelihood approaches are considered in Breidt, Davis, Lii, and Rosenblatt (1991) for noncausal autoregressive processes, in Lii and Rosenblatt (1992) for noninvertible moving average processes, and in Lii and Rosenblatt (1996) for general ARMA processes. Although all-pass models are ARMA models, their special parameterization makes the results of Lii and Rosenblatt (1996) inapplicable.

In Section 2.2, we give an approximate likelihood for all-pass model parameters. Asymptotic normality and strong consistency are established for ML estimators under general conditions and order selection is considered in Section 2.3. Proofs of the lemmas used to establish the results of Section 2.3 can be found in Section 2.5. The behavior of the estimators for finite samples is studied via simulation in Section 2.4.1. Finally, a two-step procedure for fitting noninvertible ARMA models is developed in Section 2.4.2 and applied to the deconvolution of a simulated water gun seismogram in Section 2.4.3.

2.2 Preliminaries

2.2.1 All-Pass Models

Let B denote the backshift operator ($B^k X_t = X_{t-k}$, $k = 0, \pm 1, \pm 2, \ldots$) and let

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$$

be a pth-order autoregressive polynomial, where $\phi(z) \neq 0$ for $|z| = 1$. The polynomial $\phi(B)$ is said to be *causal* if all its roots are outside the unit circle in the complex plane. In this case, for a sequence $\{W_t\}$,

$$\phi^{-1}(B)W_t = \left(\sum_{j=0}^{\infty} \psi_j B^j\right) W_t = \sum_{j=0}^{\infty} \psi_j W_{t-j},$$

a function of only the past and present $\{W_t\}$. Note that if $\phi(B)$ is causal, the polynomial $B^p\phi(B^{-1})$ is *purely noncausal* in the sense that all its roots are inside the unit circle, and hence

$$B^{-p}\phi^{-1}(B^{-1})W_t = \left(\sum_{j=0}^{\infty} \psi_j B^{-p-j}\right) W_t = \sum_{j=0}^{\infty} \psi_j W_{t+p+j},$$

a function of only the present and future $\{W_t\}$. See, for example, Chapter 3 of Brockwell and Davis (1991).

Let

$$\phi_0(z) = 1 - \phi_{01}z - \cdots - \phi_{0p}z^p,$$

where $\phi_0(z) \neq 0$ for $|z| \leq 1$. Define $\phi_{00} = 1$, and suppose $\phi_{0r} \neq 0$ for some $r \in \{0, 1, \ldots, p\}$ and $\phi_{0j} = 0$ for $j = r+1, \ldots, p$. Then, a causal all-pass time series is the ARMA series $\{X_t\}$ which satisfies the difference equations

$$\phi_0(B)X_t = \frac{B^r\phi_0(B^{-1})}{-\phi_{0r}} Z_t^* \tag{2.1}$$

or

$$X_t - \phi_{01}X_{t-1} - \cdots - \phi_{0r}X_{t-r} = Z_t^* + \frac{\phi_{0,r-1}}{\phi_{0r}} Z_{t-1}^* + \cdots + \frac{\phi_{01}}{\phi_{0r}} Z_{t-r+1}^* - \frac{1}{\phi_{0r}} Z_{t-r}^*.$$

The true order of the all-pass model is r, and the series $\{Z_t^*\}$ is an independent and identically distributed (iid) sequence of random variables with mean 0 and variance $\sigma_0^2 \in (0, \infty)$. We assume throughout that Z_1^* has probability density function $f_{\sigma_0}(z; \boldsymbol{\theta}_0) = \sigma_0^{-1} f(\sigma_0^{-1} z; \boldsymbol{\theta}_0)$, where f is a density function symmetric about zero and $\boldsymbol{\theta}$ is a parameter of the density f. We also assume that the true value of $\boldsymbol{\theta}$, $\boldsymbol{\theta}_0 = (\theta_{01}, \ldots, \theta_{0d})'$, lies in the interior of a parameter space $\Theta \subseteq \mathbb{R}^d$, $d \geq 1$. Note that the roots of the autoregressive polynomial $\phi_0(z)$ are reciprocals of the roots of the moving average polynomial $-\phi_{0r}^{-1} z^r \phi_0(z^{-1})$ and vice versa.

The spectral density for $\{X_t\}$ in (2.1) is

$$\frac{|e^{-ir\omega}|^2|\phi_0(e^{i\omega})|^2}{\phi_{0r}^2|\phi_0(e^{-i\omega})|^2}\frac{\sigma_0^2}{2\pi}=\frac{\sigma_0^2}{\phi_{0r}^2 2\pi},$$

which is constant for $\omega \in [-\pi,\pi]$, and thus $\{X_t\}$ is an uncorrelated sequence. In the case of Gaussian $\{Z_t^*\}$, this implies that $\{X_t\}$ is iid $\mathrm{N}(0,\sigma_0^2\phi_{0r}^{-2})$, but independence does not hold in the non-Gaussian case if $r \geq 1$ (e.g., Breidt and Davis, 1991). The model (2.1) is called all-pass because the power transfer function of the all-pass filter passes all the power for every frequency in the spectrum. In other words, an all-pass filter does not change the distribution of power over the spectrum.

We can express (2.1) as

$$\phi_0(B)X_t = \frac{B^p\phi_0(B^{-1})}{-\phi_{0r}}Z_t, \tag{2.2}$$

where $\{Z_t\} = \{Z_{t+p-r}^*\}$ is an iid sequence of random variables with mean 0, variance σ_0^2, and probability density function $f_{\sigma_0}(z;\boldsymbol{\theta}_0)$. Rearranging (2.2) and setting $z_t = \phi_{0r}^{-1}Z_t$, we have the backward recursion

$$z_{t-p} = \phi_{01}z_{t-p+1}+\cdots+\phi_{0p}z_t-(X_t-\phi_{01}X_{t-1}-\cdots-\phi_{0p}X_{t-p}).$$

An analogous recursion for an arbitrary, causal autoregressive polynomial $\phi(z)=1-\phi_1 z-\cdots-\phi_p z^p$ can be defined as follows:

$$z_{t-p}(\boldsymbol{\phi}) = \begin{cases} 0, & t=n+p,\ldots,n+1, \\ \phi_1 z_{t-p+1}(\boldsymbol{\phi})+\cdots+\phi_p z_t(\boldsymbol{\phi})-\phi(B)X_t, & t=n,\ldots,p+1, \end{cases} \tag{2.3}$$

where $\boldsymbol{\phi} := (\phi_1,\ldots,\phi_p)'$. If $\boldsymbol{\phi}_0 := (\phi_{01},\ldots,\phi_{0p})' = (\phi_{01},\ldots,\phi_{0r},0,\ldots,0)'$, note that $\{z_t(\boldsymbol{\phi}_0)\}_{t=1}^{n-p}$ closely approximates $\{z_t\}_{t=1}^{n-p}$; the error is due to the initialization with zeros. Although $\{z_t\}$ is iid, $\{z_t(\boldsymbol{\phi}_0)\}_{t=1}^{n-p}$ is not iid if $r \geq 1$.

2.2.2 Approximating the Likelihood

In this subsection, we ignore the effect of the recursion initialization in (2.3), and write

$$-\phi(B^{-1})B^p z_t(\boldsymbol{\phi}) = \phi(B)X_t.$$

Let $q = \max\{0 \leq j \leq p : \phi_j \neq 0\}$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{p+d+1})' = (\phi_1, \ldots, \phi_p, \sigma/|\phi_q|, \theta_1, \ldots, \theta_d)'$, and $\boldsymbol{\alpha}_0 = (\alpha_{01}, \ldots, \alpha_{0,p+d+1})' = (\phi_{01}, \ldots, \phi_{0p}, \sigma_0/|\phi_{0r}|, \theta_{01}, \ldots, \theta_{0d})'$. Following equation (2.7) of Breidt, Davis, and Trindade (2001), we approximate the log-likelihood of $\boldsymbol{\alpha}$ given a realization of length n, $\{X_t\}_{t=1}^n$, from model (2.1) with

$$\begin{aligned} \mathcal{L}(\boldsymbol{\alpha}) &= \sum_{t=1}^{n-p} \ln f_\sigma\left(\phi_q z_t(\boldsymbol{\phi}); \boldsymbol{\theta}\right) + (n-p)\ln|\phi_q| \\ &= \sum_{t=1}^{n-p} \left\{\ln f(z_t(\boldsymbol{\phi})/\alpha_{p+1}; \boldsymbol{\theta}) - \ln \alpha_{p+1}\right\} \\ &=: \sum_{t=1}^{n-p} g_t(\boldsymbol{\alpha}), \end{aligned} \tag{2.4}$$

where $\{z_t(\boldsymbol{\phi})\}$ can be computed recursively from (2.3).

2.3 Asymptotic Results

2.3.1 Parameter Estimation

In order to establish asymptotic normality for the MLE of $\boldsymbol{\alpha}$, we make the following additional, yet still fairly general, assumptions on f:

- **A1** *For all $s \in \mathbb{R}$ and all $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)' \in \Theta$, $f(s;\boldsymbol{\theta}) > 0$ and $f(s;\boldsymbol{\theta})$ is twice continuously differentiable with respect to $(s, \theta_1, \ldots, \theta_d)'$.*
- **A2** *For all $\boldsymbol{\theta}$ in some neighborhood of $\boldsymbol{\theta}_0$, $\int s f'(s;\boldsymbol{\theta})\, ds = s f(s;\boldsymbol{\theta})|_{-\infty}^{\infty} - \int f(s;\boldsymbol{\theta})\, ds = -1$.*
- **A3** $\int f''(s;\boldsymbol{\theta}_0)\, ds = f'(s;\boldsymbol{\theta}_0)|_{-\infty}^{\infty} = 0$.
- **A4** $\int s^2 f''(s;\boldsymbol{\theta}_0)\, ds = s^2 f'(s;\boldsymbol{\theta}_0)|_{-\infty}^{\infty} - 2\int s f'(s;\boldsymbol{\theta}_0)\, ds = 2$.
- **A5** $1 < \int \left(f'(s;\boldsymbol{\theta}_0)\right)^2 / f(s;\boldsymbol{\theta}_0)\, ds$.
- **A6** *The matrix* $\begin{bmatrix} \tilde{K} & L' \\ L & I \end{bmatrix}$ *is positive definite if* $\tilde{K} := \alpha_{0,p+1}^{-2}\left\{\int \frac{s^2(f'(s;\boldsymbol{\theta}_0))^2}{f(s;\boldsymbol{\theta}_0)}\, ds - 1\right\}$,

$$L := \left[-\alpha_{0,p+1}^{-1} \int \frac{s f'(s;\boldsymbol{\theta}_0)}{f(s;\boldsymbol{\theta}_0)} \frac{\partial f(s;\boldsymbol{\theta}_0)}{\partial \theta_j}\, ds\right]_{j=1}^{d}, \text{ and } I := \left[\int \frac{1}{f(s;\boldsymbol{\theta}_0)} \frac{\partial f(s;\boldsymbol{\theta}_0)}{\partial \theta_j} \frac{\partial f(s;\boldsymbol{\theta}_0)}{\partial \theta_k}\, ds\right]_{j,k=1}^{d}.$$

- **A7** *For* $j, k = 1, \ldots, d$ *and all* $\boldsymbol{\theta}$ *in some neighborhood of* $\boldsymbol{\theta}_0$,
  - $f(s;\boldsymbol{\theta})$ *is dominated by some function* $f_1(s)$,
  - $(1+s^2)\frac{(f'(s;\boldsymbol{\theta}))^2}{f^2(s;\boldsymbol{\theta})}$, $(1+s^2)\left|\frac{f''(s;\boldsymbol{\theta})}{f(s;\boldsymbol{\theta})}\right|$, $(1+|s|)\frac{1}{f(s;\boldsymbol{\theta})}\left|\frac{\partial}{\partial\theta_j}f'(s;\boldsymbol{\theta})\right|$, $\frac{1}{f^2(s;\boldsymbol{\theta})}\left(\frac{\partial}{\partial\theta_j}f(s;\boldsymbol{\theta})\right)^2$, *and* $\frac{1}{f(s;\boldsymbol{\theta})}\left|\frac{\partial^2}{\partial\theta_j\partial\theta_k}f(s;\boldsymbol{\theta})\right|$ *are dominated by* $a_1 + a_2|s|^{c_1}$, *where* a_1, a_2, c_1 *are non-negative constants and* $\int(a_1 + a_2|s|^{c_1})f_1(s)\,ds < \infty$.

Theorem 1 *If* f *satisfies A1–A7, then there exists a sequence of maximizers*

$$\hat{\boldsymbol{\alpha}}_{ML} = (\hat{\boldsymbol{\phi}}'_{ML}, \hat{\alpha}_{p+1,ML}, \hat{\boldsymbol{\theta}}'_{ML})'$$

of $\mathcal{L}(\boldsymbol{\alpha})$ *in (2.4) such that*

$$n^{1/2}(\hat{\boldsymbol{\alpha}}_{ML} - \boldsymbol{\alpha}_0) \xrightarrow{d} \mathbf{Y} \sim N(\mathbf{0}, \boldsymbol{\Sigma}^{-1}), \tag{2.5}$$

where

$$\boldsymbol{\Sigma}^{-1} := \begin{bmatrix} \frac{\sigma_0^2}{2(\sigma_0^2\tilde{J}-1)}\boldsymbol{\Gamma}_p^{-1} & \mathbf{0}_{p\times 1} & \mathbf{0}_{p\times d} \\ \mathbf{0}_{1\times p} & (\tilde{K} - L'I^{-1}L)^{-1} & -\tilde{K}^{-1}L'(I - L\tilde{K}^{-1}L')^{-1} \\ \mathbf{0}_{d\times p} & -(I - L\tilde{K}^{-1}L')^{-1}L\tilde{K}^{-1} & (I - L\tilde{K}^{-1}L')^{-1} \end{bmatrix},$$

$\tilde{J} := \sigma_0^{-2}\int(f'(s;\boldsymbol{\theta}_0))^2/f(s;\boldsymbol{\theta}_0)\,ds$, $\boldsymbol{\Gamma}_p := [\gamma(j-k)]_{j,k=1}^p$, *and* $\gamma(\cdot)$ *is the autocovariance function of the autoregressive process* $\{(1/\phi_0(B))Z_t\}$.

Proof: $\mathcal{L}(\boldsymbol{\alpha}) - \mathcal{L}(\boldsymbol{\alpha}_0) = S_n(\sqrt{n}\,(\boldsymbol{\alpha} - \boldsymbol{\alpha}_0))$, where $S_n(\cdot)$ is defined in Lemma 3 of Section 2.5. Because $\mathbf{Y} := \boldsymbol{\Sigma}^{-1}\mathbf{N}$ maximizes the limit process $S(\cdot)$ in Lemma 3, the result (2.5) follows by Remark 1 of Davis, Knight, and Liu (1992). □

Remark 1: Note that $\hat{\boldsymbol{\phi}}_{ML}$ is asymptotically independent of $(\hat{\alpha}_{p+1,ML}, \hat{\boldsymbol{\theta}}'_{ML})'$. Given a sample of n observations from $\{Z_t\}$, there exist local maximizers $(\hat{\sigma}_Z, \hat{\boldsymbol{\theta}}'_Z)'$ of the log-likelihood

$$\sum_{t=1}^{n}\{\ln f(Z_t/\sigma;\boldsymbol{\theta}) - \ln\sigma\}$$

such that

$$n^{1/2}[(\hat{\sigma}_Z, \hat{\boldsymbol{\theta}}'_Z)' - (\sigma_0, \boldsymbol{\theta}'_0)'] \xrightarrow{d} \mathbf{Y}_Z \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_Z^{-1}),$$

where

$$\Sigma_Z^{-1} := \begin{bmatrix} \phi_{0r}^2(\tilde{K} - L'I^{-1}L)^{-1} & -|\phi_{0r}|\tilde{K}^{-1}L'(I - L\tilde{K}^{-1}L')^{-1} \\ -|\phi_{0r}|(I - L\tilde{K}^{-1}L')^{-1}L\tilde{K}^{-1} & (I - L\tilde{K}^{-1}L')^{-1} \end{bmatrix}$$

and Σ_Z^{-1} does not depend on ϕ_{0r}.

Remark 2: Using A2 and the Cauchy-Schwartz inequality,

$$\begin{aligned} 1 &= \left\{\int s\frac{f'(s;\boldsymbol{\theta}_0)}{f(s;\boldsymbol{\theta}_0)} f(s;\boldsymbol{\theta}_0)\,ds\right\}^2 \\ &\leq \left\{\int s^2 f(s;\boldsymbol{\theta}_0)\,ds\right\}\left\{\int \left(\frac{f'(s;\boldsymbol{\theta}_0)}{f(s;\boldsymbol{\theta}_0)}\right)^2 f(s;\boldsymbol{\theta}_0)\,ds\right\} \qquad (2.6) \\ &= \sigma_0^2\tilde{J}, \end{aligned}$$

with equality in (2.6) if and only if f is Gaussian. Thus, A5 holds for non-Gaussian f. Further,

$$\begin{aligned} 1 &= \left\{\int s\frac{f'(s;\boldsymbol{\theta}_0)}{f(s;\boldsymbol{\theta}_0)} f(s;\boldsymbol{\theta}_0)\,ds\right\}^2 \\ &< \left\{\int s^2 \left(\frac{f'(s;\boldsymbol{\theta}_0)}{f(s;\boldsymbol{\theta}_0)}\right)^2 f(s;\boldsymbol{\theta}_0)\,ds\right\}\left\{\int f(s;\boldsymbol{\theta}_0)\,ds\right\} \qquad (2.7) \\ &= \alpha_{0,p+1}^2\tilde{K} + 1, \end{aligned}$$

so that $\tilde{K} > 0$. We do not have equality in (2.7) because, by Cauchy-Schwartz, there is equality if and only if $sf'(s;\theta_0)/f(s;\theta_0) = -1$ for all $s \in \mathbb{R}$ which cannot ever be the case.

Remark 3: The asymptotic covariance matrix for the estimators of the p autoregressive parameters is a scalar multiple of the asymptotic covariance matrix for the Gaussian likelihood estimators for the corresponding pth-order autoregressive process. The same property holds for LAD estimators of all-pass model parameters, as shown in Breidt, Davis, and Trindade (2001). The LAD estimators are obtained by maximizing the likelihood of an all-pass model with Laplace noise. This yields a modified LAD criterion, which can be used even if the underlying noise distribution is not Laplace. The constant in (2.5) is

$$\frac{1}{2(\sigma_0^2\tilde{J} - 1)}, \qquad (2.8)$$

while, in the LAD case, the appropriate constant is

$$\frac{\mathrm{Var}\,(|Z_1|)}{2\left(2\sigma_0^2 f_{\sigma_0}(0;\boldsymbol{\theta}_0) - \mathrm{E}|Z_1|\right)^2}. \qquad (2.9)$$

(Breidt, Davis, and Trindade (2001) contains an error in the calculation of the asymptotic variance; see the Appendix for the correction.) Although the Laplace density,

$$f(s) = \frac{1}{\sqrt{2}} \exp\left(-\sqrt{2}|s|\right),$$

does not meet assumptions A1–A7, $\mathrm{E}|Z_1| = \sigma_0/\sqrt{2}$, $f_{\sigma_0}(0) = 1/(\sqrt{2}\sigma_0)$, and $\tilde{J} = 2\sigma_0^{-2}$, so that (2.8) and (2.9) are both 1/2 for this density.

Remark 4: We obtain the asymptotic relative efficiency (ARE) of ML to LAD for the autoregressive parameters by dividing (2.9) by (2.8):

$$\mathrm{ARE} = (\sigma_0^2 \tilde{J} - 1) \frac{\mathrm{Var}\,(|Z_1|)}{\left(2\sigma_0^2 f_{\sigma_0}(0; \theta_0) - \mathrm{E}|Z_1|\right)^2}.$$

The density function

$$f(s;\theta) = \sqrt{\frac{\theta}{\theta-2}} \frac{\Gamma((\theta+1)/2)}{\Gamma(\theta/2)} \frac{1}{\sqrt{\theta\pi}} \frac{1}{(1+s^2/(\theta-2))^{(\theta+1)/2}} \tag{2.10}$$

is symmetric about zero, has variance one, and satisfies assumptions A1–A7 with $c_1 = 2$ when $\Theta \subseteq (2, \infty)$. If $\sigma_0 = (\theta_0/(\theta_0 - 2))^{1/2}$, then $f_{\sigma_0}(s; \theta_0) = \sigma_0^{-1} f(\sigma_0^{-1} s; \theta_0)$ is the Students' t-density with θ_0 degrees of freedom. In this case,

$$\mathrm{E}|Z_1| = 2\sigma_0 \frac{\sqrt{\theta_0 - 2}}{\theta_0 - 1} \frac{\Gamma((\theta_0+1)/2)}{\Gamma(\theta_0/2)\sqrt{\pi}},$$

$$f_{\sigma_0}(0; \theta_0) = \sigma_0^{-1} \frac{\Gamma((\theta_0+1)/2)}{\Gamma(\theta_0/2)\sqrt{(\theta_0-2)\pi}},$$

and

$$\tilde{J} = \sigma_0^{-2} \frac{\theta_0(\theta_0+1)}{(\theta_0-2)(\theta_0+3)},$$

so the ARE of ML to LAD is

$$\mathrm{ARE} = \frac{6}{(\theta_0+3)} \left\{ \frac{\pi}{4}(\theta_0 - 1)^2 \left(\frac{\Gamma(\theta_0/2)}{\Gamma((\theta_0+1)/2)} \right)^2 - (\theta_0 - 2) \right\}. \tag{2.11}$$

When $\theta_0 = 3$, the value of (2.11) is $2(0.7337) = 1.4674$, and thus ML is nearly 50% more efficient than LAD for the Students' t-distribution with three degrees of freedom. If $\theta_0 > 3$, (2.11) is even larger.

θ_{01}	θ_{02}	ARE
0.1	0.2	1.5506
0.1	0.8	2.1506
0.4	0.4	1.4988
0.4	0.6	1.7124
0.6	0.4	1.6012
0.6	0.6	1.8327
0.9	0.2	1.6395
0.9	0.8	2.9997

Table 2.1: AREs for ML to LAD when f is the Gaussian scale mixture density.

Remark 5: The Gaussian scale mixture density

$$f(s;(\theta_1,\theta_2)') = \frac{\theta_1}{\theta_2\sqrt{2\pi}} \exp\left(\frac{-s^2}{2\theta_2^2}\right) + \frac{(1-\theta_1)^{3/2}}{\sqrt{1-\theta_1\theta_2^2}\sqrt{2\pi}} \exp\left(\frac{-s^2(1-\theta_1)}{2(1-\theta_1\theta_2^2)}\right) \tag{2.12}$$

is symmetric about 0 and satisfies A1–A7 with $c_1 = 4$ when $\Theta \subseteq (0,1)\times(0,1)$. In this case, Z_1 is $\mathrm{N}(0,\sigma_0^2\theta_{02}^2)$ with probability θ_{01} and $\mathrm{N}(0,\sigma_0^2(1-\theta_{01}\theta_{02}^2)/(1-\theta_{01}))$ with probability $1-\theta_{01}$. Some values of ARE for ML to LAD for the autoregressive parameters are given in Table 2.1.

Remark 6: By A7 and the dominated convergence theorem,

$$\frac{1}{2(\sigma_0^2\tilde{J}-1)} = \frac{1}{2}\left(\int \frac{(f'(s;\boldsymbol{\theta}_0))^2}{f(s;\boldsymbol{\theta}_0)}\,ds - 1\right)^{-1},$$

$\tilde{K}$, L, and I are continuous with respect to $\boldsymbol{\theta}$ at $\boldsymbol{\theta}_0$. Thus, because $\hat{\boldsymbol{\theta}}_{ML} \xrightarrow{P} \boldsymbol{\theta}_0$,

$$\frac{1}{2}\left(\int \frac{(f'(s;\hat{\boldsymbol{\theta}}_{ML}))^2}{f(s;\hat{\boldsymbol{\theta}}_{ML})}\,ds - 1\right)^{-1}, \tag{2.13}$$

$$\frac{1}{\hat{\alpha}_{p+1,ML}^2}\left(\int \frac{s^2(f'(s;\hat{\boldsymbol{\theta}}_{ML}))^2}{f(s;\hat{\boldsymbol{\theta}}_{ML})}\,ds - 1\right), \tag{2.14}$$

$$\left[\frac{-1}{\hat{\alpha}_{p+1,ML}}\int \frac{sf'(s;\hat{\boldsymbol{\theta}}_{ML})}{f(s;\hat{\boldsymbol{\theta}}_{ML})}\frac{\partial f(s;\hat{\boldsymbol{\theta}}_{ML})}{\partial\theta_j}\,ds\right]_{j=1}^{d}, \tag{2.15}$$

and

$$\left[\int \frac{1}{f(s;\hat{\boldsymbol{\theta}}_{ML})}\frac{\partial f(s;\hat{\boldsymbol{\theta}}_{ML})}{\partial\theta_j}\frac{\partial f(s;\hat{\boldsymbol{\theta}}_{ML})}{\partial\theta_k}\,ds\right]_{j,k=1}^{d} \tag{2.16}$$

are consistent estimators of $[2(\sigma_0^2\tilde{J}-1)]^{-1}$, $\tilde{K}$, L, and I respectively.

If we restrict $\boldsymbol{\alpha}$ to a compact, convex space Ξ with $\boldsymbol{\alpha}_0$ in the interior, then the MLE of $\boldsymbol{\alpha}$ is almost surely consistent and asymptotically normal as shown in the following theorems.

Theorem 2 *If*

- *f satisfies A1,*
- *the parameter space $\boldsymbol{\Xi}$ is compact and convex with $\boldsymbol{\alpha}_0$ in the interior,*
- *$\boldsymbol{\phi}$ forms a causal polynomial, $\alpha_{p+1} > 0$, and $\boldsymbol{\theta} \in \Theta$ for all $\boldsymbol{\alpha} = (\boldsymbol{\phi}', \alpha_{p+1}, \boldsymbol{\theta}')' \in \boldsymbol{\Xi}$,*
- *$(1+|s|)|f'(s;\boldsymbol{\theta})|/f(s;\boldsymbol{\theta})$ and $|\frac{\partial}{\partial\theta_j}f(s;\boldsymbol{\theta})|/f(s;\boldsymbol{\theta})$, $j = 1,\ldots,p$, are dominated by $a_3 + a_4|s|^{c_2}$ for all $\boldsymbol{\alpha} \in \boldsymbol{\Xi}$, where a_3, a_4 and c_2 are non-negative constants such that $\int |s|^{c_2} f(s;\boldsymbol{\theta}_0) < \infty$,*
- *$E\{\ln f(\tilde{z}_1(\boldsymbol{\phi})/\alpha_{p+1};\boldsymbol{\theta}) - \ln \alpha_{p+1}\}$ has a unique maximum at $\boldsymbol{\alpha}_0$, where $\tilde{z}_1(\boldsymbol{\phi}) := -(\phi(B)/\phi(B^{-1}))X_{1+p}$,*

then

$$\hat{\boldsymbol{\alpha}} := argmax_{\boldsymbol{\alpha}\in\boldsymbol{\Xi}}\mathcal{L}(\boldsymbol{\alpha}) \xrightarrow{a.s.} \boldsymbol{\alpha}_0.$$

Proof: By Lemma 4 in Section 2.5,

$$n^{-1}\mathcal{L}(\boldsymbol{\alpha}) \xrightarrow{a.s.} \mathrm{E}\{\ln f(\tilde{z}_1(\boldsymbol{\phi})/\alpha_{p+1};\boldsymbol{\theta}) - \ln \alpha_{p+1}\}$$

uniformly on $\boldsymbol{\Xi}$. Since $\boldsymbol{\Xi}$ is compact and the limit has a unique maximum at $\boldsymbol{\alpha}_0$, the result follows.

□

Theorem 3 *If f satisfies A1–A7 and the conditions of Theorem 2 hold, then*

$$n^{1/2}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0) \xrightarrow{d} \mathbf{Y} \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}^{-1}\right).$$

Proof: Because $\mathcal{L}$ is differentiable on $\boldsymbol{\Xi}$ and $\hat{\boldsymbol{\alpha}}$ maximizes $\mathcal{L}(\boldsymbol{\alpha})$, $\frac{\partial}{\partial\boldsymbol{\alpha}}\mathcal{L}(\hat{\boldsymbol{\alpha}}) = \mathbf{0}$ for n sufficiently large almost surely. Thus, we have

$$\begin{aligned}
\mathbf{0} &= n^{-1}\sum_{t=1}^{n-p}\frac{\partial g_t(\hat{\boldsymbol{\alpha}})}{\partial\boldsymbol{\alpha}} \\
&= n^{-1}\sum_{t=1}^{n-p}\frac{\partial g_t(\boldsymbol{\alpha}_0)}{\partial\boldsymbol{\alpha}} + n^{-1}\sum_{t=1}^{n-p}\frac{\partial^2 g_t(\boldsymbol{\alpha}_n^*)}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}'}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0),
\end{aligned}$$

where $\boldsymbol{\alpha}_n^*$ is between $\hat{\boldsymbol{\alpha}}$ and $\boldsymbol{\alpha}_0$, and so

$$n^{1/2}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0) = -\left(n^{-1}\sum_{t=1}^{n-p}\frac{\partial^2 g_t(\boldsymbol{\alpha}_n^*)}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}'}\right)^{-1}\left(n^{-1/2}\sum_{t=1}^{n-p}\frac{\partial g_t(\boldsymbol{\alpha}_0)}{\partial\boldsymbol{\alpha}}\right).$$

By Lemma 1,

$$n^{-1/2}\sum_{t=1}^{n-p}\frac{\partial g_t(\boldsymbol{\alpha}_0)}{\partial\boldsymbol{\alpha}} \xrightarrow{d} \mathbf{N} \sim N(\mathbf{0}, \boldsymbol{\Sigma}),$$

and, by Lemma 2,

$$n^{-1}\sum_{t=1}^{n-p}\frac{\partial^2 g_t(\boldsymbol{\alpha}_0)}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}'} \xrightarrow{P} -\boldsymbol{\Sigma}.$$

Because $\hat{\boldsymbol{\alpha}} \xrightarrow{a.s.} \boldsymbol{\alpha}_0$ and f satisfies A7,

$$n^{-1}\sum_{t=1}^{n-p}\frac{\partial^2 g_t(\boldsymbol{\alpha}_n^*)}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}'} - n^{-1}\sum_{t=1}^{n-p}\frac{\partial^2 g_t(\boldsymbol{\alpha}_0)}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}'} \xrightarrow{P} \mathbf{0},$$

and the result follows. □

Remark 7: If f satisfies A7, the fourth condition of Theorem 2 holds. The fifth condition can be much more of a challenge to verify without resorting to simulation, but it is straightforward to establish, for example, when the parameter $\boldsymbol{\theta}$ is dropped and

$$f(s) = \frac{2^{1/4}\Gamma^2(3/4)}{\pi^{3/2}}\exp\left(\frac{-s^4\Gamma^4(3/4)}{2\pi^2}\right).$$

2.3.2 Order Selection

In practice, the order r of an all-pass model is usually unknown. Therefore, we present the following corollary to Theorem 1 for use in order selection.

Corollary 1 *Assume f satisfies A1–A7. If the true order of the all-pass model is r and the order of the fitted model is $p > r$, then*

$$n^{1/2}\hat{\phi}_{p,ML} \xrightarrow{d} N\left(0, \frac{1}{2(\sigma_0^2\tilde{\mathcal{J}} - 1)}\right).$$

Proof: By Problem 8.15 in Brockwell and Davis (1991), the pth diagonal element of $\boldsymbol{\Gamma}_p^{-1}$ is σ_0^{-2} if $p > r$, and so the result follows from (2.5). □

A practical approach to order determination using a large sample follows:

1. For some large P, fit all-pass models of order p, $p = 1, 2, \ldots, P$, via ML and obtain the pth coefficient, $\hat{\phi}_{p,ML}$, for each.

2. Let the model order r be the smallest order beyond which the estimated coefficients are statistically insignificant; that is,

$$r = \min\{0 \le p \le P : |\hat{\phi}_{j,ML}| < 1.96\hat{\lambda}\, n^{-1/2} \text{ for } j > p\},$$

where $\hat{\lambda} := \left(2\int (f'(s;\hat{\theta}_{ML}))^2/f(s;\hat{\theta}_{ML})\, ds - 2\right)^{-1/2}$ and $\hat{\theta}_{ML}$ is the MLE from the fitted Pth-order model.

2.4 Numerical Results

2.4.1 Simulation Study

In this section, we describe a simulation experiment to assess the quality of the asymptotic approximations for finite samples. We used both the rescaled Students' t-density (2.10) and the Gaussian scale mixture density (2.12). For the rescaled Students' t-density, we let $\sigma_0 = (\theta_0/(\theta_0 - 2))^{1/2}$, so Z_1 followed the Students' t-distribution with θ_0 degrees of freedom.

To diminish the possibility of the optimizer being trapped at local maxima, we used 250 starting values for each of the 1000 replicates. The initial values for $\phi_1, \ldots, \phi_p$ were uniformly distributed in the space of partial autocorrelations and then mapped to the space of autoregressive coefficients using the Durbin-Levinson algorithm (Brockwell and Davis, 1991, Proposition 5.2.1). That is, for a model of order p, the kth starting value $(\phi_{p1}^{(k)}, \ldots, \phi_{pp}^{(k)})'$ was computed recursively as follows:

1. Draw $\phi_{11}^{(k)}, \phi_{22}^{(k)}, \ldots, \phi_{pp}^{(k)}$ iid uniform$(-1, 1)$.

2. For $j = 2, \ldots, p$, compute

$$\begin{bmatrix} \phi_{j1}^{(k)} \\ \vdots \\ \phi_{j,j-1}^{(k)} \end{bmatrix} = \begin{bmatrix} \phi_{j-1,1}^{(k)} \\ \vdots \\ \phi_{j-1,j-1}^{(k)} \end{bmatrix} - \phi_{jj}^{(k)} \begin{bmatrix} \phi_{j-1,j-1}^{(k)} \\ \vdots \\ \phi_{j-1,1}^{(k)} \end{bmatrix}.$$

With $(\phi_{p1}^{(k)}, \ldots, \phi_{pp}^{(k)})'$ and a realization of length n, we obtained residuals using (2.3). To get the kth starting value $\alpha_{p+1}^{(k)}$, we divided the standard deviation of the residuals by $|\phi_{pq}^{(k)}|$, where $q := \max\{0 \leq j \leq p : \phi_{pj}^{(k)} \neq 0\}$. Finally, we randomly chose the starting values for $\boldsymbol{\theta}$.

The log-likelihood was evaluated at each of the 250 candidate values. When Z_1 follows the Students' t-distribution, the likelihood function is almost constant with respect to $(\alpha_{p+1}, \theta)'$ near $(\boldsymbol{\phi}_0', \alpha_{0,p+1}, \theta_0)'$, and so the maximum can be difficult to find. This is not the case when Z_1 is a Gaussian scale mixture. Therefore, when using (2.10), the collection of initial values was reduced to the nine with the highest likelihoods plus $\boldsymbol{\alpha}_0$, and, when using (2.12), the collection of initial values was reduced to the two with the highest likelihoods plus $\boldsymbol{\alpha}_0$. We found optimized values by implementing the Hooke and Jeeves (1961) algorithm and using the ten or three values as starting points. The optimized value with the greatest likelihood was selected to be $\hat{\boldsymbol{\alpha}}_{ML}$. We constructed confidence intervals for the elements of $\boldsymbol{\alpha}_0$ using (2.5) and the estimators (2.13)–(2.16).

Results of the simulations appear in Tables 2.2 and 2.3. In the tables, we see that the MLEs are approximately unbiased and the confidence interval coverages are fairly close to the nominal 95% level, particularly when $n = 5000$. For the Students' t-distribution, the asymptotic standard deviations tend to understate the true variability of the MLEs when $n = 500$, but are more accurate when $n = 5000$. Normal probability plots of the MLEs show that approximate normality is achieved when $n = 5000$.

2.4.2 Noninvertible ARMA Modeling

As mentioned in the introduction, all-pass models can be used to fit causal, noninvertible ARMA models. Suppose the series $\{X_t\}$ follows the model

$$\phi(B)X_t = \theta_i(B)\theta_{ni}(B)Z_t, \tag{2.17}$$

where $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ is a causal AR(p) polynomial (all the roots of $\phi(B)$ fall outside the unit circle), $\theta_i(B) = 1 + \theta_{i,1}B + \cdots + \theta_{i,q}B^q$ is an invertible MA(q) polynomial (all the roots of $\theta_i(B)$ fall outside the unit circle), $\theta_{ni}(B) = 1 + \theta_{ni,1}B + \cdots + \theta_{ni,r}B^r$ is a purely noninvertible

	Asymptotic		Empirical		
n	mean	std.dev.	mean (c.i.)	std.dev. (c.i.)	% coverage (c.i.)
500	$\phi_1 = 0.5$	0.0274	0.4971 (0.4951,0.4990)	0.0315 (0.0301,0.0329)	93.0 (91.4,94.6)
	$\alpha_2 = 3.4641$	0.4177	3.5533 (3.5144,3.5922)	0.6277 (0.5995,0.6546)	90.0 (88.1,91.9)
	$\theta = 3.0$	0.4480	3.1123 (3.0812,3.1433)	0.5008 (0.4783,0.5223)	95.8 (94.6,97.0)
5000	$\phi_1 = 0.5$	0.0087	0.4997 (0.4991,0.5003)	0.0091 (0.0087,0.0095)	93.4 (91.9,94.9)
	$\alpha_2 = 3.4641$	0.1321	3.4787 (3.4699,3.4876)	0.1427 (0.1363,0.1489)	94.0 (92.5,95.5)
	$\theta = 3.0$	0.1417	3.0084 (2.9989,3.0179)	0.1533 (0.1464,0.1599)	94.0 (92.5,95.5)
500	$\phi_1 = 0.3$	0.0290	0.2993 (0.2971,0.3014)	0.0345 (0.0330,0.0360)	90.6 (88.8,92.4)
	$\phi_2 = 0.4$	0.0290	0.3964 (0.3942,0.3986)	0.0350 (0.0335,0.0365)	90.1 (88.2,92.0)
	$\alpha_3 = 4.3301$	0.5222	4.4842 (4.4256,4.5428)	0.9460 (0.9036,0.9866)	94.0 (92.5,95.5)
	$\theta = 3.0$	0.4480	3.0789 (3.0497,3.1082)	0.4722 (0.4510,0.4925)	94.8 (93.4,96.2)
5000	$\phi_1 = 0.3$	0.0092	0.2999 (0.2993,0.3005)	0.0095 (0.0091,0.0099)	94.0 (92.5,95.5)
	$\phi_2 = 0.4$	0.0092	0.3999 (0.3993,0.4005)	0.0094 (0.0090,0.0098)	94.6 (93.2,96.0)
	$\alpha_3 = 4.3301$	0.1651	4.3421 (4.3313,4.3528)	0.1740 (0.1662,0.1815)	94.6 (93.2,96.0)
	$\theta = 3.0$	0.1417	3.0079 (2.9989,3.0169)	0.1458 (0.1393,0.1521)	95.2 (93.9,96.5)

Table 2.2: Empirical means, standard deviations, and percent coverages of nominal 95% confidence intervals for maximum likelihood estimates of all-pass model parameters when f is the rescaled Students' t-density. For each sample size, n, empirical confidence intervals were computed using standard asymptotic theory for 1000 iid replicates. Asymptotic means and standard deviations were computed using Theorem 1.

	Asymptotic		Empirical		
n	mean	std.dev.	mean (c.i.)	std.dev. (c.i.)	% coverage (c.i.)
500	$\phi_1 = 0.5$	0.0218	0.4989	0.0232	93.3
			(0.4975,0.5004)	(0.0222,0.0242)	(91.8,94.8)
	$\alpha_2 = 4.0$	0.2045	3.9984	0.1928	96.2
			(3.9865,4.0104)	(0.1841,0.2010)	(95.0,97.4)
	$\theta_1 = 0.6$	0.0476	0.6001	0.0492	92.5
			(0.5970,0.6031)	(0.0470,0.0513)	(90.9,94.1)
	$\theta_2 = 0.4$	0.0370	0.3995	0.0378	94.1
			(0.3972,0.4019)	(0.0361,0.0395)	(92.6,95.6)
5000	$\phi_1 = 0.5$	0.0069	0.5000	0.0070	94.1
			(0.4995,0.5004)	(0.0067,0.0073)	(92.6,95.6)
	$\alpha_2 = 4.0$	0.0646	3.9976	0.0643	94.3
			(3.9936,4.0016)	(0.0614,0.0671)	(92.9,95.7)
	$\theta_1 = 0.6$	0.0150	0.6001	0.0141	95.7
			(0.5992,0.6010)	(0.0135,0.0147)	(94.4,97.0)
	$\theta_2 = 0.4$	0.0117	0.3998	0.0114	95.5
			(0.3991,0.4005)	(0.0109,0.0119)	(94.2,96.8)
500	$\phi_1 = 0.3$	0.0230	0.2989	0.0239	93.7
			(0.2974,0.3004)	(0.0228,0.0249)	(92.2,95.2)
	$\phi_2 = 0.4$	0.0230	0.3990	0.0233	95.2
			(0.3975,0.4004)	(0.0223,0.0243)	(93.9,96.5)
	$\alpha_3 = 5.0$	0.2555	4.9902	0.2591	93.3
			(4.9742,5.0063)	(0.2475,0.2702)	(91.8,94.8)
	$\theta_1 = 0.6$	0.0476	0.5972	0.0483	94.5
			(0.5942,0.6001)	(0.0461,0.0503)	(93.1,95.9)
	$\theta_2 = 0.4$	0.0370	0.3977	0.0367	94.7
			(0.3954,0.4000)	(0.0351,0.0383)	(93.3,96.1)
5000	$\phi_1 = 0.3$	0.0073	0.3000	0.0074	95.1
			(0.2995,0.3004)	(0.0070,0.0077)	(93.8,96.4)
	$\phi_2 = 0.4$	0.0073	0.3996	0.0072	95.5
			(0.3991,0.4000)	(0.0069,0.0075)	(94.2,96.8)
	$\alpha_3 = 5.0$	0.0806	4.9960	0.0795	94.9
			(4.9911,5.0010)	(0.0759,0.0829)	(93.5,96.3)
	$\theta_1 = 0.6$	0.0150	0.6005	0.0147	94.5
			(0.5996,0.6014)	(0.0141,0.0154)	(93.1,95.9)
	$\theta_2 = 0.4$	0.0117	0.4006	0.0117	95.3
			(0.3999,0.4013)	(0.0112,0.0122)	(94.0,96.6)

Table 2.3: Empirical means, standard deviations, and percent coverages of nominal 95% confidence intervals for maximum likelihood estimates of all-pass model parameters when f is the Gaussian scale mixture density. For each sample size, n, empirical confidence intervals were computed using standard asymptotic theory for 1000 iid replicates. Asymptotic means and standard deviations were computed using Theorem 1.

MA(r) polynomial such that $\theta_{ni}(z) \neq 0$ for $|z| = 1$ (all the roots of $\theta_{ni}(B)$ fall inside the unit circle), and $\{Z_t\}$ is iid. If $\theta_{ni}^{(i)}(B)$ is the invertible rth order polynomial with roots that are the reciprocals of the roots of $\theta_{ni}(B)$ and $\{X_t\}$ is mistakenly modeled as the causal, invertible ARMA

$$\phi(B)X_t = \theta_i(B)\theta_{ni}^{(i)}(B)W_t,$$

then $\{W_t\}$ satisfies

$$\begin{aligned} W_t &= \frac{\theta_{ni}(B)}{\theta_{ni}^{(i)}(B)} Z_t \\ &= \frac{B^r \theta_{ni}^{(i)}(B^{-1})}{\theta_{ni,r}^{(i)}\theta_{ni}^{(i)}(B)} Z_t, \end{aligned}$$

where $\theta_{ni,r}^{(i)}$ is the coefficient of B^r in $\theta_{ni}^{(i)}(B)$. So, $\{W_t\}$ follows the causal all-pass model

$$\theta_{ni}^{(i)}(B)W_t = \frac{B^r \theta_{ni}^{(i)}(B^{-1})}{\theta_{ni,r}^{(i)}} Z_t.$$

Therefore, we have a way to avoid looking at all possible configurations of roots inside and outside the unit circle when fitting (2.17). First, fit a causal, invertible ARMA($p, q + r$) model to the data using a standard method such as Gaussian maximum likelihood, and obtain estimates of $\phi(B)$ and $\theta_i(B)\theta_{ni}^{(i)}(B)$ and the residuals $\{\hat{W}_t\}$. Then fit a causal all-pass model of order r to $\{\hat{W}_t\}$ and obtain $\hat{\theta}_{ni}^{(i)}(B)$, an estimate of $\theta_{ni}^{(i)}(B)$. The rth order polynomial with roots that are reciprocals of the roots of $\hat{\theta}_{ni}^{(i)}(B)$ is an estimate of $\theta_{ni}(B)$. An estimate of $\theta_i(B)$ can be obtained by canceling the roots of the invertible MA($q + r$) polynomial from the Gaussian likelihood fit which correspond to roots of $\hat{\theta}_{ni}^{(i)}(B)$.

2.4.3 Deconvolution

In this example, we simulate a seismogram $\{X_t\}_{t=1}^{1000}$ via

$$X_t = \sum_k \beta_k Z_{t-k},$$

where $\{\beta_k\}$ is the water gun wavelet sequence shown in Figure 8(2) of Lii and Rosenblatt (1988) and $\{Z_t\}$ is a reflectivity sequence simulated here as iid noise from the Students' t distribution with five degrees of freedom. It is assumed that the seismogram is observed, but the wavelet and reflectivity sequences are unknown, as would be the case in a real deconvolution problem. We model the seismogram as a possibly noninvertible ARMA using the procedure described in Section 2.4.2 and attempt to reconstruct the wavelet and reflectivity sequences. This problem is of interest because, for an observed water gun seismogram, the reflectivity sequence corresponds to reflection coefficients for layers of the earth.

The simulated seismogram $\{X_t\}$ is shown in Figure 2.1(a). The corrected Akaike information criterion indicates that an ARMA(12, 13) model is appropriate for the data, and the causal, invertible ARMA fit to $\{X_t\}$ using Gaussian maximum likelihood is $X_t = \phi^{-1}(B)\theta(B)W_t$, where

$$\begin{aligned}\phi(B) &= 1 - 0.255B + 0.451B^2 - 0.319B^3 + 0.141B^4 - 0.050B^5 - 0.214B^7 + 0.085B^8 \\ &\quad +0.212B^9 - 0.061B^{11} + 0.124B^{12}\end{aligned}$$

and

$$\begin{aligned}\theta(B) &= 1 - 0.221B + 0.124B^2 - 0.090B^3 - 0.178B^4 + 0.109B^5 - 0.339B^6 + 0.204B^7 \\ &\quad +0.081B^8 + 0.471B^9 - 0.184B^{10} + 0.234B^{12} + 0.322B^{13}.\end{aligned}$$

The residuals from this fitted model are denoted $\{\hat{W}_t\}$. From the sample autocorrelation functions of $\{\hat{W}_t\}$, $\{\hat{W}_t^2\}$, and $\{|\hat{W}_t|\}$ in Figure 2.1(b)–(d), it appears the ARMA residuals are uncorrelated but dependent, suggesting the inappropriateness of a causal, invertible ARMA model.

The all-pass order selection procedure indicates that an all-pass model of order two provides a good fit for $\{\hat{W}_t\}$ and, when the rescaled Students' t-density (2.10) is used, the MLEs for this fitted all-pass model are

$$\begin{aligned}\hat{\alpha}_{ML} &= (\hat{\phi}_1, \hat{\phi}_2, \hat{\alpha}_3, \hat{\theta})' \\ &= (1.531, -0.593, 1077930.125, 4.745)',\end{aligned}$$

with standard errors 0.034, 0.034, 43519.012, and 0.712 respectively. The sample autocorrelation functions for the squares and absolute values of $\{\hat{Z}_t\}$, the residuals from the fitted all-pass model,

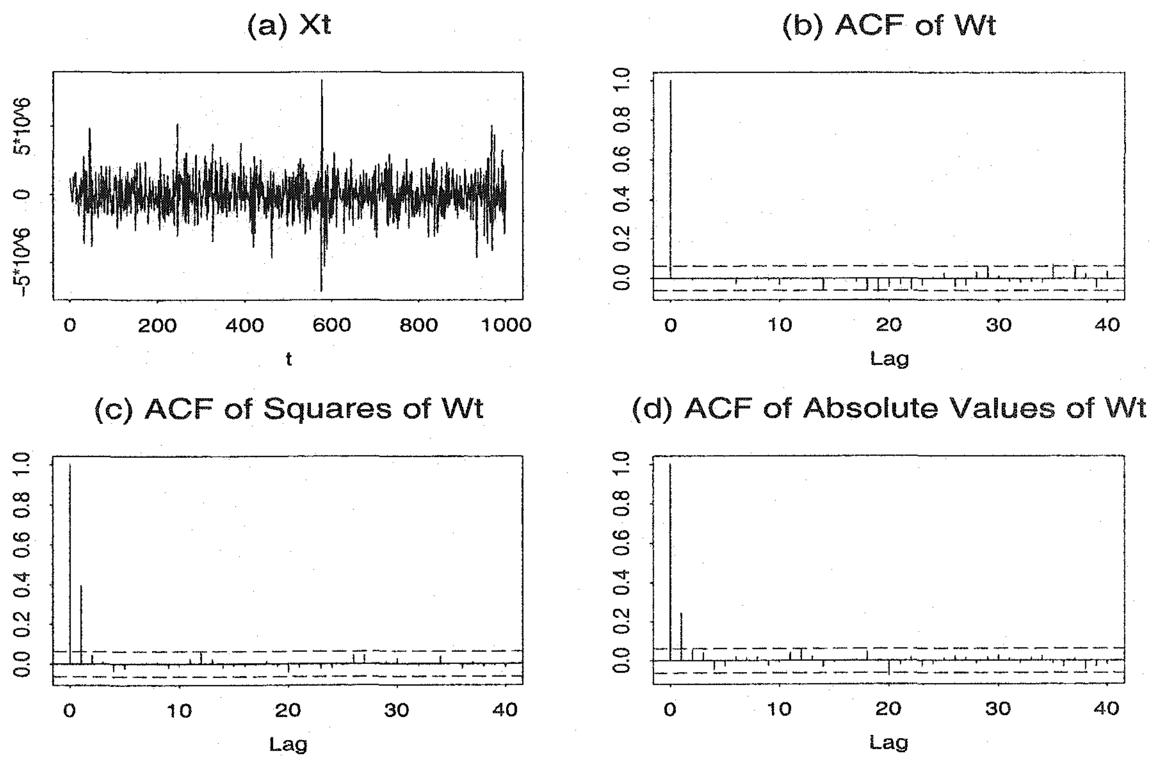

Figure 2.1: (a) The simulated seismogram of length 1000, $\{X_t\}$, and the sample autocorrelation functions with bounds $\pm 1.96/\sqrt{1000}$ for (b) $\{\hat{W}_t\}$, (c) $\{\hat{W}_t^2\}$, and (d) $\{|\hat{W}_t|\}$.

are shown in Figure 2.2. Because the series $\{\hat{Z}_t\}$ appears independent,

$$X_t = \frac{B^2(1 - 1.531B^{-1} + 0.593B^{-2})}{0.593(1 - 1.531B + 0.593B^2)} \frac{\theta(B)}{\phi(B)} Z_t \tag{2.18}$$

seems to be a more appropriate model for $\{X_t\}$.

In an effort to reconstruct the wavelet and reflectivity sequences, note that, since the simulated reflectivity sequence has variance 5/3, the right hand side of (2.18) is equal in distribution to

$$-(0.593)(1077930.125)\sqrt{3/5}\frac{B^2(1 - 1.531B^{-1} + 0.593B^{-2})}{0.593(1 - 1.531B + 0.593B^2)} \frac{\theta(B)}{\phi(B)} \tilde{Z}_t, \tag{2.19}$$

where $\{\tilde{Z}_t\}$ is iid with density $\sqrt{3/5}\, f(\sqrt{3/5}\, s; 5)$. Note, also, that no roots of the polynomial in the denominator of (2.19) cancel exactly with roots of the polynomial in the numerator. So, for further model accuracy, we can directly fit a causal, noninvertible ARMA(12, 13) with two roots of

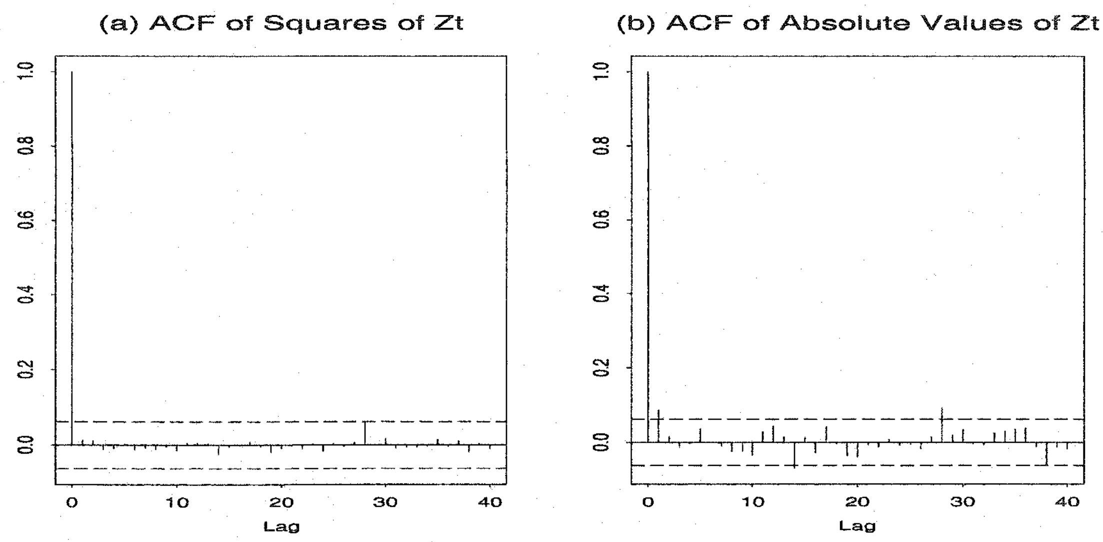

Figure 2.2: Diagnostics for the all-pass model of order two fit to the seismogram residuals. The sample autocorrelation functions with bounds $\pm 1.96/\sqrt{1000}$ for (a) $\{\hat{Z}_t^2\}$ and (b) $\{|\hat{Z}_t|\}$.

the moving average polynomial inside the unit circle. Using maximum likelihood estimation with the Students' t-density and five degrees of freedom, this yields

$$X_t = -(0.593)(1077930.125)\sqrt{3/5}\,\tilde{\phi}^{-1}(B)\tilde{\theta}(B)\tilde{Z}_t,$$

where

$$\begin{aligned}\tilde{\phi}(B) &= 1 - 0.053B + 0.502B^2 - 0.066B^3 + 0.315B^4 - 0.051B^5 + 0.218B^6 - 0.236B^7 \\ &\quad +0.039B^8 - 0.031B^9 - 0.063B^{10} - 0.195B^{11} - 0.067B^{12}\end{aligned}$$

and

$$\begin{aligned}\tilde{\theta}(B) &= 1 - 1.177B - 0.302B^2 - 0.357B^3 - 0.143B^4 + 0.023B^5 + 0.060B^6 + 0.629B^7 \\ &\quad +0.267B^8 + 0.718B^9 - 0.026B^{10} + 0.541B^{11} + 0.524B^{12} + 0.827B^{13}.\end{aligned}$$

As shown in Figures 2.3 and 2.4,

$$-(0.593)(1077930.125)\sqrt{3/5}\,\tilde{\phi}^{-1}(B)\tilde{\theta}(B)$$

and the residuals provide good estimates of the water gun wavelet and reflectivity sequences respectively. The estimates that can be obtained from the causal, invertible Gaussian maximum likelihood fit are not as accurate.

2.5 Additional Results

This section contains proofs of the lemmas used to establish the results of Section 2.3. First, for an arbitrary, causal autoregressive polynomial $\phi(z)$, define $\theta(z) = \phi_1 z + \cdots + \phi_p z^p = 1 - \phi(z)$, and define $\theta_0(z) = 1 - \phi_0(z)$. Note that, for $t = 1, \ldots, n-p$,

$$\phi(B)X_{t+p} = -z_t(\phi) + \varphi(B^{-1})z_t(\phi),$$

so, if $j = 1, \ldots, p$, then

$$\frac{\partial}{\partial \phi_j}\left\{\varphi(B^{-1})z_t(\phi)\right\} = -X_{t+p-j} + \frac{\partial z_t(\phi)}{\partial \phi_j}. \tag{2.20}$$

Also, if $j = 1, \ldots, p$, then

$$\begin{aligned}\frac{\partial}{\partial \phi_j}\left\{\varphi(B^{-1})z_t(\phi)\right\} &= \frac{\partial}{\partial \phi_j}\left\{\phi_1 z_{t+1}(\phi) + \cdots + \phi_p z_{t+p}(\phi)\right\} \\ &= \varphi(B^{-1})\frac{\partial z_t(\phi)}{\partial \phi_j} + z_{t+j}(\phi).\end{aligned} \tag{2.21}$$

Equating (2.20) and (2.21) and solving for $\partial z_t(\phi)/\partial \phi_j$, we obtain

$$\frac{\partial z_t(\phi)}{\partial \phi_j} = \frac{1}{\phi(B^{-1})}\left\{X_{t+p-j} + z_{t+j}(\phi)\right\}. \tag{2.22}$$

Evaluating (2.22) at the true value of ϕ and ignoring the effect of recursion initialization, we have

$$\begin{aligned}\frac{\partial z_t(\phi_0)}{\partial \phi_j} &= \frac{1}{\phi_0(B^{-1})}\left\{\frac{-\phi_0(B^{-1})B^p z_{t+p-j}}{\phi_0(B)} + z_{t+j}(\phi_0)\right\} \\ &\simeq \frac{-z_{t-j}}{\phi_0(B)} + \frac{z_{t+j}}{\phi_0(B^{-1})},\end{aligned} \tag{2.23}$$

where the first term is an element of $\sigma(z_{t-1}, z_{t-2}, \ldots)$ and the second term is an element of $\sigma(z_{t+1}, z_{t+2}, \ldots)$ because $\phi_0(B)$ is a causal operator and $\phi_0(B^{-1})$ is a purely noncausal operator. It follows that (2.23) is independent of $z_t = \phi_{0r}^{-1} Z_t$.

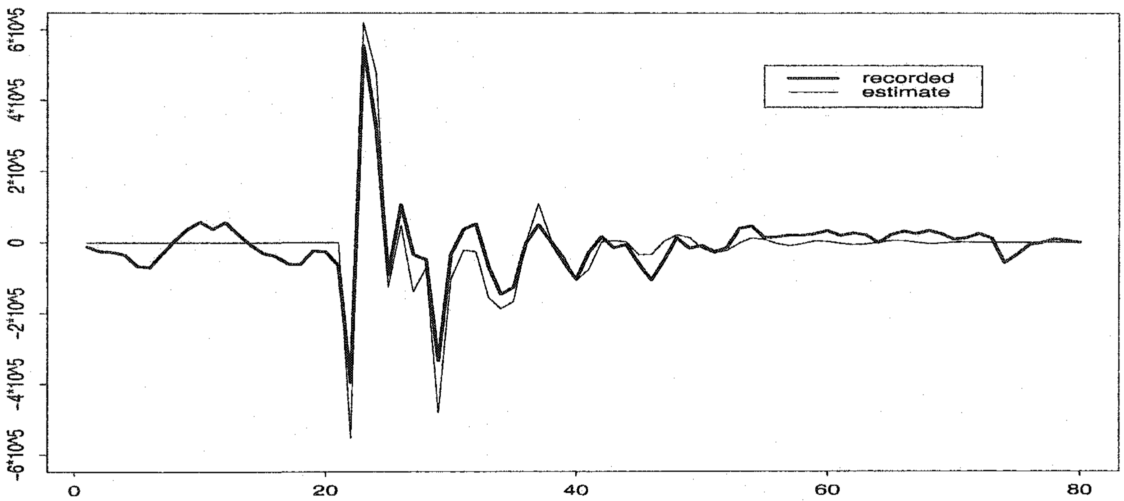

Figure 2.3: The recorded water gun wavelet and its estimate.

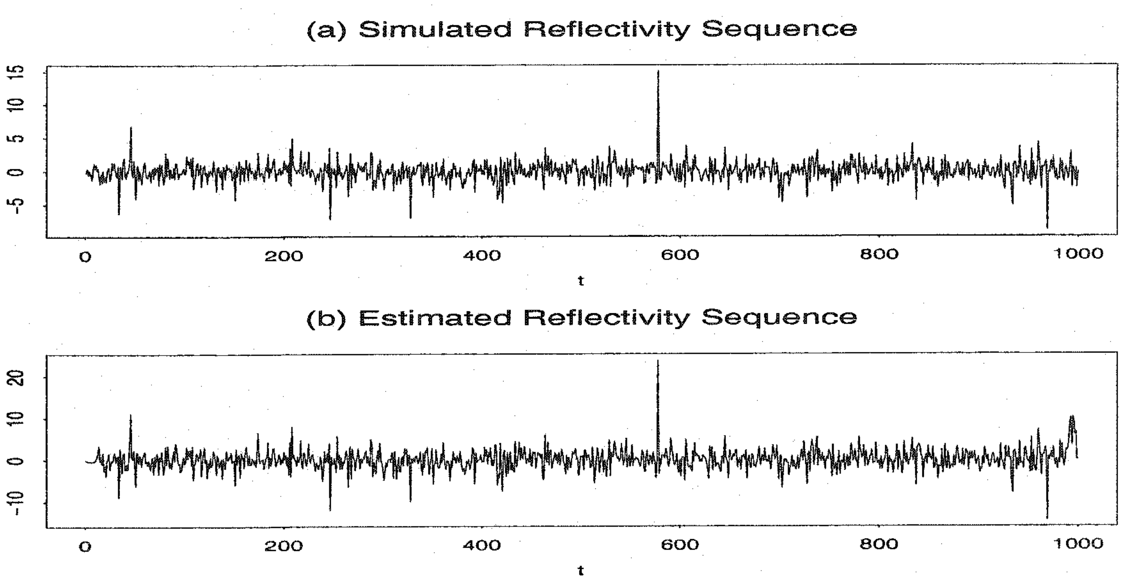

Figure 2.4: The simulated reflectivity sequence and its estimate.

Thus, for $j = 1, \ldots, p$,

$$\begin{aligned}\frac{\partial g_t(\boldsymbol{\alpha}_0)}{\partial \alpha_j} &= \frac{f'\left(z_t(\boldsymbol{\phi}_0)/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}{f\left(z_t(\boldsymbol{\phi}_0)/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}\frac{1}{\alpha_{0,p+1}}\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial \phi_j} \\ &\simeq \frac{f'\left(z_t/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}{f\left(z_t/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}\frac{1}{\alpha_{0,p+1}}\left\{\frac{-z_{t-j}}{\phi_0(B)} + \frac{z_{t+j}}{\phi_0(B^{-1})}\right\} \\ &=: \frac{\partial g_t^*(\boldsymbol{\alpha}_0)}{\partial \alpha_j}.\end{aligned} \tag{2.24}$$

The expectation of (2.24) is zero by the independence of its two terms.

We now compute the autocovariance function $\gamma^\dagger(h)$ of the zero-mean, stationary process $\left\{\mathbf{u}_p' \left[\partial g_t^*(\boldsymbol{\alpha}_0)/\partial\alpha_j\right]_{j=1}^p\right\}$ for $\mathbf{u}_p \in \mathbb{R}^p$:

$$\begin{aligned}\gamma^\dagger(h) &= \mathrm{E}\left\{\mathbf{u}_p'\left[\frac{\partial g_t^*(\boldsymbol{\alpha}_0)}{\partial \alpha_j}\right]_{j=1}^p\left(\left[\frac{\partial g_{t+h}^*(\boldsymbol{\alpha}_0)}{\partial \alpha_k}\right]_{k=1}^p\right)'\mathbf{u}_p\right\} \\ &= \mathbf{u}_p'\left[\nu_{jk}(h)\right]_{j,k=1}^p\mathbf{u}_p,\end{aligned}$$

where

$$\nu_{jk}(h) := \begin{cases} 2\gamma(j-k)\tilde{J}, & h = 0, \\ -\psi_{|h|-j}\psi_{|h|-k}, & h \neq 0, \end{cases}$$

and the ψ_ℓ are given by $\sum_{\ell=0}^\infty \psi_\ell z^\ell = 1/\phi_0(z)$ with $\psi_\ell = 0$ for $\ell < 0$. Thus,

$$\begin{aligned}\gamma^\dagger(0) + 2\sum_{h=1}^\infty \gamma^\dagger(h) &= \mathbf{u}_p'\left\{\left[2\tilde{J}\gamma(j-k)\right]_{j,k=1}^p - 2\left[\sum_{h=1}^\infty \psi_{h-j}\psi_{h-k}\right]_{j,k=1}^p\right\}\mathbf{u}_p \\ &= 2(\sigma_0^2\tilde{J} - 1)\mathbf{u}_p'\sigma_0^{-2}\boldsymbol{\Gamma}_p\mathbf{u}_p.\end{aligned}$$

By A5, $\sigma_0^2\tilde{J} - 1 > 0$.

Next,

$$\begin{aligned}\frac{\partial g_t(\boldsymbol{\alpha}_0)}{\partial \alpha_{p+1}} &= \frac{-f'\left(z_t(\boldsymbol{\phi}_0)/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}{f\left(z_t(\boldsymbol{\phi}_0)/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}\frac{z_t(\boldsymbol{\phi}_0)}{\alpha_{0,p+1}^2} - \frac{1}{\alpha_{0,p+1}} \\ &\simeq \frac{-f'\left(z_t/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}{f\left(z_t/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}\frac{z_t}{\alpha_{0,p+1}^2} - \frac{1}{\alpha_{0,p+1}} \\ &=: \frac{\partial g_t^*(\boldsymbol{\alpha}_0)}{\partial \alpha_{p+1}}.\end{aligned} \tag{2.25}$$

The expectation of (2.25) is zero and the variance is $\tilde{K}$. Also, the sequence (2.25) is iid and orthogonal to the corresponding partials for α_j, $j = 1, \ldots, p$, in (2.24).

For $j = p+2, \ldots, p+d+1$,

$$\begin{aligned} \frac{\partial g_t(\boldsymbol{\alpha}_0)}{\partial \alpha_j} &= \frac{1}{f\left(z_t(\boldsymbol{\phi}_0)/\alpha_{0,p+1}; \boldsymbol{\theta}_0\right)} \frac{\partial f\left(z_t(\boldsymbol{\phi}_0)/\alpha_{0,p+1}; \boldsymbol{\theta}_0\right)}{\partial \theta_{j-p-1}} \\ &\simeq \frac{1}{f\left(z_t/\alpha_{0,p+1}; \boldsymbol{\theta}_0\right)} \frac{\partial f\left(z_t/\alpha_{0,p+1}; \boldsymbol{\theta}_0\right)}{\partial \theta_{j-p-1}} \\ &=: \frac{\partial g_t^*(\boldsymbol{\alpha}_0)}{\partial \alpha_j}. \end{aligned} \tag{2.26}$$

By A7 and the dominated convergence theorem, the expectation of (2.26) is zero. In addition, the series $\left\{[\partial g_t^*(\boldsymbol{\alpha}_0)/\partial \alpha_j]_{j=p+2}^{p+d+1}\right\}$ is iid, has covariance matrix I, and is orthogonal to the partials for α_j, $j = 1, \ldots, p$, in (2.24). The expectation of $(\partial g_t^*(\boldsymbol{\alpha}_0)/\partial \alpha_{p+1}) [\partial g_t^*(\boldsymbol{\alpha}_0)/\partial \alpha_j]_{j=p+2}^{p+d+1}$ is L.

The preceding calculations lead directly to the following lemma.

Lemma 1 *If f satisfies A1–A7, then, as $n \to \infty$,*

$$n^{-1/2} \sum_{t=1}^{n-p} \frac{\partial g_t(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}} \xrightarrow{d} \mathbf{N} \sim N(\mathbf{0}, \boldsymbol{\Sigma}),$$

where

$$\boldsymbol{\Sigma} = \begin{bmatrix} 2(\sigma_0^2 \tilde{J} - 1)\sigma_0^{-2} \boldsymbol{\Gamma}_p & \mathbf{0}_{p\times 1} & \mathbf{0}_{p\times d} \\ \mathbf{0}_{1\times p} & \tilde{K} & L' \\ \mathbf{0}_{d\times p} & L & I \end{bmatrix}.$$

Proof: Note that, for $t = 0, \ldots, n-p-1$,

$$z_{n-p-t} = \sum_{l=0}^{\infty} \psi_l \left(\phi_0(B^{-1}) z_{n-p-t+l}\right) \quad \text{and} \quad z_{n-p-t}(\boldsymbol{\phi}_0) = \sum_{l=0}^{t} \psi_l \left(\phi_0(B^{-1}) z_{n-p-t+l}\right).$$

Because there exist constants $C \in (0, \infty)$ and $D \in (0,1)$ such that $|\psi_l| < CD^l$ for all $l \in \{0, 1, \ldots\}$ (see Brockwell and Davis, 1991, Section 3.1), using A7 and the mean value theorem we can show that

$$n^{-1/2} \sum_{t=1}^{n-p} \frac{\partial g_t(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}} - n^{-1/2} \sum_{t=1}^{n-p} \frac{\partial g_t^*(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}} \to \mathbf{0}$$

in L_1 and hence in probability.

Let $\mathbf{u} = (\mathbf{u}_p', u_1, \mathbf{u}_d')' \in \mathbb{R}^{p+d+1}$. By the Cramér-Wold device, it suffices to show

$$n^{-1/2} \sum_{t=1}^{n-p} V_t \xrightarrow{d} \mathrm{N}\left(0, 2(\sigma_0^2 \tilde{J} - 1)\mathbf{u}_p' \sigma_0^{-2} \boldsymbol{\Gamma}_p \mathbf{u}_p + u_1^2 \tilde{K} + 2u_1 \mathbf{u}_d' L + \mathbf{u}_d' I \mathbf{u}_d\right),$$

where $V_t := \mathbf{u}'\partial g_t^*(\boldsymbol{\alpha}_0)/\partial\boldsymbol{\alpha}$. Elements of the infinite order moving average stationary sequence $\{V_t\}$ can be truncated to create a finite order moving average stationary sequence. By applying a central limit theorem (Brockwell and Davis, 1991, Theorem 6.4.2) to each truncation level, asymptotic normality can be deduced. The details are omitted. □

Now consider the mixed partials of $g_t(\boldsymbol{\alpha})$. For $j, k = 1, \ldots, p$,

$$\begin{aligned}\frac{\partial^2 g_t(\boldsymbol{\alpha}_0)}{\partial\alpha_j\partial\alpha_k} &= \frac{f'\left(z_t(\boldsymbol{\phi}_0)/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}{f\left(z_t(\boldsymbol{\phi}_0)/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}\frac{1}{\alpha_{0,p+1}}\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial\phi_j\partial\phi_k}\\ &\quad+\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\phi_j}\frac{f''\left(z_t(\boldsymbol{\phi}_0)/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}{\alpha_{0,p+1}^2 f\left(z_t(\boldsymbol{\phi}_0)/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\phi_k}\\ &\quad-\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\phi_j}\frac{\left(f'\left(z_t(\boldsymbol{\phi}_0)/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)\right)^2}{\alpha_{0,p+1}^2 f^2\left(z_t(\boldsymbol{\phi}_0)/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\phi_k}.\end{aligned} \tag{2.27}$$

Because

$$\begin{aligned}\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial\phi_j\partial\phi_k} &= \frac{1}{\phi_0^2(B^{-1})}\left\{X_{t+p+j-k}+X_{t+p+k-j}+2z_{t+j+k}(\boldsymbol{\phi}_0)\right\}\\ &\simeq \frac{-z_{t+j-k}-z_{t+k-j}}{\phi_0(B^{-1})\phi_0(B)}+\frac{2z_{t+j+k}}{\phi_0^2(B^{-1})}\\ &= -\sum_{m=0}^{\infty}\sum_{\ell=0}^{\infty}\psi_m\psi_\ell(z_{t+j-k-\ell+m}+z_{t+k-j-\ell+m})+\frac{2z_{t+j+k}}{\phi_0^2(B^{-1})},\end{aligned}$$

(2.27) is approximately

$$\begin{aligned}\frac{\partial^2 g_t^*(\boldsymbol{\alpha}_0)}{\partial\alpha_j\partial\alpha_k} := &\frac{f'\left(z_t/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}{f\left(z_t/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}\frac{1}{\alpha_{0,p+1}}\\ &\times\left\{-\sum_{m=0}^{\infty}\sum_{\ell=0}^{\infty}\psi_m\psi_\ell(z_{t+j-k-\ell+m}+z_{t+k-j-\ell+m})+\frac{2z_{t+j+k}}{\phi_0^2(B^{-1})}\right\}\\ &+\frac{f\left(z_t/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)f''\left(z_t/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)-\left(f'\left(z_t/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)\right)^2}{\alpha_{0,p+1}^2 f^2\left(z_t/\alpha_{0,p+1};\boldsymbol{\theta}_0\right)}\\ &\times\left\{\frac{-z_{t-j}}{\phi_0(B)}+\frac{z_{t+j}}{\phi_0(B^{-1})}\right\}\left\{\frac{-z_{t-k}}{\phi_0(B)}+\frac{z_{t+k}}{\phi_0(B^{-1})}\right\},\end{aligned}$$

which has expectation $-2\sigma_0^{-2}\gamma(j-k)(\sigma_0^2\tilde{J}-1)$. Similar arguments show that the approximations of the mixed partials evaluated at the true parameter values have expectation zero for $j = 1, \ldots, p$, $k = p+1, \ldots, p+d+1$, $-\tilde{K}$ for $j = k = p+1$,

$$\frac{1}{\alpha_{0,p+1}}\int\frac{sf'(s;\boldsymbol{\theta}_0)}{f(s;\boldsymbol{\theta}_0)}\frac{\partial f(s;\boldsymbol{\theta}_0)}{\partial\theta_{k-p-1}}ds$$

for $j = p+1$, $k = p+2, \ldots, p+d+1$, and

$$-\int \frac{1}{f(s;\boldsymbol{\theta}_0)} \frac{\partial f(s;\boldsymbol{\theta}_0)}{\partial \theta_{j-p-1}} \frac{\partial f(s;\boldsymbol{\theta}_0)}{\partial \theta_{k-p-1}} ds$$

for $j, k = p+2, \ldots, p+d+1$.

Lemma 2 *If f satisfies A1–A7, then, as $n \to \infty$,*

$$n^{-1} \sum_{t=1}^{n-p} \frac{\partial^2 g_t(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}\, \partial \boldsymbol{\alpha}'} \xrightarrow{P} \begin{bmatrix} -2(\sigma_0^2 \tilde{J} - 1)\sigma_0^{-2}\boldsymbol{\Gamma}_p & \mathbf{0}_{p\times 1} & \mathbf{0}_{p\times d} \\ \mathbf{0}_{1\times p} & -\tilde{K} & -L' \\ \mathbf{0}_{d\times p} & -L & -I \end{bmatrix} = -\boldsymbol{\Sigma}.$$

Proof: By A7,

$$n^{-1} \sum_{t=1}^{n-p} \frac{\partial^2 g_t(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}\, \partial \boldsymbol{\alpha}'} - n^{-1} \sum_{t=1}^{n-p} \frac{\partial^2 g_t^*(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}\, \partial \boldsymbol{\alpha}'} \xrightarrow{P} \mathbf{0},$$

and, by the ergodic theorem and the computations preceding the lemma,

$$n^{-1} \sum_{t=1}^{n-p} \frac{\partial^2 g_t^*(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}\, \partial \boldsymbol{\alpha}'} \xrightarrow{P} -\boldsymbol{\Sigma}.$$

□

Lemma 3 *For $\mathbf{u} \in \mathbb{R}^{p+d+1}$, define*

$$S_n^{\dagger}(\mathbf{u}) = n^{-1/2} \sum_{t=1}^{n-p} \mathbf{u}' \frac{\partial g_t(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}} + \frac{1}{2} n^{-1} \sum_{t=1}^{n-p} \mathbf{u}' \frac{\partial^2 g_t(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}\, \partial \boldsymbol{\alpha}'} \mathbf{u}$$

and

$$S_n(\mathbf{u}) = \sum_{t=1}^{n-p} \left[g_t\left(\boldsymbol{\alpha}_0 + n^{-1/2}\mathbf{u}\right) - g_t(\boldsymbol{\alpha}_0) \right].$$

If f satisfies A1–A7,

1. *$S_n^{\dagger} \xrightarrow{d} S$ on $C(\mathbb{R}^{p+d+1})$, where*

$$S(\mathbf{u}) := \mathbf{u}'\mathbf{N} - \frac{1}{2}\mathbf{u}'\boldsymbol{\Sigma}\mathbf{u},$$

$\mathbf{N} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, *and $C(\mathbb{R}^{p+d+1})$ is the space of continuous functions on $\mathbb{R}^{p+d+1}$ where convergence is equivalent to uniform convergence on every compact set.*

2. $S_n \xrightarrow{d} S$ *on* $C(\mathbb{R}^{p+d+1})$.

Proof:

1. The finite dimensional distributions of $S_n^\dagger$ converge to those of S by Lemmas 1 and 2. Since $S_n^\dagger$ is quadratic in $\mathbf{u}$, $\{S_n^\dagger\}$ is tight on $C(K)$ for any compact set $K \subset \mathbb{R}^{p+d+1}$. Therefore, $S_n^\dagger$ converges to S on $C(\mathbb{R}^{p+d+1})$ by Theorem 7.1 in Billingsley (1999).

2. By a Taylor series expansion,

$$\begin{aligned} S_n(\mathbf{u}) &= \sum_{t=1}^{n-p}\left[g_t\left(\boldsymbol{\alpha}_0 + n^{-1/2}\mathbf{u}\right) - g_t\left(\boldsymbol{\alpha}_0\right)\right] \\ &= n^{-1/2}\sum_{t=1}^{n-p}\mathbf{u}'\frac{\partial g_t(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}} \\ &\quad + \frac{1}{2}n^{-1}\sum_{t=1}^{n-p}\mathbf{u}'\frac{\partial^2 g_t(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}\,\partial \boldsymbol{\alpha}'}\mathbf{u} \\ &\quad + \frac{1}{2}n^{-1}\sum_{t=1}^{n-p}\mathbf{u}'\left(\frac{\partial^2 g_t(\boldsymbol{\alpha}_n^*(\mathbf{u}))}{\partial \boldsymbol{\alpha}\,\partial \boldsymbol{\alpha}'} - \frac{\partial^2 g_t(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}\,\partial \boldsymbol{\alpha}'}\right)\mathbf{u} \end{aligned}$$

for some $\boldsymbol{\alpha}_n^*(\mathbf{u})$ on the line segment connecting $\boldsymbol{\alpha}_0$ and $\boldsymbol{\alpha}_0 + n^{-1/2}\mathbf{u}$. If $||\cdot||$ measures Euclidean distance,

$$\sup_{\mathbf{u}\in K} ||\boldsymbol{\alpha}_n^*(\mathbf{u}) - \boldsymbol{\alpha}_0|| \to 0$$

for any compact set $K \subset \mathbb{R}^{p+d+1}$, and so using A7 we can show that

$$\frac{1}{2}n^{-1}\sum_{t=1}^{n-p}\mathbf{u}'\left(\frac{\partial^2 g_t(\boldsymbol{\alpha}_n^*(\mathbf{u}))}{\partial \boldsymbol{\alpha}\,\partial \boldsymbol{\alpha}'} - \frac{\partial^2 g_t(\boldsymbol{\alpha}_0)}{\partial \boldsymbol{\alpha}\,\partial \boldsymbol{\alpha}'}\right)\mathbf{u} \xrightarrow{P} \mathbf{0}$$

on $C(\mathbb{R}^{p+d+1})$. Thus, $\{S_n\}$ must have the same limiting distribution as $\{S_n^\dagger\}$ on $C(\mathbb{R}^{p+d+1})$.

□

Now let $\tilde{z}_t(\boldsymbol{\phi}) = -(\phi(B)/\phi(B^{-1}))X_{t+p}$ and consider $\boldsymbol{\Xi}$, any compact, convex parameter space such that $\boldsymbol{\phi}$ forms a causal polynomial, $\alpha_{p+1} > 0$, and $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ for all $\boldsymbol{\alpha} = (\boldsymbol{\phi}', \alpha_{p+1}, \boldsymbol{\theta}')' \in \boldsymbol{\Xi}$.

Lemma 4 *If* f *satisfies A1, and* $(1+|s|)|f'(s;\boldsymbol{\theta})|/f(s;\boldsymbol{\theta})$ *and* $|\frac{\partial}{\partial\theta_j}f(s;\boldsymbol{\theta})|/f(s;\boldsymbol{\theta})$, $j = 1,\ldots,p$, *are dominated by* $a_3 + a_4|s|^{c_2}$ *for all* $\boldsymbol{\alpha} \in \boldsymbol{\Xi}$, *where* a_3, a_4 *and* c_2 *are non-negative constants such that*

$\int |s|^{c_2} f(s;\boldsymbol{\theta}_0) < \infty$, *then, as* $n \to \infty$,

$$n^{-1}\mathcal{L}(\boldsymbol{\alpha}) \stackrel{a.s.}{\to} E\{\ln f(\tilde{z}_1(\boldsymbol{\phi})/\alpha_{p+1};\boldsymbol{\theta}) - \ln \alpha_{p+1}\}$$

uniformly on Ξ.

Proof: For any $\boldsymbol{\alpha} \in \Xi$,

$$\mathrm{E}|\ln f(\tilde{z}_1(\boldsymbol{\phi})/\alpha_{p+1};\boldsymbol{\theta}) - \ln \alpha_{p+1}| < \infty$$

because $a_3 + a_4|s|^{c_2}$ dominates $(1+|s|)|\partial \ln f(s;\boldsymbol{\theta})/\partial s| = (1+|s|)|f'(s;\boldsymbol{\theta})|/f(s;\boldsymbol{\theta})$, and so

$$n^{-1}\mathcal{L}(\boldsymbol{\alpha}) \stackrel{a.s.}{\to} \mathrm{E}\{\ln f(\tilde{z}_1(\boldsymbol{\phi})/\alpha_{p+1};\boldsymbol{\theta}) - \ln \alpha_{p+1}\}$$

by the ergodic theorem. Therefore, the lemma follows by the Arzela-Ascoli theorem if $n^{-1}\mathcal{L}(\boldsymbol{\alpha})$ is equicontinuous and uniformly bounded on Ξ almost surely.

If $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \Xi$, then, for some $\boldsymbol{\alpha}^*$ on the line segment connecting $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$,

$$\begin{aligned}
&|n^{-1}\mathcal{L}(\boldsymbol{\alpha}_1) - n^{-1}\mathcal{L}(\boldsymbol{\alpha}_2)| \\
\leq\ & ||\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2|| n^{-1} \left\| \sum_{t=1}^{n-p} \frac{\partial}{\partial \boldsymbol{\alpha}} g_t(\boldsymbol{\alpha}^*) \right\|.
\end{aligned}$$

Because $z_t(\boldsymbol{\phi})$ and $\partial z_t(\boldsymbol{\phi})/\partial\boldsymbol{\phi}$ are continuous with respect to $\boldsymbol{\phi}$, there exist coefficients $\pi_k \geq 0$, $k = 0, \pm 1, \ldots$, decaying at a geometric rate such that

$$\sup_{\boldsymbol{\alpha}\in\Xi} |z_t(\boldsymbol{\phi})| \leq \sum_{k=-\infty}^{\infty} \pi_k |z_{t-k}| \quad \text{and} \quad \sup_{\boldsymbol{\alpha}\in\Xi} \left|\frac{\partial z_t(\boldsymbol{\phi})}{\partial \phi_j}\right| \leq \sum_{k=-\infty}^{\infty} \pi_k |z_{t-k}|,\ j = 1, \ldots, p,$$

for $t = 1, \ldots, n-p$ and $n = 1, 2, \ldots$. Also, because $\alpha_{p+1} > 0$ for every value of α_{p+1} in Ξ, there exists a constant $M > 0$ such that

$$\sup_{\boldsymbol{\alpha}\in\Xi} \frac{1}{\alpha_{p+1}} \leq M.$$

Consequently, we have

$$\begin{aligned}
\left|\frac{\partial g_t(\boldsymbol{\alpha}^*)}{\partial \alpha_j}\right| &= \left| \frac{f'(z_t(\boldsymbol{\phi}^*)/\alpha^*_{p+1};\boldsymbol{\theta}^*)}{f(z_t(\boldsymbol{\phi}^*)/\alpha^*_{p+1};\boldsymbol{\theta}^*)} \frac{1}{\alpha^*_{p+1}} \frac{\partial z_t(\boldsymbol{\phi}^*)}{\partial \phi_j} \right| \\
&\leq \left[a_3 + a_4 \left(\frac{z_t(\boldsymbol{\phi}^*)}{\alpha^*_{p+1}} \right)^{c_2 - 1} \right] \frac{1}{\alpha^*_{p+1}} \left| \frac{\partial z_t(\boldsymbol{\phi}^*)}{\partial \phi_j} \right| \\
&\leq M \left[a_3 \left(\sum_{k=-\infty}^{\infty} \pi_k |z_{t-k}| \right) + a_4 \left(\sum_{k=-\infty}^{\infty} \pi_k |z_{t-k}| \right)^{c_2} \right]
\end{aligned}$$

for $j = 1, \ldots, p$,

$$\begin{aligned}\left|\frac{\partial g_t(\boldsymbol{\alpha}^*)}{\partial \alpha_{p+1}}\right| &= \left|\frac{f'(z_t(\boldsymbol{\phi}^*)/\alpha^*_{p+1};\boldsymbol{\theta}^*)}{f(z_t(\boldsymbol{\phi}^*)/\alpha^*_{p+1};\boldsymbol{\theta}^*)}\frac{z_t(\boldsymbol{\phi}^*)}{(\alpha^*_{p+1})^2}+\frac{1}{\alpha^*_{p+1}}\right| \\ &\leq \frac{1}{\alpha^*_{p+1}}\left[a_3 + a_4\left(\frac{z_t(\boldsymbol{\phi}^*)}{\alpha^*_{p+1}}\right)^{c_2}\right] + \frac{1}{\alpha^*_{p+1}} \\ &\leq M\left[a_3 + a_4\left(M\sum_{k=-\infty}^{\infty}\pi_k|z_{t-k}|\right)^{c_2}\right] + M,\end{aligned}$$

and

$$\begin{aligned}\left|\frac{\partial g_t(\boldsymbol{\alpha}^*)}{\partial \alpha_j}\right| &= \left|\frac{1}{f(z_t(\boldsymbol{\phi}^*)/\alpha^*_{p+1};\boldsymbol{\theta}^*)}\frac{\partial f(z_t(\boldsymbol{\phi}^*)/\alpha^*_{p+1};\boldsymbol{\theta}^*)}{\partial \theta_{j-p-1}}\right| \\ &\leq a_3 + a_4\left(\frac{z_t(\boldsymbol{\phi}^*)}{\alpha^*_{p+1}}\right)^{c_2} \\ &\leq a_3 + a_4\left(M\sum_{k=-\infty}^{\infty}\pi_k|z_{t-k}|\right)^{c_2}\end{aligned}$$

for $j = p+2, \ldots, p+d+1$. It follows that almost surely, for n sufficiently large,

$$n^{-1}\left\|\sum_{t=1}^{n-p}\frac{\partial}{\partial \boldsymbol{\alpha}}g_t(\boldsymbol{\alpha}^*)\right\| \leq constant,$$

and so $n^{-1}\mathcal{L}(\boldsymbol{\alpha})$ is equicontinuous on Ξ almost surely. It can be shown similarly that $n^{-1}\mathcal{L}(\boldsymbol{\alpha})$ is uniformly bounded on Ξ almost surely. □

Chapter 3

Rank Estimation for All-Pass Time Series Models

3.1 Introduction

All-pass models are autoregressive-moving average (ARMA) models in which the roots of the autoregressive polynomial are reciprocals of roots of the moving average polynomial and vice versa. These models generate uncorrelated (white noise) time series that are not independent in the non-Gaussian case. As discussed in Chapter 2, an all-pass series can be obtained by fitting a causal, invertible ARMA model (all the roots of the autoregressive and moving average polynomials are outside the unit circle) to a series generated by a causal, noninvertible ARMA model (all the roots of the autoregressive polynomial are outside the unit circle and at least one root of the moving average polynomial is inside the unit circle). The residuals follow an all-pass model of order r, where r is the number of roots of the true moving average polynomial inside the unit circle. Noninvertible ARMA models have been used, for example, in vocal tract filters (Chi and Kung, 1995; Chien, Yang, and Chi, 1997), in the analysis of unemployment rates (Huang and Pawitan, 2000), and in seismogram deconvolution (Chapter 2; Lii and Rosenblatt, 1988). Similarly, an all-pass series can be obtained by fitting a causal autoregressive model to a series generated by a noncausal autoregressive model (Breidt, Davis, and Trindade, 2001). See Chapter 1 for a list of applications for noncausal models.

Estimation methods based on Gaussian likelihood, least-squares, or related second-order moment techniques cannot identify all-pass models because Gaussian all-pass series are independent. Therefore, cumulant-based estimators, using cumulants of order greater than two, are often used to estimate such models (Giannakis and Swami, 1990; Chi and Kung, 1995; Chien, Yang, and Chi, 1997). Breidt, Davis, and Trindade (2001) consider a least absolute deviations (LAD) approach motivated by approximating the likelihood of an all-pass model with Laplace (two-sided exponential) noise, and a maximum likelihood (ML) approach is considered in Chapter 2. Under general conditions, LAD and ML estimators are asymptotically normal. Because the objective functions for LAD and ML estimation are complicated functions of the model parameters, they tend to be quite bumpy and can, therefore, be hard to minimize or maximize. In addition, ML estimation is difficult to implement when the noise distribution is complicated or unknown. Rank (R) estimation can often overcome these limitations of LAD and ML estimation.

We consider a R-estimator first proposed by Jaeckel (1972) for estimating linear regression parameters. Jaeckel's estimator minimizes the sum of model residuals weighted by a function of residual rank. In this chapter, the asymptotic properties of this R-estimator are studied for all-pass model parameters. If the weight function is properly chosen, R-estimators can be nearly as asymptotically efficient as ML estimators, and the objective function for R-estimation can be fairly smooth and hence easy to minimize. Because the objective function involves not only the residual ranks but also the residual values, this is not pure R-estimation. Koul and Saleh (1993), Terpstra, McKean, and Naranjo (2001), and Mukherjee and Bai (2002) consider related estimation approaches for autoregressive model parameters. Also, Allal, Kaaouachi, and Paindaveine (2001) examine a pure R-estimator for ARMA model parameters based on correlations of weighted residual ranks. Their results are not applicable to all-pass models because the parameters in the autoregressive polynomial of an all-pass model are functions of parameters in the moving average polynomial and vice versa.

In Section 3.2, we consider Jaeckel's R-function in the context of all-pass parameter estimation. Asymptotic normality for R-estimators is established under general conditions and order selection is discussed in Section 3.3. Proofs of the lemmas used to confirm the results of Section 3.3 can be found in Section 3.5. We study the behavior of the estimators for finite samples via simulation in Section 3.4.1 and apply the estimation procedure to exchange rate log returns in Section 3.4.2.

3.2 Preliminaries

3.2.1 All-Pass Models

Let B denote the backshift operator ($B^k X_t = X_{t-k}$, $k = 0, \pm 1, \pm 2, \ldots$) and let

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$$

be a pth order autoregressive polynomial, where $\phi(z) \neq 0$ for $|z| = 1$. The polynomial $\phi(B)$ is said to be *causal* if all its roots are outside the unit circle in the complex plane. In this case, for a sequence $\{W_t\}$,

$$\phi^{-1}(B)W_t = \left(\sum_{j=0}^{\infty} \psi_j B^j\right) W_t = \sum_{j=0}^{\infty} \psi_j W_{t-j},$$

a function of only the past and present $\{W_t\}$. Note that if $\phi(B)$ is causal, the polynomial $B^p \phi(B^{-1})$ is *purely noncausal* in the sense that all its roots are inside the unit circle, and hence

$$B^{-p}\phi^{-1}(B^{-1})W_t = \left(\sum_{j=0}^{\infty} \psi_j B^{-p-j}\right) W_t = \sum_{j=0}^{\infty} \psi_j W_{t+p+j},$$

a function of only the present and future $\{W_t\}$. See, for example, Chapter 3 of Brockwell and Davis (1991).

Let

$$\phi_0(z) = 1 - \phi_{01} z - \cdots - \phi_{0p} z^p,$$

where $\phi_0(z) \neq 0$ for $|z| \leq 1$. Define $\phi_{00} = 1$ and suppose $\phi_{0r} \neq 0$ for some $r \in \{0, 1, \ldots, p\}$ and $\phi_{0j} = 0$ for $j = r+1, \ldots, p$. Then, a causal all-pass time series is the ARMA series $\{X_t\}$ which satisfies the difference equations

$$\phi_0(B)X_t = \frac{B^r \phi_0(B^{-1})}{-\phi_{0r}} Z_t^* \tag{3.1}$$

or

$$X_t - \phi_{01}X_{t-1} - \cdots - \phi_{0r}X_{t-r} = Z_t^* + \frac{\phi_{0,r-1}}{\phi_{0r}}Z_{t-1}^* + \cdots + \frac{\phi_{01}}{\phi_{0r}}Z_{t-r+1}^* - \frac{1}{\phi_{0r}}Z_{t-r}^*,$$

where the series $\{Z_t^*\}$ is an independent and identically distributed (iid) sequence of random variables with mean 0 and variance $\sigma^2 \in (0,\infty)$. The true order of the all-pass model is r. We assume throughout that Z_1^* has a distribution function F that is strictly increasing and continuously differentiable on $\mathbb{R}$ with density f. Observe that the roots of the autoregressive polynomial $\phi_0(z)$ are reciprocals of the roots of the moving average polynomial $-\phi_{0r}^{-1}z^r\phi_0(z^{-1})$ and vice versa.

The spectral density for $\{X_t\}$ in (3.1) is

$$\frac{|e^{-ir\omega}|^2|\phi_0(e^{i\omega})|^2}{\phi_{0r}^2|\phi_0(e^{-i\omega})|^2}\frac{\sigma^2}{2\pi} = \frac{\sigma^2}{\phi_{0r}^2 2\pi},$$

which is constant for $\omega \in [-\pi,\pi]$, and thus $\{X_t\}$ is an uncorrelated sequence. In the case of Gaussian $\{Z_t^*\}$, this implies that $\{X_t\}$ is iid $\mathrm{N}(0,\sigma^2\phi_{0r}^{-2})$, but independence does not hold in the non-Gaussian case if $r \geq 1$ (e.g., Breidt and Davis, 1991). The model (3.1) is called all-pass because the power transfer function of the all-pass filter passes all the power for every frequency in the spectrum. In other words, an all-pass filter does not change the distribution of power over the spectrum.

We can express (3.1) as

$$\phi_0(B)X_t = \frac{B^p\phi_0(B^{-1})}{-\phi_{0r}}Z_t, \tag{3.2}$$

where $\{Z_t\} = \{Z_{t+p-r}^*\}$ is an iid sequence of random variables with mean 0, variance σ^2, and distribution function F. Rearranging (3.2) and setting $z_t = \phi_{0r}^{-1}Z_t$, we have the backward recursion

$$z_{t-p} = \phi_{01}z_{t-p+1} + \cdots + \phi_{0p}z_t - (X_t - \phi_{01}X_{t-1} - \cdots - \phi_{0p}X_{t-p}).$$

An analogous recursion for an arbitrary, causal autoregressive polynomial $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ can be defined as follows:

$$z_{t-p}(\phi) = \begin{cases} 0, & t = n+p,\ldots,n+1, \\ \phi_1 z_{t-p+1}(\phi) + \cdots + \phi_p z_t(\phi) - \phi(B)X_t, & t = n,\ldots,p+1, \end{cases} \tag{3.3}$$

where $\phi := (\phi_1,\ldots,\phi_p)'$. Let $\phi_0 = (\phi_{01},\ldots,\phi_{0p})' = (\phi_{01},\ldots,\phi_{0r},0,\ldots,0)'$ denote the true parameter vector. Note that $\{z_t(\phi_0)\}_{t=1}^{n-p}$ closely approximates $\{z_t\}_{t=1}^{n-p}$; the error is due to the initialization with zeros. Although $\{z_t\}$ is iid, $\{z_t(\phi_0)\}_{t=1}^{n-p}$ is not iid if $r \geq 1$.

3.2.2 Jaeckel's Rank Function

Suppose we have a realization of length n, $\{X_1, \ldots, X_n\}$, from (3.1). Let φ be a differentiable, strictly increasing function from $(0,1)$ to $\mathbb{R}$ such that

$$\varphi(s) = -\varphi(1-s) \;\; \forall s \in (0,1). \tag{3.4}$$

If ϕ forms a causal autoregressive, pth order polynomial and $\{R_t(\phi)\}_{t=1}^{n-p}$ contains the ranks of $\{z_t(\phi)\}_{t=1}^{n-p}$ from (3.3), then the R-function attributed to Jaeckel evaluated at ϕ with weight function φ is

$$D(\phi) := \sum_{t=1}^{n-p} \varphi\left(\frac{R_t(\phi)}{n-p+1}\right) z_t(\phi). \tag{3.5}$$

Because it tends to be near zero when the elements of $\{z_t(\phi)\}$ are similar, (3.5) is a measure of the dispersion of the residuals $\{z_t(\phi)\}$. When $\{z_{(t)}(\phi)\}_{t=1}^{n-p}$ is the series $\{z_t(\phi)\}_{t=1}^{n-p}$ ordered from smallest to largest, (3.5) can also be written as

$$D(\phi) = \sum_{t=1}^{n-p} \varphi\left(\frac{t}{n-p+1}\right) z_{(t)}(\phi).$$

A popular choice for the weight function is $\varphi(s) = s - 1/2$. In this case, the weights $\{\varphi\left(t/(n-p+1)\right)\}_{t=1}^{n-p}$ are known as the Wilcoxon scores.

We give some properties for D in the following two theorems. Jaeckel (1972) shows that these same properties hold for the rank function in the linear regression case.

Theorem 1 *For any $\phi \in \mathbb{R}^p$, if*

$$\{P_1(\phi), \ldots, P_{(n-p)!}(\phi)\} = \{\{z_{1,1}(\phi), \ldots, z_{1,n-p}(\phi)\}, \ldots, \{z_{(n-p)!,1}(\phi), \ldots, z_{(n-p)!,n-p}(\phi)\}\}$$

contains the $(n-p)!$ permutations of the sequence $\{z_t(\phi)\}_{t=1}^{n-p}$, then

$$D(\phi) = \sup_{j \in \{1,\ldots,(n-p)!\}} \sum_{t=1}^{n-p} \varphi\left(\frac{t}{n-p+1}\right) z_{j,t}(\phi).$$

Proof: See the proof of Theorem 1 in Jaeckel (1972). □

Theorem 2 *D is a non-negative, continuous function on $\mathbb{R}^p$. Also, $D(\phi) = 0$ if and only if the elements of $\{z_t(\phi)\}_{t=1}^{n-p}$ are all equal.*

Proof: This argument is similar to one given in the proof of Theorem 1 in Jaeckel (1972). Suppose $\phi \in \mathbb{R}^p$ and

$$\varphi\left(\frac{t_0}{n-p+1}\right) \leq 0 < \varphi\left(\frac{t_0+1}{n-p+1}\right),$$

with $t_0 \in \{1, \ldots, n-p-1\}$. Then

$$\begin{aligned} D(\phi) &= \sum_{t=1}^{n-p} \varphi\left(\frac{R_t(\phi)}{n-p+1}\right) z_t(\phi) \\ &= \sum_{t=1}^{n-p} \varphi\left(\frac{t}{n-p+1}\right) z_{(t)}(\phi) - z_{(t_0)}(\phi) \sum_{t=1}^{n-p} \varphi\left(\frac{t}{n-p+1}\right) \\ &= \sum_{t=1}^{n-p} \varphi\left(\frac{t}{n-p+1}\right) \left(z_{(t)}(\phi) - z_{(t_0)}(\phi)\right) \end{aligned} \tag{3.6}$$

because, by (3.4),

$$\sum_{t=1}^{n-p} \varphi\left(\frac{t}{n-p+1}\right) = 0.$$

Since φ is strictly increasing, all terms in the sum (3.6) are non-negative, and so $D(\phi)$ is non-negative. Also, it must be the case that $D(\phi)$ equals zero if and only if the elements of $\{z_t(\phi)\}_{t=1}^{n-p}$ are all equal.

As a function of ϕ, $z_{n-p}(\phi)$ is a polynomial of degree one, $z_{n-p-1}(\phi)$ is a polynomial of degree two, ..., and $z_1(\phi)$ is a polynomial of degree $n-p$. Hence,

$$\sum_{t=1}^{n-p} \varphi\left(\frac{t}{n-p+1}\right) z_{j,t}(\cdot)$$

is continuous on $\mathbb{R}^p$ for any $j \in \{1, \ldots, (n-p)!\}$, and it follows that

$$D(\cdot) = \sup_{j \in \{1,\ldots,(n-p)!\}} \sum_{t=1}^{n-p} \varphi\left(\frac{t}{n-p+1}\right) z_{j,t}(\cdot)$$

must also be continuous on $\mathbb{R}^p$. □

3.3 Asymptotic Results

3.3.1 Parameter Estimation

Suppose f is uniformly continuous on $\mathbb{R}$, $\sup_{s\in\mathbb{R}} |s|f(s) < \infty$, and φ' is uniformly continuous on $(0,1)$. Also, let $\tilde{J} = \int_0^1 \varphi^2(s)\,ds$, $\tilde{K} = \int_0^1 F^{-1}(s)\varphi(s)\,ds$, and $\tilde{L} = \int_0^1 f(F^{-1}(s))\varphi'(s)\,ds$, and assume $\sigma^2\tilde{L} > \tilde{K}$.

Theorem 3 *There exists a sequence of minimizers $\hat{\phi}_R$ of $D(\cdot)$ in (3.5) such that*

$$n^{1/2}(\hat{\phi}_R - \phi_0) \xrightarrow{d} \mathbf{Y} \sim N(\mathbf{0}, \mathbf{\Sigma}), \tag{3.7}$$

where

$$\mathbf{\Sigma} := \frac{\sigma^2\tilde{J} - \tilde{K}^2}{2(\sigma^2\tilde{L} - \tilde{K})^2}\sigma^2\mathbf{\Gamma}_p^{-1},$$

$\mathbf{\Gamma}_p := [\gamma(j-k)]_{j,k=1}^p$, *and $\gamma(\cdot)$ is the autocovariance function for the autoregressive process $\{(1/\phi_0(B))Z_t\}$.*

Proof: See Lemma 14 in Section 3.5. □

Remark 1: Using the Cauchy-Schwartz inequality,

$$\begin{aligned}
\sigma^2\tilde{J} - \tilde{K}^2 &= \sigma^2\mathrm{E}\left\{\varphi^2(F(Z_1))\right\} - \left(\mathrm{E}\left\{Z_1\varphi(F(Z_1))\right\}\right)^2 \\
&\geq \sigma^2\mathrm{E}\left\{\varphi^2(F(Z_1))\right\} - \mathrm{E}\left\{Z_1^2\right\}\mathrm{E}\left\{\varphi^2(F(Z_1))\right\} \qquad (3.8)\\
&= 0,
\end{aligned}$$

with equality in (3.8) if and only if φ is proportional to F^{-1}, which is not possible since $F^{-1}(0) = -\infty$, $F^{-1}(1) = \infty$, and φ is bounded on $(0,1)$. Also,

$$\tilde{K} = \int_0^1 F^{-1}(s)\varphi(s)\,ds > 0$$

since F^{-1} and φ are strictly increasing functions on $(0,1)$ and φ is odd about $1/2$.

Remark 2: The asymptotic covariance matrix in (3.7) is a scalar multiple of the asymptotic covariance matrix for Gaussian likelihood estimators of the parameters of the corresponding pth order autoregressive process. The same property holds for LAD and ML estimators, as shown in Breidt, Davis, and Trindade (2001) and Chapter 2 respectively. The LAD estimators are obtained by maximizing the likelihood of an all-pass model with Laplace noise. This yields a modified LAD criterion, which can be used even if the underlying noise distribution is not Laplace. The appropriate scalar multiple is

$$\frac{\mathrm{Var}|Z_1|}{2\left(2\sigma^2 f(0) - \mathrm{E}|Z_1|\right)^2} \tag{3.9}$$

in the LAD case (Breidt, Davis, and Trindade (2001) contains an error in the calculation of the asymptotic variance; see the Appendix for the correction) and

$$\frac{1}{2}\left(\sigma^2 \int \frac{(f'(s))^2}{f(s)}\, ds - 1\right)^{-1} \tag{3.10}$$

in the ML case, while the multiple in (3.7) is

$$\frac{\sigma^2 \tilde{J} - \tilde{K}^2}{2(\sigma^2 \tilde{L} - \tilde{K})^2}. \tag{3.11}$$

Consider the sequence of weight functions $\{\varphi_m\}$ such that

$$\varphi_m(s) = \frac{2}{\pi} \arctan\left(m\left(s - \frac{1}{2}\right)\right).$$

Note that $\varphi_m(s) \to \varphi(s)$ pointwise as $m \to \infty$, when

$$\varphi(s) = \begin{cases} -1, & s < 1/2, \\ 0, & s = 1/2, \\ 1, & s > 1/2. \end{cases}$$

Note, also, that

$$\tilde{J}_m = \int_0^1 \varphi_m^2(s)\, ds \to 1.$$

If Z_1 has median zero,

$$\tilde{K}_m = \int_0^1 F^{-1}(s)\varphi_m(s)\, ds = \mathrm{E}\{Z_1 \varphi_m(F(Z_1))\} \to \mathrm{E}|Z_1|$$

and

$$\tilde{L}_m = \int_0^1 f(F^{-1}(s))\varphi'_m(s)\, ds = \mathrm{E}\{f(Z_1)\varphi'_m(F(Z_1))\} \to 2f(0).$$

Hence, if Z_1 has median zero,

$$\frac{\sigma^2 \tilde{J}_m - \tilde{K}_m^2}{2(\sigma^2 \tilde{L}_m - \tilde{K}_m)^2} \to \frac{\sigma^2 - \mathrm{E}^2|Z_1|}{2\left(2\sigma^2 f(0) - \mathrm{E}|Z_1|\right)^2} = \frac{\mathrm{Var}|Z_1|}{2\left(2\sigma^2 f(0) - \mathrm{E}|Z_1|\right)^2},$$

and so, using the weight function φ_m, we can obtain the same asymptotic efficiency with LAD and R-estimation when $m \to \infty$.

Remark 3: If Z_1 has a Laplace or two-sided exponential distribution,

$$f(s) = \frac{1}{\sqrt{2}\sigma} \exp\left(\frac{-\sqrt{2}|s|}{\sigma}\right),$$

$\mathrm{E}|Z_1| = \sigma/\sqrt{2}$, $f(0) = 1/(\sqrt{2}\sigma)$, and thus (3.9) equals 1/2. The constant (3.10) also equals 1/2 because, when the noise distribution is Laplace, LAD estimation corresponds to ML estimation. For the Laplace distribution,

$$F^{-1}(s) = \begin{cases} \frac{\sigma}{\sqrt{2}} \ln(2s), & 0 < s < \frac{1}{2}, \\ -\frac{\sigma}{\sqrt{2}} \ln(2-2s), & \frac{1}{2} \le s < 1, \end{cases} \quad \text{and} \quad f(F^{-1}(s)) = \begin{cases} \frac{\sqrt{2}}{\sigma} s, & 0 < s < \frac{1}{2}, \\ \frac{\sqrt{2}}{\sigma}(1-s), & \frac{1}{2} \le s < 1. \end{cases}$$

Therefore, if we use the Wilcoxon weight function $\varphi(s) = s - 1/2$, $\sigma^2 \tilde{L} > \tilde{K}$ and (3.11) equals 5/6. Consequently, the asymptotic relative efficiency (ARE) of R to ML is 0.6. Even though ML estimation is 40% asymptotically more efficient than R-estimation in this case, R-estimation can be useful because $D(\cdot)$ tends to be smoother than $\sum_{t=1}^{n-p} |z_t(\cdot)|$ and hence easier to minimize. Figure 3.1 shows ML and R (with Wilcoxon weights) objective functions for a realization of length $n = 50$ from an all-pass model with $p = 1$, $\phi_{01} = 0.5$, and Laplace noise with variance one. Note that the ML objective function has a large number of local minima and thus could be difficult to minimize using numerical optimization techniques.

Remark 4: R-estimation with the Wilcoxon weight function performs more efficiently when the noise distribution is the Students' t with three degrees of freedom. In this case, because $\mathrm{E}|Z_1| = 2\sqrt{3}/\pi$ and $f(0) = 2/(\sqrt{3}\pi)$, (3.9) equals $(\pi^2 - 4)/8$. Also, the constant (3.10) equals 1/2. The distribution function for the noise is given by

$$F(s) = \frac{1}{2} + \frac{\sqrt{3}s}{\pi(s^2+3)} + \frac{1}{\pi} \arctan\left(\frac{s}{\sqrt{3}}\right),$$

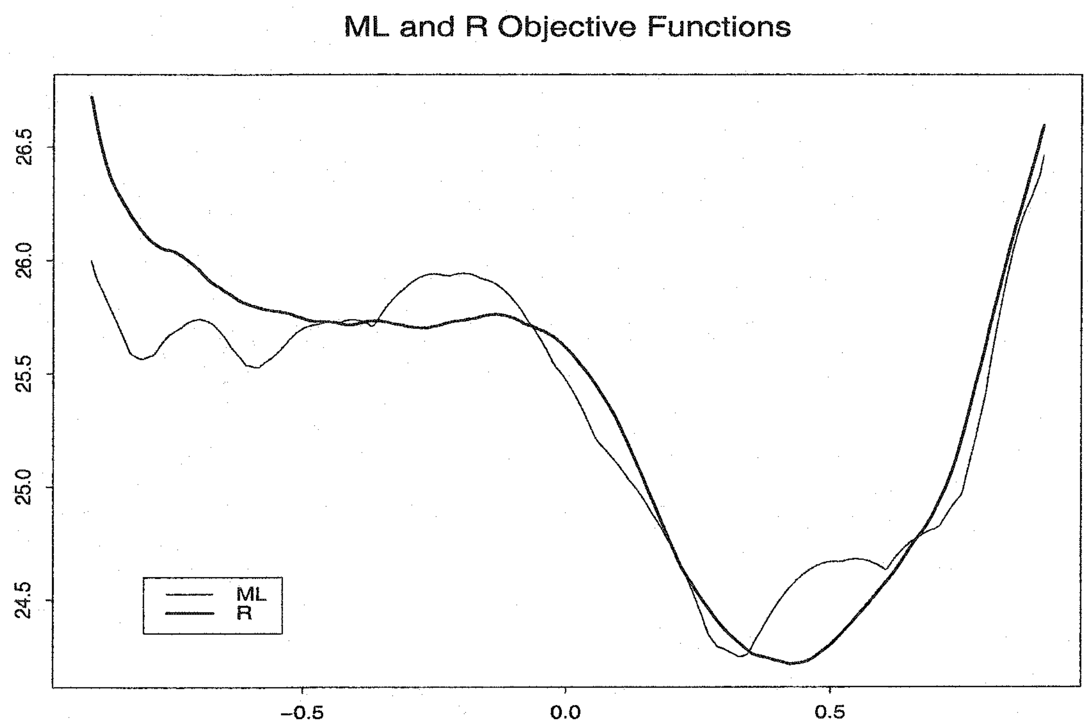

Figure 3.1: ML and R (with Wilcoxon weights) objective functions for a realization of length $n = 50$ from an all-pass model with $p = 1$, $\phi_{01} = 0.5$, and Laplace noise with variance one.

df	ARE (R to LAD)	ARE (R to ML)
3	1.4111	0.9617
6	2.0675	0.9968
9	2.3543	0.9796
12	2.5100	0.9636
15	2.6072	0.9515
20	2.7072	0.9374
30	2.8099	0.9213

Table 3.1: AREs for R (with Wilcoxon weights) to LAD and R to ML for the Students' t-distribution with several different degrees of freedom.

and so

$$\tilde{K} = \mathrm{E}\{Z_1\varphi(F(Z_1))\} = \frac{3\sqrt{3}}{4\pi} \quad \text{and} \quad \tilde{L} = \mathrm{E}\{f(Z_1)\varphi'(F(Z_1))\} = \frac{5}{4\sqrt{3}\pi}.$$

Hence, $\sigma^2\tilde{L} > \tilde{K}$ and (3.11) equals $(4\pi^2 - 27)/24$. Because the ARE of R to LAD is 1.41 and the ARE of R to ML is 0.96, R-estimation is asymptotically more efficient than LAD and nearly as asymptotically efficient as ML. Table 3.1 gives the AREs for R to LAD and R to ML for the Students' t-distribution with several different degrees of freedom. Note that, in Table 3.1, all values of ARE for R to ML are very close to one.

3.3.2 Order Selection

In practice, the true order r of an all-pass model is usually unknown and must be estimated. Therefore, in this section, we consider an order selection procedure. First note that

$$\frac{\sigma^2\tilde{J} - \tilde{K}^2}{2(\sigma^2\tilde{L} - \tilde{K})^2} = \frac{\tilde{J} - \left(\frac{|\phi_{0r}|}{\sigma}\tilde{K}_z\right)^2}{2\left(\frac{\sigma}{|\phi_{0r}|}\tilde{L}_z - \frac{|\phi_{0r}|}{\sigma}\tilde{K}_z\right)^2},$$

where $\tilde{K}_z := \int_0^1 F_z^{-1}(s)\varphi(s)\,ds$, $\tilde{L}_z := \int_0^1 f_z\left(F_z^{-1}(s)\right)\varphi'(s)\,ds$, and f_z and F_z are the density and distribution functions respectively for $z_1 = \phi_{0r}^{-1}Z_1$. Because $\hat{\boldsymbol{\phi}}_R \xrightarrow{P} \boldsymbol{\phi}_0$,

$$\hat{s} := \left(\frac{1}{n}\sum_{t=1}^{n-p} z_t^2(\hat{\boldsymbol{\phi}}_R)\right)^{1/2} \xrightarrow{P} \left(\mathrm{E}\{z_1^2\}\right)^{1/2} = \frac{\sigma}{|\phi_{0r}|} \tag{3.12}$$

and $\hat{K}_z := n^{-1}D(\hat{\phi}_R) \xrightarrow{P} \tilde{K}_z$ by Lemma 15 in Section 3.5. Corollary 1 provides a consistent estimator of $\tilde{L}_z$.

Corollary 1 *Consider the kernel density estimator of* f_z

$$\hat{f}_n(s) := \frac{1}{b_n n} \sum_{t=1}^{n-p} \kappa\left(\frac{s - z_t(\hat{\phi}_R)}{b_n}\right), \tag{3.13}$$

where κ *is a uniformly continuous, differentiable kernel density function on* $\mathbb{R}$ *such that* $\int |s \ln |s||^{1/2} |\kappa'(s)|\, ds < \infty$ *and* κ' *is uniformly continuous on* $\mathbb{R}$, *and the bandwidth* b_n *is chosen so that* $b_n \xrightarrow{P} 0$ *and* $b_n^2\sqrt{n} \xrightarrow{P} \infty$ *as* $n \to \infty$. *Then*

$$\hat{L}_z := n^{-1} \sum_{t=1}^{n-p} \varphi'\left(\frac{t}{n-p}\right) \hat{f}_n\left(z_{(t)}(\hat{\phi}_R)\right) \xrightarrow{P} \tilde{L}_z.$$

Proof: If $I\{\cdot\}$ is the indicator function,

$$\hat{F}_n(s) := \frac{1}{n-p} \sum_{t=1}^{n-p} I\left\{z_t(\hat{\phi}_R) \le s\right\},$$

$\hat{F}_n^{-1}(s) := \inf\{x : \hat{F}_n(x) \ge s\}$, and

$$\varphi_n'(s) := \varphi'\left(\frac{t}{n-p}\right) \text{ for } s \in \left(\frac{t-1}{n-p}, \frac{t}{n-p}\right], \ t = 1, \ldots, n-p,$$

then $n\hat{L}_z/(n-p) = \int_0^1 \hat{f}_n\left(\hat{F}_n^{-1}(s)\right) \varphi_n'(s)\, ds$. By the uniform continuity of φ',

$$\sup_{s\in(0,1)} |\varphi_n'(s) - \varphi'(s)| \to 0 \qquad \text{and} \qquad \sup_{s\in(0,1)} |\varphi'(s)| < \infty,$$

and, by the uniform continuity of f_z, $f_z(F_z^{-1}(\cdot))$ is bounded on $(0,1)$. Consequently, the proof is complete if

$$\begin{aligned} &\sup_{s\in(0,1)} \left|\hat{f}_n\left(\hat{F}_n^{-1}(s)\right) - f_z\left(F_z^{-1}(s)\right)\right| \\ \le\ &\sup_{s\in(0,1)} \left|\hat{f}_n\left(\hat{F}_n^{-1}(s)\right) - f_z\left(\hat{F}_n^{-1}(s)\right)\right| + \sup_{s\in(0,1)} \left|f_z\left(\hat{F}_n^{-1}(s)\right) - f_z\left(F_z^{-1}(s)\right)\right| \\ \xrightarrow{P}\ &0. \end{aligned} \tag{3.14}$$

The first term in (3.14) converges in probability to zero by Lemma 16 in Section 3.5, so we now consider the second term and use an argument similar to one found in the proof of Lemma 4 in Koul, Sievers, and McKean (1987). Note that

$$\begin{aligned}\sup_{s\in(0,1)}\left|F_z\left(\hat{F}_n^{-1}(s)\right)-s\right| &= \sup_{t\in\{1,\ldots,n-p\}}\left(\max\left\{\left|F_z(z_{(t)}(\hat{\phi}_R))-\frac{t-1}{n-p}\right|,\left|F_z(z_{(t)}(\hat{\phi}_R))-\frac{t}{n-p}\right|\right\}\right)\\ &= \sup_{s\in\mathbb{R}}\left|\hat{F}_n(s)-F_z(s)\right|,\end{aligned}$$

and, using the Glivenko-Cantelli theorem, it can be shown that

$$\sup_{s\in\mathbb{R}}\left|\hat{F}_n(s)-F_z(s)\right|\xrightarrow{P}0.$$

Because $f_z\left(F_z^{-1}(\cdot)\right)$ is uniformly continuous on $(0,1)$ and $F_z^{-1}(F_z(s))=s$ for all $s\in\mathbb{R}$ since F_z is strictly increasing on $\mathbb{R}$,

$$\sup_{s\in(0,1)}\left|f_z\left(\hat{F}_n^{-1}(s)\right)-f_z\left(F_z^{-1}(s)\right)\right|=\sup_{s\in(0,1)}\left|f_z\left(F_z^{-1}\left[F_z\left\{\hat{F}_n^{-1}(s)\right\}\right]\right)-f_z\left(F_z^{-1}(s)\right)\right|\xrightarrow{P}0,$$

and the proof is complete. □

It follows that

$$\frac{\tilde{J}-(\hat{s}^{-1}\hat{K}_z)^2}{2(\hat{s}\hat{L}_z-\hat{s}^{-1}\hat{K}_z)^2}\xrightarrow{P}\frac{\sigma^2\tilde{J}-\tilde{K}^2}{2(\sigma^2\tilde{L}-\tilde{K})^2}. \tag{3.15}$$

Note that the Gaussian and the Students' t densities satisfy the conditions for the kernel density function in Corollary 1.

We now give the following corollary for use in order selection.

Corollary 2 *If the true order of the all-pass model is r and the order of the fitted model is $p>r$, then*

$$n^{1/2}\hat{\phi}_{p,R}\xrightarrow{d}N\left(0,\frac{\sigma^2\tilde{J}-\tilde{K}^2}{2(\sigma^2\tilde{L}-\tilde{K})^2}\right).$$

Proof: By Problem 8.15 in Brockwell and Davis (1991), the pth diagonal element of $\mathbf{\Gamma}_p^{-1}$ is σ^{-2} if $p>r$, and so the result follows from (3.7). □

A practical approach to order determination using a large sample follows:

1. For some large P, fit all-pass models of order p, $p = 1, 2, \ldots, P$, via R-estimation and obtain the pth coefficient, $\hat{\phi}_{p,R}$, for each.

2. Let the model order r be the smallest order beyond which the estimated coefficients are statistically insignificant; that is,

$$r = \min\{0 \leq p \leq P : |\hat{\phi}_{j,R}| < 1.96\hat{\lambda}\, n^{-1/2} \text{ for } j > p\},$$

   where

$$\hat{\lambda} := \left(\frac{\tilde{J} - (\hat{s}^{-1}\hat{K}_z)^2}{2(\hat{s}\hat{L}_z - \hat{s}^{-1}\hat{K}_z)^2} \right)^{1/2}$$

   and the estimates $\hat{s}$, $\hat{K}_z$, and $\hat{L}_z$ are from the fitted Pth order model.

3.4 Numerical Results

3.4.1 Simulation Study

In this section, we describe a simulation experiment to assess the quality of the asymptotic approximations for finite samples. For each of the 1000 replicates, we simulated all-pass data and found $\hat{\phi}_R$. To diminish the possibility of the optimizer being trapped at local minima, we used 250 starting values for each of the replicates. These initial values for $\phi_1, \ldots, \phi_p$ were uniformly distributed in the space of partial autocorrelations and then mapped to the space of autoregressive coefficients using the Durbin-Levinson algorithm (Brockwell and Davis, 1991, Proposition 5.2.1). That is, for a model of order p, the kth starting value $(\phi_{p1}^{(k)}, \ldots, \phi_{pp}^{(k)})'$ was computed recursively as follows:

1. Draw $\phi_{11}^{(k)}, \phi_{22}^{(k)}, \ldots, \phi_{pp}^{(k)}$ iid uniform$(-1, 1)$.

2. For $j = 2, \ldots, p$, compute

$$\begin{bmatrix} \phi_{j1}^{(k)} \\ \vdots \\ \phi_{j,j-1}^{(k)} \end{bmatrix} = \begin{bmatrix} \phi_{j-1,1}^{(k)} \\ \vdots \\ \phi_{j-1,j-1}^{(k)} \end{bmatrix} - \phi_{jj}^{(k)} \begin{bmatrix} \phi_{j-1,j-1}^{(k)} \\ \vdots \\ \phi_{j-1,1}^{(k)} \end{bmatrix}.$$

We evaluated D at each of the 250 candidate values and reduced the collection of initial values to the nine with the smallest values of D plus $\boldsymbol{\phi}_0$. We found optimized values by implementing the Hooke and Jeeves (1961) algorithm and using the ten initial values as starting points. The optimized value for which D was smallest was chosen to be $\hat{\boldsymbol{\phi}}_R$. Confidence intervals for the elements of $\boldsymbol{\phi}_0$ were constructed using (3.7) and the estimator in (3.15). For the kernel density estimator (3.13), we used the standard Gaussian kernel density and, because of its recommendation in Silverman (1986) (page 48), we used

$$b_n = 0.9n^{-1/5} \min\{\hat{s}, IQR/1.34\},$$

where $\hat{s}$, defined in (3.12), is the sample standard deviation for $\{z_t(\hat{\boldsymbol{\phi}}_R)\}$ and IQR is the interquartile range for $\{z_t(\hat{\boldsymbol{\phi}}_R)\}$.

Results of the simulations appear in Tables 3.2 and 3.3. We show the empirical means, standard deviations, and percent coverages of nominal 95% confidence intervals for rank estimates of all-pass model parameters. Wilcoxon scores and Laplace and Students' t noise distributions were used. For each sample size n, empirical confidence intervals were computed using standard asymptotic theory for 1000 iid replicates. Asymptotic means and standard deviations were obtained using Theorem 3. Note that the rank estimates appear nearly unbiased, particularly when $n = 5000$, and the confidence interval coverages are close to the nominal 95% level. The asymptotic standard deviations tend to understate the true variability of the estimates when $n = 500$, but are fairly accurate when $n = 5000$. Normal probability plots show that the estimates are approximately normal in all cases.

3.4.2 Exchange Rate Modeling

Log returns for daily exchange rates tend to be uncorrelated but dependent and exhibit volatility clustering. Therefore, generalized autoregressive conditionally heteroscedastic (GARCH) and stochastic volatility models, which generate uncorrelated data and model conditional variances, are usually used to describe these series. All-pass models also generate uncorrelated data and can frequently provide a good fit for log returns. In Section 4.2 of Breidt, Davis, and Trindade (2001), an

	Asymptotic		Empirical		
n	mean	std.dev.	mean (c.i.)	std.dev. (c.i.)	% coverage (c.i.)
500	$\phi_1 = 0.5$	0.0354	0.4971 (0.4946,0.4995)	0.0397 (0.0379,0.0414)	95.4 (94.1,96.7)
5000	$\phi_1 = 0.5$	0.0112	0.5003 (0.4996,0.5010)	0.0114 (0.0109,0.0119)	96.3 (95.1,97.5)
500	$\phi_1 = 0.3$	0.0374	0.2971 (0.2947,0.2996)	0.0399 (0.0381,0.0416)	96.2 (95.0,97.4)
	$\phi_2 = 0.4$	0.0374	0.3958 (0.3932,0.3985)	0.0432 (0.0412,0.0450)	94.2 (92.8,95.6)
5000	$\phi_1 = 0.3$	0.0118	0.2991 (0.2983,0.2998)	0.0126 (0.0120,0.0131)	95.4 (94.1,96.7)
	$\phi_2 = 0.4$	0.0118	0.3997 (0.3990,0.4005)	0.0123 (0.0117,0.0128)	95.6 (94.3,96.9)

Table 3.2: Empirical means, standard deviations, and percent coverages of nominal 95% confidence intervals for rank estimates of all-pass model parameters. Wilcoxon scores and the Laplace noise distribution with variance one were used.

	Asymptotic		Empirical		
n	mean	std.dev.	mean (c.i.)	std.dev. (c.i.)	% coverage (c.i.)
500	$\phi_1 = 0.5$	0.0279	0.4982 (0.4962,0.5002)	0.0321 (0.0307,0.0335)	96.7 (95.6,97.8)
5000	$\phi_1 = 0.5$	0.0088	0.5000 (0.4995,0.5006)	0.0093 (0.0088,0.0097)	95.9 (94.7,97.1)
500	$\phi_1 = 0.3$	0.0296	0.2983 (0.2961,0.3005)	0.0361 (0.0345,0.0377)	95.0 (93.6,96.4)
	$\phi_2 = 0.4$	0.0296	0.3964 (0.3941,0.3986)	0.0356 (0.0340,0.0372)	94.8 (93.4,96.2)
5000	$\phi_1 = 0.3$	0.0093	0.2999 (0.2993,0.3005)	0.0098 (0.0094,0.0102)	95.5 (94.2,96.8)
	$\phi_2 = 0.4$	0.0093	0.3995 (0.3989,0.4001)	0.0097 (0.0093,0.0101)	96.0 (94.8,97.2)

Table 3.3: Empirical means, standard deviations, and percent coverages of nominal 95% confidence intervals for rank estimates of all-pass model parameters. Wilcoxon scores and the Students' t noise distribution with three degrees of freedom were used.

all-pass model was fit to log returns for the New Zealand/United States exchange rate. Here we consider the daily log returns for the Canada/United States exchange rate for 1992 through 1995 shown in Figure 3.2(a). The data were recorded at noon Pacific time and are available online courtesy of the Pacific Exchange Rate Service. From the sample autocorrelation functions for the log returns and their squares and absolute values in Figure 3.2(b)–(d), it appears this series is uncorrelated but dependent. The all-pass order selection procedure indicates that a model of order twelve is appropriate, and the R-estimates of model parameters obtained using the Wilcoxon scores are

$$\hat{\phi}_R = (-0.109, 0.473, -0.041, -0.246, -0.019, 0.236, 0.325, 0.153, -0.008, -0.009, 0.127, 0.115)',$$

with standard errors 0.050, 0.050, 0.056, 0.056, 0.056, 0.054, 0.054, 0.056, 0.056, 0.056, 0.050, and 0.050 respectively. From the sample autocorrelation functions for the squares and absolute values of the model residuals in Figure 3.3, it appears the residuals are independent. Hence, the fitted all-pass model seems to capture the dependence in the data.

Applications for all-pass models are not limited to exchange rates. As discussed and demonstrated in Breidt, Davis, and Trindade (2001) and Chapter 2, all-pass models can be used to identify and estimate noncausal or noninvertible ARMA models. This is a somewhat more natural application for all-pass models.

3.5 Additional Results

This section contains proofs of the lemmas used to confirm the results of Section 3.3. First, for an arbitrary, causal autoregressive polynomial $\phi(z)$, define $\theta(z) = \phi_1 z + \cdots + \phi_p z^p = 1 - \phi(z)$, and define $\theta_0(z) = 1 - \phi_0(z)$. Note that, for $t \in \{1, \ldots, n-p\}$,

$$\phi(B)X_{t+p} = -z_t(\phi) + \theta(B^{-1})z_t(\phi),$$

so, if $j \in \{1, \ldots, p\}$, then

$$\frac{\partial}{\partial \phi_j}\left\{\theta(B^{-1})z_t(\phi)\right\} = -X_{t+p-j} + \frac{\partial z_t(\phi)}{\partial \phi_j}. \tag{3.16}$$

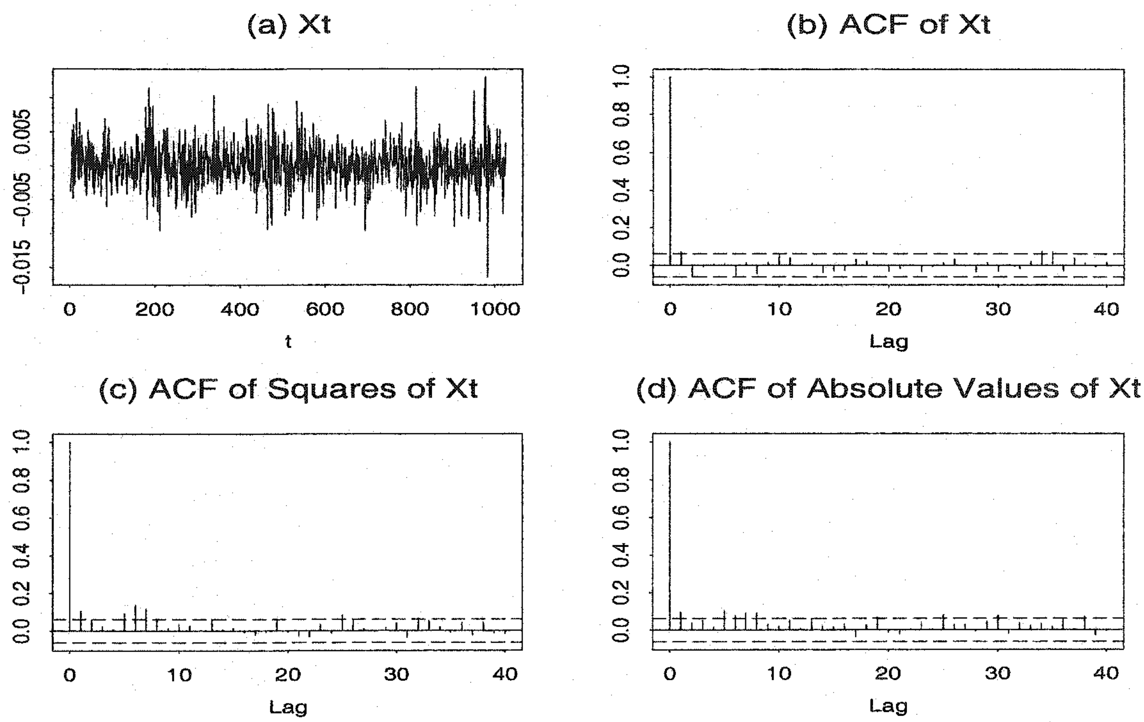

Figure 3.2: (a) The Canada/United States exchange rate log returns, $\{X_t\}$, and the sample autocorrelation functions with bounds $\pm 1.96/\sqrt{n}$ for (b) $\{X_t\}$, (c) $\{X_t^2\}$, and (d) $\{|X_t|\}$.

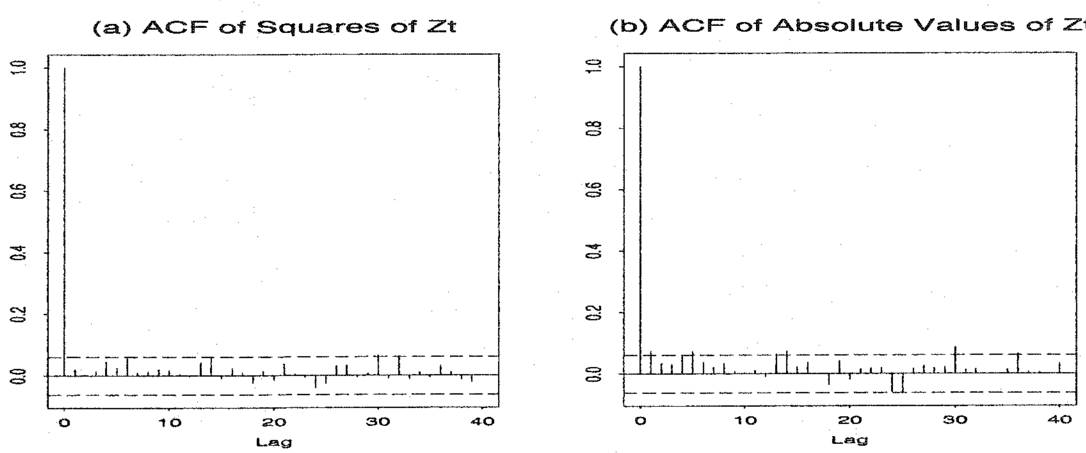

Figure 3.3: Diagnostics for the fitted all-pass model of order twelve. The sample autocorrelation functions with bounds $\pm 1.96/\sqrt{n}$ for (a) the squared residuals $\{Z_t^2\}$ and (b) the absolute residuals $\{|Z_t|\}$.

Also, if $j \in \{1,\ldots,p\}$, then

$$\begin{aligned}\frac{\partial}{\partial\phi_j}\{\theta(B^{-1})z_t(\phi)\} &= \frac{\partial}{\partial\phi_j}\{\phi_1 z_{t+1}(\phi)+\cdots+\phi_p z_{t+p}(\phi)\}\\ &= \theta(B^{-1})\frac{\partial z_t(\phi)}{\partial\phi_j}+z_{t+j}(\phi).\end{aligned} \tag{3.17}$$

Equating (3.16) and (3.17) and solving for $\partial z_t(\phi)/\partial\phi_j$, we obtain

$$\frac{\partial z_t(\phi)}{\partial\phi_j}=\frac{1}{\phi(B^{-1})}\{X_{t+p-j}+z_{t+j}(\phi)\}. \tag{3.18}$$

Evaluating (3.18) at the true value of ϕ and ignoring the effect of recursion initialization, we have

$$\begin{aligned}\frac{\partial z_t(\phi_0)}{\partial\phi_j} &= \frac{1}{\phi_0(B^{-1})}\left\{\frac{-\phi_0(B^{-1})B^p z_{t+p-j}}{\phi_0(B)}+z_{t+j}(\phi_0)\right\}\\ &\simeq \frac{-z_{t-j}}{\phi_0(B)}+\frac{z_{t+j}}{\phi_0(B^{-1})},\end{aligned} \tag{3.19}$$

where the first term is an element of $\sigma(z_{t-1},z_{t-2},\ldots)$ and the second term is an element of $\sigma(z_{t+1},z_{t+2},\ldots)$ because $\phi_0(B)$ is a causal operator and $\phi_0(B^{-1})$ is a purely noncausal operator. It follows that (3.19) is independent of $z_t=\phi_{0r}^{-1}Z_t$. Thus, if F_z is the distribution function of z_1 and $g_t(\phi):=\varphi(F_z(z_t))z_t(\phi)$, then, for $j\in\{1,\ldots,p\}$,

$$\begin{aligned}\frac{\partial g_t(\phi_0)}{\partial\phi_j} &= \varphi(F_z(z_t))\frac{\partial z_t(\phi_0)}{\partial\phi_j}\\ &\simeq \varphi(F_z(z_t))\left\{\frac{-z_{t-j}}{\phi_0(B)}+\frac{z_{t+j}}{\phi_0(B^{-1})}\right\}\\ &=: \frac{\partial g_t^*(\phi_0)}{\partial\phi_j}.\end{aligned} \tag{3.20}$$

The expectation of (3.20) is zero by the independence of its two terms.

We now compute the autocovariance function $\gamma^\dagger(h)$ of the zero-mean, stationary process $\{\mathbf{u}'\partial g_t^*(\phi_0)/\partial\phi\}$ for $\mathbf{u}\in\mathbb{R}^p$:

$$\begin{aligned}\gamma^\dagger(h) &= \mathrm{E}\left\{\mathbf{u}'\frac{\partial g_t^*(\phi_0)}{\partial\phi}\left(\frac{\partial g_{t+h}^*(\phi_0)}{\partial\phi}\right)'\mathbf{u}\right\}\\ &= \mathbf{u}'[\nu_{jk}(h)]_{j,k=1}^p\mathbf{u},\end{aligned}$$

where

$$\nu_{jk}(h):=\begin{cases}2\phi_{0r}^{-2}\tilde{J}\gamma(j-k), & \text{if } h=0,\\ -\psi_{|h|-j}\psi_{|h|-k}\phi_{0r}^{-2}\tilde{K}^2, & \text{if } h\neq 0,\end{cases}$$

and the ψ_l are given by $\sum_{l=0}^{\infty} \psi_l z^l = 1/\phi_0(z)$ with $\psi_l = 0$ for $l < 0$. Thus,

$$\begin{aligned}\gamma^\dagger(0) + 2\sum_{h=1}^{\infty}\gamma^\dagger(h) &= \mathbf{u}'\left\{[2\phi_{0r}^{-2}\tilde{J}\gamma(j-k)]_{j,k=1}^{p} - 2\phi_{0r}^{-2}\tilde{K}^2\left[\sum_{h=1}^{\infty}\psi_{h-j}\psi_{h-k}\right]_{j,k=1}^{p}\right\}\mathbf{u} \\ &= 2\phi_{0r}^{-2}(\sigma^2\tilde{J} - \tilde{K}^2)\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u}.\end{aligned}$$

The preceding calculations lead directly to the following lemma.

Lemma 1 *As $n \to \infty$,*

$$n^{-1/2}\sum_{t=1}^{n-p}\frac{\partial g_t(\phi_0)}{\partial \phi} \xrightarrow{d} \mathbf{N} \sim N\left(\mathbf{0}, 2\phi_{0r}^{-2}\{\sigma^2\tilde{J} - \tilde{K}^2\}\sigma^{-2}\boldsymbol{\Gamma}_p\right).$$

Proof: Note that, for $t \in \{0, \ldots, n-p-1\}$,

$$z_{n-p-t} = \sum_{l=0}^{\infty}\psi_l\left(\phi_0(B^{-1})z_{n-p-t+l}\right) \quad \text{and} \quad z_{n-p-t}(\phi_0) = \sum_{l=0}^{t}\psi_l\left(\phi_0(B^{-1})z_{n-p-t+l}\right). \qquad (3.21)$$

Because there exist constants $c > 0$ and $0 < d < 1$ such that $|\psi_l| < cd^l$ for all $l \in \{0, 1, \ldots\}$ (see Brockwell and Davis, 1991, Section 3.1),

$$\begin{aligned}\sum_{t=1}^{n-p}\mathrm{E}\left|\frac{\partial g_t(\phi_0)}{\partial \phi_j} - \frac{\partial g_t^*(\phi_0)}{\partial \phi_j}\right| &= \sum_{t=1}^{n-p}\mathrm{E}\left|\varphi(F_z(z_t))\left\{\frac{z_{t+j}(\phi_0)}{\phi_0(B^{-1})} - \frac{z_{t+j}}{\phi_0(B^{-1})}\right\}\right| \\ &= O(1)\end{aligned}$$

for $j \in \{1, \ldots, p\}$. Consequently,

$$n^{-1/2}\sum_{t=1}^{n-p}\left(\frac{\partial g_t(\phi_0)}{\partial \phi} - \frac{\partial g_t^*(\phi_0)}{\partial \phi}\right) \to \mathbf{0}$$

in L_1 and hence in probability.

Let $\mathbf{u} \in \mathbb{R}^p$. By the Cramér-Wold device, it suffices to show

$$n^{-1/2}\sum_{t=1}^{n-p}\mathbf{u}'\frac{\partial g_t^*(\phi_0)}{\partial \phi} \xrightarrow{d} \mathbf{u}'\mathbf{N} \sim \mathrm{N}(0, 2\phi_{0r}^{-2}\{\sigma^2\tilde{J} - \tilde{K}^2\}\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u}).$$

Elements of the infinite order moving average stationary sequence $\{\mathbf{u}'\partial g_t^*(\phi_0)/\partial\phi\}$ can be truncated to create a finite order moving average stationary sequence. By applying a central limit theorem

(Brockwell and Davis, 1991, Theorem 6.4.2) to each truncation level, asymptotic normality can be deduced. The details are omitted. □

Now consider the mixed partials of $g_t(\boldsymbol{\phi})$. For $j,k \in \{1,\ldots,p\}$,

$$\begin{aligned}
\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial\phi_j\partial\phi_k} &= \frac{1}{\phi_0^2(B^{-1})}\{X_{t+p+j-k} + X_{t+p+k-j} + 2z_{t+j+k}(\boldsymbol{\phi}_0)\} \\
&\simeq \frac{-z_{t+j-k} - z_{t+k-j}}{\phi_0(B^{-1})\phi_0(B)} + \frac{2z_{t+j+k}}{\phi_0^2(B^{-1})} \\
&= -\sum_{m=0}^{\infty}\sum_{\ell=0}^{\infty}\psi_m\psi_\ell(z_{t+j-k-\ell+m} + z_{t+k-j-\ell+m}) + \frac{2z_{t+j+k}}{\phi_0^2(B^{-1})},
\end{aligned}$$

and so

$$\begin{aligned}
\frac{\partial^2 g_t(\boldsymbol{\phi}_0)}{\partial\phi_j\partial\phi_k} &= \varphi(F_z(z_t))\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial\phi_j\partial\phi_k} \\
&\simeq \varphi(F_z(z_t))\left\{-\sum_{m=0}^{\infty}\sum_{\ell=0}^{\infty}\psi_m\psi_\ell(z_{t+j-k-\ell+m} + z_{t+k-j-\ell+m}) + \frac{2z_{t+j+k}}{\phi_0^2(B^{-1})}\right\} \qquad (3.22) \\
&=: \frac{\partial^2 g_t^*(\boldsymbol{\phi}_0)}{\partial\phi_j\partial\phi_k}.
\end{aligned}$$

(3.22) has expectation

$$-2\sigma^{-2}\gamma(j-k)\int_0^1 F_z^{-1}(s)\varphi(s)\,ds = -2|\phi_{0r}|^{-1}\tilde{K}\sigma^{-2}\gamma(j-k).$$

Lemma 2 *As* $n \to \infty$,

$$n^{-1}\sum_{t=1}^{n-p}\frac{\partial^2 g_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'} \xrightarrow{P} -2|\phi_{0r}|^{-1}\tilde{K}\sigma^{-2}\boldsymbol{\Gamma}_p.$$

Proof: For $j,k \in \{1,\ldots,p\}$,

$$\begin{aligned}
\sum_{t=1}^{n-p}\mathrm{E}\left|\frac{\partial^2 g_t(\boldsymbol{\phi}_0)}{\partial\phi_j\partial\phi_k} - \frac{\partial^2 g_t^*(\boldsymbol{\phi}_0)}{\partial\phi_j\partial\phi_k}\right| &= 2\sum_{t=1}^{n-p}\mathrm{E}\left|\varphi(F_z(z_t))\left\{\frac{z_{t+j+k}(\boldsymbol{\phi}_0)}{\phi_0^2(B^{-1})} - \frac{z_{t+j+k}}{\phi_0^2(B^{-1})}\right\}\right| \\
&= O(1),
\end{aligned}$$

and so

$$n^{-1}\sum_{t=1}^{n-p}\left(\frac{\partial^2 g_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'} - \frac{\partial^2 g_t^*(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\right) \to \mathbf{0}$$

in L_1 and probability. Because (3.22) has expectation $-2|\phi_{0r}|^{-1}\tilde{K}\sigma^{-2}\gamma(j-k)$,

$$n^{-1}\sum_{t=1}^{n-p}\frac{\partial^2 g_t^*(\phi_0)}{\partial\phi\partial\phi'}\xrightarrow{P}-2|\phi_{0r}|^{-1}\tilde{K}\sigma^{-2}\mathbf{\Gamma}_p$$

by the ergodic theorem. □

Let $F_n(\cdot)$ denote

$$F_n(x)=n^{-1}\sum_{t=1}^{n-p}I\{z_t\leq x\},$$

where $I\{\cdot\}$ is the indicator function.

Lemma 3 *For any $T\in(0,\infty)$, as $n\to\infty$,*

$$\sup_{\|\mathbf{u}\|\leq T,t\in\{1,\ldots,n-p\}}\sqrt{n}\left|\frac{R_t\left(\phi_0+n^{-1/2}\mathbf{u}\right)}{n-p+1}-F_n\left(z_t\left(\phi_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)\right|\xrightarrow{P}0.$$

Proof: Observe that

$$\begin{aligned}&\frac{R_t\left(\phi_0+n^{-1/2}\mathbf{u}\right)}{n-p+1}-F_n\left(z_t\left(\phi_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)\\&=\frac{1}{n-p+1}\left[\sum_{j=1}^{n-p}I\left\{z_j\left(\phi_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\leq z_t\left(\phi_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right\}-\sum_{j=t+1}^{n-p}I\left\{z_j\left(\phi_0+\frac{\mathbf{u}}{\sqrt{n}}\right)=z_t\left(\phi_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right\}\right]\\&\quad-\frac{1}{n}\sum_{j=1}^{n-p}I\left\{z_j\leq z_t\left(\phi_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right\}.\end{aligned}$$

Because z_1 has a continuous distribution,

$$\mathrm{P}\left(\sup_{\|\mathbf{u}\|\leq T,t\in\{1,\ldots,n-p\}}\sum_{j=t+1}^{n-p}I\left\{z_j\left(\phi_0+\frac{\mathbf{u}}{\sqrt{n}}\right)=z_t\left(\phi_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right\}>1\right)=0,$$

and so

$$\sup_{\|\mathbf{u}\|\leq T,t\in\{1,\ldots,n-p\}}\frac{\sqrt{n}}{n-p+1}\sum_{j=t+1}^{n-p}I\left\{z_j\left(\phi_0+\frac{\mathbf{u}}{\sqrt{n}}\right)=z_t\left(\phi_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right\}\xrightarrow{P}0.$$

Since

$$\sqrt{n}\left|\frac{n}{n-p+1}-1\right|\to0,$$

it suffices to show

$$\sup_{\|\mathbf{u}\|\leq T,x\in\mathbb{R}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j\left(\phi_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\leq x\right\}-I\left\{z_j\leq x\right\}\right]\right|\xrightarrow{P}0.$$

For $\|\mathbf{u}\| \leq T$, define

$$\begin{aligned} e_{j,n}(\mathbf{u}) &= z_j - z_j\left(\boldsymbol{\phi}_0 + \frac{\mathbf{u}}{\sqrt{n}}\right) \\ &= z_j - z_j(\boldsymbol{\phi}_0) - \frac{\mathbf{u}'}{\sqrt{n}}\frac{\partial z_j(\boldsymbol{\phi}_0)}{\partial \boldsymbol{\phi}} - \frac{\mathbf{u}'}{2n}\frac{\partial^2 z_j(\boldsymbol{\phi}^*_{j,n}(\mathbf{u}))}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\mathbf{u} \\ &= z_j - z_j(\boldsymbol{\phi}_0) - \frac{\mathbf{u}'}{\sqrt{n}}\frac{\partial z_j(\boldsymbol{\phi}_0)}{\partial \boldsymbol{\phi}} - \frac{1}{n}\sum_{k=-\infty}^{\infty}\psi^*_{k,j,n}(\mathbf{u})z_{j-k}, \end{aligned}$$

where $\boldsymbol{\phi}^*_{j,n}(\mathbf{u})$ lies between $\boldsymbol{\phi}_0$ and $\boldsymbol{\phi}_0 + n^{-1/2}\mathbf{u}$ and $\{\psi^*_{k,j,n}(\mathbf{u})\}_{k=-\infty}^{\infty}$ is a real-valued sequence such that

$$\sum_{k=-\infty}^{\infty}\psi^*_{k,j,n}(\mathbf{u})z_{j-k} = \frac{\mathbf{u}'}{2}\frac{\partial^2 z_j(\boldsymbol{\phi}^*_{j,n}(\mathbf{u}))}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\mathbf{u}. \tag{3.23}$$

Set

$$\epsilon_{j,n}(\mathbf{u}) = z_j - z_j(\boldsymbol{\phi}_0) - \frac{\mathbf{u}'}{\sqrt{n}}\frac{\partial z_j(\boldsymbol{\phi}_0)}{\partial \boldsymbol{\phi}} - \frac{1}{n}\sum_{k\neq 0}\psi^*_{k,j,n}(\mathbf{u})z_{j-k}. \tag{3.24}$$

Because $\partial^2 z_j(\boldsymbol{\phi})/(\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}')$ is continuous with respect to $\boldsymbol{\phi}$,

$$\sup_{\|\mathbf{u}\|\leq T, j\in\{1,\ldots,n-p\}} |\psi^*_{0,j,n}(\mathbf{u})| = O(1).$$

Thus, for all n sufficiently large,

$$\inf_{\|\mathbf{u}\|\leq T, j\in\{1,\ldots,n-p\}} \left(1 + \frac{\psi^*_{0,j,n}(\mathbf{u})}{n}\right) > 0$$

and

$$\begin{aligned} &\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j\left(\boldsymbol{\phi}_0 + \frac{\mathbf{u}}{\sqrt{n}}\right) \leq x\right\} - I\{z_j \leq x\}\right]\right| \\ &= \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\{z_j \leq x + e_{j,n}(\mathbf{u})\} - I\{z_j \leq x\}\right]\right| \\ &= \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \leq \frac{x + \epsilon_{j,n}(\mathbf{u})}{1 + \psi^*_{0,j,n}(\mathbf{u})/n}\right\} - I\{z_j \leq x\}\right]\right| \\ &\leq \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(\frac{x + \epsilon_{j,n}(\mathbf{u})}{1 + \psi^*_{0,j,n}(\mathbf{u})/n}\right) - F_z(x)\right]\right| \\ &\quad + \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \leq \frac{x + \epsilon_{j,n}(\mathbf{u})}{1 + \psi^*_{0,j,n}(\mathbf{u})/n}\right\} - F_z\left(\frac{x + \epsilon_{j,n}(\mathbf{u})}{1 + \psi^*_{0,j,n}(\mathbf{u})/n}\right) - I\{z_j \leq x\} + F_z(x)\right]\right|. \end{aligned}$$

By Lemma 4,

$$\sup_{\|\mathbf{u}\|\le T, x\in\mathbb{R}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} \left[F_z\left(\frac{x+\epsilon_{j,n}(\mathbf{u})}{1+\psi^*_{0,j,n}(\mathbf{u})/n} \right) - F_z(x) \right] \right| \xrightarrow{P} 0,$$

and, by Lemma 5,

$$\sup_{\|\mathbf{u}\|\le T, x\in\mathbb{R}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} \left[I\left\{ z_j \le \frac{x+\epsilon_{j,n}(\mathbf{u})}{1+\psi^*_{0,j,n}(\mathbf{u})/n} \right\} - F_z\left(\frac{x+\epsilon_{j,n}(\mathbf{u})}{1+\psi^*_{0,j,n}(\mathbf{u})/n} \right) - I\{z_j \le x\} + F_z(x) \right] \right| \xrightarrow{P} 0.$$

□

Lemma 4 *For any $T \in (0,\infty)$, as $n \to \infty$,*

$$\sup_{\|\mathbf{u}\|\le T, x\in\mathbb{R}} \left| n^{-1/2} \sum_{j=1}^{n-p} \left[F_z\left(\frac{x+\epsilon_{j,n}(\mathbf{u})}{1+\psi^*_{0,j,n}(\mathbf{u})/n} \right) - F_z(x) \right] \right| \xrightarrow{P} 0,$$

where $\psi^*_{0,j,n}(\mathbf{u})$ and $\epsilon_{j,n}(\mathbf{u})$ are defined in (3.23) and (3.24) respectively.

Proof: For all n sufficiently large,

$$\inf_{\|\mathbf{u}\|\le T, j\in\{1,\ldots,n-p\}} \left(1 + \frac{\psi^*_{0,j,n}(\mathbf{u})}{n} \right) > 0,$$

and so

$$\begin{aligned}
&\sup_{\|\mathbf{u}\|\le T, x\in\mathbb{R}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} \left[F_z\left(\frac{x+\epsilon_{j,n}(\mathbf{u})}{1+\psi^*_{0,j,n}(\mathbf{u})/n} \right) - F_z(x) \right] \right| \\
\le\ & \sup_{\|\mathbf{u}\|\le T, x\in\mathbb{R}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} \left[F_z\left(\frac{x+\epsilon_{j,n}(\mathbf{u})}{1+\psi^*_{0,j,n}(\mathbf{u})/n} \right) - F_z(x+\epsilon_{j,n}(\mathbf{u})) \right] \right| \\
&+ \sup_{\|\mathbf{u}\|\le T, x\in\mathbb{R}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} [F_z(x+\epsilon_{j,n}(\mathbf{u})) - F_z(x)] \right| \\
=\ & \sup_{\|\mathbf{u}\|\le T, x\in\mathbb{R}} \left| \frac{1}{n^{3/2}} \sum_{j=1}^{n-p} \psi^*_{0,j,n}(\mathbf{u}) \frac{x+\epsilon_{j,n}(\mathbf{u})}{1+\psi^*_{0,j,n}(\mathbf{u})/n} f_z(x^*_{j,n}(\mathbf{u})) \right| \\
&+ \sup_{\|\mathbf{u}\|\le T, x\in\mathbb{R}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} f_z(x^\dagger_{j,n}(\mathbf{u}))\epsilon_{j,n}(\mathbf{u}) \right| \\
\le\ & \frac{1}{n^{3/2}} \left(\sup_{x\in\mathbb{R}} |x| f_z(x) \right) \sum_{j=1}^{n-p} \left| \frac{\psi^*_{0,j,n}(\mathbf{u})}{x^*_{j,n}(\mathbf{u})} \frac{x+\epsilon_{j,n}(\mathbf{u})}{1+\psi^*_{0,j,n}(\mathbf{u})/n} \right| \\
&+ \left(\sup_{x\in\mathbb{R}} f_z(x) \right) \sup_{\|\mathbf{u}\|\le T} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} \epsilon_{j,n}(\mathbf{u}) \right| + \sup_{\|\mathbf{u}\|\le T, x\in\mathbb{R}} \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} |\epsilon_{j,n}(\mathbf{u})| |f_z(x^\dagger_{j,n}(\mathbf{u})) - f_z(x)|
\end{aligned}$$

$$\leq \frac{1}{\sqrt{n}}\left(\sup_{x\in\mathbb{R}}|x|f_z(x)\right)\max\left\{\sup_{\|\mathbf{u}\|\leq T, j\in\{1,\ldots,n-p\}}|\psi^*_{0,j,n}(\mathbf{u})|, \sup_{\|\mathbf{u}\|\leq T, j\in\{1,\ldots,n-p\}}\left|\frac{\psi^*_{0,j,n}(\mathbf{u})}{1+\psi^*_{0,j,n}(\mathbf{u})/n}\right|\right\}$$
$$+\left(\sup_{x\in\mathbb{R}} f_z(x)\right)\sup_{\|\mathbf{u}\|\leq T}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\epsilon_{j,n}(\mathbf{u})\right| + \sup_{\|\mathbf{u}\|\leq T, x\in\mathbb{R}}\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}|\epsilon_{j,n}(\mathbf{u})||f_z(x^\dagger_{j,n}(\mathbf{u}))-f_z(x)|,$$

where $x^*_{j,n}(\mathbf{u})$ lies between $x+\epsilon_{j,n}(\mathbf{u})$ and $(x+\epsilon_{j,n}(\mathbf{u}))/(1+\psi^*_{0,j,n}(\mathbf{u})/n)$, $x^\dagger_{j,n}(\mathbf{u})$ lies between x and $x+\epsilon_{j,n}(\mathbf{u})$, and f_z is the density function for z_1. Because $\sup_{x\in\mathbb{R}}|x|f_z(x)<\infty$, $\sup_{x\in\mathbb{R}}f_z(x)<\infty$ by the uniform continuity of f_z, and

$$\sup_{\|\mathbf{u}\|\leq T, j\in\{1,\ldots,n-p\}}|\psi^*_{0,j,n}(\mathbf{u})| = O(1),$$

we need to show that

$$\sup_{\|\mathbf{u}\|\leq T}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\epsilon_{j,n}(\mathbf{u})\right| \tag{3.25}$$

and

$$\sup_{\|\mathbf{u}\|\leq T, x\in\mathbb{R}}\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}|\epsilon_{j,n}(\mathbf{u})||f_z(x^\dagger_{j,n}(\mathbf{u}))-f_z(x)| \tag{3.26}$$

are both $o_p(1)$ to complete the proof.

Observe that

$$\sup_{\|\mathbf{u}\|\leq T}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\epsilon_{j,n}(\mathbf{u})\right|$$
$$\leq \frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}|z_j - z_j(\boldsymbol{\phi}_0)| + \sup_{\|\mathbf{u}\|\leq T}\left|\frac{\mathbf{u}'}{n}\sum_{j=1}^{n-p}\frac{\partial z_j(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}}\right| + \sup_{\|\mathbf{u}\|\leq T}\left|\frac{1}{n^{3/2}}\sum_{j=1}^{n-p}\sum_{k\neq 0}\psi^*_{k,j,n}(\mathbf{u})z_{j-k}\right|.$$

By (3.21),

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}|z_j - z_j(\boldsymbol{\phi}_0)| \xrightarrow{P} 0,$$

and, using the ergodic theorem,

$$\sup_{\|\mathbf{u}\|\leq T}\left|\frac{\mathbf{u}'}{n}\sum_{j=1}^{n-p}\frac{\partial z_j(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}}\right| \leq T\left\|\frac{1}{n}\sum_{j=1}^{n-p}\frac{\partial z_j(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}}\right\| \xrightarrow{P} 0$$

since (3.19) has expectation zero. Finally, because $\partial^2 z_j(\boldsymbol{\phi})/(\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}')$ is continuous with respect to $\boldsymbol{\phi}$, there exists a geometrically decaying, non-negative, real-valued sequence $\{\pi_k\}_{k=-\infty}^{\infty}$ such that

$$\sup_{\|\mathbf{u}\|\leq T}\left|\sum_{k\neq 0}\psi^*_{k,j,n}(\mathbf{u})z_{j-k}\right| \leq \sum_{k\neq 0}\pi_k|z_{j-k}| \quad \forall j\in\{1,\ldots,n-p\}$$

for all n sufficiently large. Thus, for n sufficiently large,

$$\sup_{\|\mathbf{u}\|\leq T}\left|\frac{1}{n^{3/2}}\sum_{j=1}^{n-p}\sum_{k\neq 0}\psi^*_{k,j,n}(\mathbf{u})z_{j-k}\right| \leq \frac{1}{n^{3/2}}\sum_{j=1}^{n-p}\sum_{k\neq 0}\pi_k|z_{j-k}| \xrightarrow{P} 0,$$

and so (3.25) is $o_p(1)$.

Let $\delta > 0$ and $\eta > 0$. By (3.21), there exists an integer $m \geq 0$ such that

$$\mathrm{P}\left(\sup_{j\in\{1,\ldots,n-p-m\}}|z_j - z_j(\phi_0)| > \frac{\eta}{3}\right) < \frac{\delta}{3}$$

for all $n > p+m$. Also, because $\mathrm{E}(z_1^2) < \infty$,

$$\mathrm{P}\left(\sup_{\|\mathbf{u}\|\leq T, j\in\{1,\ldots,n-p\}}\left|\frac{\mathbf{u}'}{\sqrt{n}}\frac{\partial z_j(\phi_0)}{\partial\phi}\right| > \frac{\eta}{3}\right) < \frac{\delta}{3}$$

and

$$\mathrm{P}\left(\sup_{\|\mathbf{u}\|\leq T, j\in\{1,\ldots,n-p\}}\left|\frac{1}{n}\sum_{k\neq 0}\psi^*_{k,j,n}(\mathbf{u})z_{j-k}\right| > \frac{\eta}{3}\right) < \frac{\delta}{3}$$

for all n sufficiently large. It follows that

$$\mathrm{P}\left(\sup_{\|\mathbf{u}\|\leq T, j\in\{1,\ldots,n-p-m\}}|\epsilon_{j,n}(\mathbf{u})| > \eta\right) < \delta$$

for all n sufficiently large. Because f_z is uniformly continuous on $\mathbb{R}$ and $|x^\dagger_{j,n}(\mathbf{u}) - x| \leq |\epsilon_{j,n}(\mathbf{u})|$, for any $\tau > 0$ there exists an $\eta > 0$ such that

$$\sup_{\|\mathbf{u}\|\leq T, x\in\mathbb{R}, j\in\{1,\ldots,n-p-m\}}|f_z(x^\dagger_{j,n}(\mathbf{u})) - f_z(x)| \leq \tau$$

on the set where

$$\sup_{\|\mathbf{u}\|\leq T, j\in\{1,\ldots,n-p-m\}}|\epsilon_{j,n}(\mathbf{u})| \leq \eta.$$

Therefore, since

$$\sup_{\|\mathbf{u}\|\leq T}\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}|\epsilon_{j,n}(\mathbf{u})| = O_p(1),$$

(3.26) must be $o_p(1)$. □

Lemma 5 *For any $T \in (0,\infty)$, as $n\to\infty$,*

$$\sup_{\|\mathbf{u}\|\leq T, x\in\mathbb{R}}\left|n^{-1/2}\sum_{j=1}^{n-p}\left[I\left\{z_j \leq \frac{x+\epsilon_{j,n}(\mathbf{u})}{1+\psi^*_{0,j,n}(\mathbf{u})/n}\right\} - F_z\left(\frac{x+\epsilon_{j,n}(\mathbf{u})}{1+\psi^*_{0,j,n}(\mathbf{u})/n}\right) - I\{z_j\leq x\} + F_z(x)\right]\right| \xrightarrow{P} 0,$$

where $\psi^*_{0,j,n}(\mathbf{u})$ and $\epsilon_{j,n}(\mathbf{u})$ are defined in (3.23) and (3.24) respectively.

Proof: This argument is similar to the one given in Boldin (1982). First of all, by (3.21), for any $\delta > 0$, there exist constants $C_\delta > 0$ and $0 < D < 1$ such that

$$\begin{aligned} &\mathrm{P}\left(\sup_{j\in\{0,\ldots,n-p-1\}}\left[|z_{n-p-j} - z_{n-p-j}(\phi_0)| - C_\delta D^j\right] \geq 0\right) \\ &= \mathrm{P}\left(\sup_{j\in\{1,\ldots,n-p\}}\left[|z_j - z_j(\phi_0)| - C_\delta D^{n-p-j}\right] \geq 0\right) \\ &< \delta \end{aligned}$$

for all n. For $\delta > 0$ chosen arbitrarily small, the proof is complete if we show that

$$\sup_{\|\mathbf{u}\|\leq T, x\in\mathbb{R}}\left|\frac{1}{\sqrt{n}}I\{A_n\}\sum_{j=1}^{n-p}\left[I\left\{z_j \leq \frac{x+\epsilon_{j,n}(\mathbf{u})}{1+\psi^*_{0,j,n}(\mathbf{u})/n}\right\} - F_z\left(\frac{x+\epsilon_{j,n}(\mathbf{u})}{1+\psi^*_{0,j,n}(\mathbf{u})/n}\right) - I\{z_j \leq x\} + F_z(x)\right]\right|$$
$$\xrightarrow{P} 0,$$

where

$$A_n := \left\{\sup_{j\in\{1,\ldots,n-p\}}\left[|z_j - z_j(\phi_0)| - C_\delta D^{n-p-j}\right] < 0\right\}.$$

Hence, we assume to be on the set A_n for all n throughout the remainder of the proof.

Now, fix nondecreasing sequences of positive integers $\{M_n\}_{n=1}^\infty$ and $\{N_n\}_{n=1}^\infty$ such that $M_n \sim n^{3/4}$ and $N_n \sim n^{1/(8p)}$ (e.g., $n^{3/4}/M_n \to 1$ and $n^{1/(8p)}/N_n \to 1$ as $n \to \infty$). For each n, let

$$-\infty = x_{0,n} < x_{1,n} < \cdots < x_{M_n,n} = \infty, \qquad F_z(x_{r,n}) = rM_n^{-1} \tag{3.27}$$

and

$$u_{i,s_i,n} = -T + 2Ts_iN_n^{-1}, \qquad i \in \{1,\ldots,p\},\ s_i \in \{0,\ldots,N_n-1\}.$$

Thus, for any n and any $x \in \mathbb{R}$, there exists $x_{r,n}$ such that $x_{r,n} \leq x < x_{r+1,n}$. Also, for any n and any $\|\mathbf{u}\| = \|(u_1,\ldots,u_p)'\| \leq T$, there exists $(u_{1,s_1,n}, u_{2,s_2,n},\ldots,u_{p,s_p,n})'$ such that

$$\begin{aligned} \theta^-_{j,n,s_1,\ldots,s_p} &:= -\sum_{i=1}^p u_{i,s_i,n}\frac{\partial z_j(\phi_0)}{\partial\phi_i}\left(1 + 2TN_n^{-1}u^{-1}_{i,s_i,n}I\left\{\frac{\partial z_j(\phi_0)}{\partial\phi_i} > 0\right\}\right) \\ &\leq -\sum_{i=1}^p u_i\frac{\partial z_j(\phi_0)}{\partial\phi_i} \\ &\leq -\sum_{i=1}^p u_{i,s_i,n}\frac{\partial z_j(\phi_0)}{\partial\phi_i}\left(1 + 2TN_n^{-1}u^{-1}_{i,s_i,n}I\left\{\frac{\partial z_j(\phi_0)}{\partial\phi_i} < 0\right\}\right) \\ &=: \theta^+_{j,n,s_1,\ldots,s_p} \end{aligned} \tag{3.28}$$

for all $j \in \{1, \ldots, n-p\}$.

Recall that

$$\begin{aligned}\epsilon_{j,n}(\mathbf{u}) &= z_j - z_j(\boldsymbol{\phi}_0) - \frac{\mathbf{u}'}{\sqrt{n}}\frac{\partial z_j(\boldsymbol{\phi}_0)}{\partial \boldsymbol{\phi}} - \frac{\mathbf{u}'}{2n}\frac{\partial^2 z_j(\boldsymbol{\phi}^*_{j,n}(\mathbf{u}))}{\partial \boldsymbol{\phi}\partial \boldsymbol{\phi}'}\mathbf{u} + \frac{1}{n}\psi^*_{0,j,n}(\mathbf{u})z_j \\ &= z_j - z_j(\boldsymbol{\phi}_0) - \frac{\mathbf{u}'}{\sqrt{n}}\frac{\partial z_j(\boldsymbol{\phi}_0)}{\partial \boldsymbol{\phi}} - \frac{1}{n}\sum_{k\neq 0}\psi^*_{k,j,n}(\mathbf{u})z_{j-k},\end{aligned}$$

where $\boldsymbol{\phi}^*_{j,n}(\mathbf{u})$ lies between $\boldsymbol{\phi}_0$ and $\boldsymbol{\phi}_0 + n^{-1/2}\mathbf{u}$ and

$$\sum_{k=-\infty}^{\infty} \psi^*_{k,j,n}(\mathbf{u})z_{j-k} = \frac{\mathbf{u}'}{2}\frac{\partial^2 z_j(\boldsymbol{\phi}^*_{j,n}(\mathbf{u}))}{\partial \boldsymbol{\phi}\partial \boldsymbol{\phi}'}\mathbf{u}.$$

Because $\partial^2 z_j(\boldsymbol{\phi})/(\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}')$ is continuous with respect to $\boldsymbol{\phi}$, there exists a geometrically decaying, non-negative, real-valued sequence $\{\pi_k\}_{k=-\infty}^{\infty}$ such that

$$\sup_{\|\mathbf{u}\|\le T}\left|\sum_{k\neq 0}\psi^*_{k,j,n}(\mathbf{u})z_{j-k}\right| \le \sum_{k\neq 0}\pi_k|z_{j-k}| \quad \forall j \in \{1,\ldots,n-p\}$$

for all n sufficiently large. Also, because

$$\sup_{\|\mathbf{u}\|\le T, j\in\{1,\ldots,n-p\}} |\psi^*_{0,j,n}(\mathbf{u})| = O(1),$$

there exists an $L > 0$ such that

$$\sup_{\|\mathbf{u}\|\le T, j\in\{1,\ldots,n-p\}} |\psi^*_{0,j,n}(\mathbf{u})| \le L \qquad \text{and} \qquad 1 - \frac{L}{n} > 0$$

for all n sufficiently large. Choose $\eta > 0$ such that

$$\sup_{\|\mathbf{u}\|\le T}\left|\sum_{k\neq 0}\psi^*_{k,j,n}(\mathbf{u})z_{j-k}\right| \le \sum_{k\neq 0}\pi_k|z_{j-k}| \quad \forall j \in \{1,\ldots,n-p\},$$

$$\sup_{\|\mathbf{u}\|\le T, j\in\{1,\ldots,n-p\}} |\psi^*_{0,j,n}(\mathbf{u})| \le L, \qquad \text{and} \qquad 1 - \frac{L}{n} > 0$$

for all $n > \eta$.

Consequently, for any $n > \eta$, any $\|\mathbf{u}\| \le T$, and any $x \in \mathbb{R}$, there exist $(u_{1,s_1,n}, \ldots, u_{p,s_p,n})'$ and $x_{r,n}$ such that

$$\begin{aligned}& x_{r,n} - C_\delta D^{n-p-j} + n^{-1/2}\theta^-_{j,n,s_1,\ldots,s_p} - n^{-1}\sum_{k\neq 0}\pi_k|z_{j-k}| \\ &< x + \epsilon_{j,n}(\mathbf{u}) \\ &< x_{r+1,n} + C_\delta D^{n-p-j} + n^{-1/2}\theta^+_{j,n,s_1,\ldots,s_p} + n^{-1}\sum_{k\neq 0}\pi_k|z_{j-k}|\end{aligned}$$

and

$$
\begin{aligned}
y^-_{r,j,n,s_1,\ldots,s_p} &:= x^-_{r,j,n} + \frac{n^{-1/2}\theta^-_{j,n,s_1,\ldots,s_p}}{1 - L/n + 2LI\{\theta^-_{j,n,s_1,\ldots,s_p} > 0\}/n} - \frac{n^{-1}\sum_{k\neq 0}\pi_k |z_{j-k}|}{1-L/n} \\
&< \frac{x + \epsilon_{j,n}(\mathbf{u})}{1 + \psi^*_{0,j,n}(\mathbf{u})/n} \\
&< x^+_{r+1,j,n} + \frac{n^{-1/2}\theta^+_{j,n,s_1,\ldots,s_p}}{1 - L/n + 2LI\{\theta^+_{j,n,s_1,\ldots,s_p} < 0\}/n} + \frac{n^{-1}\sum_{k\neq 0}\pi_k |z_{j-k}|}{1-L/n} \\
&=: y^+_{r+1,j,n,s_1,\ldots,s_p} \qquad (3.29)
\end{aligned}
$$

for all $j \in \{1, \ldots, n-p\}$, where

$$
x^-_{r,j,n} := \frac{x_{r,n}}{1 - L/n + 2LI\{x_{r,n} > 0\}/n} - \frac{C_\delta D^{n-p-j}}{1 - L/n} \qquad (3.30)
$$

and

$$
x^+_{r+1,j,n} := \frac{x_{r+1,n}}{1 - L/n + 2LI\{x_{r+1,n} < 0\}/n} + \frac{C_\delta D^{n-p-j}}{1 - L/n}. \qquad (3.31)
$$

By (3.29) and the monotonicity of $I\{z_j \leq \cdot\}$ and $F_z(\cdot)$,

$$
\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \leq y^-_{r,j,n,s_1,\ldots,s_p}\right\} - F_z\left(y^+_{r+1,j,n,s_1,\ldots,s_p}\right) - I\{z_j \leq x\} + F_z(x)\right] \\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \leq y^-_{r,j,n,s_1,\ldots,s_p}\right\} - F_z\left(y^-_{r,j,n,s_1,\ldots,s_p}\right) - I\left\{z_j \leq x^-_{r,j,n}\right\} + F_z\left(x^-_{r,j,n}\right)\right] \\
&\quad + \frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \leq x^-_{r,j,n}\right\} - F_z\left(x^-_{r,j,n}\right) - I\left\{z_j \leq x_{r,n}\right\} + F_z\left(x_{r,n}\right)\right] \\
&\quad + \frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \leq x_{r,n}\right\} - F_z\left(x_{r,n}\right) - I\left\{z_j \leq x\right\} + F_z\left(x\right)\right] \\
&\quad + \frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(y^-_{r,j,n,s_1,\ldots,s_p}\right) - F_z\left(y^+_{r+1,j,n,s_1,\ldots,s_p}\right)\right] \\
&\leq \frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \leq \frac{x + \epsilon_{j,n}(\mathbf{u})}{1 + \psi^*_{0,j,n}(\mathbf{u})/n}\right\} - F_z\left(\frac{x + \epsilon_{j,n}(\mathbf{u})}{1 + \psi^*_{0,j,n}(\mathbf{u})/n}\right) - I\{z_j \leq x\} + F_z(x)\right] \\
&\leq \frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \leq y^+_{r+1,j,n,s_1,\ldots,s_p}\right\} - F_z\left(y^-_{r,j,n,s_1,\ldots,s_p}\right) - I\{z_j \leq x\} + F_z(x)\right] \\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \leq y^+_{r+1,j,n,s_1,\ldots,s_p}\right\} - F_z\left(y^+_{r+1,j,n,s_1,\ldots,s_p}\right) - I\left\{z_j \leq x^+_{r+1,j,n}\right\} + F_z\left(x^+_{r+1,j,n}\right)\right] \\
&\quad + \frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \leq x^+_{r+1,j,n}\right\} - F_z\left(x^+_{r+1,j,n}\right) - I\left\{z_j \leq x_{r+1,n}\right\} + F_z\left(x_{r+1,n}\right)\right]
\end{aligned}
$$

$$+\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \le x_{r+1,n}\right\} - F_z\left(x_{r+1,n}\right) - I\left\{z_j \le x\right\} + F_z\left(x\right)\right]$$

$$+\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(y^{+}_{r+1,j,n,s_1,\ldots,s_p}\right) - F_z\left(y^{-}_{r,j,n,s_1,\ldots,s_p}\right)\right].$$

Therefore, the proof is complete if

$$\sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \le y^{\pm}_{r,j,n,s_1,\ldots,s_p}\right\} - F_z\left(y^{\pm}_{r,j,n,s_1,\ldots,s_p}\right) - I\left\{z_j \le x^{\pm}_{r,j,n}\right\} + F_z\left(x^{\pm}_{r,j,n}\right)\right]\right| \xrightarrow{P} 0,$$

$$\sup_{r\in\{0,\ldots,M_n\}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \le x^{\pm}_{r,j,n}\right\} - F_z\left(x^{\pm}_{r,j,n}\right) - I\left\{z_j \le x_{r,n}\right\} + F_z\left(x_{r,n}\right)\right]\right| \xrightarrow{P} 0,$$

$$\sup_{|F_z(x)-F_z(y)|\le M_n^{-1}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \le x\right\} - F_z\left(x\right) - I\left\{z_j \le y\right\} + F_z\left(y\right)\right]\right| \xrightarrow{P} 0,$$

and

$$\sup_{\substack{r\in\{0,\ldots,M_n-1\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(y^{+}_{r+1,j,n,s_1,\ldots,s_p}\right) - F_z\left(y^{-}_{r,j,n,s_1,\ldots,s_p}\right)\right]\right| \xrightarrow{P} 0.$$

These four results are confirmed in Lemmas 6, 7, 8, and 9 respectively. □

Lemma 6 *As $n \to \infty$,*

$$\sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}}\left|n^{-1/2}\sum_{j=1}^{n-p}\left[I\left\{z_j \le y^{\pm}_{r,j,n,s_1,\ldots,s_p}\right\} - F_z\left(y^{\pm}_{r,j,n,s_1,\ldots,s_p}\right) - I\left\{z_j \le x^{\pm}_{r,j,n}\right\} + F_z\left(x^{\pm}_{r,j,n}\right)\right]\right| \xrightarrow{P} 0,$$

where $y^{-}_{r,j,n,s_1,\ldots,s_p}$, $y^{+}_{r,j,n,s_1,\ldots,s_p}$, $x^{-}_{r,j,n}$, and $x^{+}_{r,j,n}$ are defined in (3.29), (3.30), and (3.31).

Proof: We will only prove

$$\sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \le y^{+}_{r,j,n,s_1,\ldots,s_p}\right\} - F_z\left(y^{+}_{r,j,n,s_1,\ldots,s_p}\right) - I\left\{z_j \le x^{+}_{r,j,n}\right\} + F_z\left(x^{+}_{r,j,n}\right)\right]\right| \xrightarrow{P} 0,$$

as the proof of

$$\sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \le y^-_{r,j,n,s_1,\ldots,s_p}\right\} - F_z\left(y^-_{r,j,n,s_1,\ldots,s_p}\right) - I\left\{z_j \le x^-_{r,j,n}\right\} + F_z\left(x^-_{r,j,n}\right)\right]\right|$$
$$\xrightarrow{P} 0$$

is nearly identical.

By (3.19), for any $i \in \{1,\ldots,p\}$, any n, and any $j \in \{1,\ldots,n-p\}$, there exists a real-valued sequence $\{\ddot{\psi}_{k,i,j,n}\}_k$ such that

$$\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k} = \frac{\partial z_j(\phi_0)}{\partial\phi_i},$$

and, for some constants $a > 0$ and $0 < b < 1$,

$$\sup_{\substack{i\in\{1,\ldots,p\},j\in\{1,\ldots,n-p\}\\n\in\{1,2,\ldots\}}} |\ddot{\psi}_{k,i,j,n}| < ab^{|k|} \quad \forall k \in \{\ldots,-2,-1,1,2,\ldots\}.$$

In addition, from Lemma 5, there exist constants $\alpha > 0$ and $0 < \beta < 1$ such that

$$\sup_{\substack{i\in\{1,\ldots,p\},j\in\{1,\ldots,n-p\}\\n\in\{1,2,\ldots\}}} \left(|\ddot{\psi}_{k,i,j,n}| + \pi_k\right) < \alpha\beta^{|k|} \quad \forall k \in \{\ldots,-2,-1,1,2,\ldots\}.$$

So, by (3.28), for any n, any $j \in \{1,\ldots,n-p\}$, and any $(s_1,\ldots,s_p)' \in \{0,\ldots,N_n-1\}^p$,

$$\theta^+_{j,n,s_1,\ldots,s_p} = -\sum_{i=1}^{p} u_{i,s_i,n}\left\{\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}\right\}\left(1 + 2TN_n^{-1}u^{-1}_{i,s_i,n}I\left\{\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k} < 0\right\}\right).$$

As in Boldin (1982), define a non-decreasing sequence of integers $\{\ell_n\}_{n=1}^{\infty}$ such that ℓ_n is the integer part of $-4\log_\beta(n)$, and let

$$\tilde{\theta}^+_{j,n,s_1,\ldots,s_p} = -\sum_{i=1}^{p} u_{i,s_i,n}\left\{\sum_{0<|k|\le\ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k}\right\}\left(1 + 2TN_n^{-1}u^{-1}_{i,s_i,n}I\left\{\sum_{0<|k|\le\ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k} < 0\right\}\right)$$

and

$$\nu^+_{r,j,n,s_1,\ldots,s_p} = x^+_{r,j,n} + \frac{n^{-1/2}\tilde{\theta}^+_{j,n,s_1,\ldots,s_p}}{1 - L/n + 2LI\{\tilde{\theta}^+_{j,n,s_1,\ldots,s_p} < 0\}/n} + \frac{n^{-1}\sum_{0<|k|\le\ell_n}\pi_k|z_{j-k}|}{1 - L/n}.$$

To complete the proof, we will show that the terms

$$\sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \le y^+_{r,j,n,s_1,\ldots,s_p}\right\} - I\left\{z_j \le \nu^+_{r,j,n,s_1,\ldots,s_p}\right\}\right]\right|, \qquad (3.32)$$

$$\sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(y^+_{r,j,n,s_1,\ldots,s_p}\right)-F_z\left(\nu^+_{r,j,n,s_1,\ldots,s_p}\right)\right]\right|, \tag{3.33}$$

and

$$\sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j\le\nu^+_{r,j,n,s_1,\ldots,s_p}\right\}-F_z\left(\nu^+_{r,j,n,s_1,\ldots,s_p}\right)-I\left\{z_j\le x^+_{r,j,n}\right\}+F_z\left(x^+_{r,j,n}\right)\right]\right| \tag{3.34}$$

are all $o_p(1)$.

Because, for any n, $j\in\{1,\ldots,n-p\}$, $r\in\{0,\ldots,M_n\}$, and $(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p$,

$$y^+_{r,j,n,s_1,\ldots,s_p},\nu^+_{r,j,n,s_1,\ldots,s_p}\in\sigma(\ldots,z_{j-2},z_{j-1},z_{j+1},z_{j+2},\ldots),$$

for any $\eta>0$,

$$\begin{aligned}
&\mathrm{P}\left(\sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j\le y^+_{r,j,n,s_1,\ldots,s_p}\right\}-I\left\{z_j\le\nu^+_{r,j,n,s_1,\ldots,s_p}\right\}\right]\right|>\eta\right)\\
\le\;&(M_n+1)N_n^p\\
&\times\sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}}\mathrm{P}\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left|I\left\{z_j\le y^+_{r,j,n,s_1,\ldots,s_p}\right\}-I\left\{z_j\le\nu^+_{r,j,n,s_1,\ldots,s_p}\right\}\right|>\eta\right)\\
\le\;&\frac{(M_n+1)N_n^p}{\eta\sqrt{n}}\sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}}\sum_{j=1}^{n-p}\mathrm{E}\left|I\left\{z_j\le y^+_{r,j,n,s_1,\ldots,s_p}\right\}-I\left\{z_j\le\nu^+_{r,j,n,s_1,\ldots,s_p}\right\}\right|\\
\le\;&\frac{(M_n+1)N_n^p}{\eta\sqrt{n}}\left(\sup_{x\in\mathbb{R}}f_z(x)\right)\sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}}\sum_{j=1}^{n-p}\mathrm{E}\left|y^+_{r,j,n,s_1,\ldots,s_p}-\nu^+_{r,j,n,s_1,\ldots,s_p}\right|,
\end{aligned} \tag{3.35}$$

where the Markov inequality was used for the second inequality. Also,

$$\begin{aligned}
&\mathrm{E}\left|y^+_{r,j,n,s_1,\ldots,s_p}-\nu^+_{r,j,n,s_1,\ldots,s_p}\right|\\
\le\;&\frac{1}{\sqrt{n}}\mathrm{E}\left|\frac{\theta^+_{j,n,s_1,\ldots,s_p}}{1-L/n+2LI\{\theta^+_{j,n,s_1,\ldots,s_p}<0\}/n}-\frac{\tilde\theta^+_{j,n,s_1,\ldots,s_p}}{1-L/n+2LI\{\tilde\theta^+_{j,n,s_1,\ldots,s_p}<0\}/n}\right|\\
&+\frac{1}{n}\frac{\mathrm{E}\left\{\sum_{|k|>\ell_n}\pi_k|z_{j-k}|\right\}}{|1-L/n|},
\end{aligned} \tag{3.36}$$

and since

$$\mathrm{E}\left\{\sum_{|k|>\ell_n}\pi_k|z_{j-k}|\right\}<\mathrm{E}|z_1|\sum_{|k|>\ell_n}\alpha\beta^{|k|}\le\frac{1}{n^4}\frac{2\alpha}{1-\beta}\mathrm{E}|z_1|,$$

it follows that, for n is sufficiently large so that $1 - L/n > 0$, (3.36) is bounded above by

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\mathrm{E}\left|\frac{(1-L/n)(\theta^+_{j,n,s_1,\ldots,s_p}-\tilde{\theta}^+_{j,n,s_1,\ldots,s_p})}{(1-L/n+2LI\{\theta^+_{j,n,s_1,\ldots,s_p}<0\}/n)(1-L/n+2LI\{\tilde{\theta}^+_{j,n,s_1,\ldots,s_p}<0\}/n)}\right| \\
&\quad+\frac{2L}{n^{3/2}}\mathrm{E}\left|\frac{\theta^+_{j,n,s_1,\ldots,s_p}I\{\tilde{\theta}^+_{j,n,s_1,\ldots,s_p}<0\}-\tilde{\theta}^+_{j,n,s_1,\ldots,s_p}I\{\theta^+_{j,n,s_1,\ldots,s_p}<0\}}{(1-L/n+2LI\{\theta^+_{j,n,s_1,\ldots,s_p}<0\}/n)(1-L/n+2LI\{\tilde{\theta}^+_{j,n,s_1,\ldots,s_p}<0\}/n)}\right| \\
&\quad+\frac{1}{n^5}\frac{2\alpha}{1-\beta}\frac{\mathrm{E}|z_1|}{1-L/n} \\
\leq\;& \frac{1}{\sqrt{n}}\frac{\mathrm{E}|\theta^+_{j,n,s_1,\ldots,s_p}-\tilde{\theta}^+_{j,n,s_1,\ldots,s_p}|}{1-L/n}+\frac{2L}{n^{3/2}(1-L/n)^2}\mathrm{E}\left\{|\theta^+_{j,n,s_1,\ldots,s_p}|+|\tilde{\theta}^+_{j,n,s_1,\ldots,s_p}|\right\} \\
&\quad+\frac{1}{n^5}\frac{2\alpha}{1-\beta}\frac{\mathrm{E}|z_1|}{1-L/n}. \qquad (3.37)
\end{aligned}$$

Note that

$$\begin{aligned}
\mathrm{E}\left\{|\theta^+_{j,n,s_1,\ldots,s_p}|+|\tilde{\theta}^+_{j,n,s_1,\ldots,s_p}|\right\} &\leq 2T(1+2N_n^{-1})\sum_{i=1}^{p}\sum_{k\neq 0}\mathrm{E}|\ddot{\psi}_{k,i,j,n}z_{j-k}| \\
&< 2pT(1+2N_n^{-1})\mathrm{E}|z_1|\sum_{k\neq 0}\alpha\beta^{|k|} \\
&= 4\frac{\alpha\beta}{1-\beta}pT(1+2N_n^{-1})\mathrm{E}|z_1|, \qquad (3.38)
\end{aligned}$$

$$\begin{aligned}
&\mathrm{E}|\theta^+_{j,n,s_1,\ldots,s_p}-\tilde{\theta}^+_{j,n,s_1,\ldots,s_p}| \\
\leq\;& T\sum_{i=1}^{p}\sum_{|k|>\ell_n}\mathrm{E}|\ddot{\psi}_{k,i,j,n}z_{j-k}|+2TN_n^{-1} \\
&\times\sum_{i=1}^{p}\mathrm{E}\left|I\left\{\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}<0\right\}\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}-I\left\{\sum_{0<|k|\leq\ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k}<0\right\}\sum_{0<|k|\leq\ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k}\right| \\
\leq\;& T(1+2N_n^{-1})\sum_{i=1}^{p}\sum_{|k|>\ell_n}\mathrm{E}|\ddot{\psi}_{k,i,j,n}z_{j-k}| \\
&+2TN_n^{-1}\sum_{i=1}^{p}\mathrm{E}\left(\left|I\left\{\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}<0\right\}-I\left\{\sum_{0<|k|\leq\ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k}<0\right\}\right|\sum_{k\neq 0}|\ddot{\psi}_{k,i,j,n}z_{j-k}|\right) \\
<\;& pT(1+2N_n^{-1})\mathrm{E}|z_1|\sum_{|k|>\ell_n}\alpha\beta^{|k|}+2TN_n^{-1} \\
&\times\sum_{i=1}^{p}\left[\mathrm{E}\left(\sum_{k\neq 0}|\ddot{\psi}_{k,i,j,n}z_{j-k}|\right)^2\mathrm{E}\left|I\left\{\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}<0\right\}-I\left\{\sum_{0<|k|\leq\ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k}<0\right\}\right|^2\right]^{1/2} \\
<\;& \frac{1}{n^4}\frac{2\alpha}{1-\beta}pT(1+2N_n^{-1})\mathrm{E}|z_1|+2TN_n^{-1}\left[\mathrm{E}\left(\sum_{k\neq 0}\alpha\beta^{|k|}|z_{j-k}|\right)^2\right]^{1/2}
\end{aligned}$$

$$\times \sum_{i=1}^{p}\left[\mathrm{E}\left|I\left\{\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}<0\right\}-I\left\{\sum_{0<|k|\leq \ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k}<0\right\}\right|\right]^{1/2}, \tag{3.39}$$

where the Cauchy-Schwartz inequality was used in the third inequality, and

$$\left[\mathrm{E}\left(\sum_{k\neq 0}\alpha\beta^{|k|}|z_{j-k}|\right)^2\right]^{1/2} \leq \sum_{k\neq 0}\alpha\beta^{|k|}\left[\mathrm{E}\{z_1^2\}\right]^{1/2} = 2\frac{\alpha\beta}{1-\beta}\left[\mathrm{E}\{z_1^2\}\right]^{1/2} \tag{3.40}$$

by the Minkowski inequality. Also, for any $i \in \{1,\ldots,p\}$, if $f_{i,j,n}$ is the density function for $\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}$,

$$\begin{aligned}
&\mathrm{E}\left|I\left\{\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}<0\right\}-I\left\{\sum_{0<|k|\leq \ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k}<0\right\}\right| \\
&= \mathrm{P}\left(\left\{\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}\geq 0\right\}\cap\left\{\sum_{0<|k|\leq \ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k}<0\right\}\right) \\
&\quad +\mathrm{P}\left(\left\{\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}< 0\right\}\cap\left\{\sum_{0<|k|\leq \ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k}\geq 0\right\}\right) \\
&\leq \mathrm{P}\left(\left\{\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}\geq \frac{1}{n^2}\right\}\cap\left\{\sum_{|k|>\ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k}>\frac{1}{n^2}\right\}\right)+\mathrm{P}\left(0\leq\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}<\frac{1}{n^2}\right) \\
&\quad +\mathrm{P}\left(\left\{\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}\leq \frac{-1}{n^2}\right\}\cap\left\{\sum_{|k|>\ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k}\leq\frac{-1}{n^2}\right\}\right)+\mathrm{P}\left(\frac{-1}{n^2}<\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}<0\right) \\
&\leq \mathrm{P}\left(\left|\sum_{|k|>\ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k}\right|\geq\frac{1}{n^2}\right)+\mathrm{P}\left(\left|\sum_{k\neq 0}\ddot{\psi}_{k,i,j,n}z_{j-k}\right|<\frac{1}{n^2}\right) \\
&\leq n^2\sum_{|k|>\ell_n}\mathrm{E}|\ddot{\psi}_{k,i,j,n}z_{j-k}|+\frac{2}{n^2}\sup_{x\in\mathbb{R}} f_{i,j,n}(x).
\end{aligned} \tag{3.41}$$

From (3.19), $\ddot{\psi}_{i,i,j,n}=-1$ and so, if $f_{i,i,j,n}$ is the density function for $\sum_{m\notin\{0,i\}}\ddot{\psi}_{m,i,j,n}z_{j-m}$, then

$$\begin{aligned}
\sup_{x\in\mathbb{R}} f_{i,j,n}(x) &= \sup_{x\in\mathbb{R}}\int_{-\infty}^{\infty}|\ddot{\psi}_{i,i,j,n}|^{-1}f_z\left(\frac{x-s}{\ddot{\psi}_{i,i,j,n}}\right)f_{i,i,j,n}(s)\,ds \\
&\leq \sup_{x\in\mathbb{R}} f_z(x).
\end{aligned}$$

Because

$$\sum_{|k|>\ell_n}\mathrm{E}|\ddot{\psi}_{k,i,j,n}z_{j-k}| < \mathrm{E}|z_1|\sum_{|k|>\ell_n}\alpha\beta^{|k|}\leq\frac{2}{n^4}\frac{\alpha}{1-\beta}\mathrm{E}|z_1|,$$

(3.41) is bounded above by

$$\frac{2}{n^2}\frac{\alpha}{1-\beta}\mathrm{E}|z_1|+\frac{2}{n^2}\sup_{x\in\mathbb{R}} f_z(x). \tag{3.42}$$

Combining (3.36)-(3.42), we have, for n sufficiently large, $j \in \{1,\ldots,n-p\}$, $r \in \{0,\ldots,M_n\}$, and $(s_1,\ldots,s_p)' \in \{0,\ldots,N_n-1\}^p$,

$$\begin{aligned}
&\mathrm{E}\left|y^+_{r,j,n,s_1,\ldots,s_p} - \nu^+_{r,j,n,s_1,\ldots,s_p}\right| \\
&< \frac{1}{n^{9/2}}\frac{2\alpha}{1-\beta}\frac{pT}{1-L/n}(1+2N_n^{-1})\mathrm{E}|z_1| \\
&\quad + \frac{1}{n^{3/2}N_n}\frac{4\sqrt{2}\,\alpha\beta}{1-\beta}\frac{pT}{1-L/n}\left(\mathrm{E}\{z_1^2\}\right)^{1/2}\left(\frac{\alpha}{1-\beta}\mathrm{E}|z_1| + \sup_{x\in\mathbb{R}} f_z(x)\right)^{1/2} \\
&\quad + \frac{1}{n^{3/2}}\frac{8\alpha\beta}{1-\beta}\frac{pLT}{(1-L/n)^2}(1+2N_n^{-1})\mathrm{E}|z_1| + \frac{1}{n^5}\frac{2\alpha}{1-\beta}\frac{1}{1-L/n}\mathrm{E}|z_1| \\
&= O(n^{-3/2}),
\end{aligned}$$

and, since $M_n \sim n^{3/4}$ and $N_n \sim n^{1/(8p)}$, (3.35) is bounded above by

$$\sqrt{n}(M_n+1)N_n^p\, O(n^{-3/2}) \to 0.$$

Because η was arbitrarily chosen, (3.32) is $o_p(1)$. It can be shown similarly that (3.33) is $o_p(1)$.

Now consider (3.34). For all n, $r \in \{0,\ldots,M_n\}$, $j \in \{1,\ldots,n-p\}$, and $(s_1,\ldots,s_p)' \in \{0,\ldots,N_n-1\}^p$, if

$$\xi_{r,j,n,s_1,\ldots,s_p} := I\left\{z_j \le \nu^+_{r,j,n,s_1,\ldots,s_p}\right\} - F_z\left(\nu^+_{r,j,n,s_1,\ldots,s_p}\right) - I\left\{z_j \le x^+_{r,j,n}\right\} + F_z\left(x^+_{r,j,n}\right),$$

then $\{\xi_{r,j,n,s_1,\ldots,s_p}\}_{j=1}^{n-p}$ is a $2\ell_n$-dependent sequence with mean zero. Thus,

$$\begin{aligned}
&\mathrm{E}\left\{\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\xi_{r,j,n,s_1,\ldots,s_p}\right\}^4 \\
&= \frac{1}{n^2}\sum_{j=1}^{n-p}\mathrm{E}\{\xi^4_{r,j,n,s_1,\ldots,s_p}\} + \frac{3}{n^2}\sum_{\substack{j=1\\ j\ne k}}^{n-p}\sum_{k=1}^{n-p}\mathrm{E}\{\xi^2_{r,j,n,s_1,\ldots,s_p}\xi^2_{r,k,n,s_1,\ldots,s_p}\} \\
&\quad + \frac{4}{n^2}\sum_{\substack{j=1\\ j\ne k}}^{n-p}\sum_{k=1}^{n-p}\mathrm{E}\{\xi^3_{r,j,n,s_1,\ldots,s_p}\xi_{r,k,n,s_1,\ldots,s_p}\} \\
&\quad + \frac{6}{n^2}\sum_{\substack{j=1\\ (j\notin\{k,m\})\cap(k\ne m)}}^{n-p}\sum_{k=1}^{n-p}\sum_{m=1}^{n-p}\mathrm{E}\{\xi^2_{r,j,n,s_1,\ldots,s_p}\xi_{r,k,n,s_1,\ldots,s_p}\xi_{r,m,n,s_1,\ldots,s_p}\} \\
&\quad + \frac{1}{n^2}\sum_{\substack{j=1\\ (j\notin\{k,m,q\})\cap(k\notin\{m,q\})\cap(m\ne q)}}^{n-p}\sum_{k=1}^{n-p}\sum_{m=1}^{n-p}\sum_{q=1}^{n-p}\mathrm{E}\{\xi_{r,j,n,s_1,\ldots,s_p}\xi_{r,k,n,s_1,\ldots,s_p}\xi_{r,m,n,s_1,\ldots,s_p}\xi_{r,q,n,s_1,\ldots,s_p}\}
\end{aligned}$$

$$
\begin{aligned}
\leq\ & \frac{1}{n^2}\sum_{j=1}^{n-p}\mathrm{E}\{\xi^4_{r,j,n,s_1,\ldots,s_p}\} + \frac{3}{n^2}\sum_{\substack{j=1\ k=1\\ 0<|j-k|\leq 2\ell_n}}^{n-p}\sum^{n-p}\mathrm{E}\{\xi^2_{r,j,n,s_1,\ldots,s_p}\xi^2_{r,k,n,s_1,\ldots,s_p}\}\\
&+\frac{3}{n^2}\sum_{\substack{j=1\ k=1\\ |j-k|>2\ell_n}}^{n-p}\sum^{n-p}\mathrm{E}\{\xi^2_{r,j,n,s_1,\ldots,s_p}\}\mathrm{E}\{\xi^2_{r,k,n,s_1,\ldots,s_p}\} + \frac{4}{n^2}\sum_{\substack{j=1\ k=1\\ 0<|j-k|\leq 2\ell_n}}^{n-p}\sum^{n-p}\mathrm{E}|\xi^3_{r,j,n,s_1,\ldots,s_p}\xi_{r,k,n,s_1,\ldots,s_p}|\\
&+\frac{6}{n^2}\sum_{\substack{j=1\ k=1\ m=1\\ (0<|j-k|\leq 4\ell_n)\cap(0<|j-m|\leq 4\ell_n)}}^{n-p}\sum^{n-p}\sum^{n-p}\mathrm{E}|\xi^2_{r,j,n,s_1,\ldots,s_p}\xi_{r,k,n,s_1,\ldots,s_p}\xi_{r,m,n,s_1,\ldots,s_p}|\\
&+\frac{6}{n^2}\sum_{\substack{j=1\ k=1\ m=1\\ 0<|k-m|\leq 2\ell_n}}^{n-p}\sum^{n-p}\sum^{n-p}\mathrm{E}\{\xi^2_{r,j,n,s_1,\ldots,s_p}\}\mathrm{E}|\xi_{r,k,n,s_1,\ldots,s_p}\xi_{r,m,n,s_1,\ldots,s_p}|\\
&+\frac{1}{n^2}\sum_{\substack{j=1\ k=1\ m=1\ q=1\\ (0<|j-k|\leq 6\ell_n)\cap(0<|j-m|\leq 6\ell_n)\cap(0<|j-q|\leq 6\ell_n)}}^{n-p}\sum^{n-p}\sum^{n-p}\sum^{n-p}\mathrm{E}|\xi_{r,j,n,s_1,\ldots,s_p}\xi_{r,k,n,s_1,\ldots,s_p}\xi_{r,m,n,s_1,\ldots,s_p}\xi_{r,q,n,s_1,\ldots,s_p}|\\
&+\frac{3}{n^2}\sum_{\substack{j=1\ k=1\ m=1\ q=1\\ (0<|j-k|\leq 2\ell_n)\cap(0<|m-q|\leq 2\ell_n)}}^{n-p}\sum^{n-p}\sum^{n-p}\sum^{n-p}\mathrm{E}|\xi_{r,j,n,s_1,\ldots,s_p}\xi_{r,k,n,s_1,\ldots,s_p}|\mathrm{E}|\xi_{r,m,n,s_1,\ldots,s_p}\xi_{r,q,n,s_1,\ldots,s_p}|.
\end{aligned}
$$

Since $\sup_{j\in\{1,\ldots,n-p\}}|\xi_{r,j,n,s_1,\ldots,s_p}|\leq 2$, the right-hand side is bounded above by

$$
\begin{aligned}
&\frac{16}{n}+3\frac{16}{n}(4\ell_n)+3(4)\left(\sup_{j\in\{1,\ldots,n-p\}}\mathrm{E}|\xi_{r,j,n,s_1,\ldots,s_p}|\right)^2+4\frac{16}{n}(4\ell_n)+6\frac{16}{n}(8\ell_n)^2\\
&\quad+6(4)(4\ell_n)\left(\sup_{j\in\{1,\ldots,n-p\}}\mathrm{E}|\xi_{r,j,n,s_1,\ldots,s_p}|\right)^2+\frac{16}{n}(12\ell_n)^3\\
&\quad+3(4)(4\ell_n)^2\left(\sup_{j\in\{1,\ldots,n-p\}}\mathrm{E}|\xi_{r,j,n,s_1,\ldots,s_p}|\right)^2\\
=\ & \frac{1}{n}O(\ell_n^3)+\left(\sup_{j\in\{1,\ldots,n-p\}}\mathrm{E}|\xi_{r,j,n,s_1,\ldots,s_p}|\right)^2 O(\ell_n^2).
\end{aligned}
$$

If n is sufficiently large so that $1-L/n>0$,

$$
\begin{aligned}
&\sup_{j\in\{1,\ldots,n-p\}}\mathrm{E}|\xi_{r,j,n,s_1,\ldots,s_p}|\\
\leq\ & \sup_{j\in\{1,\ldots,n-p\}} 2\mathrm{E}\left|F_z\left(\nu^+_{r,j,n,s_1,\ldots,s_p}\right)-F_z\left(x^+_{r,j,n}\right)\right|\\
\leq\ & 2\left(\sup_{x\in\mathbb{R}} f_z(x)\right)\sup_{j\in\{1,\ldots,n-p\}}\mathrm{E}\left|\nu^+_{r,j,n,s_1,\ldots,s_p}-x^+_{r,j,n}\right|\\
=\ & 2\left(\sup_{x\in\mathbb{R}} f_z(x)\right)\sup_{j\in\{1,\ldots,n-p\}}\mathrm{E}\left|\frac{n^{-1/2}\tilde{\theta}^+_{j,n,s_1,\ldots,s_p}}{1-L/n+2LI\{\tilde{\theta}^+_{j,n,s_1,\ldots,s_p}<0\}/n}+\frac{n^{-1}\sum_{0<|k|\leq\ell_n}\pi_k|z_{j-k}|}{1-L/n}\right|
\end{aligned}
$$

$$\begin{aligned}
&\le \frac{2}{1-L/n}\left(\sup_{x\in\mathbb{R}} f_z(x)\right) \\
&\quad\times \sup_{j\in\{1,\ldots,n-p\}}\left[\frac{1}{\sqrt{n}}T(1+2N_n^{-1})\sum_{i=1}^{p}\mathrm{E}\left|\sum_{0<|k|\le\ell_n}\ddot{\psi}_{k,i,j,n}z_{j-k}\right| + \frac{1}{n}\mathrm{E}|z_1|\left|\sum_{k\neq 0}\pi_k\right|\right] \\
&< \frac{2}{1-L/n}\left(\sup_{x\in\mathbb{R}} f_z(x)\right)\mathrm{E}|z_1|\sum_{k\neq 0}\alpha\beta^{|k|}\left[\frac{1}{\sqrt{n}}pT(1+2N_n^{-1})+\frac{1}{n}\right] \\
&= \frac{4\alpha\beta}{(1-\beta)(1-L/n)}\left(\sup_{x\in\mathbb{R}} f_z(x)\right)\mathrm{E}|z_1|\left[\frac{1}{\sqrt{n}}pT(1+2N_n^{-1})+\frac{1}{n}\right].
\end{aligned}$$

Consequently, for any $\eta > 0$ and large n,

$$\begin{aligned}
&\mathrm{P}\left(\sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\xi_{r,j,n,s_1,\ldots,s_p}\right| > \eta\right) \\
&\le (M_n+1)N_n^p \sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \mathrm{P}\left(\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\xi_{r,j,n,s_1,\ldots,s_p}\right|^4 > \eta^4\right) \\
&\le \frac{(M_n+1)N_n^p}{\eta^4}\sup_{\substack{r\in\{0,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \mathrm{E}\left\{\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\xi_{r,j,n,s_1,\ldots,s_p}\right\}^4 \\
&< \frac{(M_n+1)N_n^p}{\eta^4}\frac{1}{n}O(\ell_n^3) + \frac{(M_n+1)N_n^p}{\eta^4}O(\ell_n^2) \\
&\quad\times\left\{\frac{4\alpha\beta}{(1-\beta)(1-L/n)}\left(\sup_{x\in\mathbb{R}} f_z(x)\right)\mathrm{E}|z_1|\left[\frac{1}{\sqrt{n}}pT(1+2N_n^{-1})+\frac{1}{n}\right]\right\}^2 \\
&\to 0,
\end{aligned}$$

and so (3.34) must be $o_p(1)$. □

Lemma 7 *As $n\to\infty$,*

$$\sup_{r\in\{0,\ldots,M_n\}}\left|n^{-1/2}\sum_{j=1}^{n-p}\left[I\left\{z_j\le x_{r,j,n}^{\pm}\right\} - F_z\left(x_{r,j,n}^{\pm}\right) - I\left\{z_j\le x_{r,n}\right\} + F_z\left(x_{r,n}\right)\right]\right| \xrightarrow{P} 0,$$

where $x_{r,n}$, $x_{r,j,n}^{-}$, and $x_{r,j,n}^{+}$ are defined in (3.27), (3.30), and (3.31) respectively.

Proof: For any $\eta > 0$,

$$\mathrm{P}\left(\sup_{r\in\{0,\ldots,M_n\}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j\le x_{r,j,n}^{+}\right\} - F_z\left(x_{r,j,n}^{+}\right) - I\left\{z_j\le x_{r,n}\right\} + F_z\left(x_{r,n}\right)\right]\right| > \eta\right)$$

$$\begin{aligned}
&\leq (M_n+1) \\
&\quad \times \sup_{r\in\{0,\ldots,M_n\}} \mathrm{P}\left(\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[I\left\{z_j \leq x^{+}_{r,j,n}\right\} - F_z\left(x^{+}_{r,j,n}\right) - I\left\{z_j \leq x_{r,n}\right\} + F_z\left(x_{r,n}\right)\right]\right|^2 > \eta^2\right) \\
&\leq \frac{M_n+1}{\eta^2 n}\sup_{r\in\{0,\ldots,M_n\}} \mathrm{E}\left|\sum_{j=1}^{n-p}\left[I\left\{z_j \leq x^{+}_{r,j,n}\right\} - F_z\left(x^{+}_{r,j,n}\right) - I\left\{z_j \leq x_{r,n}\right\} + F_z\left(x_{r,n}\right)\right]\right|^2 \\
&= \frac{M_n+1}{\eta^2 n}\sup_{r\in\{0,\ldots,M_n\}} \sum_{j=1}^{n-p}\mathrm{E}\left[I\left\{z_j \leq x^{+}_{r,j,n}\right\} - F_z\left(x^{+}_{r,j,n}\right) - I\left\{z_j \leq x_{r,n}\right\} + F_z\left(x_{r,n}\right)\right]^2 \qquad (3.43)
\end{aligned}$$

because, for any n and any $r \in \{0,\ldots,M_n\}$,

$$\left\{I\left\{z_j \leq x^{+}_{r,j,n}\right\} - F_z\left(x^{+}_{r,j,n}\right) - I\left\{z_j \leq x_{r,n}\right\} + F_z\left(x_{r,n}\right)\right\}_j$$

is a sequence of independent random variables with mean zero. For all n sufficiently large so that $1 - L/n > 0$,

$$\begin{aligned}
&\mathrm{E}\left[I\left\{z_j \leq x^{+}_{r,j,n}\right\} - F_z\left(x^{+}_{r,j,n}\right) - I\left\{z_j \leq x_{r,n}\right\} + F_z\left(x_{r,n}\right)\right]^2 \\
&\quad = \mathrm{E}\left[I\left\{x_{r,n} < z_j \leq x^{+}_{r,j,n}\right\}\right]^2 + \left(F_z\left(x^{+}_{r,j,n}\right) - F_z\left(x_{r,n}\right)\right)^2 \\
&\quad\quad -2\left(F_z\left(x^{+}_{r,j,n}\right) - F_z\left(x_{r,n}\right)\right)\mathrm{E}\left[I\left\{x_{r,n} < z_j \leq x^{+}_{r,j,n}\right\}\right] \\
&\quad = \mathrm{E}\left[I\left\{x_{r,n} < z_j \leq x^{+}_{r,j,n}\right\}\right] - \left(F_z\left(x^{+}_{r,j,n}\right) - F_z(x_{r,n})\right)^2 \\
&\quad \leq F_z\left(x^{+}_{r,j,n}\right) - F_z(x_{r,n}),
\end{aligned}$$

(3.43) is bounded above by

$$2\frac{M_n+1}{\eta^2 n}\sup_{r\in\{0,\ldots,M_n\}}\sum_{j=1}^{n-p}\left(F_z\left(x^{+}_{r,j,n}\right) - F_z\left(x_{r,n}\right)\right),$$

and, from (3.31), this can be bounded above by

$$\begin{aligned}
&2\frac{M_n+1}{\eta^2 n}\sup_{r\in\{0,\ldots,M_n\}}\sum_{j=1}^{n-p}\left[\left(\sup_{x\in\mathbb{R}} f_z(x)\right)\frac{C_\delta D^{n-p-j}}{1-L/n} + \frac{1}{n}f_z\left(x^{*}_{r,j,n}\right)\left|\frac{x_{r,n}\left(L - 2LI\{x_{r,n}<0\}\right)}{1-L/n+2LI\{x_{r,n}<0\}/n}\right|\right] \\
&\leq 2\frac{M_n+1}{\eta^2 n}\sum_{j=1}^{n-p}\left[\left(\sup_{x\in\mathbb{R}} f_z(x)\right)\frac{C_\delta D^{n-p-j}}{1-L/n} + \frac{1}{n}\left(\sup_{x\in\mathbb{R}}|x|f_z(x)\right)\frac{L}{1-L/n}\right], \qquad (3.44)
\end{aligned}$$

where $x^{*}_{r,j,n}$ lies between $x_{r,n}$ and

$$\frac{x_{r,n}}{1-L/n+2LI\{x_{r,n}<0\}/n}.$$

Because $M_n \sim n^{3/4}$ and $0 < D < 1$, (3.44) converges to zero. It follows that

$$\sup_{r\in\{0,\ldots,M_n\}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} \left[I\left\{z_j \le x^+_{r,j,n}\right\} - F_z\left(x^+_{r,j,n}\right) - I\left\{z_j \le x_{r,n}\right\} + F_z\left(x_{r,n}\right) \right] \right| \xrightarrow{P} 0,$$

and, similarly,

$$\sup_{r\in\{0,\ldots,M_n\}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} \left[I\left\{z_j \le x^-_{r,j,n}\right\} - F_z\left(x^-_{r,j,n}\right) - I\left\{z_j \le x_{r,n}\right\} + F_z\left(x_{r,n}\right) \right] \right| \xrightarrow{P} 0.$$

□

Lemma 8 *If $M_n \sim n^{3/4}$, then, as $n \to \infty$,*

$$\sup_{|F_z(x)-F_z(y)|\le M_n^{-1}} \left| n^{-1/2} \sum_{j=1}^{n-p} \left[I\{z_j \le x\} - F_z(x) - I\{z_j \le y\} + F_z(y) \right] \right| \xrightarrow{P} 0. \tag{3.45}$$

Proof: Observe that the left-hand side of (3.45) equals

$$\sup_{|F_z(x)-F_z(y)|\le M_n^{-1}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} \left[I\{F_z(z_j) \le F_z(x)\} - F_z(x) - I\{F_z(z_j) \le F_z(y)\} + F_z(y) \right] \right|$$
$$= \sup_{s,t\in[0,1],|s-t|\le M_n^{-1}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} \left[I\{F_z(z_j) \le s\} - s - I\{F_z(z_j) \le t\} + t \right] \right|.$$

If $W(\cdot)$ denotes a Brownian Bridge on $[0,1]$, by Theorem 14.3 in Billingsley (1999),

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} \left(I\{F_z(z_j) \le \cdot\} - \cdot \right) \xrightarrow{d} W(\cdot)$$

on $D([0,1])$, the space of right-continuous functions on $[0,1]$ with left-hand limits. Note that the sample paths of $W(\cdot)$ are continuous on $[0,1]$ almost surely. By the continuous mapping theorem, for any $\delta > 0$,

$$\sup_{s,t\in[0,1],|s-t|\le\delta} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p} \left[I\{F_z(z_j) \le s\} - s - I\{F_z(z_j) \le t\} + t \right] \right| \xrightarrow{d} \sup_{s,t\in[0,1],|s-t|\le\delta} \left| W(s) - W(t) \right|.$$

Since $M_n^{-1} \to 0$, we can let $\delta \to 0$ to obtain the result. □

Lemma 9 *As $n \to \infty$,*

$$\sup_{\substack{r\in\{0,\ldots,M_n-1\} \\ (s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left| n^{-1/2} \sum_{j=1}^{n-p} \left[F_z\left(y^+_{r+1,j,n,s_1,\ldots,s_p}\right) - F_z\left(y^-_{r,j,n,s_1,\ldots,s_p}\right) \right] \right| \xrightarrow{P} 0,$$

where $y^+_{r+1,j,n,s_1,\ldots,s_p}$ and $y^-_{r,j,n,s_1,\ldots,s_p}$ are defined in (3.29).

Proof: For any $r \in \{0, \ldots, M_n - 1\}$ and any $(s_1, \ldots, s_p)' \in \{0, \ldots, N_n - 1\}^p$,

$$\begin{aligned}
&\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(y^+_{r+1,j,n,s_1,\ldots,s_p}\right) - F_z\left(y^-_{r,j,n,s_1,\ldots,s_p}\right)\right]\right| \\
&\leq \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(y^+_{r+1,j,n,s_1,\ldots,s_p}\right) - F_z\left(x^+_{r+1,j,n}\right)\right]\right| + \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(x^+_{r+1,j,n}\right) - F_z\left(x_{r+1,n}\right)\right]\right| \\
&\quad + \frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(x_{r+1,n}\right) - F_z\left(x_{r,n}\right)\right] + \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(x_{r,n}\right) - F_z\left(x^-_{r,j,n}\right)\right]\right| \\
&\quad + \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(x^-_{r,j,n}\right) - F_z\left(y^-_{r,j,n,s_1,\ldots,s_p}\right)\right]\right|, \qquad (3.46)
\end{aligned}$$

where $x_{r,n}$, $x_{r+1,n}$, $x^-_{r,j,n}$, and $x^+_{r+1,j,n}$ are defined in (3.27), (3.30), and (3.31).

By the triangle inequality,

$$\begin{aligned}
&\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(y^+_{r,j,n,s_1,\ldots,s_p}\right) - F_z\left(x^+_{r,j,n}\right)\right]\right| \\
&\leq \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(y^+_{r,j,n,s_1,\ldots,s_p}\right) - F_z\left(y^+_{r,j,n,s_1,\ldots,s_p} - \frac{n^{-1}\sum_{k\neq 0}\pi_k|z_{j-k}|}{1-L/n}\right)\right]\right| \\
&\quad + \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(y^+_{r,j,n,s_1,\ldots,s_p} - \frac{n^{-1}\sum_{k\neq 0}\pi_k|z_{j-k}|}{1-L/n}\right) - F_z\left(x^+_{r,j,n} + n^{-1/2}\theta^+_{j,n,s_1,\ldots,s_p}\right)\right]\right| \\
&\quad + \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(x^+_{r,j,n} + n^{-1/2}\theta^+_{j,n,s_1,\ldots,s_p}\right) - F_z\left(x^+_{r,j,n}\right)\right]\right|,
\end{aligned}$$

where L, $\{\pi_k\}$, and $\theta^+_{j,n,s_1,\ldots,s_p}$ are given in Lemma 5. Observe that

$$\begin{aligned}
&\sup_{\substack{r\in\{1,\ldots,M_n\} \\ (s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(y^+_{r,j,n,s_1,\ldots,s_p}\right) - F_z\left(y^+_{r,j,n,s_1,\ldots,s_p} - \frac{n^{-1}\sum_{k\neq 0}\pi_k|z_{j-k}|}{1-L/n}\right)\right]\right| \\
&\leq \left(\sup_{x\in\mathbb{R}} f_z(x)\right)\frac{1}{n^{3/2}|1-L/n|}\sum_{j=1}^{n-p}\sum_{k\neq 0}\pi_k|z_{j-k}| \\
&\xrightarrow{P} 0
\end{aligned}$$

by the ergodic theorem because $\{\pi_k\}$ is a non-negative, geometrically decaying sequence. From (3.29), for all n sufficiently large so that $1 - L/n > 0$,

$$\sup_{\substack{r\in\{1,\ldots,M_n\} \\ (s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(y^+_{r,j,n,s_1,\ldots,s_p} - \frac{n^{-1}\sum_{k\neq 0}\pi_k|z_{j-k}|}{1-L/n}\right) - F_z\left(x^+_{r,j,n} + n^{-1/2}\theta^+_{j,n,s_1,\ldots,s_p}\right)\right]\right|$$

$$\leq \left(\sup_{x\in\mathbb{R}} f_z(x)\right) \sup_{(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p} \frac{1}{n}\sum_{j=1}^{n-p}\left|\frac{\theta^+_{j,n,s_1,\ldots,s_p}}{1-L/n+2LI\{\theta^+_{j,n,s_1,\ldots,s_p}<0\}/n} - \theta^+_{j,n,s_1,\ldots,s_p}\right|$$

$$\leq \left(\sup_{x\in\mathbb{R}} f_z(x)\right) \frac{L}{1-L/n} \sup_{(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p} \frac{1}{n^2}\sum_{j=1}^{n-p}\left|\theta^+_{j,n,s_1,\ldots,s_p}\right|,$$

and, from (3.28), this is bounded above by

$$\left(\sup_{x\in\mathbb{R}} f_z(x)\right) \frac{L}{1-L/n} T(1+2N_n^{-1})\frac{1}{n^2}\sum_{i=1}^{p}\sum_{j=1}^{n-p}\left|\frac{\partial z_j(\boldsymbol{\phi}_0)}{\partial\phi_i}\right| \xrightarrow{P} 0$$

since $N_n \to \infty$ and

$$\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{n-p}\left|\frac{\partial z_j(\boldsymbol{\phi}_0)}{\partial\phi_i}\right| = O_p(1).$$

Now consider

$$\sup_{\substack{r\in\{1,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(x^+_{r,j,n}+n^{-1/2}\theta^+_{j,n,s_1,\ldots,s_p}\right)-F_z\left(x^+_{r,j,n}\right)\right]\right|$$

$$= \sup_{\substack{r\in\{1,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left|\frac{1}{n}\sum_{j=1}^{n-p}\left[f_z\left(\tilde{x}^+_{r,j,n,s_1,\ldots,s_p}\right)\theta^+_{j,n,s_1,\ldots,s_p}\right]\right|$$

$$\leq \sup_{\substack{r\in\{1,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left|\frac{1}{n}\sum_{j=1}^{n-p} f_z\left(\frac{x_{r,n}}{1-L/n+2LI\{x_{r,n}<0\}/n}\right)\theta^+_{j,n,s_1,\ldots,s_p}\right|$$

$$+ \sup_{\substack{r\in\{1,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left|\frac{1}{n}\sum_{j=1}^{n-p}\left[f_z\left(\tilde{x}^+_{r,j,n,s_1,\ldots,s_p}\right)-f_z\left(\frac{x_{r,n}}{1-L/n+2LI\{x_{r,n}<0\}/n}\right)\right]\theta^+_{j,n,s_1,\ldots,s_p}\right|$$

$$\leq \left(\sup_{x\in\mathbb{R}} f_z(x)\right) \sup_{(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}\left|\frac{1}{n}\sum_{j=1}^{n-p}\theta^+_{j,n,s_1,\ldots,s_p}\right|$$

$$+\left(\sup_{(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}\frac{1}{n}\sum_{j=1}^{n-p}\left|\theta^+_{j,n,s_1,\ldots,s_p}\right|\right)$$

$$\times \sup_{\substack{j\in\{1,\ldots,n-p\},r\in\{1,\ldots,M_n\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}}\left|f_z\left(\tilde{x}^+_{r,j,n,s_1,\ldots,s_p}\right)-f_z\left(\frac{x_{r,n}}{1-L/n+2LI\{x_{r,n}<0\}/n}\right)\right|,$$

where $\tilde{x}^+_{r,j,n,s_1,\ldots,s_p}$ lies between $x^+_{r,j,n}$ and $x^+_{r,j,n}+n^{-1/2}\theta^+_{j,n,s_1,\ldots,s_p}$. From (3.28),

$$\sup_{(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}\left|\frac{1}{n}\sum_{j=1}^{n-p}\theta^+_{j,n,s_1,\ldots,s_p}\right| \leq T\sum_{i=1}^{p}\left|\frac{1}{n}\sum_{j=1}^{n-p}\frac{\partial z_j(\boldsymbol{\phi}_0)}{\partial\phi_i}\right| + \frac{2T}{nN_n}\sum_{i=1}^{p}\sum_{j=1}^{n-p}\left|\frac{\partial z_j(\boldsymbol{\phi}_0)}{\partial\phi_i}\right| \xrightarrow{P} 0$$

and

$$\sup_{(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}\frac{1}{n}\sum_{j=1}^{n-p}\left|\theta^+_{j,n,s_1,\ldots,s_p}\right| \leq \frac{T(1+2N_n^{-1})}{n}\sum_{i=1}^{p}\sum_{j=1}^{n-p}\left|\frac{\partial z_j(\boldsymbol{\phi}_0)}{\partial\phi_i}\right| = O_p(1).$$

In addition, because $\mathrm{E}\{z_1^2\} < \infty$, for any $\epsilon > 0$ and $\eta > 0$,

$$\mathrm{P}\left(\sup_{j\in\{1,\ldots,n-p\}} \frac{T(1+2N_n^{-1})}{\sqrt{n}} \sum_{i=1}^{p} \left|\frac{\partial z_j(\phi_0)}{\partial \phi_i}\right| > \frac{\eta}{2}\right) < \epsilon$$

for all n sufficiently large, and, because $C_\delta > 0$ and $0 < D < 1$, there exists an integer m such that

$$\sup_{j\in\{1,\ldots,n-p-m\}} \frac{C_\delta D^{n-p-j}}{|1-L/n|} \le \frac{\eta}{2}$$

for all n sufficiently large. Since, by (3.31),

$$\begin{aligned}\left|\tilde{x}^+_{r,j,n,s_1,\ldots,s_p} - \frac{x_{r,n}}{1-L/n+2LI\{x_{r,n}<0\}/n}\right| &\le \frac{1}{\sqrt{n}}\left|\theta^+_{j,n,s_1,\ldots,s_p}\right| + \frac{CD_\delta^{n-p-j}}{|1-L/n|} \\ &\le \frac{T(1+2N_n^{-1})}{\sqrt{n}} \sum_{i=1}^{p} \left|\frac{\partial z_j(\phi_0)}{\partial \phi_i}\right| + \frac{C_\delta D^{n-p-j}}{|1-L/n|},\end{aligned}$$

we have

$$\mathrm{P}\left(\sup_{\substack{j\in\{1,\ldots,n-p-m\},r\in\{1,\ldots,M_n\}\\ (s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left|\tilde{x}^+_{r,j,n,s_1,\ldots,s_p} - \frac{x_{r,n}}{1-L/n+2LI\{x_{r,n}<0\}/n}\right| > \eta\right) < \epsilon$$

for all n sufficiently large. Hence, because f_z is uniformly continuous on $\mathbb{R}$ and $\epsilon > 0$ and $\eta > 0$ were arbitrarily chosen,

$$\begin{aligned}&\left(\sup_{(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p} \frac{1}{n}\sum_{j=1}^{n-p}\left|\theta^+_{j,n,s_1,\ldots,s_p}\right|\right) \\ &\quad\times \sup_{\substack{j\in\{1,\ldots,n-p\},r\in\{1,\ldots,M_n\}\\ (s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left|f_z\left(\tilde{x}^+_{r,j,n,s_1,\ldots,s_p}\right) - f_z\left(\frac{x_{r,n}}{1-L/n+2LI\{x_{r,n}<0\}/n}\right)\right| \\ &\xrightarrow{P} 0,\end{aligned}$$

and so

$$\sup_{\substack{r\in\{1,\ldots,M_n\}\\ (s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(x^+_{r,j,n} + n^{-1/2}\theta^+_{j,n,s_1,\ldots,s_p}\right) - F_z\left(x^+_{r,j,n}\right)\right]\right| \xrightarrow{P} 0.$$

Consequently,

$$\sup_{\substack{r\in\{1,\ldots,M_n\}\\ (s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(y^+_{r,j,n,s_1,\ldots,s_p}\right) - F_z\left(x^+_{r,j,n}\right)\right]\right| \xrightarrow{P} 0.$$

Similarly, it can be shown that

$$\sup_{\substack{r\in\{0,\ldots,M_n-1\}\\(s_1,\ldots,s_p)'\in\{0,\ldots,N_n-1\}^p}} \left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(x^-_{r,j,n}\right)-F_z\left(y^-_{r,j,n,s_1,\ldots,s_p}\right)\right]\right| \xrightarrow{P} 0.$$

From the proof of Lemma 7, for all n large,

$$\begin{aligned}
&\sup_{r\in\{1,\ldots,M_n\}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(x^+_{r,j,n}\right)-F_z\left(x_{r,n}\right)\right]\right| \\
\leq\ & \frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[\left(\sup_{x\in\mathbb{R}} f_z(x)\right)\frac{C_\delta D^{n-p-j}}{1-L/n}+\frac{1}{n}\left(\sup_{x\in\mathbb{R}}|x|f_z(x)\right)\frac{L}{1-L/n}\right] \\
\to\ & 0,
\end{aligned}$$

and, similarly,

$$\sup_{r\in\{0,\ldots,M_n-1\}}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(x_{r,n}\right)-F_z\left(x^-_{r,j,n}\right)\right]\right| \to 0.$$

Finally, from (3.27), because $M_n \sim n^{3/4}$,

$$\sup_{r\in\{0,\ldots,M_n-1\}}\frac{1}{\sqrt{n}}\sum_{j=1}^{n-p}\left[F_z\left(x_{r+1,n}\right)-F_z\left(x_{r,n}\right)\right]=\frac{n-p}{\sqrt{n}}M_n^{-1}\to 0.$$

Therefore, (3.46) is $o_p(1)$, and so the lemma holds. □

Lemma 10 *For any $T\in(0,\infty)$, as $n\to\infty$,*

$$\sup_{\|\mathbf{u}\|\leq T}\left|n^{-1/2}\sum_{t=1}^{n-p}\mathbf{u}'\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}}\left[F_n\left(z_t\left(\boldsymbol{\phi}_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)-F_z\left(z_t\left(\boldsymbol{\phi}_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)\right]\right|\xrightarrow{P}0.$$

Proof: It suffices to show that

$$\begin{aligned}
&\sup_{\|\mathbf{u}\|\leq T}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\phi_j}\left[F_n\left(z_t\left(\boldsymbol{\phi}_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)-F_z\left(z_t\left(\boldsymbol{\phi}_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)\right]\right| \\
\leq\ & \sup_{\|\mathbf{u}\|\leq T}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\phi_j}\left[F_n\left(z_t\left(\boldsymbol{\phi}_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)-F_z\left(z_t\left(\boldsymbol{\phi}_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)-F_n(z_t)+F_z(z_t)\right]\right| \qquad (3.47)\\
&+\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\phi_j}\left[F_n(z_t)-F_z(z_t)\right]\right| \qquad (3.48)\\
\xrightarrow{P}\ & 0
\end{aligned}$$

for $j \in \{1, \ldots, p\}$.

For any non-negative integer m, (3.47) is bounded above by

$$\sup_{\|\mathbf{u}\| \le T} \left| \frac{1}{\sqrt{n}} \sum_{t=n-p-m}^{n-p} \varphi'(F_z(z_t)) \frac{\partial z_t(\phi_0)}{\partial \phi_j} \left[F_n\left(z_t\left(\phi_0 + \frac{\mathbf{u}}{\sqrt{n}}\right)\right) - F_z\left(z_t\left(\phi_0 + \frac{\mathbf{u}}{\sqrt{n}}\right)\right) - F_n(z_t) + F_z(z_t) \right] \right| \tag{3.49}$$

$$+ \sup_{\|\mathbf{u}\| \le T, t \in \{1, \ldots, n-p-m\}} \sqrt{n} \left| F_n\left(z_t\left(\phi_0 + \frac{\mathbf{u}}{\sqrt{n}}\right)\right) - F_z\left(z_t\left(\phi_0 + \frac{\mathbf{u}}{\sqrt{n}}\right)\right) - F_n(z_t) + F_z(z_t) \right| \tag{3.50}$$

$$\times \frac{1}{n} \sum_{t=1}^{n-p-m} \left| \varphi'(F_z(z_t)) \frac{\partial z_t(\phi_0)}{\partial \phi_j} \right|. \tag{3.51}$$

Given $\epsilon > 0$ and $\eta > 0$, we show that there exists an integer m_0 such that

$$\mathrm{P}\left(\sup_{\|\mathbf{u}\| \le T, t \in \{1, \ldots, n-p-m_0\}} \sqrt{n} \left| F_n\left(z_t\left(\phi_0 + \frac{\mathbf{u}}{\sqrt{n}}\right)\right) - F_z\left(z_t\left(\phi_0 + \frac{\mathbf{u}}{\sqrt{n}}\right)\right) - F_n(z_t) + F_z(z_t) \right| > \eta \right) < \epsilon$$

for all n sufficiently large. Since, for any m, (3.49) is $o_p(1)$ and (3.51) is $O_p(1)$, we conclude that (3.47) is $o_p(1)$.

Because the distribution function F_z is continuous on $\mathbb{R}$, it is also uniformly continuous on $\mathbb{R}$. Therefore, if $W(\cdot)$ denotes a Brownian Bridge on $[0, 1]$, there exists a $\delta > 0$ such that

$$\mathrm{P}\left(\sup_{x \in \mathbb{R}, |y| \le \delta} \left| W(F_z(x+y)) - W(F_z(x)) \right| > \frac{\eta}{2} \right) < \frac{\epsilon}{3}.$$

Also, there exists an integer m_0 such that

$$\begin{aligned}
&\mathrm{P}\left(\sup_{\|\mathbf{u}\| \le T, t \in \{1, \ldots, n-p-m_0\}} \left| z_t\left(\phi_0 + \frac{\mathbf{u}}{\sqrt{n}}\right) - z_t \right| > \delta \right) \\
&\le \mathrm{P}\left(\sup_{t \in \{1, \ldots, n-p-m_0\}} |z_t(\phi_0) - z_t| > \frac{\delta}{3} \right) + \mathrm{P}\left(\sup_{\|\mathbf{u}\| \le T, t \in \{1, \ldots, n-p\}} \left| \frac{\mathbf{u}'}{\sqrt{n}} \frac{\partial z_t(\phi_0)}{\partial \phi} \right| > \frac{\delta}{3} \right) \\
&\quad + \mathrm{P}\left(\sup_{\|\mathbf{u}\| \le T, t \in \{1, \ldots, n-p\}} \left| \frac{\mathbf{u}'}{2n} \frac{\partial^2 z_t(\phi^*_{t,n})}{\partial \phi \partial \phi'} \mathbf{u} \right| > \frac{\delta}{3} \right) \\
&< \frac{\epsilon}{3}
\end{aligned}$$

for all n sufficiently large, where $\phi^*_{t,n}$ lies between ϕ_0 and $\phi_0 + n^{-1/2}\mathbf{u}$. By Theorem 14.3 in Billingsley (1999),

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n-p} \left(I\{F_z(z_k) \le \cdot\} - \cdot \right) \xrightarrow{d} W(\cdot)$$

on $D([0,1])$. Because

$$\sqrt{n}\left[F_n(x) - \frac{n-p}{n}F_z(x)\right] = \frac{1}{\sqrt{n}}\sum_{k=1}^{n-p}\left(I\{F_z(z_k) \le F_z(x)\} - F_z(x)\right),$$

$$\mathrm{P}\left(\sup_{x\in\mathbb{R},|y|\le\delta}\sqrt{n}\left|F_n(x+y) - \frac{n-p}{n}F_z(x+y) - F_n(x) + \frac{n-p}{n}F_z(x)\right| > \frac{\eta}{2}\right)$$
$$\overset{n\to\infty}{\to} \mathrm{P}\left(\sup_{x\in\mathbb{R},|y|\le\delta}\left|W\left(F_z(x+y)\right) - W\left(F_z(x)\right)\right| > \frac{\eta}{2}\right) < \frac{\epsilon}{3},$$

and

$$\sqrt{n}\left|\frac{n-p}{n} - 1\right| \to 0,$$

we have

$$\begin{aligned}
&\mathrm{P}\left(\sup_{\|\mathbf{u}\|\le T, t\in\{1,\ldots,n-p-m_0\}}\sqrt{n}\left|F_n\left(z_t\left(\boldsymbol{\phi}_0 + \frac{\mathbf{u}}{\sqrt{n}}\right)\right) - F_z\left(z_t\left(\boldsymbol{\phi}_0 + \frac{\mathbf{u}}{\sqrt{n}}\right)\right) - F_n(z_t) + F_z(z_t)\right| > \eta\right) \\
\le\;& \mathrm{P}\left(\sup_{x\in\mathbb{R},|y|\le\delta}\sqrt{n}\left|F_n(x+y) - \frac{n-p}{n}F_z(x+y) - F_n(x) + \frac{n-p}{n}F_z(x)\right| > \frac{\eta}{2}\right) \\
&+\mathrm{P}\left(\sup_{\|\mathbf{u}\|\le T, t\in\{1,\ldots,n-p-m_0\}}\left|z_t\left(\boldsymbol{\phi}_0 + \frac{\mathbf{u}}{\sqrt{n}}\right) - z_t\right| > \delta\right) \\
&+\mathrm{P}\left(\sup_{x,y\in\mathbb{R}}\sqrt{n}\left|\frac{n-p}{n}F_z(x+y) - \frac{n-p}{n}F_z(x) - F_z(x+y) + F_z(x)\right| > \frac{\eta}{2}\right) \\
<\;& \epsilon
\end{aligned}$$

for all n sufficiently large.

Turning to (3.48),

$$\begin{aligned}
&\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\phi_j}\left[F_n(z_t) - F_z(z_t)\right]\right| \\
\le\;& \left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\phi_j}\left[\frac{1}{n}\sum_{s=1}^{n-p}I\{z_s\le z_t\} - \frac{n-p}{n}F_z(z_t)\right]\right| \\
&+\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\phi_j}F_z(z_t)\left[\frac{n-p}{n} - 1\right]\right| \\
=\;& \left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))\left\{\frac{-z_{t-j}}{\phi_0(B)} + \frac{z_{t+j}}{\phi_0(B^{-1})}\right\}\left[\frac{1}{n}\sum_{s=1}^{n-p}I\{z_s\le z_t\} - \frac{n-p}{n}F_z(z_t)\right]\right| + o_p(1) \\
=\;& \left|\frac{1}{n^{3/2}}\sum_{s=1}^{n-p}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))\left\{-\sum_{k=0}^{\infty}\psi_k z_{t-j-k} + \sum_{k=0}^{\infty}\psi_k z_{t+j+k}\right\}\left[I\{z_s\le z_t\} - F_z(z_t)\right]\right| + o_p(1),
\end{aligned}$$

where $\sum_{k=0}^{\infty} \psi_k z^k = 1/\phi_0(z)$. For any $k \in \{0,1,\ldots\}$, if

$$Y^k_{q,r,s,t} := \varphi'(F_z(z_r))\varphi'(F_z(z_t))z_{r-j-k}z_{t-j-k}\left[I\{z_q \le z_r\} - F_z(z_r)\right]\left[I\{z_s \le z_t\} - F_z(z_t)\right],$$

then, because the series $\{z_t\}$ is iid with mean zero,

$$\begin{aligned}
&\mathrm{E}\left|\frac{1}{n^{3/2}}\sum_{s=1}^{n-p}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))z_{t-j-k}\left[I\{z_s \le z_t\} - F_z(z_t)\right]\right|^2 \\
&= \frac{1}{n^3}\sum_{q=1}^{n-p}\sum_{r=1}^{n-p}\sum_{s=1}^{n-p}\sum_{t=1}^{n-p}\mathrm{E}\left\{Y^k_{q,r,s,t}\right\} \\
&= \frac{1}{n^3}\sum_{\substack{q=1\ s=1\ t=1 \\ (q\notin\{t,t-j-k\})\cap(s\notin\{t,t-j-k\})\cap(q\neq s)}}^{n-p\ n-p\ n-p}\mathrm{E}\left\{Y^k_{q,t,s,t}\right\} + \frac{1}{n^3}\sum_{\substack{q=1\ s=1\ t=1 \\ (q\in\{t,t-j-k\})\cup(s\in\{t,t-j-k\})\cup(q=s)}}^{n-p\ n-p\ n-p}\mathrm{E}\left(Y^k_{q,t,s,t}\right) \\
&\quad + \frac{1}{n^3}\sum_{\substack{q=1\ r=1\ s=1\ t=1 \\ (q=r-j-k)\cap(s=t-j-k)\cap(r\neq t)}}^{n-p\ n-p\ n-p\ n-p}\mathrm{E}\left\{Y^k_{q,r,s,t}\right\} + \frac{1}{n^3}\sum_{\substack{q=1\ r=1\ s=1\ t=1 \\ (q=t-j-k)\cap(s=r-j-k)\cap(r\neq t)}}^{n-p\ n-p\ n-p\ n-p}\mathrm{E}\left\{Y^k_{q,r,s,t}\right\}. \qquad (3.52)
\end{aligned}$$

Since $\sup_{q,r,s,t}\mathrm{E}|Y^k_{q,r,s,t}| < \infty$ and the number of terms in the three rightmost sums of (3.52) is $O(n^2)$, we have

$$\begin{aligned}
&\mathrm{E}\left|\frac{1}{n^{3/2}}\sum_{s=1}^{n-p}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))z_{t-j-k}\left[I\{z_s \le z_t\} - F_z(z_t)\right]\right|^2 \\
&= \frac{1}{n^3}\sum_{\substack{q=1\ s=1\ t=1 \\ (q\notin\{t,t-j-k\})\cap(s\notin\{t,t-j-k\})\cap(q\neq s)}}^{n-p\ n-p\ n-p}\mathrm{E}\left\{Y^k_{q,t,s,t}\right\} + o(1).
\end{aligned}$$

If $q \notin \{t, t-j-k\}$, $s \notin \{t, t-j-k\}$, and $q \neq s$, then

$$\begin{aligned}
\mathrm{E}\left\{Y^k_{q,t,s,t}\right\} &= \mathrm{E}\{z_1^2\}\mathrm{E}\left\{\left[\varphi'(F_z(z_t))\right]^2\left[I\{z_s \le z_t\} - F_z(z_t)\right]\mathrm{E}\left(I\{z_q \le z_t\} - F_z(z_t)\middle|\sigma(z_s, z_t)\right)\right\} \\
&= \mathrm{E}\{z_1^2\}\mathrm{E}\left\{\left[\varphi'(F_z(z_t))\right]^2\left[I\{z_s \le z_t\} - F_z(z_t)\right]\left[F_z(z_t) - F_z(z_t)\right]\right\} \\
&= 0.
\end{aligned}$$

Thus,

$$\left|\frac{1}{n^{3/2}}\sum_{s=1}^{n-p}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))z_{t-j-k}\left[I\{z_s \le z_t\} - F_z(z_t)\right]\right| \to 0$$

in L_2 and hence in probability. Consequently, for any non-negative integer m,

$$\left|\frac{1}{n^{3/2}}\sum_{s=1}^{n-p}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))\left\{-\sum_{k=0}^{m}\psi_k z_{t-j-k} + \sum_{k=0}^{m}\psi_k z_{t+j+k}\right\}\left[I\{z_s \le z_t\} - F_z(z_t)\right]\right| \xrightarrow{P} 0.$$

Because

$$\left|\frac{1}{n^{3/2}}\sum_{s=1}^{n-p}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))\left\{-\sum_{k=m+1}^{\infty}\psi_k z_{t-j-k}+\sum_{k=m+1}^{\infty}\psi_k z_{t+j+k}\right\}[I\{z_s\le z_t\}-F_z(z_t)]\right|$$
$$\le\left(\sup_{x\in\mathbb{R}}\left|\frac{1}{\sqrt{n}}\sum_{s=1}^{n-p}[I\{z_s\le x\}-F_z(x)]\right|\right)$$
$$\times\left(\frac{1}{n}\sum_{t=1}^{n-p}|\varphi'(F_z(z_t))|\left\{\sum_{k=m+1}^{\infty}|\psi_k z_{t-j-k}|+\sum_{k=m+1}^{\infty}|\psi_k z_{t+j+k}|\right\}\right),$$

$$\sup_{x\in\mathbb{R}}\left|\frac{1}{\sqrt{n}}\sum_{s=1}^{n-p}[I\{z_s\le x\}-F_z(x)]\right|=\sup_{y\in[0,1]}\left|\frac{1}{\sqrt{n}}\sum_{s=1}^{n-p}[I\{F_z(z_s)\le y\}-y]\right|=O_p(1),$$

and $\sum_{k=m+1}^{\infty}|\psi_k|\overset{m\to\infty}{\to}0$,

$$\left|\frac{1}{n^{3/2}}\sum_{s=1}^{n-p}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))\left\{-\sum_{k=0}\psi_k z_{t-j-k}+\sum_{k=0}\psi_k z_{t+j+k}\right\}[I\{z_s\le z_t\}-F_z(z_t)]\right|\overset{P}{\to}0,$$

and thus (3.48) equals $o_p(1)$. □

Lemma 11 *For any $T\in(0,\infty)$, as $n\to\infty$,*

$$\sup_{\|\mathbf{u}\|\le T}\left|n^{-1/2}\sum_{t=1}^{n-p}\mathbf{u}'\left[\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{u}\right)}{n-p+1}\right)-\varphi(F_z(z_t))\right]\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}}-2|\phi_{0r}|^{-1}\tilde{L}\mathbf{u}'\boldsymbol{\Gamma}_p\mathbf{u}\right|\overset{P}{\to}0.$$

Proof: Observe that

$$\sup_{\|\mathbf{u}\|\le T}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\mathbf{u}'\left[\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{u}\right)}{n-p+1}\right)-\varphi(F_z(z_t))\right]\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}}-2|\phi_{0r}|^{-1}\tilde{L}\mathbf{u}'\boldsymbol{\Gamma}_p\mathbf{u}\right|$$
$$\le\sup_{\|\mathbf{u}\|\le T}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\mathbf{u}'\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}}\left[\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{u}\right)}{n-p+1}-F_z(z_t)\right]-2|\phi_{0r}|^{-1}\tilde{L}\mathbf{u}'\boldsymbol{\Gamma}_p\mathbf{u}\right|\quad(3.53)$$
$$+\sup_{\|\mathbf{u}\|\le T}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\mathbf{u}'\left[\varphi'(F^*_{t,n}(\mathbf{u}))-\varphi'(F_z(z_t))\right]\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}}\left[\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{u}\right)}{n-p+1}-F_z(z_t)\right]\right|,\quad(3.54)$$

where $F^*_{t,n}(\mathbf{u})$ is between $F_z(z_t)$ and $R_t(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{u})/(n-p+1)$. An upper bound for (3.53) is

$$\sup_{\|\mathbf{u}\|\le T}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\mathbf{u}'\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}}\left[\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{u}\right)}{n-p+1}-F_n\left(z_t\left(\boldsymbol{\phi}_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)\right]\right|\quad(3.55)$$
$$+\sup_{\|\mathbf{u}\|\le T}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\mathbf{u}'\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}}\left[F_n\left(z_t\left(\boldsymbol{\phi}_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)-F_z\left(z_t\left(\boldsymbol{\phi}_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)\right]\right|\quad(3.56)$$
$$+\sup_{\|\mathbf{u}\|\le T}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\mathbf{u}'\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}}\left[F_z\left(z_t\left(\boldsymbol{\phi}_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)-F_z(z_t)\right]-2|\phi_{0r}|^{-1}\tilde{L}\mathbf{u}'\boldsymbol{\Gamma}_p\mathbf{u}\right|.\quad(3.57)$$

By Lemma 3,

$$\sup_{\|\mathbf{u}\|\le T, t\in\{1,\ldots,n-p\}} \sqrt{n}\left|\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{u}\right)}{n-p+1} - F_n\left(z_t\left(\boldsymbol{\phi}_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)\right| \xrightarrow{P} 0.$$

Thus, because

$$\sup_{\|\mathbf{u}\|\le T} \frac{1}{n}\sum_{t=1}^{n-p}\left|\mathbf{u}'\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial \boldsymbol{\phi}}\right| = O_p(1),$$

(3.55) is $o_p(1)$. By Lemma 10, (3.56) is also $o_p(1)$. Finally,

$$\sup_{\|\mathbf{u}\|\le T}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\mathbf{u}'\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial \boldsymbol{\phi}}\left[F_z\left(z_t\left(\boldsymbol{\phi}_0+\frac{\mathbf{u}}{\sqrt{n}}\right)\right)-F_z(z_t)\right] - 2|\phi_{0r}|^{-1}\tilde{L}\mathbf{u}'\boldsymbol{\Gamma}_p\mathbf{u}\right|$$

$$= \sup_{\|\mathbf{u}\|\le T}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\mathbf{u}'\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial \boldsymbol{\phi}}f_z(z^*_{t,n}(\mathbf{u}))\left(z_t\left(\boldsymbol{\phi}_0+\frac{\mathbf{u}}{\sqrt{n}}\right)-z_t\right) - 2|\phi_{0r}|^{-1}\tilde{L}\mathbf{u}'\boldsymbol{\Gamma}_p\mathbf{u}\right|$$

$$\le \sup_{\|\mathbf{u}\|\le T}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\mathbf{u}'\varphi'(F_z(z_t))\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial \boldsymbol{\phi}}f_z(z^*_{t,n}(\mathbf{u}))(z_t(\boldsymbol{\phi}_0)-z_t)\right|$$

$$+ \sup_{\|\mathbf{u}\|\le T}\left|\frac{1}{n}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))f_z(z^*_{t,n}(\mathbf{u}))\left(\mathbf{u}'\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial \boldsymbol{\phi}}\right)^2 - 2|\phi_{0r}|^{-1}\tilde{L}\mathbf{u}'\boldsymbol{\Gamma}_p\mathbf{u}\right|$$

$$+ \sup_{\|\mathbf{u}\|\le T}\left|\frac{1}{2n\sqrt{n}}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))f_z(z^*_{t,n}(\mathbf{u}))\mathbf{u}'\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial \boldsymbol{\phi}}\mathbf{u}'\frac{\partial^2 z_t(\boldsymbol{\phi}^*_{t,n}(\mathbf{u}))}{\partial \boldsymbol{\phi}\partial \boldsymbol{\phi}'}\mathbf{u}\right|,$$

where $z^*_{t,n}(\mathbf{u})$ is between z_t and $z_t(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{u})$ and $\boldsymbol{\phi}^*_{t,n}(\mathbf{u})$ is between $\boldsymbol{\phi}_0$ and $\boldsymbol{\phi}_0+n^{-1/2}\mathbf{u}$. Because f_z is uniformly continuous on $\mathbb{R}$, from (3.21), the first term on the right-hand side is $o_p(1)$, and, since there exists a geometrically decaying, non-negative sequence $\{\ddot{\pi}_k\}_{k=-\infty}^{\infty}$ such that

$$\sup_{\|\mathbf{u}\|\le T}\left|\mathbf{u}'\frac{\partial^2 z_t(\boldsymbol{\phi}^*_{t,n}(\mathbf{u}))}{\partial \boldsymbol{\phi}\partial \boldsymbol{\phi}'}\mathbf{u}\right| \le \sum_{k=-\infty}^{\infty}\ddot{\pi}_k|z_{t-k}|$$

for all n sufficiently large and all $t\in\{1,\ldots,n-p\}$, the third term is also $o_p(1)$. The middle term equals

$$\sup_{\|\mathbf{u}\|\le T}\left|\frac{1}{n}\sum_{t=1}^{n-p}\varphi'(F_z(z_t))f_z(z_t)\left(\mathbf{u}'\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial \boldsymbol{\phi}}\right)^2 - 2\phi_{0r}^{-2}\left(\int_0^1 f_z(F_z^{-1}(s))\varphi'(s)\,ds\right)\mathbf{u}'\boldsymbol{\Gamma}_p\mathbf{u}\right| + o_p(1),$$

which is $o_p(1)$ by the ergodic theorem. Therefore, (3.57) is $o_p(1)$. Similarly, using the uniform continuity of φ', it can be shown that (3.54) is $o_p(1)$. □

Lemma 12 *For any $T\in(0,\infty)$, as $n\to\infty$,*

$$\sup_{\|\mathbf{u}\|,\|\mathbf{v}\|\le T}\left|n^{-1}\sum_{t=1}^{n-p}\mathbf{u}'\left[\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{v}\right)}{n-p+1}\right)-\varphi(F_z(z_t))\right]\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial \boldsymbol{\phi}\partial \boldsymbol{\phi}'}\mathbf{u}\right| \xrightarrow{P} 0.$$

Proof: Using Lemma 3 and the Glivenko-Cantelli theorem, for any $\epsilon, \eta > 0$, it can be shown that there exists an integer m such that

$$\mathrm{P}\left(\sup_{\|\mathbf{v}\|\leq T, t\in\{1,\ldots,n-p-m\}}\left|\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{v}\right)}{n-p+1}-F_z(z_t)\right|>\eta\right)<\epsilon$$

for all n sufficiently large. Hence,

$$\begin{aligned}
&\sup_{\|\mathbf{u}\|,\|\mathbf{v}\|\leq T}\left|\frac{1}{n}\sum_{t=1}^{n-p}\mathbf{u}'\left[\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{v}\right)}{n-p+1}\right)-\varphi(F_z(z_t))\right]\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\mathbf{u}\right| \\
&= \sup_{\|\mathbf{u}\|,\|\mathbf{v}\|\leq T}\left|\frac{1}{n}\sum_{t=1}^{n-p}\mathbf{u}'\varphi'(F^*_{t,n}(\mathbf{v}))\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\left[\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{v}\right)}{n-p+1}-F_z(z_t)\right]\mathbf{u}\right| \\
&\xrightarrow{P} 0,
\end{aligned}$$

where $F^*_{t,n}(\mathbf{v})$ is between $F_z(z_t)$ and $R_t(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{v})/(n-p+1)$, since

$$\sup_{\|\mathbf{u}\|,\|\mathbf{v}\|\leq T}\frac{1}{n}\sum_{t=1}^{n-p}\left|\mathbf{u}'\varphi'(F^*_{t,n}(\mathbf{v}))\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\mathbf{u}\right|=O_p(1).$$

□

For $\mathbf{u}\in\mathbb{R}^p$ and $\delta_1,\delta_2\in[0,1]$, let

$$U_n(\mathbf{u},\delta_1,\delta_2)=\sum_{t=1}^{n-p}\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_1\mathbf{u}\right)}{n-p+1}\right)\left[z_t\left(\boldsymbol{\phi}_0+\frac{\delta_2\mathbf{u}}{\sqrt{n}}\right)-z_t\left(\boldsymbol{\phi}_0+\frac{\delta_1\mathbf{u}}{\sqrt{n}}\right)\right]$$

and

$$V_n(\mathbf{u},\delta_1,\delta_2)=\sum_{t=1}^{n-p}\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_2\mathbf{u}\right)}{n-p+1}\right)\left[z_t\left(\boldsymbol{\phi}_0+\frac{\delta_2\mathbf{u}}{\sqrt{n}}\right)-z_t\left(\boldsymbol{\phi}_0+\frac{\delta_1\mathbf{u}}{\sqrt{n}}\right)\right].$$

By Taylor series expansions of $z_t(\cdot)$,

$$\begin{aligned}
U_n(\mathbf{u},\delta_1,\delta_2) &= \sum_{t=1}^{n-p}\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_1\mathbf{u}\right)}{n-p+1}\right)\left[z_t\left(\boldsymbol{\phi}_0+\frac{\delta_2\mathbf{u}}{\sqrt{n}}\right)-z_t(\boldsymbol{\phi}_0)\right] \\
&\quad -\sum_{t=1}^{n-p}\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_1\mathbf{u}\right)}{n-p+1}\right)\left[z_t\left(\boldsymbol{\phi}_0+\frac{\delta_1\mathbf{u}}{\sqrt{n}}\right)-z_t(\boldsymbol{\phi}_0)\right] \\
&= \frac{\delta_2-\delta_1}{\sqrt{n}}\sum_{t=1}^{n-p}\mathbf{u}'\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_1\mathbf{u}\right)}{n-p+1}\right)\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}} \\
&\quad +\frac{1}{2}\frac{\delta_2^2-\delta_1^2}{n}\sum_{t=1}^{n-p}\mathbf{u}'\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_1\mathbf{u}\right)}{n-p+1}\right)\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\mathbf{u}
\end{aligned}$$

$$+\frac{1}{2}\frac{\delta_2^2}{n}\sum_{t=1}^{n-p}\mathbf{u}'\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_1\mathbf{u}\right)}{n-p+1}\right)\left[\frac{\partial^2 z_t(\boldsymbol{\phi}_n^*(\mathbf{u},\delta_1,\delta_2))}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}-\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\right]\mathbf{u}$$
$$-\frac{1}{2}\frac{\delta_1^2}{n}\sum_{t=1}^{n-p}\mathbf{u}'\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_1\mathbf{u}\right)}{n-p+1}\right)\left[\frac{\partial^2 z_t(\boldsymbol{\phi}_n^*(\mathbf{u},\delta_1,\delta_1))}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}-\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\right]\mathbf{u} \tag{3.58}$$

and

$$\begin{aligned}
V_n(\mathbf{u},\delta_1,\delta_2) &= \sum_{t=1}^{n-p}\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_2\mathbf{u}\right)}{n-p+1}\right)\left[z_t\left(\boldsymbol{\phi}_0+\frac{\delta_2\mathbf{u}}{\sqrt{n}}\right)-z_t(\boldsymbol{\phi}_0)\right]\\
&\quad-\sum_{t=1}^{n-p}\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_2\mathbf{u}\right)}{n-p+1}\right)\left[z_t\left(\boldsymbol{\phi}_0+\frac{\delta_1\mathbf{u}}{\sqrt{n}}\right)-z_t(\boldsymbol{\phi}_0)\right]\\
&= \frac{\delta_2-\delta_1}{\sqrt{n}}\sum_{t=1}^{n-p}\mathbf{u}'\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_2\mathbf{u}\right)}{n-p+1}\right)\frac{\partial z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}}\\
&\quad+\frac{1}{2}\frac{\delta_2^2-\delta_1^2}{n}\sum_{t=1}^{n-p}\mathbf{u}'\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_2\mathbf{u}\right)}{n-p+1}\right)\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\mathbf{u}\\
&\quad+\frac{1}{2}\frac{\delta_2^2}{n}\sum_{t=1}^{n-p}\mathbf{u}'\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_2\mathbf{u}\right)}{n-p+1}\right)\left[\frac{\partial^2 z_t(\boldsymbol{\phi}_n^*(\mathbf{u},\delta_2,\delta_2))}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}-\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\right]\mathbf{u}\\
&\quad-\frac{1}{2}\frac{\delta_1^2}{n}\sum_{t=1}^{n-p}\mathbf{u}'\varphi\left(\frac{R_t\left(\boldsymbol{\phi}_0+n^{-1/2}\delta_2\mathbf{u}\right)}{n-p+1}\right)\left[\frac{\partial^2 z_t(\boldsymbol{\phi}_n^*(\mathbf{u},\delta_2,\delta_1))}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}-\frac{\partial^2 z_t(\boldsymbol{\phi}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\right]\mathbf{u},
\end{aligned} \tag{3.59}$$

where values of $\boldsymbol{\phi}_n^*(\mathbf{u},\cdot,\cdot)$ lie between $\boldsymbol{\phi}_0$ and $\boldsymbol{\phi}_0+n^{-1/2}\mathbf{u}$.

Lemma 13 *For* $\mathbf{u}\in\mathbb{R}^p$, *let*

$$S_n(\mathbf{u})=D\left(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{u}\right)-D(\boldsymbol{\phi}_0)$$

and

$$S(\mathbf{u})=\mathbf{u}'\mathbf{N}+|\phi_{0r}|^{-1}(\sigma^2\tilde{L}-\tilde{K})\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u},$$

where $\mathbf{N}\sim N\left(\mathbf{0},2\phi_{0r}^{-2}\{\sigma^2\tilde{J}-\tilde{K}^2\}\sigma^{-2}\boldsymbol{\Gamma}_p\right)$. *Then* $S_n(\cdot)\xrightarrow{d}S(\cdot)$ *on* $C(\mathbb{R}^p)$, *the space of continuous functions on* $\mathbb{R}^p$ *where convergence is equivalent to uniform convergence on every compact set.*

Proof: Let $\mathbf{u}\in\mathbb{R}^p$ and suppose m is any positive integer. Because

$$D\left(\boldsymbol{\phi}_0+n^{-1/2}\mathbf{u}\right)-D(\boldsymbol{\phi}_0)=\sum_{k=1}^{m}\left[D\left(\boldsymbol{\phi}_0+\frac{k\mathbf{u}}{m\sqrt{n}}\right)-D\left(\boldsymbol{\phi}_0+\frac{(k-1)\mathbf{u}}{m\sqrt{n}}\right)\right],$$

we have

$$\sum_{k=1}^{m} U_n\left(\mathbf{u}, \frac{k-1}{m}, \frac{k}{m}\right) \le D\left(\boldsymbol{\phi}_0 + n^{-1/2}\mathbf{u}\right) - D(\boldsymbol{\phi}_0) \le \sum_{k=1}^{m} V_n\left(\mathbf{u}, \frac{k-1}{m}, \frac{k}{m}\right) \tag{3.60}$$

by Theorem 1. From (3.58), (3.59), Lemmas 1, 2, 11, and 12, and because $\partial^2 z_t(\boldsymbol{\phi})/(\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}')$ is continuous with respect to $\boldsymbol{\phi}$,

$$\begin{bmatrix} U_n\left(\mathbf{u}, 0, \frac{1}{m}\right) \\ U_n\left(\mathbf{u}, \frac{1}{m}, \frac{2}{m}\right) \\ \vdots \\ U_n\left(\mathbf{u}, \frac{m-1}{m}, 1\right) \\ V_n\left(\mathbf{u}, 0, \frac{1}{m}\right) \\ V_n\left(\mathbf{u}, \frac{1}{m}, \frac{2}{m}\right) \\ \vdots \\ V_n\left(\mathbf{u}, \frac{m-1}{m}, 1\right) \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \frac{1}{m}\mathbf{u}'\mathbf{N} - |\phi_{0r}|^{-1}\left[\left(\frac{1}{m}\right)^2 - \left(\frac{0}{m}\right)^2\right]\tilde{K}\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u} \\ \frac{1}{m}\mathbf{u}'\mathbf{N} + |\phi_{0r}|^{-1}\left(2\frac{1}{m^2}\sigma^2\tilde{L} - \left[\left(\frac{2}{m}\right)^2 - \left(\frac{1}{m}\right)^2\right]\tilde{K}\right)\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u} \\ \vdots \\ \frac{1}{m}\mathbf{u}'\mathbf{N} + |\phi_{0r}|^{-1}\left(2\frac{m-1}{m^2}\sigma^2\tilde{L} - \left[\left(\frac{m}{m}\right)^2 - \left(\frac{m-1}{m}\right)^2\right]\tilde{K}\right)\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u} \\ \frac{1}{m}\mathbf{u}'\mathbf{N} + |\phi_{0r}|^{-1}\left(2\frac{1}{m^2}\sigma^2\tilde{L} - \left[\left(\frac{1}{m}\right)^2 - \left(\frac{0}{m}\right)^2\right]\tilde{K}\right)\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u} \\ \frac{1}{m}\mathbf{u}'\mathbf{N} + |\phi_{0r}|^{-1}\left(2\frac{2}{m^2}\sigma^2\tilde{L} - \left[\left(\frac{2}{m}\right)^2 + \left(\frac{1}{m}\right)^2\right]\tilde{K}\right)\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u} \\ \vdots \\ \frac{1}{m}\mathbf{u}'\mathbf{N} + |\phi_{0r}|^{-1}\left(2\frac{m}{m^2}\sigma^2\tilde{L} - \left[\left(\frac{m}{m}\right)^2 - \left(\frac{m-1}{m}\right)^2\right]\tilde{K}\right)\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u} \end{bmatrix}$$

on $\mathbb{R}^{2m}$, and so

$$\begin{bmatrix} \sum_{k=1}^{m} U_n\left(\mathbf{u}, \frac{k-1}{m}, \frac{k}{m}\right) \\ \sum_{k=1}^{m} V_n\left(\mathbf{u}, \frac{k-1}{m}, \frac{k}{m}\right) \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \mathbf{u}'\mathbf{N} + |\phi_{0r}|^{-1}\left(\frac{m-1}{m}\sigma^2\tilde{L} - \tilde{K}\right)\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u} \\ \mathbf{u}'\mathbf{N} + |\phi_{0r}|^{-1}\left(\frac{m+1}{m}\sigma^2\tilde{L} - \tilde{K}\right)\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u} \end{bmatrix}$$

on $\mathbb{R}^2$. For any $\epsilon > 0$, there exists an integer m such that

$$\mathbf{u}'\mathbf{N} + |\phi_{0r}|^{-1}\left(\frac{m-1}{m}\sigma^2\tilde{L} - \tilde{K}\right)\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u} \quad \text{and} \quad \mathbf{u}'\mathbf{N} + |\phi_{0r}|^{-1}\left(\frac{m+1}{m}\sigma^2\tilde{L} - \tilde{K}\right)\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u}$$

are both in an ϵ-neighborhood of $\mathbf{u}'\mathbf{N} + |\phi_{0r}|^{-1}(\sigma^2\tilde{L} - \tilde{K})\mathbf{u}'\sigma^{-2}\boldsymbol{\Gamma}_p\mathbf{u}$. Thus, by (3.60), for any $\mathbf{u} \in \mathbb{R}^p$, $S_n(\mathbf{u}) \xrightarrow{d} S(\mathbf{u})$. Similarly, it can be shown that all finite-dimensional distributions of $S_n(\cdot)$ converge to those of $S(\cdot)$.

Let $\eta > 0$ and suppose K is any compact subset of $\mathbb{R}^p$. We now show that

$$\lim_{\delta\to 0^+} \limsup_{n\to\infty} \mathrm{P}\left(\sup_{\mathbf{u},\mathbf{v}\in K, \|\mathbf{u}-\mathbf{v}\|\le\delta} |S_n(\mathbf{u}) - S_n(\mathbf{v})| > \eta\right) = 0.$$

From (3.60), for any positive integer m,

$$\mathrm{P}\left(\sup_{\mathbf{u},\mathbf{v}\in K, \|\mathbf{u}-\mathbf{v}\|\le\delta} |S_n(\mathbf{u}) - S_n(\mathbf{v})| > \eta\right)$$

$$\leq \mathrm{P}\left(\sup_{\mathbf{u},\mathbf{v}\in K,\|\mathbf{u}-\mathbf{v}\|\leq\delta}\left|\sum_{k=1}^{m}\left[U_n\left(\mathbf{u},\frac{k-1}{m},\frac{k}{m}\right)-V_n\left(\mathbf{v},\frac{k-1}{m},\frac{k}{m}\right)\right]\right|>\eta\right),$$

and, from (3.58) and (3.59), for any $\mathbf{u},\mathbf{v}\in K$,

$$\left|\sum_{k=1}^{m}\left[U_n\left(\mathbf{u},\frac{k-1}{m},\frac{k}{m}\right)-V_n\left(\mathbf{v},\frac{k-1}{m},\frac{k}{m}\right)\right]\right|$$

$$\leq \left|\frac{(\mathbf{u}-\mathbf{v})'}{\sqrt{n}}\sum_{t=1}^{n-p}\varphi\left(F_z(z_t)\right)\frac{\partial z_t(\phi_0)}{\partial\phi}\right| \tag{3.61}$$

$$+|\phi_{0r}|^{-1}\tilde{L}\left|\frac{m-1}{m}\mathbf{u}'\mathbf{\Gamma}_p\mathbf{u}-\frac{m+1}{m}\mathbf{v}'\mathbf{\Gamma}_p\mathbf{v}\right| \tag{3.62}$$

$$+\left|\frac{\mathbf{u}'}{n}\sum_{t=1}^{n-p}\varphi\left(F_z(z_t)\right)\frac{\partial^2 z_t(\phi_0)}{\partial\phi\partial\phi'}\mathbf{u}-\frac{\mathbf{v}'}{n}\sum_{t=1}^{n-p}\varphi\left(F_z(z_t)\right)\frac{\partial^2 z_t(\phi_0)}{\partial\phi\partial\phi'}\mathbf{v}\right| \tag{3.63}$$

$$+2m\sup_{\mathbf{u}\in K}\left|\frac{\mathbf{u}'}{\sqrt{n}}\sum_{t=1}^{n-p}\left[\varphi\left(\frac{R_t(\phi_0+n^{-1/2}\mathbf{u})}{n-p+1}\right)-\varphi\left(F_z(z_t)\right)\right]\frac{\partial z_t(\phi_0)}{\partial\phi}-2|\phi_{0r}|^{-1}\tilde{L}\mathbf{u}'\mathbf{\Gamma}_p\mathbf{u}\right| \tag{3.64}$$

$$+\sup_{\mathbf{u}\in K}\left|\frac{\mathbf{u}'}{\sqrt{n}}\sum_{t=1}^{n-p}\left[\varphi\left(\frac{R_t(\phi_0)}{n-p+1}\right)-\varphi\left(F_z(z_t)\right)\right]\frac{\partial z_t(\phi_0)}{\partial\phi}\right| \tag{3.65}$$

$$+2m\sup_{\mathbf{u},\mathbf{v}\in K}\left|\frac{\mathbf{u}'}{n}\sum_{t=1}^{n-p}\left[\varphi\left(\frac{R_t\left(\phi_0+n^{-1/2}\mathbf{v}\right)}{n-p+1}\right)-\varphi\left(F_z(z_t)\right)\right]\frac{\partial^2 z_t(\phi_0)}{\partial\phi\partial\phi'}\mathbf{u}\right| \tag{3.66}$$

$$+4m\sup_{\mathbf{u},\mathbf{v},\mathbf{w}\in K}\left|\frac{\mathbf{u}'}{n}\sum_{t=1}^{n-p}\varphi\left(\frac{R_t\left(\phi_0+n^{-1/2}\mathbf{v}\right)}{n-p+1}\right)\left[\frac{\partial^2 z_t\left(\phi_0+n^{-1/2}\mathbf{w}\right)}{\partial\phi\partial\phi'}-\frac{\partial^2 z_t\left(\phi_0\right)}{\partial\phi\partial\phi'}\right]\mathbf{u}\right|. \tag{3.67}$$

Choose a $\delta_0>0$ sufficiently small and choose a positive integer m_0 sufficiently large so that, when $m=m_0$, (3.62) is less than $\eta/7$ for all $\mathbf{u},\mathbf{v}\in K$ such that $\|\mathbf{u}-\mathbf{v}\|\leq\delta_0$. By Lemmas 11 and 12 and the continuity of $\partial^2 z_t(\phi)/(\partial\phi\partial\phi')$, (3.64)-(3.67) are $o_p(1)$ when $m=m_0$. Also, because

$$\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-p}\varphi\left(F_z(z_t)\right)\frac{\partial z_t(\phi_0)}{\partial\phi}\right|=O_p(1)\qquad\text{and}\qquad\left|\frac{1}{n}\sum_{t=1}^{n-p}\varphi\left(F_z(z_t)\right)\frac{\partial^2 z_t(\phi_0)}{\partial\phi\partial\phi'}\right|=O_p(1)$$

by Lemmas 1 and 2,

$$\lim_{\delta\to0^+}\limsup_{n\to\infty}\mathrm{P}\left(\sup_{\mathbf{u},\mathbf{v}\in K,\|\mathbf{u}-\mathbf{v}\|\leq\delta}\left|\frac{(\mathbf{u}-\mathbf{v})'}{\sqrt{n}}\sum_{t=1}^{n-p}\varphi\left(F_z(z_t)\right)\frac{\partial z_t(\phi_0)}{\partial\phi}\right|>\frac{\eta}{7}\right)=0$$

and

$$\lim_{\delta\to0^+}\limsup_{n\to\infty}\mathrm{P}\left(\sup_{\mathbf{u},\mathbf{v}\in K,\|\mathbf{u}-\mathbf{v}\|\leq\delta}\left|\frac{\mathbf{u}'}{n}\sum_{t=1}^{n-p}\varphi\left(F_z(z_t)\right)\frac{\partial^2 z_t(\phi_0)}{\partial\phi\partial\phi'}\mathbf{u}-\frac{\mathbf{v}'}{n}\sum_{t=1}^{n-p}\varphi\left(F_z(z_t)\right)\frac{\partial^2 z_t(\phi_0)}{\partial\phi\partial\phi'}\mathbf{v}\right|>\frac{\eta}{7}\right)=0.$$

It follows that

$$\lim_{\delta\to0^+}\limsup_{n\to\infty}\mathrm{P}\left(\sup_{\mathbf{u},\mathbf{v}\in K,\|\mathbf{u}-\mathbf{v}\|\leq\delta}|S_n(\mathbf{u})-S_n(\mathbf{v})|>\eta\right)$$

$$\leq \lim_{\delta\to 0+} \limsup_{n\to\infty} \mathrm{P}\left(\sup_{\mathbf{u},\mathbf{v}\in K, \|\mathbf{u}-\mathbf{v}\|\leq\delta} \left| \sum_{k=1}^{m_0} \left[U_n\left(\mathbf{u}, \frac{k-1}{m_0}, \frac{k}{m_0}\right) - V_n\left(\mathbf{v}, \frac{k-1}{m_0}, \frac{k}{m_0}\right)\right] \right| > \eta \right)$$
$$= 0,$$

and so $S_n(\cdot)$ must be tight on $C(K)$ for any compact set $K \subset \mathbb{R}^p$. Therefore, $S_n(\cdot) \xrightarrow{d} S(\cdot)$ on $C(\mathbb{R}^p)$ by Theorem 7.1 in Billingsley (1999). □

Lemma 14 *There exists a sequence of minimizers $\hat{\phi}_R$ of $D(\cdot)$ such that*

$$n^{1/2}(\hat{\phi}_R - \phi_0) \xrightarrow{d} \mathbf{Y} \sim N(\mathbf{0}, \mathbf{\Sigma}),$$

where

$$\mathbf{\Sigma} = \frac{\sigma^2 \tilde{J} - \tilde{K}^2}{2(\sigma^2 \tilde{L} - \tilde{K})^2} \sigma^2 \mathbf{\Gamma}_p^{-1}.$$

Proof: $D(\phi) - D(\phi_0) = S_n(\sqrt{n}(\phi - \phi_0))$, where $S_n(\cdot)$ is defined in Lemma 13. Because $\mathbf{Y} := -|\phi_{0r}|\sigma^2 \mathbf{\Gamma}_p^{-1}\mathbf{N}/[2(\sigma^2\tilde{L} - \tilde{K})]$ minimizes the limit process $S(\cdot)$ in Lemma 13, the result follows by Remark 1 in Davis, Knight, and Liu (1992). □

Lemma 15 *If $\epsilon > 0$ is sufficiently small so that ϕ forms a causal polynomial for all $\phi \in \mathbf{\Phi} := \{\phi \in \mathbb{R}^p : \|\phi - \phi_0\| \leq \epsilon\}$, then*

$$n^{-1}\sum_{t=1}^{n-p} z_t^2(\phi) \xrightarrow{a.s.} E\{\tilde{z}_1^2(\phi)\} \qquad \text{and} \qquad n^{-1}D(\phi) \xrightarrow{a.s.} \int_0^1 F_{\tilde{z}(\phi)}^{-1}(s)\varphi(s)\,ds$$

uniformly on $\mathbf{\Phi}$, where $\tilde{z}_t(\phi) := -\phi^{-1}(B^{-1})\phi(B)X_{t+p}$ and $F_{\tilde{z}(\phi)}(\cdot)$ is the distribution function for $\tilde{z}_1(\phi)$.

Proof: For any $\phi \in \mathbf{\Phi}$,

$$n^{-1}\sum_{t=1}^{n-p} z_t^2(\phi) \xrightarrow{a.s.} \mathrm{E}\{\tilde{z}_1^2(\phi)\}$$

and

$$n^{-1}D(\phi) \xrightarrow{a.s.} \mathrm{E}\left\{\varphi\left(F_{\tilde{z}(\phi)}(\tilde{z}_1(\phi))\right)\tilde{z}_1(\phi)\right\} = \int_0^1 F_{\tilde{z}(\phi)}^{-1}(s)\varphi(s)\,ds$$

by the ergodic theorem. Consequently, the lemma follows by the Arzela-Ascoli theorem if $n^{-1}\sum_{t=1}^{n-p} z_t^2(\cdot)$ and $n^{-1}D(\cdot)$ are equicontinuous and uniformly bounded on $\mathbf{\Phi}$ almost surely.

If $\phi_1, \phi_2 \in \mathbf{\Phi}$, then

$$\left| n^{-1}\sum_{t=1}^{n-p}\left(z_t^2(\phi_1) - z_t^2(\phi_2)\right)\right| \le \|\phi_1 - \phi_2\| n^{-1}\sum_{t=1}^{n-p}\left\| 2z_t(\ddot{\phi}_{t,n})\frac{\partial z_t(\ddot{\phi}_{t,n})}{\partial \phi}\right\|$$

and, by Theorem 1,

$$\begin{aligned}
&\left| n^{-1}D(\phi_1) - n^{-1}D(\phi_2)\right| \\
&= n^{-1}\max\left\{\pm\sum_{t=1}^{n-p}\left[\varphi\left(\frac{R_t(\phi_1)}{n-p+1}\right)z_t(\phi_1) - \varphi\left(\frac{R_t(\phi_2)}{n-p+1}\right)z_t(\phi_2)\right]\right\} \\
&\le n^{-1}\max\left\{\sum_{t=1}^{n-p}\varphi\left(\frac{R_t(\phi_1)}{n-p+1}\right)[z_t(\phi_1) - z_t(\phi_2)], \sum_{t=1}^{n-p}\varphi\left(\frac{R_t(\phi_2)}{n-p+1}\right)[z_t(\phi_2) - z_t(\phi_1)]\right\} \\
&\le \|\phi_1 - \phi_2\|\left(\sup_{s\in(0,1)}|\varphi(s)|\right) n^{-1}\sum_{t=1}^{n-p}\left\|\frac{\partial z_t(\tilde{\phi}_{t,n})}{\partial \phi}\right\|,
\end{aligned}$$

where $\ddot{\phi}_{t,n}$ and $\tilde{\phi}_{t,n}$ lie on the line segment connecting ϕ_1 and ϕ_2. Because φ' is uniformly continuous on $(0,1)$, $\sup_{s\in(0,1)}|\varphi(s)| < \infty$. Also, because $z_t(\phi)$ and $\partial z_t(\phi)/\partial\phi$ are continuous with respect to ϕ on $\mathbf{\Phi}$, there exist coefficients $\tilde{\pi}_k \ge 0$, $k \in \{\ldots,-1,0,1,\ldots\}$, decaying at a geometric rate such that

$$\sup_{\phi\in\mathbf{\Phi}}|z_t(\phi)| \le \sum_{k=-\infty}^{\infty}\tilde{\pi}_k|z_{t-k}| \quad \forall t \in \{1,\ldots,n-p\}$$

and

$$\sup_{\phi\in\mathbf{\Phi}}\left|\frac{\partial z_t(\phi)}{\partial \phi_j}\right| \le \sum_{k=-\infty}^{\infty}\tilde{\pi}_k|z_{t-k}| \quad \forall j \in \{1,\ldots,p\}, \forall t \in \{1,\ldots,n-p\}$$

for all n. Thus, almost surely, for all n sufficiently large,

$$n^{-1}\sum_{t=1}^{n-p}\left\|2z_t(\ddot{\phi}_{t,n})\frac{\partial z_t(\ddot{\phi}_{t,n})}{\partial \phi}\right\| \le constant$$

and

$$\left(\sup_{s\in(0,1)}|\varphi(s)|\right) n^{-1}\sum_{t=1}^{n-p}\left\|\frac{\partial z_t(\tilde{\phi}_{t,n})}{\partial \phi}\right\| \le constant,$$

and so $n^{-1}\sum_{t=1}^{n-p} z_t^2(\cdot)$ and $n^{-1}D(\cdot)$ are equicontinuous on $\mathbf{\Phi}$ almost surely. It can be shown similarly that $n^{-1}\sum_{t=1}^{n-p} z_t^2(\cdot)$ and $n^{-1}D(\cdot)$ are uniformly bounded on $\mathbf{\Phi}$ almost surely. □

Lemma 16 *If κ is a uniformly continuous, differentiable kernel density function on $I\!R$ such that $\int |s\ln|s||^{1/2}|\kappa'(s)|\,ds < \infty$ and κ' is uniformly continuous on $I\!R$, and the bandwidth b_n is chosen so that $b_n \xrightarrow{P} 0$ and $b_n^2\sqrt{n} \xrightarrow{P} \infty$ as $n \to \infty$, then*

$$\sup_{s\in I\!R}\left|\hat{f}_n(s) - f_z(s)\right| \xrightarrow{P} 0,$$

where the kernel density estimator $\hat{f}_n(\cdot)$ is defined in (3.13).

Proof: If

$$f_n(s) := \frac{1}{b_n n}\sum_{t=1}^{n-p}\kappa\left(\frac{s-z_t}{b_n}\right),$$

then

$$\sup_{s\in\mathbb{R}}\left|f_n(s) - f_z(s)\right| \xrightarrow{P} 0$$

by Theorem A in Silverman (1978). Hence, the proof is complete if

$$\begin{aligned}
\sup_{s\in\mathbb{R}}|\hat{f}_n(s) - f_n(s)| &\leq \sup_{s\in\mathbb{R}}\left|\frac{1}{b_n n}\sum_{t=1}^{n-p}\left[\kappa\left(\frac{s-z_t(\hat{\phi}_R)}{b_n}\right) - \kappa\left(\frac{s-z_t(\phi_0)}{b_n}\right)\right]\right| \\
&\quad + \sup_{s\in\mathbb{R}}\left|\frac{1}{b_n n}\sum_{t=1}^{n-p}\left[\kappa\left(\frac{s-z_t(\phi_0)}{b_n}\right) - \kappa\left(\frac{s-z_t}{b_n}\right)\right]\right| \\
&= \sup_{s\in\mathbb{R}}\left|\frac{(\hat{\phi}_R-\phi_0)'}{b_n^2 n}\sum_{t=1}^{n-p}\frac{\partial z_t(\phi_n^*)}{\partial\phi}\kappa'\left(\frac{s-z_t(\phi_n^*)}{b_n}\right)\right| \qquad (3.68)\\
&\quad + \sup_{s\in\mathbb{R}}\left|\frac{1}{b_n^2 n}\sum_{t=1}^{n-p}(z_t - z_t(\phi_0))\kappa'\left(\frac{s-z_{t,n}^*}{b_n}\right)\right| \qquad (3.69)\\
&\xrightarrow{P} 0,
\end{aligned}$$

where ϕ_n^* is between ϕ_0 and $\hat{\phi}_R$ and $z_{t,n}^*$ is between z_t and $z_t(\phi_0)$. Because $\sqrt{n}(\hat{\phi}_R - \phi_0) = O_p(1)$, $b_n^2\sqrt{n} \xrightarrow{P} \infty$, $\sup_{s\in\mathbb{R}}|\kappa'(s)| < \infty$ by the uniform continuity of κ', and

$$\frac{1}{n}\sum_{t=1}^{n-p}\left|\frac{\partial z_t(\phi_n^*)}{\partial\phi}\right| = O_p(1),$$

it follows that (3.68) is $o_p(1)$. By (3.21), (3.69) is also $o_p(1)$. □

Chapter 4

Least Absolute Deviations and Maximum Likelihood Estimation for All-Pass Time Series Processes with Infinite Variance

4.1 Introduction

All-pass models are autoregressive-moving average (ARMA) models in which the roots of the autoregressive polynomial are reciprocals of roots of the moving average polynomial and vice versa. These models generate time series that are dependent in the non-Gaussian case. As discussed in Breidt, Davis, and Trindade (2001), an all-pass series can be obtained by fitting a causal autoregressive model (all the roots of the autoregressive polynomial are outside the unit circle) to a series generated by a noncausal autoregressive model. The residuals follow an all-pass model of order r, where r is the number of roots of the true autoregressive polynomial inside the unit circle. Calder (1998) uses infinite variance, noncausal autoregressive models in the analysis of stock market trading volume data. Additional applications for noncausal models can be found in Chapter 1. Similarly, an all-pass series can be obtained by fitting a causal, invertible ARMA model (all the roots of the moving average polynomial are outside the unit circle) to a series generated by a causal, noninvertible ARMA model (at least one root of the moving average polynomial is inside the unit circle) (see

Chapter 2). Noninvertible ARMA models have been used, for example, in vocal tract filters (Chi and Kung, 1995; Chien, Yang, and Chi, 1997), in the analysis of unemployment rates (Huang and Pawitan, 2000), and in seismogram deconvolution (Lii and Rosenblatt, 1988; Chapter 2).

For all-pass processes with finite variance, cumulant-based estimators, using cumulants of order greater than two, are often used for parameter estimation (Giannakis and Swami, 1990; Chi and Kung, 1995; Chien, Yang, and Chi, 1997). Breidt, Davis, and Trindade (2001) consider a least absolute deviations (LAD) approach motivated by approximating the likelihood of an all-pass model with Laplace (two-sided exponential) noise and a general maximum likelihood (ML) approach is considered in Chapter 2. Also, in Chapter 3, all-pass model parameters are estimated by minimizing the sum of model residuals weighted by a function of residual rank. For all-pass series with finite variance, the LAD, ML, and rank estimators are asymptotically normal under general conditions.

In this chapter, we consider LAD and ML estimation for all-pass processes with infinite variance. In particular, we assume the noise distribution belongs to the domain of attraction of a stable distribution with exponent $\alpha \in (0,2)$. These all-pass series tend to be very "spiky" due to the appearance of observations unusually large in absolute value in the noise sequence, and so they are often called *heavy-tailed.* For the case when the noise distribution is in the domain of attraction of a stable distribution with $\alpha \in (0,2)$, Davis, Knight, and Liu (1992) examine M-estimation, which includes LAD estimation, for causal autoregressive processes and Davis (1996) examines M-estimation for causal, invertible ARMA processes. General ARMA parameter estimation results are not applicable to all-pass models because the autoregressive and moving average parameters are dependent. In Davis (1996), LAD estimation is compared to least squares estimation and it is found that LAD estimation is more efficient. Because LAD estimation appears to be useful for ARMA processes with infinite variance, we study this type of estimator for all-pass series with infinite variance. For ML estimation, stronger restrictions are required for the noise distribution. We assume that it is stable with exponent $\alpha \in (0,2)$, and not simply in the domain of attraction of a stable distribution. A related likelihood approach is considered in Calder (1998) for noncausal autoregressive processes.

In Section 4.2, we introduce all-pass models and discuss stable distributions and distributions in domains of attraction of stable distributions. Under general conditions, we establish consistency for LAD and ML estimators and show that these estimators have nondegenerate limiting distributions in Section 4.3. Proofs of the lemmas used to establish the results of Section 4.3 can be found in Section 4.5. The behavior of LAD estimators for finite samples is studied via simulation in Section 4.4.1, and the estimation procedure is used to fit a noncausal autoregressive model to stock market trading volume data in Section 4.4.2.

4.2 Preliminaries

Let B denote the backshift operator ($B^k X_t = X_{t-k}$, $k = 0, \pm 1, \pm 2, \ldots$) and let

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$$

be a pth order autoregressive polynomial, where $\phi(z) \neq 0$ for $|z| = 1$. The polynomial $\phi(B)$ is said to be *causal* if all its roots are outside the unit circle in the complex plane. In this case, for a sequence $\{W_t\}$,

$$\phi^{-1}(B)W_t = \left(\sum_{j=0}^{\infty} \psi_j B^j\right) W_t = \sum_{j=0}^{\infty} \psi_j W_{t-j},$$

a function of only the past and present $\{W_t\}$. Note that if $\phi(B)$ is causal, the polynomial $B^p \phi(B^{-1})$ is *purely noncausal* in the sense that all its roots are inside the unit circle, and hence

$$B^{-p}\phi^{-1}(B^{-1})W_t = \left(\sum_{j=0}^{\infty} \psi_j B^{-p-j}\right) W_t = \sum_{j=0}^{\infty} \psi_j W_{t+p+j},$$

a function of only the present and future $\{W_t\}$. See, for example, Chapter 3 of Brockwell and Davis (1991).

Let

$$\phi_0(z) = 1 - \phi_{01} z - \cdots - \phi_{0p} z^p,$$

where $\phi_0(z) \neq 0$ for $|z| \leq 1$. Define $\phi_{00} = 1$ and suppose $\phi_{0r} \neq 0$ for some $r \in \{0, 1, \ldots, p\}$ and $\phi_{0j} = 0$ for $j = r+1, \ldots, p$. Then, a causal all-pass time series is the ARMA series $\{X_t\}$ which satisfies the difference equations

$$\phi_0(B)X_t = \frac{B^r \phi_0(B^{-1})}{-\phi_{0r}} Z_t^* \tag{4.1}$$

or

$$X_t - \phi_{01}X_{t-1} - \cdots - \phi_{0r}X_{t-r} = Z_t^* + \frac{\phi_{0,r-1}}{\phi_{0r}} Z_{t-1}^* + \cdots + \frac{\phi_{01}}{\phi_{0r}} Z_{t-r+1}^* - \frac{1}{\phi_{0r}} Z_{t-r}^*,$$

where the series $\{Z_t^*\}$ is an independent and identically distributed (iid) sequence of random variables. The true order of the all-pass model is r. Observe that the roots of the autoregressive polynomial $\phi_0(z)$ are reciprocals of the roots of the moving average polynomial $-\phi_{0r}^{-1} z^r \phi_0(z^{-1})$ and vice versa. In this chapter, we assume that the distribution of Z_1^* belongs to the domain of attraction of a stable distribution with exponent $\alpha \in (0, 2)$ ($Z_1^* \in D(\alpha)$).

By definition, nondegenerate, iid random variables $\{S_t\}$ have a stable distribution if there exist positive constants $\{b_n\}$ and constants $\{c_n\}$ such that

$$b_n (S_1 + \cdots + S_n) + c_n \stackrel{d}{=} S_1$$

for all n. The distribution of Z_1^* is in the domain of attraction of the distribution of S_1 if there exist constants $\{\tilde{b}_n\}$ and $\{\tilde{c}_n\}$ such that

$$\tilde{b}_n (Z_1^* + \cdots + Z_n^*) + \tilde{c}_n \stackrel{d}{\to} S_1.$$

Hence, a stable distribution is in its own domain of attraction. In general, stable distributions can be indexed by an exponent in $(0, 2]$. When the exponent is two, the stable distribution is Gaussian. Since $\alpha \in (0, 2)$, the conditions

$$\lim_{x \to \infty} \frac{\mathrm{P}(Z_1^* > x)}{\mathrm{P}(|Z_1^*| > x)} = q \tag{4.2}$$

for some $0 \leq q \leq 1$ and

$$\mathrm{P}(|Z_1^*| > x) = x^{-\alpha} L(x) \tag{4.3}$$

for all $x > 0$, where $L(x)$ is slowly varying at ∞ (e.g., $\lim_{x\to\infty} L(sx)/L(x) = 1 \ \forall s > 0$), are necessary and sufficient for $Z_1^* \in D(\alpha)$ (see, for example, Feller, 1971, page 312). By (4.3),

$$\mathrm{E}|Z_1^*|^\delta < \infty \ \forall \delta \in [0, \alpha) \quad \text{and} \quad \mathrm{E}|Z_1^*|^\delta = \infty \ \forall \delta > \alpha,$$

and so Z_1^* has infinite variance. In addition, if $\{\psi_j\}_{j=0}^\infty$ is given by $\sum_{j=0}^\infty \psi_j z^j = 1/\phi_0(z)$, then $X_1 = \sum_{j=0}^\infty \psi_j Z_{1-j}^* < \infty$ almost surely and

$$\lim_{x\to\infty} \frac{\mathrm{P}(|X_1| > x)}{\mathrm{P}(|Z_1^*| > x)} = \lim_{x\to\infty} \frac{\mathrm{P}\left(|\sum_{j=0}^\infty \psi_j Z_{1-j}^*| > x\right)}{\mathrm{P}(|Z_1^*| > x)} = \sum_{j=0}^\infty |\psi_j|^\alpha \tag{4.4}$$

by Cline (1983). Because there exist constants $a > 0$ and $0 < b < 1$ such that $|\psi_j| \le ab^j$ for all $j \in \{0, 1, \ldots\}$ (see Brockwell and Davis, 1991, Section 3.1), $\sum_{j=0}^\infty |\psi_j|^\alpha < \infty$. To give an example, if Z_1^* has a Students' t-distribution with $0 < \nu < 2$ degrees of freedom, then $Z_1^* \in D(\nu)$ with $q = 1/2$.

If, instead of having infinite variance, Z_1^* had finite variance σ^2, then the spectral density for $\{X_t\}$ in (4.1) would be

$$\frac{|e^{-ir\omega}|^2 |\phi_0(e^{i\omega})|^2}{\phi_{0r}^2 |\phi_0(e^{-i\omega})|^2} \frac{\sigma^2}{2\pi} = \frac{\sigma^2}{\phi_{0r}^2 2\pi},$$

which is constant for $\omega \in [-\pi, \pi]$, and thus $\{X_t\}$ would be an uncorrelated sequence. In the case of Gaussian $\{Z_t^*\}$, this implies that $\{X_t\}$ is iid $\mathrm{N}(0, \sigma^2\phi_{0r}^{-2})$. Independence, however, does not hold in the non-Gaussian case if $r \ge 1$, whether or not Z_1^* has finite variance (e.g., Breidt and Davis, 1991). The model (4.1) is called all-pass because the power transfer function of the all-pass filter passes all the power for every frequency in the spectrum. In other words, an all-pass filter does not change the distribution of power over the spectrum.

We can express (4.1) as

$$\phi_0(B)X_t = \frac{B^p \phi_0(B^{-1})}{-\phi_{0r}} Z_t, \tag{4.5}$$

where $\{Z_t\} = \{Z_{t+p-r}^*\}$. Rearranging (4.5) and setting $z_t = \phi_{0r}^{-1} Z_t$, we have the backward recursion

$$z_{t-p} = \phi_{01} z_{t-p+1} + \cdots + \phi_{0p} z_t - (X_t - \phi_{01} X_{t-1} - \cdots - \phi_{0p} X_{t-p}).$$

An analogous recursion for an arbitrary, causal autoregressive polynomial $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$

can be defined as follows:

$$z_{t-p}(\phi) = \begin{cases} 0, & t = n+p, \ldots, n+1, \\ \phi_1 z_{t-p+1}(\phi) + \cdots + \phi_p z_t(\phi) - \phi(B)X_t, & t = n, \ldots, p+1, \end{cases} \tag{4.6}$$

where $\phi := (\phi_1, \ldots, \phi_p)'$. Let $\phi_0 = (\phi_{01}, \ldots, \phi_{0p})' = (\phi_{01}, \ldots, \phi_{0r}, 0, \ldots, 0)'$ denote the true parameter vector. Note that $\{z_t(\phi_0)\}_{t=1}^{n-p}$ closely approximates $\{z_t\}_{t=1}^{n-p}$; the error is due to the initialization with zeros. Although $\{z_t\}$ is iid, $\{z_t(\phi_0)\}_{t=1}^{n-p}$ is not iid if $r \geq 1$. Given a realization of length n, $\{X_t\}_{t=1}^{n}$, we estimate ϕ_0 by minimizing the LAD criterion $\sum_{t=1}^{n-p} |z_t(\phi)|$ with respect to ϕ, where $\{z_t(\phi)\}$ can be computed recursively from (4.6). Also, when the noise distribution is stable, we consider estimating ϕ_0 by maximum likelihood. LAD and ML estimation for all-pass model parameters are considered in Breidt, Davis, and Trindade (2001) and Chapter 2 respectively, but these results apply to all-pass processes with finite variance.

4.3 Asymptotic Results

4.3.1 LAD Estimation when the Noise Distribution is in a Stable Domain of Attraction

We begin this section by defining a limiting process $V(\cdot)$. For $\mathbf{u} = (u_1, \ldots, u_p)' \in \mathbb{R}^p$, let $\{c_j(\mathbf{u})\}_{j \neq 0}$ contain the coefficients of the z's in $\mathbf{u}'[-(z_{1-i}/\phi_0(B)) + (z_{1+i}/\phi_0(B^{-1}))]_{i=1}^{p}$. Thus,

$$\begin{aligned} c_1(\mathbf{u}) &= -u_1\psi_0, \\ c_2(\mathbf{u}) &= -u_1\psi_1 - u_2\psi_0, \\ c_3(\mathbf{u}) &= -u_1\psi_2 - u_2\psi_1 - u_3\psi_0, \\ &\vdots \\ c_p(\mathbf{u}) &= -u_1\psi_{p-1} - u_2\psi_{p-2} - \cdots - u_p\psi_0, \\ c_{p+1}(\mathbf{u}) &= -u_1\psi_p - u_2\psi_{p-1} - \cdots - u_p\psi_1, \\ c_{p+2}(\mathbf{u}) &= -u_1\psi_{p+1} - u_2\psi_p - \cdots - u_p\psi_2, \\ &\vdots \end{aligned}$$

and

$$\begin{aligned} c_{-1}(\mathbf{u}) &= -c_1(\mathbf{u}), \\ c_{-2}(\mathbf{u}) &= -c_2(\mathbf{u}), \\ &\vdots \end{aligned} \tag{4.7}$$

where $\{\psi_j\}_{j=0}^{\infty}$ is given by $\sum_{j=0}^{\infty} \psi_j z^j = 1/\phi_0(z)$. Then, let

$$\begin{aligned} V(\mathbf{u}) &= |\phi_{0r}|^{-1} \sum_{j\neq 0} \sum_{k=1}^{\infty} \left(\left| Z_{k,j} + c_j(\mathbf{u})\delta_k \Gamma_k^{-1/\alpha} \right| - |Z_{k,j}| \right) \\ &= |\phi_{0r}|^{-1} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \left(\left| Z_{k,j} + c_j(\mathbf{u})\delta_k \Gamma_k^{-1/\alpha} \right| + \left| Z_{k,-j} - c_j(\mathbf{u})\delta_k \Gamma_k^{-1/\alpha} \right| - |Z_{k,j}| - |Z_{k,-j}| \right), \end{aligned} \tag{4.8}$$

where

- $\{Z_{k,j}\}_{k,j}$ is an iid sequence with $Z_{k,j} \stackrel{d}{=} Z_1$,
- $\{\delta_k\}$ is iid with $\mathrm{P}(\delta_k = 1) = q$ and $\mathrm{P}(\delta_k = -1) = 1 - q$, where q is defined in (4.2),
- $\Gamma_k = E_1 + \ldots + E_k$, where $\{E_k\}$ is an iid series of exponential random variables with mean one, and
- $\{Z_{k,j}\}$, $\{\delta_k\}$, and $\{E_k\}$ are mutually independent.

Define

$$a_n = \inf\{x : \mathrm{P}(|Z_1| > x) \leq n^{-1}\} \tag{4.9}$$

for all n. Note that

$$n\mathrm{P}(|Z_1| > a_n x) \to x^{-\alpha} \tag{4.10}$$

for any $x > 0$, and that minimizing $\sum_{t=1}^{n-p} |z_t(\boldsymbol{\phi})|$ with respect to $\boldsymbol{\phi}$ is equivalent to minimizing

$$V_n(\mathbf{u}) := \sum_{t=1}^{n-p} \left(|z_t(\boldsymbol{\phi}_0 + a_n^{-1}\mathbf{u})| - |z_t(\boldsymbol{\phi}_0)| \right) \tag{4.11}$$

with respect to $\mathbf{u}$ if $\mathbf{u} = a_n(\boldsymbol{\phi} - \boldsymbol{\phi}_0)$. Using techniques from the proofs of Theorem 4.1 in Davis, Knight, and Liu (1992) and Theorem 3.4 in Davis (1996), we now show that $V_n(\cdot) \stackrel{d}{\to} V(\cdot)$ on $C(\mathbb{R}^p)$, the space of continuous functions on $\mathbb{R}^p$ where convergence is equivalent to uniform convergence on every compact subset.

Theorem 1 *If $Z_1 \in D(\alpha)$ with $\alpha \in (0,2)$ and*

- *$\alpha < 1$, or*
- *$\alpha \geq 1$ and Z_1 has a distribution that is symmetric about zero and continuously differentiable on $[-\kappa, \kappa]$ for some $\kappa > 0$,*

then $V(\mathbf{u})$ is finite for all $\mathbf{u} \in \mathbb{R}^p$ almost surely and

$$V_n(\cdot) \xrightarrow{d} V(\cdot)$$

on $C(\mathbb{R}^p)$.

Proof: Let $\mathbf{u} \in \mathbb{R}^p$. If $\alpha < 1$, choose $k^* \in \{1, 2, \ldots\}$ such that $k^* > 1/\alpha$. Hence,

$$\begin{aligned} \mathrm{E}\left\{\sum_{j\neq 0}\sum_{k=k^*}^{\infty} |c_j(\mathbf{u})|\Gamma_k^{-1/\alpha}\right\} &= \sum_{j\neq 0}|c_j(\mathbf{u})| \sum_{k=k^*}^{\infty} \mathrm{E}\left\{\Gamma_k^{-1/\alpha}\right\} \\ &= \sum_{j\neq 0}|c_j(\mathbf{u})| \sum_{k=k^*}^{\infty} \frac{\Gamma(k-1/\alpha)}{\Gamma(k)} \\ &< (constant)\sum_{j\neq 0}|c_j(\mathbf{u})| \sum_{k=k^*}^{\infty} k^{-1/\alpha} \\ &< \infty. \end{aligned}$$

Consequently, since $\sum_{k=1}^{k^*-1}\Gamma_k^{-1/\alpha} < \infty$ almost surely,

$$|V(\mathbf{u})| \leq |\phi_{0r}|^{-1}\sum_{j\neq 0}\sum_{k=1}^{\infty}|c_j(\mathbf{u})|\Gamma_k^{-1/\alpha} < \infty$$

almost surely. If $\alpha \geq 1$, $|V(\mathbf{u})| < \infty$ almost surely by Proposition A.3 in Davis, Knight, Liu (1992). Therefore, $V(\mathbf{u})$ is finite for all rationals $\mathbf{u} \in \mathbb{R}^p$ almost surely. Because the sample paths of $V(\cdot)$ are convex, it follows that $V(\mathbf{u})$ is finite for all $\mathbf{u} \in \mathbb{R}^p$ almost surely.

We now show that $V_n(\cdot) \xrightarrow{d} V(\cdot)$ on $C(\mathbb{R}^p)$. Let

$$V_n^*(\mathbf{u}) = \sum_{t=1}^{n-p}\left(\left|z_t + a_n^{-1}\sum_{j\neq 0} c_j(\mathbf{u})z_{t-j}\right| - |z_t|\right) \tag{4.12}$$

and

$$V_n^\dagger(\mathbf{u}) = \sum_{t=1}^{n-p} \left(\left| z_t(\phi_0) + a_n^{-1} \sum_{j \neq 0} c_j(\mathbf{u}) z_{t-j} \right| - |z_t(\phi_0)| \right). \tag{4.13}$$

To complete the proof, we show that $V_n^*(\cdot) \xrightarrow{d} V(\cdot)$, $V_n^*(\cdot) - V_n^\dagger(\cdot) = o_p(1)$, and $V_n^\dagger(\cdot) - V_n(\cdot) = o_p(1)$ on $C(\mathbb{R}^p)$.

Let $\mathbf{u} \in \mathbb{R}^p$. We begin by showing $V_n^*(\mathbf{u}) \xrightarrow{d} V(\mathbf{u})$. For each $x \in \mathbb{R} \times (\overline{\mathbb{R}}^2 \setminus \{0\})$, where $\overline{\mathbb{R}} = [-\infty, \infty]$, define a set function $\varepsilon_x(\cdot)$ as follows:

$$\varepsilon_x(A) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{otherwise,} \end{cases}$$

where A is an element of the σ-algebra generated by the open subsets of $\mathbb{R} \times (\overline{\mathbb{R}}^2 \setminus \{0\})$. A point measure m is a measure of the form

$$m(\cdot) = \sum_{j \in J} \varepsilon_{x_j}(\cdot),$$

where m is finite on all relatively compact subsets of $\mathbb{R} \times (\overline{\mathbb{R}}^2 \setminus \{0\})$ (subsets A for which the closure $\overline{A}$ is compact; note that a compact subset of $\overline{\mathbb{R}}^2 \setminus \{0\}$ is closed and bounded away from the origin). The class of all such point measures is denoted $M_p(\mathbb{R} \times (\overline{\mathbb{R}}^2 \setminus \{0\}))$. If

$$U_{t,n}^- := -a_n^{-1} \sum_{j=1}^{\infty} c_j(\mathbf{u}) z_{t+j} \quad \text{and} \quad U_{t,n}^+ := a_n^{-1} \sum_{j=1}^{\infty} c_j(\mathbf{u}) z_{t-j}, \tag{4.14}$$

then, by a straightforward extension of results in Section 4 of Feigin and Resnick (1994),

$$\sum_{t=1}^{n-p} \varepsilon_{(z_t, U_{t,n}^-, U_{t,n}^+)} \xrightarrow{d} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \left(\varepsilon_{(\phi_{0r}^{-1} Z_{k,-j}, -\phi_{0r}^{-1} c_j(\mathbf{u}) \delta_k \Gamma_k^{-1/\alpha}, 0)} + \varepsilon_{(\phi_{0r}^{-1} Z_{k,j}, 0, \phi_{0r}^{-1} c_j(\mathbf{u}) \delta_k \Gamma_k^{-1/\alpha})} \right)$$

in $M_p(\mathbb{R} \times (\overline{\mathbb{R}}^2 \setminus \{0\}))$. Also, as discussed in the Appendix of Davis, Knight, and Liu (1992), for all continuous functions g on $\mathbb{R} \times (\overline{\mathbb{R}}^2 \setminus \{0\})$ with compact support,

$$\sum_{t=1}^{n-p} g(z_t, U_{t,n}^-, U_{t,n}^+) \xrightarrow{d} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \left(g(\phi_{0r}^{-1} Z_{k,-j}, -\phi_{0r}^{-1} c_j(\mathbf{u}) \delta_k \Gamma_k^{-1/\alpha}, 0) + g(\phi_{0r}^{-1} Z_{k,j}, 0, \phi_{0r}^{-1} c_j(\mathbf{u}) \delta_k \Gamma_k^{-1/\alpha}) \right).$$

Suppose

$$A_0 := \{|x| : \text{the distribution of } z_1 \text{ is discontinuous at } x\}$$

and observe that A_0 contains at most countably many elements. If $M > 0$, $M \notin A_0$, and $\lambda > 0$, then the function

$$g_0(x,y,z) = (|x+y+z| - |x|)I\{|x| \le M\}I\{(|y| > \lambda) \cup (|z| > \lambda)\},$$

where $I\{\cdot\}$ is the indicator function, is almost everywhere continuous on $\mathbb{R} \times (\overline{\mathbb{R}}^2 \setminus \{\mathbf{0}\})$ with compact support. Because

$$\mathrm{P}\left(\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\left(\varepsilon_{(\phi_{0r}^{-1}Z_{k,-j},-\phi_{0r}^{-1}c_j(\mathbf{u})\delta_k\Gamma_k^{-1/\alpha},0)}(A) + \varepsilon_{(\phi_{0r}^{-1}Z_{k,j},0,\phi_{0r}^{-1}c_j(\mathbf{u})\delta_k\Gamma_k^{-1/\alpha})}(A)\right) > 0\right) = 0$$

when $A = \{(x,y,z) : |x| = M, \ |y| = \lambda, \ \text{or } |z| = \lambda\}$,

$$\sum_{t=1}^{n-p}\left(|z_t + U_{t,n}^- + U_{t,n}^+| - |z_t|\right) I_{t,n}^{\lambda,\lambda,M} \xrightarrow{d} |\phi_{0r}|^{-1}\sum_{j\ne 0}\sum_{k=1}^{\infty}\left(\left|Z_{k,j} + c_j(\mathbf{u})\delta_k\Gamma_k^{-1/\alpha}\right| - |Z_{k,j}|\right) I_{k,j}^{\lambda,M},$$

where

$$I_{t,n}^{\lambda,\lambda,M} := I\{|z_t| \le M\}I\{(|U_{t,n}^-| > \lambda) \cup (|U_{t,n}^+| > \lambda)\}$$

and

$$I_{k,j}^{\lambda,M} := I\{|\phi_{0r}^{-1}Z_{k,j}| \le M\}I\left\{\left|\phi_{0r}^{-1}c_j(\mathbf{u})\delta_k\Gamma_k^{-1/\alpha}\right| > \lambda\right\}.$$

By Theorem 3.2 in Billingsley (1999), it follows that

$$V_n^*(\mathbf{u}) = \sum_{t=1}^{n-p}(|z_t + U_{t,n}^- + U_{t,n}^+| - |z_t|) \xrightarrow{d} V(\mathbf{u})$$

if, for all $\eta > 0$,

$$\lim_{\lambda\to 0^+}\lim_{M\to\infty}\limsup_{n\to\infty}\mathrm{P}\left(\left|\sum_{t=1}^{n-p}\left(|z_t + U_{t,n}^- + U_{t,n}^+| - |z_t|\right)\left(1 - I_{t,n}^{\lambda,\lambda,M}\right)\right| > \eta\right) = 0 \tag{4.15}$$

and

$$|\phi_{0r}|^{-1}\sum_{j\ne 0}\sum_{k=1}^{\infty}\left(\left|Z_{k,j} + c_j(\mathbf{u})\delta_k\Gamma_k^{-1/\alpha}\right| - |Z_{k,j}|\right)\left(1 - I_{k,j}^{\lambda,M}\right) \xrightarrow{P} 0 \tag{4.16}$$

as $\lambda \to 0^+$ and $M \to \infty$.

We first consider the case $\alpha < 1$. Note that

$$\begin{aligned} 1 - I_{t,n}^{\lambda,\lambda,M} &= I\{|U_{t,n}^-| \le \lambda\}I\{|U_{t,n}^+| \le \lambda\} + I\{|z_t| > M\}I\{|U_{t,n}^-| > \lambda\} + I\{|z_t| > M\}I\{|U_{t,n}^+| > \lambda\} \\ &\quad - I\{|z_t| > M\}I\{|U_{t,n}^-| > \lambda\}I\{|U_{t,n}^+| > \lambda\}, \end{aligned} \tag{4.17}$$

and thus, for any $\eta > 0$,

$$\begin{aligned}
&\mathrm{P}\left(\left|\sum_{t=1}^{n-p}\left(|z_t + U_{t,n}^- + U_{t,n}^+| - |z_t|\right)\left(1 - I_{t,n}^{\lambda,\lambda,M}\right)\right| > \eta\right) \\
\leq\ & \mathrm{P}\left(\sum_{t=1}^{n-p}\left(|U_{t,n}^-| + |U_{t,n}^+|\right) I\{|U_{t,n}^-| \leq \lambda\} I\{|U_{t,n}^+| \leq \lambda\} > \frac{\eta}{4}\right) \\
&+\mathrm{P}\left(\bigcup_{t=1}^{n-p}\{(|z_t| > M) \cap (|U_{t,n}^-| > \lambda)\}\right) + \mathrm{P}\left(\bigcup_{t=1}^{n-p}\{(|z_t| > M) \cap (|U_{t,n}^+| > \lambda)\}\right) \\
&+\mathrm{P}\left(\bigcup_{t=1}^{n-p}\{(|z_t| > M) \cap (|U_{t,n}^-| > \lambda) \cap (|U_{t,n}^+| > \lambda)\}\right).
\end{aligned}$$

By Lemma 1, $\lim_{\lambda\to 0+}\lim_{M\to\infty}\limsup_{n\to\infty}$ of the three rightmost terms equals zero, and, by Proposition A.2(a) in Davis, Knight, and Liu (1992),

$$\begin{aligned}
&\lim_{\lambda\to 0+}\limsup_{n\to\infty}\mathrm{P}\left(\sum_{t=1}^{n-p}\left(|U_{t,n}^-| + |U_{t,n}^+|\right) I\{|U_{t,n}^-| \leq \lambda\} I\{|U_{t,n}^+| \leq \lambda\} > \frac{\eta}{4}\right) \\
\leq\ & \lim_{\lambda\to 0+}\limsup_{n\to\infty}\mathrm{P}\left(\sum_{t=1}^{n-p}|U_{t,n}^-| I\{|U_{t,n}^-| \leq \lambda\} > \frac{\eta}{8}\right) + \lim_{\lambda\to 0+}\limsup_{n\to\infty}\mathrm{P}\left(\sum_{t=1}^{n-p}|U_{t,n}^+| I\{|U_{t,n}^+| \leq \lambda\} > \frac{\eta}{8}\right) \\
\leq\ & (constant)\eta^{-1}\lim_{\lambda\to 0+}\lambda^{1-\alpha} \\
=\ & 0.
\end{aligned}$$

Therefore, (4.15) must hold when $\alpha < 1$. Also, since

$$|\phi_{0r}|^{-1}\sum_{j\neq 0}\sum_{k=1}^{\infty}\left|\left|Z_{k,j} + c_j(\mathbf{u})\delta_k\Gamma_k^{-1/\alpha}\right| - |Z_{k,j}|\right| \leq |\phi_{0r}|^{-1}\sum_{j\neq 0}|c_j(\mathbf{u})|\sum_{k=1}^{\infty}\Gamma_k^{-1/\alpha},$$

which is finite almost surely, (4.16) holds in this case.

Now consider the case when $\alpha \geq 1$. From (4.17), because

$$|x+y| - |x| = y(I\{x>0\} - I\{x<0\}) + 2(x+y)(I\{-y<x<0\} - I\{-y>x>0\}) \tag{4.18}$$

if $x \neq 0$ and $\mathrm{P}(Z_1 = 0) = 0$,

$$\begin{aligned}
&\mathrm{P}\left(\left|\sum_{t=1}^{n-p}\left(|z_t + U_{t,n}^- + U_{t,n}^+| - |z_t|\right)\left(1 - I_{t,n}^{\lambda,\lambda,M}\right)\right| > \eta\right) \\
\leq\ & \mathrm{P}\left(\left|\sum_{t=1}^{n-p}U_{t,n}(I\{z_t>0\} - I\{z_t<0\}) I\{|U_{t,n}^-| \leq \lambda\} I\{|U_{t,n}^+| \leq \lambda\}\right| > \frac{\eta}{8}\right) \\
&+\mathrm{P}\left(\left|\sum_{t=1}^{n-p}(z_t + U_{t,n})(I\{-U_{t,n} < z_t < 0\} - I\{-U_{t,n} > z_t > 0\}) I\{|U_{t,n}^-| \leq \lambda\} I\{|U_{t,n}^+| \leq \lambda\}\right| > \frac{\eta}{16}\right)
\end{aligned}$$

$$+2\mathrm{P}\left(\bigcup_{t=1}^{n-p}\{(|z_t|>M)\cap(|U_{t,n}^{-}|>\lambda)\}\right)+2\mathrm{P}\left(\bigcup_{t=1}^{n-p}\{(|z_t|>M)\cap(|U_{t,n}^{+}|>\lambda)\}\right)$$
$$+2\mathrm{P}\left(\bigcup_{t=1}^{n-p}\{(|z_t|>M)\cap(|U_{t,n}^{-}|>\lambda)\cap(|U_{t,n}^{+}|>\lambda)\}\right)$$

for any $\eta>0$, where $U_{t,n}:=U_{t,n}^{-}+U_{t,n}^{+}$. By Lemmas 1-3 in Section 4.5, $\lim_{\lambda\to 0^+}\lim_{M\to\infty}\limsup_{n\to\infty}$ of each probability on the right-hand side of the above equation equals zero, and so (4.15) holds. To show (4.16), we note that

$$1-I_{k,j}^{\lambda,M}=I\left\{\left|\phi_{0r}^{-1}c_j(\mathbf{u})\delta_k\Gamma_k^{-1/\alpha}\right|\le\lambda\right\}+I\{|\phi_{0r}^{-1}Z_{k,j}|>M\}I\left\{\left|\phi_{0r}^{-1}c_j(\mathbf{u})\delta_k\Gamma_k^{-1/\alpha}\right|>\lambda\right\},$$

and so

$$\mathrm{P}\left(\left|\phi_{0r}^{-1}\sum_{j\neq 0}\sum_{k=1}^{\infty}\left(\left|Z_{k,j}+c_j(\mathbf{u})\delta_k\Gamma_k^{-1/\alpha}\right|-|Z_{k,j}|\right)\left(1-I_{k,j}^{\lambda,M}\right)\right|>\eta\right)$$
$$\le\ \mathrm{P}\left(\left|\phi_{0r}^{-1}\sum_{j\neq 0}\sum_{k=1}^{\infty}\gamma_{k,j}(I\{Z_{k,j}>0\}-I\{Z_{k,j}<0\})I\{|\phi_{0r}^{-1}\gamma_{k,j}|\le\lambda\}\right|>\frac{\eta}{4}\right)$$
$$+\mathrm{P}\left(\left|\phi_{0r}^{-1}\sum_{j\neq 0}\sum_{k=1}^{\infty}(Z_{k,j}+\gamma_{k,j})(I\{-\gamma_{k,j}<Z_{k,j}<0\}-I\{-\gamma_{k,j}>Z_{k,j}>0\})I\{|\phi_{0r}^{-1}\gamma_{k,j}|\le\lambda\}\right|>\frac{\eta}{8}\right)$$
$$+\mathrm{P}\left(\left|\phi_{0r}^{-1}\sum_{j\neq 0}\sum_{k=1}^{\infty}\gamma_{k,j}(I\{Z_{k,j}>0\}-I\{Z_{k,j}<0\})\tilde{I}_{k,j}^{\lambda,M}\right|>\frac{\eta}{4}\right)$$
$$+\mathrm{P}\left(\left|\phi_{0r}^{-1}\sum_{j\neq 0}\sum_{k=1}^{\infty}(Z_{k,j}+\gamma_{k,j})(I\{-\gamma_{k,j}<Z_{k,j}<0\}-I\{-\gamma_{k,j}>Z_{k,j}>0\})\tilde{I}_{k,j}^{\lambda,M}\right|>\frac{\eta}{8}\right)$$

for any $\eta>0$, where $\gamma_{k,j}:=c_j(\mathbf{u})\delta_k\Gamma_k^{-1/\alpha}$ and

$$\tilde{I}_{k,j}^{\lambda,M}:=I\{|\phi_{0r}^{-1}Z_{k,j}|>M\}I\{|\phi_{0r}^{-1}\gamma_{k,j}|>\lambda\}.$$

By Lemmas 4-6, $\lim_{\lambda\to 0^+}\lim_{M\to\infty}$ of each probability on the right-hand side of the above equation equals zero.

We have therefore shown that $V_n^*(\mathbf{u})\xrightarrow{d}V(\mathbf{u})$ for any $\mathbf{u}\in\mathbb{R}^p$. Using the Cramér-Wold device, it can be shown similarly that all finite-dimensional distributions of $V_n^*(\cdot)$ converge to those of $V(\cdot)$. By Lemma 7, $V_n^*(\cdot)$ is tight on $C(K)$ for any compact set $K\subset\mathbb{R}^p$, and thus $V_n^*(\cdot)\xrightarrow{d}V(\cdot)$ on $C(\mathbb{R}^p)$ by Theorem 7.1 in Billingsley (1999). Because $V_n^*(\cdot)-V_n^\dagger(\cdot)=o_p(1)$ on $C(\mathbb{R}^p)$ by Lemma 8 and $V_n^\dagger(\cdot)-V_n(\cdot)=o_p(1)$ on $C(\mathbb{R}^p)$ by Lemma 9, the proof is complete. □

Corollary 1 *If the conditions of Theorem 1 hold and*

- $\alpha < 1$ *and there exist constants* $C_0 > 0$ *and* $\epsilon > 0$ *such that* $P(x < Z_1 < y) \geq C_0(y-x)^\alpha$ *for all* $-\epsilon < x < y < \epsilon$, *or*
- $\alpha \geq 1$ *and the distribution function for* Z_1 *is strictly increasing in a neighborhood of zero,*

then $V(\cdot)$ *has a unique minimum almost surely.*

Proof: When $\alpha \geq 1$, because the distribution of Z_1 is strictly increasing and continuously differentiable in a neighborhood of zero, $\mathrm{P}(x < Z_1 < y) \geq C_1(y-x)$ for some constant $C_1 > 0$ and all $x < y$ in some neighborhood of zero. Therefore, the result follows by the remark on page 159 of Davis, Knight, and Liu (1992) where a proof is given for the case $\alpha \geq 1$. The proof for the case $\alpha < 1$, which uses the fact that $\sum_{k=1}^{\infty} k^{-1} = \infty$, is similar. □

We can now give conditions for the existence of consistent LAD estimators of ϕ_0 with a nondegenerate limiting distribution.

Corollary 2 *If the conditions of Corollary 1 hold, then there exists a sequence of minimizers* $\hat{\phi}_{LAD}$ *of* $\sum_{t=1}^{n-p} |z_t(\cdot)|$ *such that*

$$a_n(\hat{\phi}_{LAD} - \phi_0) \xrightarrow{d} \zeta,$$

where ζ *is the unique minimizer of* $V(\cdot)$.

Proof: Note that $\sum_{t=1}^{n-p} |z_t(\phi)| - \sum_{t=1}^{n-p} |z_t(\phi_0)| = V_n(a_n(\phi - \phi_0))$, where $V_n(\cdot)$ is defined in (4.11). Because ζ uniquely minimizes $V(\cdot)$, the limiting process for $V_n(\cdot)$, the result follows by Remark 1 in Davis, Knight, and Liu (1992). □

Remark 1: Suppose $\alpha < 1$, the distribution function for Z_1 is continuously differentiable in some neighborhood of zero, and there exist constants $C_0 > 0$ and $\epsilon > 0$ such that

$$\mathrm{P}(x < Z_1 < y) \geq C_0(y-x)^\alpha$$

for all $-\epsilon < x < y < \epsilon$. Then,

$$\frac{\mathrm{P}(-y < Z_1 < y)}{2y} \geq \frac{C_0(2y)^\alpha}{2y} \tag{4.19}$$

if $-\epsilon < -y < y < \epsilon$. Because the distribution function is continuously differentiable near zero,

$$\lim_{y \to 0+} \frac{\mathrm{P}(-y < Z_1 < y)}{2y} = constant < \infty.$$

However,

$$\lim_{y \to 0+} \frac{C_0(2y)^\alpha}{2y} = \infty,$$

and so we have contradicted (4.19). Consequently, the conditions of Corollary 2 can never hold when $\alpha < 1$ and the distribution function for Z_1 is continuously differentiable in a neighborhood of zero.

Remark 2: If Z_1 has a Students' t-distribution with $0 < \nu < 2$ degrees of freedom, then $Z_1 \in D(\nu)$ and the distribution of Z_1 is symmetric about zero, strictly increasing on $\mathbb{R}$, and continuously differentiable everywhere on $\mathbb{R}$. Therefore, when an all-pass model has t noise with $1 \leq \nu < 2$, there exist consistent LAD estimators of model parameters with a nondegenerate limiting distribution. Note that the Students' t-distribution with one degree of freedom is Cauchy.

4.3.2 ML Estimation when the Noise Distribution is Stable

We now impose further restrictions on the noise distribution and assume it is stable with exponent $\alpha \in (0, 2)$. A stable distribution is indexed by not only an exponent, but also a parameter of symmetry ($|\beta| \leq \min\{\alpha, 2-\alpha\}$), a parameter of location ($-\infty < \mu < \infty$), and a scale parameter ($0 < \sigma < \infty$). If $\beta = 0$, the stable distribution is symmetric about μ, and, if $\alpha = 1$ and $\beta = 0$, the distribution is Cauchy. Also, when $\alpha = 1/2$ and $\beta = 1$, the distribution is Lévy. We consider the case when Z_1 has a stable distribution that is symmetric about zero ($\beta = 0$ and $\mu = 0$). In this case, the distribution of Z_1 is indexed by two parameters, α and σ, and the density of Z_1 can be expressed as

$$f_\sigma(z; \alpha) = \sigma^{-1} f(z\sigma^{-1}; \alpha)$$

for some density function f. Although, when $\alpha \in (0, 2)$, $\beta = 0$, and $\mu = 0$, the characteristic function for Z_1 is given by

$$\mathrm{E}\{\exp(itZ_1)\} = \exp\left(-|\sigma t|^\alpha\right),$$

no general, closed-form expression is known for f when Z_1 is not Cauchy. However, there are computational formulas that can be used to evaluate f (see Nolan, 1997). It is also known that $f(z;\theta_{p+2})$ is unimodal on $\mathbb{R}$ and infinitely differentiable with respect to both z and θ_{p+2} on $\mathbb{R}\times(0,2)$ (see Zolotarev, 1986).

Let $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_{p+2})' = (\phi_1,\ldots,\phi_p,\theta_{p+1},\theta_{p+2})'$ and $\boldsymbol{\theta}_0 = (\theta_{01},\ldots,\theta_{0,p+2})' = (\phi_{01},\ldots,\phi_{0p},\sigma/|\phi_{0r}|,\alpha)'$. From Breidt, Davis, and Trindade (2001), the log-likelihood of $\boldsymbol{\theta}$ given a realization $\{X_t\}_{t=1}^n$ is approximately

$$\mathcal{L}(\boldsymbol{\theta}) := \sum_{t=1}^{n-p}\{\ln f(z_t(\boldsymbol{\phi})/\theta_{p+1};\theta_{p+2}) - \ln\theta_{p+1}\}.$$

ML estimates of $\boldsymbol{\phi}_0$, α, and σ can be obtained by maximizing $\mathcal{L}(\cdot)$. Because they are based on specific information about the noise distribution, ML estimates tend to be less disperse than LAD estimates. Using asymptotic expansions for the partial derivatives of f in DuMouchel (1973), Calder (1998) (pages 14-15) found asymptotic expansions for the first and second order partial derivatives of $\ln f$ as $z\to\infty$. Calder used these expansions to obtain limiting results for ML estimators of the parameters in an autoregressive model with stable noise. We now give similar limiting results for ML estimators of all-pass model parameters.

For $\mathbf{u}\in\mathbb{R}^p$ and $\mathbf{v}\in\mathbb{R}^2$, let

$$W_n(\mathbf{u},\mathbf{v}) = \mathcal{L}(\boldsymbol{\theta}_0 + (n^{-1/\alpha}\mathbf{u}', n^{-1/2}\mathbf{v}')') - \mathcal{L}(\boldsymbol{\theta}_0) \tag{4.20}$$

and, for $\mathbf{u}\in\mathbb{R}^p$, let

$$W(\mathbf{u}) = \sum_{j\neq 0}\sum_{k=1}^{\infty}\left\{\ln f_\sigma\left(Z_{k,j} + c_j(\mathbf{u})\delta_k\Gamma_k^{-1/\alpha};\alpha\right) - \ln f_\sigma\left(Z_{k,j};\alpha\right)\right\},$$

where $\{c_j(\mathbf{u})\}$, $\{Z_{k,j}\}$, $\{\delta_k\}$, and $\{E_k\}$ were defined previously. In this case, however, $\mathrm{P}(\delta_k = 1) = 1/2$ because Z_1 is symmetric about zero. From Calder (1998), $W(\mathbf{u})$ is finite for all $\mathbf{u}\in\mathbb{R}^p$ almost surely. Observe that maximizing $\mathcal{L}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ is equivalent to maximizing $W_n(\mathbf{u},\mathbf{v})$ with respect to $\mathbf{u}$ and $\mathbf{v}$ if $\mathbf{u} = n^{1/\alpha}(\boldsymbol{\phi}-\boldsymbol{\phi}_0)$ and $\mathbf{v} = n^{1/2}(\theta_{p+1}-\theta_{0,p+1},\theta_{p+2}-\theta_{0,p+2})'$.

Theorem 2 *If Z_1 has a stable distribution that is symmetric about zero with $\alpha\in(0,2)$, then*

$$W_n(\mathbf{u},\mathbf{v}) \xrightarrow{d} W(\mathbf{u}) + \mathbf{v}'\mathbf{N} - \frac{1}{2}\mathbf{v}'\mathbf{I}\mathbf{v}$$

on $C(\mathbb{R}^{p+2})$, *where* $\mathbf{N} \sim N(\mathbf{0}, \mathbf{I})$ *is independent of* $W(\cdot)$ *and*

$$\mathbf{I} := \begin{bmatrix} -\phi_{0r}^2 E\left\{\frac{\partial^2}{\partial\sigma^2}\ln f_\sigma(Z_1;\alpha)\right\} & -|\phi_{0r}|E\left\{\frac{\partial^2}{\partial\sigma\partial\alpha}\ln f_\sigma(Z_1;\alpha)\right\} \\ -|\phi_{0r}|E\left\{\frac{\partial^2}{\partial\sigma\partial\alpha}\ln f_\sigma(Z_1;\alpha)\right\} & -E\left\{\frac{\partial^2}{\partial\alpha^2}\ln f_\sigma(Z_1;\alpha)\right\} \end{bmatrix}.$$

Proof: For $\mathbf{u} \in \mathbb{R}^p$ and $\mathbf{v} = (v_1, v_2)' \in \mathbb{R}^2$, let

$$\begin{aligned} W_n^*(\mathbf{u},\mathbf{v}) &= \sum_{t=1}^{n-p} \ln\left[\left(\frac{\sigma}{|\phi_{0r}|} + \frac{v_1}{n^{1/2}}\right)^{-1} f\left(\frac{z_t + n^{-1/\alpha}\sum_{j\neq 0} c_j(\mathbf{u}) z_{t-j}}{\sigma/|\phi_{0r}| + v_1/n^{1/2}}; \alpha + \frac{v_2}{n^{1/2}}\right)\right] \\ &\quad - \sum_{t=1}^{n-p} \ln\left[\left(\frac{\sigma}{|\phi_{0r}|}\right)^{-1} f\left(\frac{z_t}{\sigma/|\phi_{0r}|}; \alpha\right)\right] \end{aligned} \tag{4.21}$$

and

$$\begin{aligned} W_n^\dagger(\mathbf{u},\mathbf{v}) &= \sum_{t=1}^{n-p} \ln\left[\left(\frac{\sigma}{|\phi_{0r}|} + \frac{v_1}{n^{1/2}}\right)^{-1} f\left(\frac{z_t(\boldsymbol{\phi}_0) + n^{-1/\alpha}\sum_{j\neq 0} c_j(\mathbf{u}) z_{t-j}}{\sigma/|\phi_{0r}| + v_1/n^{1/2}}; \alpha + \frac{v_2}{n^{1/2}}\right)\right] \\ &\quad - \sum_{t=1}^{n-p} \ln\left[\left(\frac{\sigma}{|\phi_{0r}|}\right)^{-1} f\left(\frac{z_t(\boldsymbol{\phi}_0)}{\sigma/|\phi_{0r}|}; \alpha\right)\right]. \end{aligned} \tag{4.22}$$

By a straightforward extension of Proposition 3 in Calder (1998),

$$W_n^*(\mathbf{u},\mathbf{v}) \xrightarrow{d} W(\mathbf{u}) + \mathbf{v}'\mathbf{N} - \frac{1}{2}\mathbf{v}'\mathbf{I}\mathbf{v}$$

on $C(\mathbb{R}^{p+2})$. Since $W_n^*(\cdot,\cdot) - W_n^\dagger(\cdot,\cdot) = o_p(1)$ on $C(\mathbb{R}^{p+2})$ by Lemma 10 and $W_n(\cdot,\cdot) - W_n^\dagger(\cdot,\cdot) = o_p(1)$ on $C(\mathbb{R}^{p+2})$ by Lemma 11, the proof is complete. □

It follows that there exist consistent ML estimators of $\boldsymbol{\theta}_0 = (\phi_{01}, \ldots, \phi_{0p}, \sigma/|\phi_{0r}|, \alpha)'$ with a nondegenerate limiting distribution.

Corollary 3 *Under the conditions of Theorem 2, there exists a sequence of maximizers*

$$\hat{\boldsymbol{\theta}}_{ML} = (\hat{\boldsymbol{\phi}}_{ML}', \hat{\theta}_{p+1,ML}, \hat{\alpha}_{ML})'$$

of $\mathcal{L}(\cdot)$ *such that*

$$n^{1/\alpha}(\hat{\boldsymbol{\phi}}_{ML} - \boldsymbol{\phi}_0) \xrightarrow{d} \boldsymbol{\zeta}^*$$

and

$$n^{1/2}\left(\begin{pmatrix}\hat{\theta}_{p+1,ML}\\ \hat{\alpha}_{ML}\end{pmatrix}-\begin{pmatrix}\sigma/|\phi_{0r}|\\ \alpha\end{pmatrix}\right)\xrightarrow{d}\mathbf{Y}\sim N(\mathbf{0},\mathbf{I}^{-1}),$$

where ζ^* *is the unique maximizer of* $W(\cdot)$ *and* ζ^* *and* $\mathbf{Y}$ *are independent.*

Proof: Since $f_\sigma(\cdot;\alpha)$ is unimodal, $\ln f_\sigma(\cdot;\alpha)$ is strictly concave in a neighborhood of zero, and so $W(\cdot)$ has a unique maximum almost surely by Remark 2 in Davis, Knight, and Liu (1992). Because

$$\mathcal{L}(\boldsymbol{\theta})-\mathcal{L}(\boldsymbol{\theta}_0)=W_n\left(n^{1/\alpha}(\boldsymbol{\phi}-\boldsymbol{\phi}_0),n^{1/2}(\theta_{p+1}-\sigma/|\phi_{0r}|,\theta_{p+2}-\alpha)'\right)$$

and $\mathbf{Y}=\mathbf{I}^{-1}\mathbf{N}$ uniquely maximizes $\mathbf{v}'\mathbf{N}-2^{-1}\mathbf{v}'\mathbf{I}\mathbf{v}$, the result follows by Remark 1 in Davis, Knight, and Liu (1992). □

Note that $n^{1/\alpha}/a_n\to$ *constant* when Z_1 is stable (see Gnedenko and Kolmogorov, 1968). Therefore, in this case, $\hat{\boldsymbol{\phi}}_{LAD}$ and $\hat{\boldsymbol{\phi}}_{ML}$ converge in distribution at the same rate.

4.4 Numerical Results

4.4.1 Simulation Study

In this section, we describe a simulation experiment to study the behavior of the LAD estimator for finite samples. For each of the 1000 replicates, we simulated all-pass data from the model (4.1) with $r=1$, $\phi_{01}=0.5$, and Students' t noise with $1\le\nu<2$ degrees of freedom and then found $\hat{\phi}_{LAD}$. To diminish the possibility of the optimizer being trapped at local minima, we used 250 starting values for each of the replicates. These initial values for ϕ were uniformly distributed on $(-1,1)$. We evaluated $\sum_{t=1}^{n-p}|z_t(\cdot)|$ at each of the 250 candidate values and reduced the collection of initial values to the nine with the smallest values of $\sum_{t=1}^{n-p}|z_t(\cdot)|$ plus 0.5. Optimized values were found by implementing the Hooke and Jeeves (1961) algorithm and using the ten initial values as starting points. The optimized value for which $\sum_{t=1}^{n-p}|z_t(\cdot)|$ was smallest was chosen to be $\hat{\phi}_{LAD}$. Results of the simulations appear in Table 4.1, where the empirical means and standard deviations for the LAD estimates are shown. Analogous results are given in Table 1 of Breidt, Davis, and Trindade (2001)

ν	n	Empirical mean	std.dev.
1.0	200	0.4981	0.0283
	1000	0.5000	0.0014
1.5	200	0.4988	0.0201
	1000	0.4997	0.0061

Table 4.1: Empirical means and standard deviations for LAD estimates of the all-pass model parameter when $r = 1$, $\phi_{01} = 0.5$, and the noise distribution is t with ν degrees of freedom.

for the case when $r = 1$, $\phi_{01} = 0.5$, and the noise distribution has three degrees of freedom. Based on a comparison of the results, it appears that LAD estimates tend to be much less disperse when the noise distribution has infinite variance.

Unfortunately, it is difficult to compare the finite sample results with the limiting distribution because a closed-form expression for the minimizer of $V(\cdot)$ in (4.8) is not available. For M-estimates for infinite variance autoregressive processes, Davis and Wu (1997) show that the bootstrap can be used to obtain consistent estimates of values of the limiting distribution. It would be beneficial to consider extending these bootstrap results to all-pass processes.

4.4.2 Noncausal Autoregressive Modeling

As discussed in Breidt, Davis, and Trindade (2001), suppose the series $\{X_t\}$ follows the autoregressive model

$$\phi_c(B)\phi_{nc}(B)X_t = Z_t,$$

where $\{Z_t\}$ is iid noise, the r_1 roots of $\phi_c(z)$ lie outside the unit circle in the complex plane, and the r_2 roots of $\phi_{nc}(z)$ lie inside the unit circle. If $\phi_{nc}^c(z)$ denotes the causal r_2th order polynomial whose roots are the reciprocals of the roots of $\phi_{nc}(z)$ and $\{X_t\}$ is mistakenly modeled with the second-order equivalent causal representation

$$\phi_c(B)\phi_{nc}^c(B)X_t = U_t, \tag{4.23}$$

then $\{U_t\}$ satisfies

$$U_t = \frac{\phi_{nc}^c(B)}{\phi_{nc}(B)} Z_t = \frac{\phi_{nc}^c(B)}{-\phi_{nc,r_2} B^{r_2} \phi_{nc}^c(B^{-1})} Z_t,$$

where ϕ_{nc,r_2} is the coefficient of $-B^{r_2}$ in $\phi_{nc}(B)$. Therefore, $\{U_t\}$ is a purely noncausal all-pass time series and $\{U_{-t}\}$ is a causal all-pass time series. It follows that it is not necessary to examine all combinations of roots inside and outside the unit circle when fitting noncausal autoregressive models. A causal model can be fit to the data, and then the number of roots inside the unit circle can be determined by fitting an all-pass model to the residuals from the causal fit.

By an application of Theorem 4.2 in Davis and Resnick (1985), if $Z_1 \in D(\alpha)$ with $\alpha \in (0,2)$ and

$$X_t = \phi_c^{-1}(B)\phi_{nc}^{-1}(B)Z_t = \sum_{j=-\infty}^{\infty} \pi_j Z_{t-j},$$

then

$$\frac{\sum_{t=h+1}^{n} X_t X_{t-h}}{\sum_{t=1}^{n} X_t^2} \xrightarrow{P} \frac{\sum_{j=-\infty}^{\infty} \pi_j \pi_{j-h}}{\sum_{j=-\infty}^{\infty} \pi_j^2} \quad \forall h \in \{0,1,\ldots\}.$$

Also, if

$$\phi_c^{-1}(z)\left(\phi_{nc}^c(z)\right)^{-1} = \sum_{j=0}^{\infty} \varsigma_j z^j,$$

then

$$\frac{\sum_{j=-\infty}^{\infty} \pi_j \pi_{j-h}}{\sum_{j=h}^{\infty} \varsigma_j \varsigma_{j-h}} = constant \quad \forall h \in \{0,1,\ldots\}$$

for some constant that does not depend on h. Consequently, even if $Z_1 \in D(\alpha)$ with $\alpha \in (0,2)$ and $\{X_t\}$ has infinite second-order moments, the model (4.23) can be fit to the data using least squares estimation.

Figure 4.1 shows the volumes of Microsoft stock traded daily from 10/13/1997 to 4/9/1998. In Calder (1998), an autoregressive model with stable, infinite variance noise was fit to the mean-corrected data, $\{X_t\}$, using maximum likelihood. For each model order considered, both causal and noncausal autoregressive models were examined, and the maximizer of the log-likelihood was found. Because the estimated first-order model, which was noncausal, minimized the Akaike information criterion for order selection, this fitted model was chosen to describe the data. To reduce the number of likelihood evaluations, all-pass models could have been used to fit the autoregressive models. For

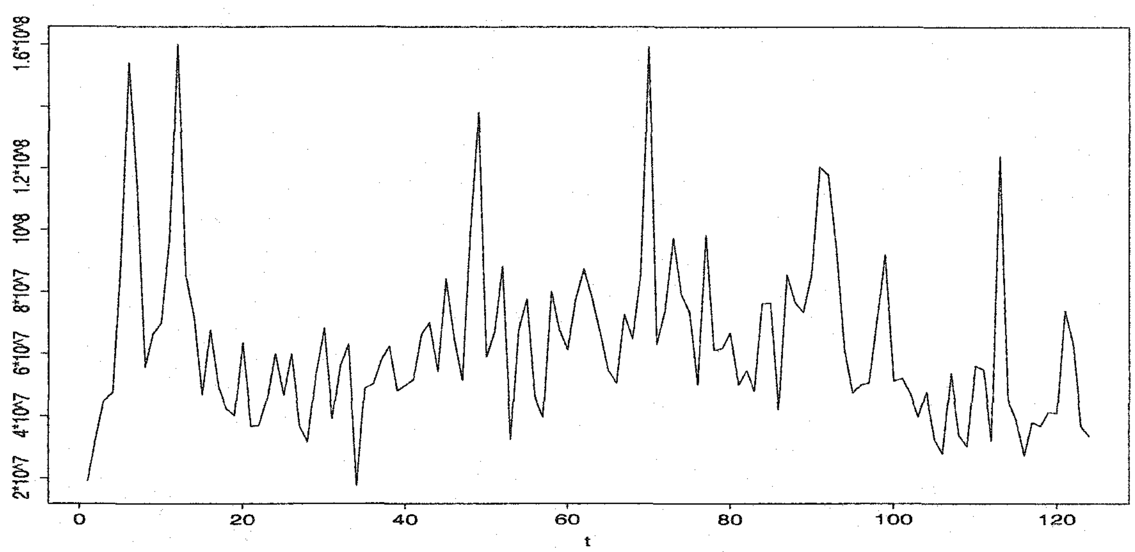

Figure 4.1: Volumes of Microsoft stock traded daily from 10/13/1997 to 4/9/1998.

example, if a causal, first-order autoregressive model is fit to $\{X_t\}$ using least squares, the parameter estimate is 0.405 and the residuals from this fitted model, $\{\hat{U}_t\}$, can be obtained. If a causal all-pass model of order one is fit to $\{\hat{U}_{-t}\}$, the parameter estimate is 0.349. Since $\{\hat{U}_{-t}\}$ seems to follow a causal all-pass model of order one, it appears that the best first-order autoregressive model for $\{X_t\}$ is noncausal. A noncausal autoregressive model of order one can then be fit directly to the data. Using maximum likelihood, Calder (1998) obtained the estimates

$$\hat{\phi} = 2.62, \quad \hat{\alpha} = 1.36, \quad \text{and} \quad \hat{\beta} = 0.48.$$

Higher order noncausal models can be fit in a similar fashion. Note that the bootstrap could be used to estimate the sampling distribution for the all-pass parameter estimator (see Davis and Wu, 1997).

4.5 Additional Results

This section contains proofs of the lemmas used to confirm the results of Section 4.3.

Lemma 1 *Under the conditions of Theorem 1,*

$$\lim_{\lambda\to 0+}\lim_{M\to\infty}\limsup_{n\to\infty} P\left(\bigcup_{t=1}^{n-p}\{(|z_t|>M)\cap(|U_{t,n}^-|>\lambda)\}\right)=0,$$

$$\lim_{\lambda\to 0+}\lim_{M\to\infty}\limsup_{n\to\infty} P\left(\bigcup_{t=1}^{n-p}\{(|z_t|>M)\cap(|U_{t,n}^+|>\lambda)\}\right)=0,$$

and

$$\lim_{\lambda\to 0+}\lim_{M\to\infty}\limsup_{n\to\infty} P\left(\bigcup_{t=1}^{n-p}\{(|z_t|>M)\cap(|U_{t,n}^-|>\lambda)\cap(|U_{t,n}^+|>\lambda)\}\right)=0,$$

where $U_{t,n}^-$ and $U_{t,n}^+$ are defined in (4.14).

Proof: The first result holds because, for any $\lambda>0$,

$$\begin{aligned}
&\lim_{M\to\infty}\limsup_{n\to\infty} \mathrm{P}\left(\bigcup_{t=1}^{n-p}\{(|z_t|>M)\cap(|U_{t,n}^-|>\lambda)\}\right)\\
&\leq \lim_{M\to\infty}\limsup_{n\to\infty} n\mathrm{P}\left(\{|z_1|>M\}\cap\{|U_{1,n}^-|>\lambda\}\right)\\
&= \left(\lim_{M\to\infty}\mathrm{P}(|z_1|>M)\right)\lim_{n\to\infty} n\mathrm{P}\left(\left|\sum_{j=1}^{\infty}c_j(\mathbf{u})z_{1+j}\right|>a_n\lambda\right)\\
&= \lambda^{-\alpha}|\phi_{0r}|^{-\alpha}\sum_{j=1}^{\infty}|c_j(\mathbf{u})|^{\alpha}\lim_{M\to\infty}\mathrm{P}(|z_1|>M)\\
&= 0.
\end{aligned}$$

Similarly,

$$\lim_{\lambda\to 0+}\lim_{M\to\infty}\limsup_{n\to\infty} \mathrm{P}\left(\bigcup_{t=1}^{n-p}\{(|z_t|>M)\cap(|U_{t,n}^+|>\lambda)\}\right)=0,$$

and the third result follows from the first two. □

Lemma 2 *Under the conditions of Theorem 1, if $\alpha\geq 1$, then, for any $\eta>0$,*

$$\lim_{\lambda\to 0+}\limsup_{n\to\infty} P\left(\left|\sum_{t=1}^{n-p}(U_{t,n}^-+U_{t,n}^+)(I\{z_t>0\}-I\{z_t<0\})I\{|U_{t,n}^-|\leq\lambda\}I\{|U_{t,n}^+|\leq\lambda\}\right|>\eta\right)=0,$$

where $U_{t,n}^-$ and $U_{t,n}^+$ are defined in (4.14).

Proof: Observe that

$$\mathrm{P}\left(\left|\sum_{t=1}^{n-p}(U_{t,n}^-+U_{t,n}^+)(I\{z_t>0\}-I\{z_t<0\})I\{|U_{t,n}^-|\leq\lambda\}I\{|U_{t,n}^+|\leq\lambda\}\right|>\eta\right)$$

$$\leq \mathrm{P}\left(\left|\sum_{t=1}^{n-p} U_{t,n}^-(I\{z_t>0\}-I\{z_t<0\})I\{|U_{t,n}^-|\leq\lambda\}I\{|U_{t,n}^+|\leq\lambda\}\right|>\frac{\eta}{2}\right)$$

$$+\mathrm{P}\left(\left|\sum_{t=1}^{n-p} U_{t,n}^+(I\{z_t>0\}-I\{z_t<0\})I\{|U_{t,n}^-|\leq\lambda\}I\{|U_{t,n}^+|\leq\lambda\}\right|>\frac{\eta}{2}\right)$$

$$\leq \mathrm{P}\left(\left|\sum_{t=1}^{n-p} U_{t,n}^-(I\{z_t>0\}-I\{z_t<0\})I\{|U_{t,n}^-|\leq\lambda\}\right|>\frac{\eta}{4}\right) \tag{4.24}$$

$$+\mathrm{P}\left(\left|\sum_{t=1}^{n-p} U_{t,n}^-(I\{z_t>0\}-I\{z_t<0\})I\{|U_{t,n}^-|\leq\lambda\}I\{|U_{t,n}^+|>\lambda\}\right|>\frac{\eta}{4}\right) \tag{4.25}$$

$$+\mathrm{P}\left(\left|\sum_{t=1}^{n-p} U_{t,n}^+(I\{z_t>0\}-I\{z_t<0\})I\{|U_{t,n}^+|\leq\lambda\}\right|>\frac{\eta}{4}\right) \tag{4.26}$$

$$+\mathrm{P}\left(\left|\sum_{t=1}^{n-p} U_{t,n}^+(I\{z_t>0\}-I\{z_t<0\})I\{|U_{t,n}^+|\leq\lambda\}I\{|U_{t,n}^-|>\lambda\}\right|>\frac{\eta}{4}\right). \tag{4.27}$$

By the Markov inequality,

$$\mathrm{P}\left(\left|\sum_{t=1}^{n-p} U_{t,n}^-(I\{z_t>0\}-I\{z_t<0\})I\{|U_{t,n}^-|\leq\lambda\}\right|>\frac{\eta}{4}\right)$$
$$\leq \frac{16}{\eta^2}\mathrm{E}\left|\sum_{t=1}^{n-p} U_{t,n}^-(I\{z_t>0\}-I\{z_t<0\})I\{|U_{t,n}^-|\leq\lambda\}\right|^2.$$

Because z_1 has median zero and is independent of $U_{1,n}^-$,

$$\mathrm{E}\left\{U_{1,n}^-(I\{z_1>0\}-I\{z_1<0\})I\{|U_{1,n}^-|\leq\lambda\}\right\}$$
$$=\mathrm{E}\left\{U_{1,n}^- I\{|U_{1,n}^-|\leq\lambda\}\right\}\mathrm{E}\left\{I\{z_1>0\}-I\{z_1<0\}\right\}$$
$$=0,$$

and so

$$\lim_{\lambda\to 0^+}\limsup_{n\to\infty}\mathrm{E}\left|\sum_{t=1}^{n-p} U_{t,n}^-(I\{z_t>0\}-I\{z_t<0\})I\{|U_{t,n}^-|\leq\lambda\}\right|^2=0$$

by Proposition A.2(c) in Davis, Knight, and Liu (1992). Therefore, $\lim_{\lambda\to 0^+}\limsup_{n\to\infty}$ of (4.24) equals zero, and, similarly, $\lim_{\lambda\to 0^+}\limsup_{n\to\infty}$ of (4.26) equals zero. Because (4.25) is bounded above by

$$\mathrm{P}\left(\sum_{t=1}^{n-p}|U_{t,n}^-|I\{|U_{t,n}^-|\leq\lambda^2\}I\{|U_{t,n}^+|>\lambda\}>\frac{\eta}{8}\right)+\mathrm{P}\left(\sum_{t=1}^{n-p}|U_{t,n}^-|I\{|U_{t,n}^-|>\lambda^2\}I\{|U_{t,n}^+|>\lambda\}>\frac{\eta}{8}\right)$$
$$\leq \frac{8}{\eta}n\mathrm{E}\left|U_{1,n}^- I\{|U_{1,n}^-|\leq\lambda^2\}\right|\mathrm{P}(|U_{1,n}^+|>\lambda)+\mathrm{P}\left(\bigcup_{t=1}^{n-p}\{(|U_{t,n}^-|>\lambda^2)\cap(|U_{t,n}^+|>\lambda)\}\right)$$

$$\leq \frac{8}{\eta}\lambda^2 n\mathrm{P}(|U_{1,n}^+| > \lambda) + n\mathrm{P}\left(|U_{1,n}^-| > \lambda^2\right)\mathrm{P}\left(|U_{1,n}^+| > \lambda\right),$$

$$\begin{aligned}
&\lim_{\lambda\to 0+}\limsup_{n\to\infty}\lambda^2 n\mathrm{P}(|U_{1,n}^+| > \lambda) \\
&= \lim_{\lambda\to 0+}\limsup_{n\to\infty}\lambda^2 n\mathrm{P}\left(\left|\sum_{j=1}^{\infty} c_j(\mathbf{u})z_{1-j}\right| > a_n\lambda\right) \\
&= |\phi_{0r}|^{-\alpha}\sum_{j=1}^{\infty}|c_j(\mathbf{u})|^{\alpha}\lim_{\lambda\to 0+}\lim_{n\to\infty}\lambda^{2-\alpha}n\mathrm{P}(|Z_1| > a_n) \\
&= 0
\end{aligned}$$

by (4.4) and (4.10), and, for any $\lambda > 0$,

$$\begin{aligned}
&\limsup_{n\to\infty} n\mathrm{P}\left(|U_{1,n}^-| > \lambda^2\right)\mathrm{P}\left(|U_{1,n}^+| > \lambda\right) \\
&= |\phi_{0r}|^{-\alpha}\sum_{j=1}^{\infty}|c_j(\mathbf{u})|^{\alpha}\lambda^{-\alpha}\lim_{n\to\infty}\mathrm{P}\left(\left|\sum_{j=1}^{\infty} c_j(\mathbf{u})z_{1+j}\right| > a_n\lambda^2\right) \\
&= 0,
\end{aligned}$$

$\lim_{\lambda\to 0+}\limsup_{n\to\infty}$ of (4.25) equals zero. Similarly, $\lim_{\lambda\to 0+}\limsup_{n\to\infty}$ of (4.27) equals zero, and thus the proof is complete. □

Lemma 3 *Under the conditions of Theorem 1, if $\alpha \geq 1$, then, for any $\eta > 0$,*

$$\begin{aligned}
&\lim_{\lambda\to 0+}\limsup_{n\to\infty} P\left(\left|\sum_{t=1}^{n-p}(z_t + U_{t,n})(I\{-U_{t,n} < z_t < 0\} - I\{-U_{t,n} > z_t > 0\})I\{|U_{t,n}^-| \leq \lambda\}I\{|U_{t,n}^+| \leq \lambda\}\right| > \eta\right) \\
&= 0,
\end{aligned}$$

where $U_{t,n}^-$ and $U_{t,n}^+$ are defined in (4.14) and $U_{t,n} = U_{t,n}^- + U_{t,n}^+$.

Proof: We first consider $\alpha > 1$ and give an argument similar to one found in the proof of Theorem 4.1 in Davis, Knight, and Liu (1992). If $F(\cdot)$ and $G(\cdot)$ are the distribution functions of $|z_1|$ and $|\sum_{j\neq 0} c_j(\mathbf{u})z_{1-j}|$ respectively, then, using the Markov inequality,

$$\begin{aligned}
&\mathrm{P}\left(\left|\sum_{t=1}^{n-p}(z_t + U_{t,n})(I\{-U_{t,n} < z_t < 0\} - I\{-U_{t,n} > z_t > 0\})I\{|U_{t,n}^-| \leq \lambda\}I\{|U_{t,n}^+| \leq \lambda\}\right| > \eta\right) \\
&\leq \mathrm{P}\left(\sum_{t=1}^{n-p}(z_t + U_{t,n})(I\{-2\lambda \leq -U_{t,n} < z_t < 0\} - I\{2\lambda \geq -U_{t,n} > z_t > 0\}) > \eta\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq \mathrm{P}\left(a_n^{-1}\sum_{t=1}^{n-p}\left|\sum_{j\neq 0} c_j(\mathbf{u})z_{t-j}\right| I\left\{2\lambda \geq a_n^{-1}\left|\sum_{j\neq 0} c_j(\mathbf{u})z_{t-j}\right| > |z_t| > 0\right\} > \eta\right) \\
&\leq \eta^{-1} n a_n^{-1} \mathrm{E}\left(\left|\sum_{j\neq 0} c_j(\mathbf{u})z_{1-j}\right| I\left\{2\lambda \geq a_n^{-1}\left|\sum_{j\neq 0} c_j(\mathbf{u})z_{1-j}\right| > |z_1| > 0\right\}\right) \\
&\leq \eta^{-1} n a_n^{-1} \int_0^{2\lambda}\int_{a_n z}^{\infty} y\, G(dy)\, F(dz).
\end{aligned}
$$

If

$$H(z) := \int_z^\infty y\, G(dy),$$

then

$$
\begin{aligned}
\lim_{n\to\infty} n a_n^{-1} H(a_n) &= \frac{\alpha}{\alpha-1}\lim_{n\to\infty} n\mathrm{P}\left(\left|\sum_{j\neq 0} c_j(\mathbf{u})z_{1-j}\right| > a_n\right) \\
&= \frac{\alpha}{\alpha-1}|\phi_{0r}|^{-\alpha}\sum_{j\neq 0}|c_j(\mathbf{u})|^\alpha
\end{aligned}
$$

by Karamata's theorem (see Feller, 1971, page 283), and so

$$\limsup_{n\to\infty} \eta^{-1} n a_n^{-1}\int_0^{2\lambda} H(a_n z)\, F(dz) = \eta^{-1}\frac{\alpha}{\alpha-1}|\phi_{0r}|^{-\alpha}\sum_{j\neq 0}|c_j(\mathbf{u})|^\alpha \limsup_{n\to\infty}\int_0^{2\lambda}\frac{H(a_n z)}{H(a_n)}\, F(dz).$$

Also by Karamata's theorem, $H(z) = z^{1-\alpha}L_0(z)$ for $z > 0$, where $L_0(\cdot)$ is slowly varying at ∞. By Potter's theorem (see Bingham, Goldie, and Teugels, 1987, page 25), there exist constants $A_1 > 0$, $0 < \xi < \min\{\alpha-1, 2-\alpha\}$, and $x_0 > 0$ such that

$$0 < \frac{H(xz)}{H(x)} \leq A_1 \max\{z^{1-\alpha-\xi}, z^{1-\alpha+\xi}\}$$

for $xz \geq x_0$ and $x \geq x_0$. In addition, if $0 < xz < x_0$ and $x \geq x_0$, then there exists a constant $A_2 > 0$ such that

$$0 < \frac{H(xz)}{H(x)} \leq \frac{H(0)}{H(x)} \leq \frac{(xz/x_0)^{1-\alpha-\xi}H(0)}{H(x)} = \frac{z^{1-\alpha-\xi}x_0^{\alpha-1+\xi}H(0)}{x^{\alpha-1+\xi}H(x)} \leq A_2 z^{1-\alpha-\xi}$$

since $H(0) < \infty$ and $x^{\alpha-1+\xi}H(x) \to \infty$ as $x \to \infty$. Therefore, because the distribution of Z_1 is continuously differentiable on $[-\kappa, \kappa]$, $\mathrm{E}|z_1|^{1-\alpha-\xi} < \infty$, and thus

$$\lim_{\lambda\to 0^+}\limsup_{n\to\infty}\mathrm{P}\left(\left|\sum_{t=1}^{n-p}(z_t + U_{t,n})(I\{-U_{t,n} < z_t < 0\} - I\{-U_{t,n} > z_t > 0\})I\{|U_{t,n}^-| \leq \lambda\}I\{|U_{t,n}^+| \leq \lambda\}\right| > \eta\right)$$

$$
\begin{aligned}
&\leq \eta^{-1}\frac{\alpha}{\alpha-1}|\phi_{0r}|^{-\alpha}\sum_{j\neq 0}|c_j(\mathbf{u})|^{\alpha}\lim_{\lambda\to 0+}\limsup_{n\to\infty}\int_0^{2\lambda}\frac{H(a_n z)}{H(a_n)}F(dz)\\
&\leq \eta^{-1}\frac{\alpha}{\alpha-1}|\phi_{0r}|^{-\alpha}\sum_{j\neq 0}|c_j(\mathbf{u})|^{\alpha}(A_1\vee A_2)\lim_{\lambda\to 0+}\int_0^{2\lambda}z^{1-\alpha-\xi}F(dz)\\
&= 0.
\end{aligned}
$$

Now, consider $\alpha = 1$ and observe that

$$
\begin{aligned}
&\mathrm{P}\left(\left|\sum_{t=1}^{n-p}(z_t+U_{t,n})(I\{-U_{t,n}<z_t<0\}-I\{-U_{t,n}>z_t>0\})I\{|U_{t,n}^-|\leq\lambda\}I\{|U_{t,n}^+|\leq\lambda\}\right|>\eta\right)\\
&\leq \mathrm{P}\left(a_n^{-1/2}\left[\sum_{t=1}^{n-p}\left|\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}\right|I\left\{2\lambda\geq a_n^{-1}\left|\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}\right|>|z_t|>0\right\}\right]^{1/2}>\eta^{1/2}\right)\\
&\leq \eta^{-1/2}na_n^{-1/2}\mathrm{E}\left(\left|\sum_{j\neq 0}c_j(\mathbf{u})z_{1-j}\right|I\left\{2\lambda\geq a_n^{-1}\left|\sum_{j\neq 0}c_j(\mathbf{u})z_{1-j}\right|>|z_1|>0\right\}\right)^{1/2}\\
&\leq \eta^{-1/2}na_n^{-1/2}\int_0^{2\lambda}\int_{a_n z}^{\infty}y^{1/2}\,G(dy)\,F(dz).
\end{aligned}
$$

If

$$H^*(z) := \int_z^{\infty} y^{1/2}\,G(dy),$$

then

$$
\begin{aligned}
\lim_{n\to\infty} na_n^{-1/2}H^*(a_n) &= 2\lim_{n\to\infty} n\mathrm{P}\left(\left|\sum_{j\neq 0}c_j(\mathbf{u})z_{1-j}\right|>a_n\right)\\
&= 2|\phi_{0r}|^{-1}\sum_{j\neq 0}|c_j(\mathbf{u})|
\end{aligned}
$$

by Karamata's theorem, and thus

$$\limsup_{n\to\infty}\eta^{-1/2}na_n^{-1/2}\int_0^{2\lambda}H^*(a_n z)\,F(dz) = 2\eta^{-1/2}|\phi_{0r}|^{-1}\sum_{j\neq 0}|c_j(\mathbf{u})|\limsup_{n\to\infty}\int_0^{2\lambda}\frac{H^*(a_n z)}{H^*(a_n)}F(dz).$$

Because $H^*(z) = z^{-1/2}L_1(z)$ for $z>0$, where $L_1(\cdot)$ is slowing varying at ∞, there exists a constant $A_3 > 0$ such that

$$
\begin{aligned}
&\lim_{\lambda\to 0+}\limsup_{n\to\infty}\mathrm{P}\left(\left|\sum_{t=1}^{n-p}(z_t+U_{t,n})(I\{-U_{t,n}<z_t<0\}-I\{-U_{t,n}>z_t>0\})I\{|U_{t,n}^-|\leq\lambda\}I\{|U_{t,n}^+|\leq\lambda\}\right|>\eta\right)\\
&\leq 2\eta^{-1/2}|\phi_{0r}|^{-1}\sum_{j\neq 0}|c_j(\mathbf{u})|\lim_{\lambda\to 0+}\limsup_{n\to\infty}\int_0^{2\lambda}\frac{H^*(a_n z)}{H^*(a_n)}F(dz)
\end{aligned}
$$

$$\begin{aligned}
&\le 2\eta^{-1/2}|\phi_{0r}|^{-1}\sum_{j\neq 0}|c_j(\mathbf{u})|A_3\lim_{\lambda\to 0+}\int_0^{2\lambda} z^{-3/4}F(dz) \\
&= 0.
\end{aligned}$$

□

Lemma 4 *Under the conditions of Theorem 1, if $\alpha \geq 1$, then, for any $\eta > 0$,*

$$\lim_{\lambda\to 0+} P\left(\left|\sum_{j\neq 0}\sum_{k=1}^{\infty}\gamma_{k,j}(I\{Z_{k,j}>0\}-I\{Z_{k,j}<0\})I\{|\gamma_{k,j}|\leq\lambda\}\right|>\eta\right)=0,$$

where $\gamma_{k,j} = c_j(\mathbf{u})\delta_k\Gamma_k^{-1/\alpha}$.

Proof: Choose $\epsilon > 0$ and $k^\dagger \in \{1,2,\ldots\}$ such that $k^\dagger > 2/\alpha$. Because $\sum_{j\neq 0}\sum_{k=1}^{k^\dagger-1}|\gamma_{k,j}| < \infty$ almost surely, it suffices to show that

$$\mathrm{P}\left(\left|\sum_{j\neq 0}\sum_{k=k^\dagger}^{\infty}\gamma_{k,j}(I\{Z_{k,j}>0\}-I\{Z_{k,j}<0\})I\{|\gamma_{k,j}|\leq\lambda\}\right|>\frac{\eta}{2}\right)<\epsilon \tag{4.28}$$

for all $\lambda > 0$ sufficiently close to zero. Choose $\lambda_0 > 0$ such that

$$\frac{16}{\eta^2}\sum_{j\neq 0}c_j^2(\mathbf{u})\sum_{k=k^\dagger}^{\infty}\mathrm{E}\left\{\Gamma_k^{-2/\alpha}I\{|c_j(\mathbf{u})|\Gamma_k^{-1/\alpha}\leq\lambda_0\}\right\}<\frac{\epsilon}{2}. \tag{4.29}$$

Possible values for λ_0 exist since

$$\begin{aligned}
\sum_{j\neq 0}c_j^2(\mathbf{u})\sum_{k=k^\dagger}^{\infty}\mathrm{E}\left\{\Gamma_k^{-2/\alpha}\right\} &= \sum_{j\neq 0}c_j^2(\mathbf{u})\sum_{k=k^\dagger}^{\infty}\frac{\Gamma(k-2/\alpha)}{\Gamma(k)} \\
&< (\textit{constant})\sum_{j\neq 0}c_j^2(\mathbf{u})\sum_{k=k^\dagger}^{\infty}k^{-2/\alpha} \\
&< \infty.
\end{aligned}$$

We will show that (4.28) holds when $\lambda = \lambda_0$.

Because $V(\mathbf{u})$ in (4.8) is finite almost surely and $\Gamma_k \to \infty$ almost surely,

$$\left|\sum_{j\neq 0}\sum_{k=k^\dagger}^{\infty}\gamma_{k,j}\left(I\{Z_{k,j}>0\}-I\{Z_{k,j}<0\}\right)I\{|\gamma_{k,j}|\leq\lambda_0\}\right|<\infty$$

almost surely, and hence there exists an $M_0 > 0$ such that $\mathrm{P}(A_\infty) < \epsilon/2$ if

$$A_\infty := \left\{\left|\sum_{j\neq 0}\sum_{k=k^\dagger}^{\infty}\gamma_{k,j}\left(I\{Z_{k,j}>0\}-I\{Z_{k,j}<0\}\right)I\{|\gamma_{k,j}|\leq\lambda_0\}\right|\geq M_0\right\}.$$

For any integer $m \geq k^\dagger$, let

$$A_m = \left\{ \left| \sum_{j \neq 0} \sum_{k=k^\dagger}^{m} \gamma_{k,j} \left(I\{Z_{k,j} > 0\} - I\{Z_{k,j} < 0\} \right) I\{|\gamma_{k,j}| \leq \lambda_0\} \right| \geq M_0 \right\},$$

and observe that, since

$$\{\gamma_{k,j}(I\{Z_{k,j} > 0\} - I\{Z_{k,j} < 0\})I\{|\gamma_{k,j}| \leq \lambda_0\}\}_{k \geq k^\dagger, j \neq 0}$$

is a sequence of uncorrelated random variables with mean zero,

$$\begin{aligned}
&\mathrm{E}\left\{ I\{A_m^c\} \sum_{j \neq 0} \sum_{k=k^\dagger}^{m} \gamma_{k,j}(I\{Z_{k,j} > 0\} - I\{Z_{k,j} < 0\})I\{|\gamma_{k,j}| \leq \lambda_0\} \right\}^2 \\
&\leq \mathrm{E}\left\{ \sum_{j \neq 0} \sum_{k=k^\dagger}^{m} \gamma_{k,j}(I\{Z_{k,j} > 0\} - I\{Z_{k,j} < 0\})I\{|\gamma_{k,j}| \leq \lambda_0\} \right\}^2 \\
&= \sum_{j \neq 0} \sum_{k=k^\dagger}^{m} \mathrm{E}\left\{ \gamma_{k,j}^2 I\{|\gamma_{k,j}| \leq \lambda_0\} \right\} \\
&= \sum_{j \neq 0} c_j^2(\mathbf{u}) \sum_{k=k^\dagger}^{m} \mathrm{E}\left\{ \Gamma_k^{-2/\alpha} I\{|c_j(\mathbf{u})|\Gamma_k^{-1/\alpha} \leq \lambda_0\} \right\},
\end{aligned}$$

where A_m^c is the complement of the set A_m. By the dominated convergence theorem,

$$\begin{aligned}
&\mathrm{E}\left\{ I\{A_\infty^c\} \sum_{j \neq 0} \sum_{k=k^\dagger}^{\infty} \gamma_{k,j}(I\{Z_{k,j} > 0\} - I\{Z_{k,j} < 0\})I\{|\gamma_{k,j}| \leq \lambda_0\} \right\}^2 \\
&= \mathrm{E}\left\{ \lim_{m \to \infty} \left[I\{A_m^c\} \sum_{j \neq 0} \sum_{k=k^\dagger}^{m} \gamma_{k,j}(I\{Z_{k,j} > 0\} - I\{Z_{k,j} < 0\})I\{|\gamma_{k,j}| \leq \lambda_0\} \right]^2 \right\} \\
&= \lim_{m \to \infty} \mathrm{E}\left\{ I\{A_m^c\} \sum_{j \neq 0} \sum_{k=k^\dagger}^{m} \gamma_{k,j}(I\{Z_{k,j} > 0\} - I\{Z_{k,j} < 0\})I\{|\gamma_{k,j}| \leq \lambda_0\} \right\}^2 \\
&\leq \sum_{j \neq 0} c_j^2(\mathbf{u}) \sum_{k=k^\dagger}^{\infty} \mathrm{E}\left\{ \Gamma_k^{-2/\alpha} I\{|c_j(\mathbf{u})|\Gamma_k^{-1/\alpha} \leq \lambda_0\} \right\},
\end{aligned}$$

and so

$$\begin{aligned}
&\mathrm{P}\left(\left| \sum_{j \neq 0} \sum_{k=k^\dagger}^{\infty} \gamma_{k,j} \left(I\{Z_{k,j} > 0\} - I\{Z_{k,j} < 0\} \right) I\{|\gamma_{k,j}| \leq \lambda_0\} \right| > \frac{\eta}{2} \right) \\
&\leq \mathrm{P}(A_\infty) + \mathrm{P}\left(I\{A_\infty^c\} \left| \sum_{j \neq 0} \sum_{k=k^\dagger}^{\infty} \gamma_{k,j} \left(I\{Z_{k,j} > 0\} - I\{Z_{k,j} < 0\} \right) I\{|\gamma_{k,j}| \leq \lambda_0\} \right| > \frac{\eta}{4} \right)
\end{aligned}$$

$$
\begin{aligned}
&< \frac{\epsilon}{2} + \mathrm{P}\left(I\{A_\infty^c\} \left| \sum_{j\neq 0} \sum_{k=k^\dagger}^{\infty} \gamma_{k,j} \left(I\{Z_{k,j} > 0\} - I\{Z_{k,j} < 0\} \right) I\{|\gamma_{k,j}| \le \lambda_0\} \right|^2 > \frac{\eta^2}{16} \right) \\
&\le \frac{\epsilon}{2} + \frac{16}{\eta^2} \mathrm{E} \left\{ I\{A_\infty^c\} \sum_{j\neq 0} \sum_{k=k^\dagger}^{\infty} \gamma_{k,j} \left(I\{Z_{k,j} > 0\} - I\{Z_{k,j} < 0\} \right) I\{|\gamma_{k,j}| \le \lambda_0\} \right\}^2 \\
&\le \frac{\epsilon}{2} + \frac{16}{\eta^2} \sum_{j\neq 0} c_j^2(\mathbf{u}) \sum_{k=k^\dagger}^{\infty} \mathrm{E}\left\{ \Gamma_k^{-2/\alpha} I\{|c_j(\mathbf{u})|\Gamma_k^{-1/\alpha} \le \lambda_0\} \right\} \\
&< \epsilon
\end{aligned}
$$

by (4.29). □

Lemma 5 *Under the conditions of Theorem 1, if $\alpha \ge 1$, then, for any $\eta > 0$,*

$$
\lim_{\lambda \to 0^+} P\left(\left| \sum_{j\neq 0} \sum_{k=1}^{\infty} (Z_{k,j} + \gamma_{k,j})(I\{-\gamma_{k,j} < Z_{k,j} < 0\} - I\{-\gamma_{k,j} > Z_{k,j} > 0\}) I\{|\gamma_{k,j}| \le \lambda\} \right| > \eta \right) = 0,
$$

where $\gamma_{k,j} = c_j(\mathbf{u})\delta_k \Gamma_k^{-1/\alpha}$.

Proof: By the law of the iterated logarithm,

$$
\limsup_{k\to\infty} \frac{|\Gamma_k - k|}{\sqrt{2k \log\log k}} = 1
$$

almost surely. Therefore, since

$$
\begin{aligned}
|\Gamma_k^{-1/\alpha} - k^{-1/\alpha}| &= |\Gamma_k - k| \alpha^{-1} \tilde{k}^{-1-1/\alpha} \\
&\le |\Gamma_k - k| \alpha^{-1} \left(\min\{k, |\, k - |\Gamma_k - k|\,|\} \right)^{-1-1/\alpha},
\end{aligned}
$$

where $\tilde{k}$ lies between Γ_k and k,

$$
\limsup_{k\to\infty} \frac{|\Gamma_k^{-1/\alpha} - k^{-1/\alpha}|}{\sqrt{2}\,\alpha^{-1} k^{-1-\epsilon}} \le \limsup_{k\to\infty} \frac{|\Gamma_k^{-1/\alpha} - k^{-1/\alpha}|}{\alpha^{-1}\sqrt{2k\log\log k}\left(k - \sqrt{2k\log\log k}\right)^{-1-1/\alpha}} \le 1
$$

almost surely for some $\epsilon > 0$. Consequently, $\sum_{k=1}^{\infty} |\Gamma_k^{-1/\alpha} - k^{-1/\alpha}| < \infty$ almost surely, and so the result holds if

$$
\sum_{j\neq 0} \sum_{k=1}^{\infty} (|c_j(\mathbf{u})| k^{-1/\alpha} - |Z_{k,j}|) I\{|c_j(\mathbf{u})| k^{-1/\alpha} > |Z_{k,j}| > 0\} < \infty \tag{4.30}
$$

almost surely. As in the proof of Proposition A.3 in Davis, Knight, and Liu (1992), if $\tilde{F}(\cdot)$ is the distribution function for $|Z_{1,1}|$, then

$$\begin{aligned}
&\mathrm{E}\left\{\sum_{j\neq 0}\sum_{k=1}^{\infty}(|c_j(\mathbf{u})|k^{-1/\alpha} - |Z_{k,j}|)I\{|c_j(\mathbf{u})|k^{-1/\alpha} > |Z_{k,j}| > 0\}\right\} \\
&= \sum_{j\neq 0}\sum_{k=1}^{\infty}\mathrm{E}\left\{(|c_j(\mathbf{u})|k^{-1/\alpha} - |Z_{k,j}|)I\{|c_j(\mathbf{u})|k^{-1/\alpha} > |Z_{k,j}| > 0\}\right\} \\
&= \sum_{j\neq 0}\sum_{k=1}^{\infty}\int_0^{|c_j(\mathbf{u})|k^{-1/\alpha}}(|c_j(\mathbf{u})|k^{-1/\alpha} - z)\,\tilde{F}(dz) \\
&\leq 2\sum_{j\neq 0}\int_{1/2}^{\infty}\int_0^{|c_j(\mathbf{u})|y^{-1/\alpha}}(|c_j(\mathbf{u})|y^{-1/\alpha} - z)\,\tilde{F}(dz)\,dy \\
&= 2\alpha\sum_{j\neq 0}|c_j(\mathbf{u})|^{\alpha}\int_0^{2^{1/\alpha}|c_j(\mathbf{u})|}\int_z^{2^{1/\alpha}|c_j(\mathbf{u})|}(v-z)v^{-\alpha-1}\,dv\,\tilde{F}(dz).
\end{aligned}$$

Because $\tilde{F}$ is continuously differentiable on $[0,\kappa]$, $\mathrm{E}|Z_{1,1}|^{1-\alpha} < \infty$ if $\alpha > 1$ and $\mathrm{E}(\ln|Z_{1,1}|) > -\infty$ if $\alpha = 1$, and thus the expectation is finite. Hence, (4.30) holds almost surely and the proof is complete. □

Lemma 6 *Under the conditions of Theorem 1, if $\alpha \geq 1$, then, for any $\eta > 0$,*

$$\lim_{\lambda\to 0+}\lim_{M\to\infty} P\left(\left|\sum_{j\neq 0}\sum_{k=1}^{\infty}\gamma_{k,j}(I\{Z_{k,j} > 0\} - I\{Z_{k,j} < 0\})\ddot{I}_{k,j}^{\lambda,M}\right| > \eta\right) = 0$$

and

$$\lim_{\lambda\to 0+}\lim_{M\to\infty} P\left(\left|\sum_{j\neq 0}\sum_{k=1}^{\infty}(Z_{k,j} + \gamma_{k,j})(I\{-\gamma_{k,j} < Z_{k,j} < 0\} - I\{-\gamma_{k,j} > Z_{k,j} > 0\})\ddot{I}_{k,j}^{\lambda,M}\right| > \eta\right) = 0,$$

where $\gamma_{k,j} = c_j(\mathbf{u})\delta_k\Gamma_k^{-1/\alpha}$ and $\ddot{I}_{k,j}^{\lambda,M} := I\{|Z_{k,j}| > M\}I\{|\gamma_{k,j}| > \lambda\}$.

Proof: For any $\lambda > 0$,

$$\mathrm{P}\left(\Gamma_k^{-1/\alpha}\sup_{j\in\{1,2,\ldots\}}|c_j(\mathbf{u})| > \lambda \text{ infinitely often}\right) = 0$$

since $\Gamma_k \to \infty$ almost surely. Hence, if

$$K := \sup\left\{k\in\{1,2,\ldots\} : \Gamma_k^{-1/\alpha}\sup_{j\in\{1,2,\ldots\}}|c_j(\mathbf{u})| > \lambda\right\},$$

then $K < \infty$ almost surely and

$$\begin{aligned}|\gamma_{k,j}| &= \left|c_j(\mathbf{u})\Gamma_k^{-1/\alpha}\right| \\ &\le \Gamma_k^{-1/\alpha} \sup_{j\in\{1,2,\ldots\}} |c_j(\mathbf{u})| \\ &\le \lambda\end{aligned}$$

for all $k > K$ and all $j \in \{1, 2, \ldots\}$. Therefore,

$$\begin{aligned}&\left|\sum_{j\neq 0}\sum_{k=1}^{\infty}\gamma_{k,j}(I\{Z_{k,j}>0\} - I\{Z_{k,j}<0\})I\{|Z_{k,j}|>M\}I\{|\gamma_{k,j}|>\lambda\}\right| \\ &\le \sum_{j\neq 0}|c_j(\mathbf{u})|\sum_{k=1}^{K}\Gamma_k^{-1/\alpha}I\{|Z_{k,j}|>M\} \\ &\to 0\end{aligned}$$

almost surely as $M \to \infty$ and

$$\begin{aligned}&\left|\sum_{j\neq 0}\sum_{k=1}^{\infty}(Z_{k,j}+\gamma_{k,j})(I\{-\gamma_{k,j}<Z_{k,j}<0\} - I\{-\gamma_{k,j}>Z_{k,j}>0\})I\{|Z_{k,j}|>M\}I\{|\gamma_{k,j}|>\lambda\}\right| \\ &\le \sum_{j\neq 0}|c_j(\mathbf{u})|\sum_{k=1}^{K}\Gamma_k^{-1/\alpha}I\{|Z_{k,j}|>M\} \\ &\to 0\end{aligned}$$

almost surely as $M \to \infty$, and so the result holds. □

Lemma 7 *Under the conditions of Theorem 1, for any $T > 0$, any $\eta > 0$, and any $\tau > 0$,*

$$\lim_{\epsilon\to 0+}\limsup_{n\to\infty} P\left(\sup_{\substack{\|\mathbf{u}-\mathbf{v}\|\le\epsilon \\ \|\mathbf{u}\|,\|\mathbf{v}\|\le T}} |V_n^*(\mathbf{u}) - V_n^*(\mathbf{v})| > \eta\right) < \tau,$$

where $V_n^*(\cdot)$ is defined in (4.12).

Proof: First consider the case $\alpha < 1$ and observe that, for any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^p$,

$$\begin{aligned}|V_n^*(\mathbf{u}) - V_n^*(\mathbf{v})| &= \left|\sum_{t=1}^{n-p}\left(\left|z_t + a_n^{-1}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}\right| - \left|z_t + a_n^{-1}\sum_{j\neq 0}c_j(\mathbf{v})z_{t-j}\right|\right)\right| \\ &\le a_n^{-1}\sum_{t=1}^{n-p}\left|\sum_{j\neq 0}c_j(\mathbf{u}-\mathbf{v})z_{t-j}\right|\end{aligned}$$

$$= a_n^{-1}\sum_{t=1}^{n-p}\left|(\mathbf{u}-\mathbf{v})'\left[\frac{-z_{t-i}}{\phi_0(B)}+\frac{z_{t+i}}{\phi_0(B^{-1})}\right]_{i=1}^{p}\right|$$
$$\leq \|\mathbf{u}-\mathbf{v}\|a_n^{-1}\sum_{i=1}^{p}\sum_{t=1}^{n-p}\left|\frac{-z_{t-i}}{\phi_0(B)}+\frac{z_{t+i}}{\phi_0(B^{-1})}\right|.$$

Recall that $\{\psi_j\}$ is given by $\sum_{j=0}^{\infty}\psi_j z^j = 1/\phi_0(z)$ with $\psi_j = 0$ if $j < 0$. By an extension of the proof of Theorem 4.2 in Davis and Resnick (1985), it can be shown that

$$a_n^{-1}\sum_{i=1}^{p}\sum_{t=1}^{n-p}\left|\frac{-z_{t-i}}{\phi_0(B)}+\frac{z_{t+i}}{\phi_0(B^{-1})}\right| \leq a_n^{-1}\sum_{t=1}^{n-p}\sum_{j=1}^{\infty}(|\psi_{j-1}|+\cdots+|\psi_{j-p}|)(|z_{t-j}|+|z_{t+j}|)$$
$$\xrightarrow{d} 2|\phi_{0r}|^{-1}\sum_{j=1}^{\infty}(|\psi_{j-1}|+\cdots+|\psi_{j-p}|)\sum_{k=1}^{\infty}\Gamma_k^{-1/\alpha}$$

since $\sum_{k=1}^{\infty}\Gamma_k^{-1/\alpha} < \infty$ almost surely. Therefore, if $\alpha < 1$,

$$|V_n^*(\mathbf{u}) - V_n^*(\mathbf{v})| = \|\mathbf{u}-\mathbf{v}\|\,O_p(1),$$

and so the result must hold.

Now consider the case $\alpha \geq 1$. From (4.18), for any $\lambda > 0$,

$$P\left(\sup_{\substack{\|\mathbf{u}-\mathbf{v}\|\leq\epsilon\\ \|\mathbf{u}\|,\|\mathbf{v}\|\leq T}}|V_n^*(\mathbf{u}) - V_n^*(\mathbf{v})| > \eta\right)$$
$$\leq P\left(\sup_{\substack{\|\mathbf{u}-\mathbf{v}\|\leq\epsilon\\ \|\mathbf{u}\|,\|\mathbf{v}\|\leq T}}\left|\sum_{t=1}^{n-p}\left(\left|z_t + a_n^{-1}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}\right| - |z_t| - \left|z_t + a_n^{-1}\sum_{j\neq 0}c_j(\mathbf{v})z_{t-j}\right| + |z_t|\right)I_{t,n}^{\lambda}\right| > \frac{\eta}{2}\right)$$
$$+P\left(\sup_{\substack{\|\mathbf{u}-\mathbf{v}\|\leq\epsilon\\ \|\mathbf{u}\|,\|\mathbf{v}\|\leq T}} a_n^{-1}\sum_{t=1}^{n-p}\left|\sum_{j\neq 0}c_j(\mathbf{u}-\mathbf{v})z_{t-j}\right|\left|1 - I_{t,n}^{\lambda}\right| > \frac{\eta}{2}\right)$$
$$\leq P\left(\sup_{\substack{\|\mathbf{u}-\mathbf{v}\|\leq\epsilon\\ \|\mathbf{u}\|,\|\mathbf{v}\|\leq T}}\left|a_n^{-1}\sum_{t=1}^{n-p}\sum_{j\neq 0}c_j(\mathbf{u}-\mathbf{v})z_{t-j}\left(I\{z_t>0\} - I\{z_t<0\}\right)I_{t,n}^{\lambda}\right| > \frac{\eta}{6}\right) \tag{4.31}$$
$$+P\left(\sup_{\|\mathbf{u}\|\leq T}\left|2\sum_{t=1}^{n-p}(z_t + U_{t,n}(\mathbf{u}))\left(I\{-U_{t,n}(\mathbf{u}) < z_t < 0\} - I\{-U_{t,n}(\mathbf{u}) > z_t > 0\}\right)I_{t,n}^{\lambda}\right| > \frac{\eta}{6}\right) \tag{4.32}$$
$$+P\left(\sup_{\|\mathbf{v}\|\leq T}\left|2\sum_{t=1}^{n-p}(z_t + U_{t,n}(\mathbf{v}))\left(I\{-U_{t,n}(\mathbf{v}) < z_t < 0\} - I\{-U_{t,n}(\mathbf{v}) > z_t > 0\}\right)I_{t,n}^{\lambda}\right| > \frac{\eta}{6}\right)$$

$$+\mathrm{P}\left(\sup_{\substack{\|\mathbf{u}-\mathbf{v}\|\le\epsilon\\ \|\mathbf{u}\|,\|\mathbf{v}\|\le T}} a_n^{-1}\sum_{t=1}^{n-p}\left|\sum_{j\neq 0}c_j(\mathbf{u}-\mathbf{v})z_{t-j}\right|\left|1-I_{t,n}^{\lambda}\right|>\frac{\eta}{2}\right), \tag{4.33}$$

where

$$\begin{aligned} I_{t,n}^{\lambda} &:= I\left\{\max_{i\in\{1,\dots,p\}}\left|\phi_0^{-1}(B)z_{t-i}\right|\le a_n\lambda\right\} I\left\{\max_{i\in\{1,\dots,p\}}\left|\phi_0^{-1}(B^{-1})z_{t+i}\right|\le a_n\lambda\right\}\\ &= I\left\{\max_{i\in\{1,\dots,p\}}\left|\sum_{j=0}^{\infty}\psi_j z_{t-i-j}\right|\le a_n\lambda\right\} I\left\{\max_{i\in\{1,\dots,p\}}\left|\sum_{j=0}^{\infty}\psi_j z_{t+i+j}\right|\le a_n\lambda\right\}\end{aligned}$$

and

$$U_{t,n}(\mathbf{u}) := a_n^{-1}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j},\ \mathbf{u}\in\mathbb{R}^p.$$

Observe that (4.31) equals

$$\mathrm{P}\left(\sup_{\substack{\|\mathbf{u}-\mathbf{v}\|\le\epsilon\\ \|\mathbf{u}\|,\|\mathbf{v}\|\le T}}\left|a_n^{-1}\sum_{t=1}^{n-p}(\mathbf{u}-\mathbf{v})'\left[\frac{-z_{t-i}}{\phi_0(B)}+\frac{z_{t+i}}{\phi_0(B^{-1})}\right]_{i=1}^{p}\left(I\{z_t>0\}-I\{z_t<0\}\right)I_{t,n}^{\lambda}\right|>\frac{\eta}{6}\right),$$

which is bounded above by

$$\sum_{i=1}^{p}\mathrm{P}\left(2Ta_n^{-1}\left|\sum_{t=1}^{n-p}\left(\frac{-z_{t-i}}{\phi_0(B)}+\frac{z_{t+i}}{\phi_0(B^{-1})}\right)\left(I\{z_t>0\}-I\{z_t<0\}\right)I_{t,n}^{\lambda}\right|>\frac{\eta}{6p}\right). \tag{4.34}$$

Also, (4.32) and (4.33) are bounded above by

$$\begin{aligned}&\mathrm{P}\left(Ta_n^{-1}\sum_{t=1}^{n-p}Y_tI\{Ta_n^{-1}Y_t>|z_t|>0\}I_{t,n}^{\lambda}>\frac{\eta}{12}\right)\\ &\le\left(\frac{12}{\eta}\right)^{\beta}T^{\beta}na_n^{-\beta}\int_0^{\infty}\int_{a_nz/T}^{\infty}y^{\beta}I_{1,n}^{\lambda}\,\tilde{G}(dy)\,F(dz),\end{aligned} \tag{4.35}$$

where

$$Y_t := \sum_{j=1}^{\infty}\left(|\psi_{j-1}|+\cdots+|\psi_{j-p}|\right)\left(|z_{t-j}|+|z_{t+j}|\right),$$

$$\beta=\begin{cases}1/2, & \alpha=1,\\ 1, & \alpha>1,\end{cases}$$

and $F(\cdot)$ and $\tilde{G}(\cdot)$ are the distribution functions for $|z_1|$ and Y_1 respectively. Using an argument similar to the proof of Lemma 2, it can be shown that $\lim_{\lambda\to 0^+}\limsup_{n\to\infty}$ of (4.34) equals zero.

By the proof of Lemma 3, $\limsup_{n\to\infty}$ of (4.35) is finite for all $\lambda > 0$, and so $\lim_{\lambda\to 0+}\limsup_{n\to\infty}$ of (4.35) equals zero. Therefore, there exists a $\lambda_0 > 0$ small enough so that, for all $\epsilon > 0$, $\limsup_{n\to\infty}$ of

$$\mathrm{P}\left(\sup_{\substack{\|\mathbf{u}-\mathbf{v}\|\le\epsilon\\ \|\mathbf{u}\|,\|\mathbf{v}\|\le T}}\left|a_n^{-1}\sum_{t=1}^{n-p}\sum_{j\neq 0}c_j(\mathbf{u}-\mathbf{v})z_{t-j}\left(I\{z_t>0\}-I\{z_t<0\}\right)I_{t,n}^{\lambda_0}\right|>\frac{\eta}{6}\right)$$
$$+\mathrm{P}\left(\sup_{\|\mathbf{u}\|\le T}\left|2\sum_{t=1}^{n-p}\left(z_t+U_{t,n}(\mathbf{u})\right)\left(I\{-U_{t,n}(\mathbf{u})<z_t<0\}-I\{-U_{t,n}(\mathbf{u})>z_t>0\}\right)I_{t,n}^{\lambda_0}\right|>\frac{\eta}{6}\right)$$
$$+\mathrm{P}\left(\sup_{\|\mathbf{v}\|\le T}\left|2\sum_{t=1}^{n-p}\left(z_t+U_{t,n}(\mathbf{v})\right)\left(I\{-U_{t,n}(\mathbf{v})<z_t<0\}-I\{-U_{t,n}(\mathbf{v})>z_t>0\}\right)I_{t,n}^{\lambda_0}\right|>\frac{\eta}{6}\right)$$

is bounded above by $\tau/2$. Consequently, to complete the proof, we show that

$$\lim_{\epsilon\to 0+}\limsup_{n\to\infty}\mathrm{P}\left(\sup_{\substack{\|\mathbf{u}-\mathbf{v}\|\le\epsilon\\ \|\mathbf{u}\|,\|\mathbf{v}\|\le T}} a_n^{-1}\sum_{t=1}^{n-p}\left|\sum_{j\neq 0}c_j(\mathbf{u}-\mathbf{v})z_{t-j}\right|\left|1-I_{t,n}^{\lambda_0}\right|>\frac{\eta}{2}\right)=0.$$

The left-hand side is bounded above by

$$\lim_{\epsilon\to 0+}\limsup_{n\to\infty}\mathrm{P}\left(\epsilon\, a_n^{-1}\sum_{t=1}^{n-p}Y_tI\{Y_t>a_n\lambda_0\}>\frac{\eta}{2}\right)\le\left(\frac{2}{\eta}\right)^{\beta}\lim_{\epsilon\to 0+}\limsup_{n\to\infty}\epsilon^{\beta}na_n^{-\beta}\mathrm{E}\left\{Y_1I\{Y_1>a_n\lambda_0\}\right\}^{\beta},$$

and, by Karamata's theorem,

$$\begin{aligned}
&\lim_{\epsilon\to 0+}\limsup_{n\to\infty}\epsilon^{\beta}na_n^{-\beta}\mathrm{E}\left\{Y_1I\{Y_1>a_n\lambda_0\}\right\}^{\beta}\\
&= \frac{\alpha}{\alpha-\beta}\lambda_0^{\beta}\lim_{\epsilon\to 0+}\epsilon^{\beta}\lim_{n\to\infty}n\mathrm{P}\left(Y_1>a_n\lambda_0\right)\\
&= 2\frac{\alpha}{\alpha-\beta}\lambda_0^{\beta-\alpha}|\phi_{0r}|^{-\alpha}\sum_{j=1}^{\infty}\left(|\psi_{j-1}|+\cdots+|\psi_{j-p}|\right)^{\alpha}\lim_{\epsilon\to 0+}\epsilon^{\beta}\\
&= 0.
\end{aligned}$$

□

Lemma 8 *Under the conditions of Theorem 1,*

$$V_n^*(\cdot)-V_n^{\dagger}(\cdot)=o_p(1)$$

on $C(\mathbb{R}^p)$, where $V_n^*(\cdot)$ and $V_n^{\dagger}(\cdot)$ are defined in (4.12) and (4.13) respectively.

Proof: Let $T > 0$. We begin by showing that $V_n^*(\cdot) - V_n^\dagger(\cdot) = o_p(1)$ on $C([-T,T]^p)$. In particular, for $\epsilon > 0$ and $\eta > 0$, we show that

$$\mathrm{P}\left(\sup_{\mathbf{u}\in[-T,T]^p} |V_n^*(\mathbf{u}) - V_n^\dagger(\mathbf{u})| > \eta\right) < \epsilon \tag{4.36}$$

for all n sufficiently large. As in the proof of Proposition 3.2 in Davis (1996), for any fixed integer $m < n-p$ and any $\mathbf{u} \in \mathbb{R}^p$,

$$\begin{aligned}
|V_n^*(\mathbf{u}) - V_n^\dagger(\mathbf{u})| &= \left|\sum_{t=1}^{n-p}\left(\left|z_t + a_n^{-1}\sum_{j\neq 0} c_j(\mathbf{u})z_{t-j}\right| - |z_t| - \left|z_t(\phi_0) + a_n^{-1}\sum_{j\neq 0} c_j(\mathbf{u})z_{t-j}\right| + |z_t(\phi_0)|\right)\right| \\
&\leq \sum_{t=1}^{n-p-m}\left|\left|z_t + a_n^{-1}\sum_{j\neq 0} c_j(\mathbf{u})z_{t-j}\right| - \left|z_t(\phi_0) + a_n^{-1}\sum_{j\neq 0} c_j(\mathbf{u})z_{t-j}\right|\right| \\
&\quad + \sum_{t=1}^{n-p-m}\Big||z_t| - |z_t(\phi_0)|\Big| + \sum_{t=n-p-m+1}^{n-p}\left|\left|z_t + a_n^{-1}\sum_{j\neq 0} c_j(\mathbf{u})z_{t-j}\right| - |z_t|\right| \\
&\quad + \sum_{t=n-p-m+1}^{n-p}\left|\left|z_t(\phi_0) + a_n^{-1}\sum_{j\neq 0} c_j(\mathbf{u})z_{t-j}\right| - |z_t(\phi_0)|\right| \\
&\leq 2\sum_{t=1}^{n-p-m}|z_t - z_t(\phi_0)| + 2a_n^{-1}\sum_{t=n-p-m+1}^{n-p}\sum_{j\neq 0}|c_j(\mathbf{u})z_{t-j}|.
\end{aligned}$$

Thus, for any $m < n-p$,

$$\begin{aligned}
&\mathrm{P}\left(\sup_{\mathbf{u}\in[-T,T]^p} |V_n^*(\mathbf{u}) - V_n^\dagger(\mathbf{u})| > \eta\right) \\
&\leq \mathrm{P}\left(2\sum_{t=1}^{n-p-m}|z_t - z_t(\phi_0)| > \frac{\eta}{2}\right) + \mathrm{P}\left(\sup_{\mathbf{u}\in[-T,T]^p} 2a_n^{-1}\sum_{t=n-p-m+1}^{n-p}\sum_{j\neq 0}|c_j(\mathbf{u})z_{t-j}| > \frac{\eta}{2}\right).
\end{aligned}$$

Choose $0 < \beta < \min\{\alpha, 1\}$ and observe that $\mathrm{E}|z_1|^\beta < \infty$. Because

$$z_{n-p-t} = \sum_{j=0}^{\infty}\psi_j\left(\phi_0(B^{-1})z_{n-p-t+j}\right) \qquad \text{and} \qquad z_{n-p-t}(\phi_0) = \sum_{j=0}^{t}\psi_j\left(\phi_0(B^{-1})z_{n-p-t+j}\right) \tag{4.37}$$

for $t \in \{0,\ldots,n-p-1\}$, where $\sum_{j=0}^{\infty}\psi_j z^j = 1/\phi_0(z)$, and there exist constants $a > 0$ and $0 < b < 1$ such that $|\psi_j| \leq ab^j$ for all $j \in \{0,1,\ldots\}$, there exists an integer m_0 large enough so that

$$\begin{aligned}
\mathrm{P}\left(2\sum_{t=1}^{n-p-m_0}|z_t - z_t(\phi_0)| > \frac{\eta}{2}\right) &\leq \left(\frac{4}{\eta}\right)^\beta \mathrm{E}\left\{\sum_{t=1}^{n-p-m_0}|z_t - z_t(\phi_0)|\right\}^\beta \\
&\leq \left(\frac{4}{\eta}\right)^\beta \mathrm{E}\left\{\sum_{t=m_0}^{n-p-1}\sum_{j=t+1}^{\infty}|\psi_j|\left|\phi_0(B^{-1})z_{n-p-t+j}\right|\right\}^\beta
\end{aligned}$$

$$
\begin{aligned}
&\le \left(\frac{4}{\eta}\right)^{\beta} \mathrm{E}\left\{\frac{a}{1-b}\sum_{t=1}^{\infty} b^{m_0+t}\left|\phi_0(B^{-1})z_{n-p+t}\right|\right\}^{\beta} \\
&\le \left(\frac{4a}{\eta(1-b)}\right)^{\beta} \mathrm{E}\left|\phi_0(B^{-1})z_1\right|^{\beta} \sum_{t=1}^{\infty} \left(b^{\beta}\right)^{m_0+t} \\
&< \frac{\epsilon}{2}
\end{aligned}
$$

for all $n > m_0 + p$. By (4.7), there exist constants $C > 0$ and $0 < D < 1$ such that

$$\sup_{\mathbf{u}\in[-T,T]^p} |c_j(\mathbf{u})| < CD^{|j|}$$

for all $j \neq 0$. Hence, because $a_n \to \infty$,

$$
\begin{aligned}
\mathrm{P}\left(\sup_{\mathbf{u}\in[-T,T]^p} 2a_n^{-1} \sum_{t=n-p-m_0+1}^{n-p} \sum_{j\neq 0} |c_j(\mathbf{u})z_{t-j}| > \frac{\eta}{2}\right) &\le \mathrm{P}\left(a_n^{-1} \sum_{t=n-p-m_0+1}^{n-p} \sum_{j\neq 0} CD^{|j|}|z_{t-j}| > \frac{\eta}{4}\right) \\
&\le \left(\frac{4}{\eta}\right)^{\beta} \mathrm{E}\left\{a_n^{-1} \sum_{t=n-p-m_0+1}^{n-p} \sum_{j\neq 0} CD^{|j|}|z_{t-j}|\right\}^{\beta} \\
&\le a_n^{-\beta}\left(\frac{4}{\eta}\right)^{\beta} m_0 \mathrm{E}|z_1|^{\beta} \sum_{j\neq 0} C^{\beta}\left(D^{\beta}\right)^{|j|} \\
&< \frac{\epsilon}{2}
\end{aligned}
$$

for all n sufficiently large, and so (4.36) holds for all n sufficiently large. Consequently, $V_n^*(\cdot) - V_n^\dagger(\cdot) = o_p(1)$ on $C([-T,T]^p)$. Because $T > 0$ was arbitrarily chosen, for any compact set $K \subset \mathbb{R}^p$, $V_n^*(\cdot) - V_n^\dagger(\cdot) = o_p(1)$ on $C(K)$, and thus $V_n^*(\cdot) - V_n^\dagger(\cdot) = o_p(1)$ on $C(\mathbb{R}^p)$. □

Now consider the partial derivatives for $z_t(\cdot)$. For an arbitrary, causal autoregressive polynomial $\phi(z)$, define $\varphi(z) = \phi_1 z + \cdots + \phi_p z^p = 1 - \phi(z)$, and define $\varphi_0(z) = 1 - \phi_0(z)$. Note that, for $t \in \{1, \ldots, n-p\}$,

$$\phi(B)X_{t+p} = -z_t(\boldsymbol{\phi}) + \varphi(B^{-1})z_t(\boldsymbol{\phi}),$$

so, if $j \in \{1, \ldots, p\}$, then

$$\frac{\partial}{\partial \phi_j}\left\{\varphi(B^{-1})z_t(\boldsymbol{\phi})\right\} = -X_{t+p-j} + \frac{\partial z_t(\boldsymbol{\phi})}{\partial \phi_j}. \tag{4.38}$$

Also, if $j \in \{1, \ldots, p\}$, then

$$\frac{\partial}{\partial \phi_j}\left\{\varphi(B^{-1})z_t(\boldsymbol{\phi})\right\} = \frac{\partial}{\partial \phi_j}\left\{\phi_1 z_{t+1}(\boldsymbol{\phi}) + \cdots + \phi_p z_{t+p}(\boldsymbol{\phi})\right\}$$

$$= \varphi(B^{-1})\frac{\partial z_t(\phi)}{\partial \phi_j} + z_{t+j}(\phi). \tag{4.39}$$

Equating (4.38) and (4.39) and solving for $\partial z_t(\phi)/\partial\phi_j$, we obtain

$$\frac{\partial z_t(\phi)}{\partial \phi_j} = \frac{1}{\phi(B^{-1})}\left\{X_{t+p-j} + z_{t+j}(\phi)\right\}. \tag{4.40}$$

Evaluating (4.40) at the true value of ϕ and ignoring the effect of recursion initialization, we have

$$\begin{aligned}\frac{\partial z_t(\phi_0)}{\partial \phi_j} &= \frac{1}{\phi_0(B^{-1})}\left\{\frac{-\phi_0(B^{-1})B^p z_{t+p-j}}{\phi_0(B)} + z_{t+j}(\phi_0)\right\}\\ &\simeq \frac{-z_{t-j}}{\phi_0(B)} + \frac{z_{t+j}}{\phi_0(B^{-1})}.\end{aligned} \tag{4.41}$$

Also, for $j, k \in \{1, \ldots, p\}$,

$$\frac{\partial^2 z_t(\phi)}{\partial\phi_j\partial\phi_k} = \frac{1}{\phi^2(B^{-1})}\left\{X_{t+p+j-k} + X_{t+p+k-j} + 2z_{t+j+k}(\phi)\right\}, \tag{4.42}$$

and so

$$\begin{aligned}\frac{\partial^2 z_t(\phi_0)}{\partial\phi_j\partial\phi_k} &= \frac{1}{\phi_0^2(B^{-1})}\left\{X_{t+p+j-k} + X_{t+p+k-j} + 2z_{t+j+k}(\phi_0)\right\}\\ &\simeq \frac{-z_{t+j-k} - z_{t+k-j}}{\phi_0(B^{-1})\phi_0(B)} + \frac{2z_{t+j+k}}{\phi_0^2(B^{-1})}\\ &= -\sum_{m=0}^{\infty}\sum_{\ell=0}^{\infty}\psi_m\psi_\ell(z_{t+j-k-\ell+m} + z_{t+k-j-\ell+m}) + \frac{2z_{t+j+k}}{\phi_0^2(B^{-1})}.\end{aligned}$$

Lemma 9 *Under the conditions of Theorem 1,*

$$V_n(\cdot) - V_n^\dagger(\cdot) = o_p(1)$$

on $C(\mathbb{R}^p)$, where $V_n(\cdot)$ and $V_n^\dagger(\cdot)$ are defined in (4.11) and (4.13) respectively.

Proof: For any $\mathbf{u} \in \mathbb{R}^p$,

$$\begin{aligned}|V_n(\mathbf{u}) - V_n^\dagger(\mathbf{u})| &= \left|\sum_{t=1}^{n-p}\left(\left|z_t(\phi_0 + a_n^{-1}\mathbf{u})\right| - \left|z_t(\phi_0) + a_n^{-1}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}\right|\right)\right|\\ &\le \sum_{t=1}^{n-p}\left|z_t(\phi_0 + a_n^{-1}\mathbf{u}) - z_t(\phi_0) - a_n^{-1}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}\right|\\ &= \sum_{t=1}^{n-p}\left|a_n^{-1}\mathbf{u}'\frac{\partial z_t(\phi_0)}{\partial\phi} + \frac{a_n^{-2}}{2}\mathbf{u}'\frac{\partial^2 z_t(\phi_{t,n}^*(\mathbf{u}))}{\partial\phi\partial\phi'}\mathbf{u} - a_n^{-1}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}\right|\\ &\le a_n^{-1}\sum_{t=1}^{n-p}\left|\mathbf{u}'\frac{\partial z_t(\phi_0)}{\partial\phi} - \sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}\right| + \frac{a_n^{-2}}{2}\sum_{t=1}^{n-p}\left|\mathbf{u}'\frac{\partial^2 z_t(\phi_{t,n}^*(\mathbf{u}))}{\partial\phi\partial\phi'}\mathbf{u}\right|,\end{aligned}$$

where $\phi^*_{t,n}(\mathbf{u})$ lies between ϕ_0 and $\phi_0 + a_n^{-1}\mathbf{u}$ for all $t \in \{1, \ldots, n-p\}$. Fix $T > 0$. In order to prove that $V_n(\cdot) - V_n^\dagger(\cdot) = o_p(1)$ on $C(K_T)$, where $K_T = \{\mathbf{u} \in \mathbb{R}^p : \|\mathbf{u}\| \leq T\}$, we show that

$$\begin{aligned}
&\mathrm{P}\left(\sup_{\|\mathbf{u}\|\leq T} |V_n(\mathbf{u}) - V_n^\dagger(\mathbf{u})| > \eta\right) \\
&\leq \mathrm{P}\left(\sup_{\|\mathbf{u}\|\leq T} a_n^{-1}\sum_{t=1}^{n-p}\left|\mathbf{u}'\frac{\partial z_t(\phi_0)}{\partial \phi} - \sum_{j\neq 0} c_j(\mathbf{u})z_{t-j}\right| > \frac{\eta}{2}\right) + \mathrm{P}\left(\sup_{\|\mathbf{u}\|\leq T} \frac{a_n^{-2}}{2}\sum_{t=1}^{n-p}\left|\mathbf{u}'\frac{\partial^2 z_t(\phi^*_{t,n}(\mathbf{u}))}{\partial\phi\partial\phi'}\mathbf{u}\right| > \frac{\eta}{2}\right) \\
&\to 0
\end{aligned}$$

for some $\eta > 0$. Choose $\beta > 0$ so that

$$\beta = \begin{cases} \frac{3}{4}\alpha, & \alpha \leq 1, \\ 1, & \alpha > 1, \end{cases}$$

and observe that $\mathrm{E}|z_1|^\beta < \infty$. Therefore, from (4.7) and (4.41),

$$\begin{aligned}
&\mathrm{P}\left(\sup_{\|\mathbf{u}\|\leq T} a_n^{-1}\sum_{t=1}^{n-p}\left|\mathbf{u}'\frac{\partial z_t(\phi_0)}{\partial \phi} - \sum_{j\neq 0} c_j(\mathbf{u})z_{t-j}\right| > \frac{\eta}{2}\right) \\
&= \mathrm{P}\left(\sup_{\|\mathbf{u}\|\leq T} a_n^{-1}\sum_{t=1}^{n-p}\left|\mathbf{u}'\frac{\partial z_t(\phi_0)}{\partial \phi} - \mathbf{u}'\left[\frac{-z_{t-i}}{\phi_0(B)} + \frac{z_{t+i}}{\phi_0(B^{-1})}\right]_{i=1}^p\right| > \frac{\eta}{2}\right) \\
&\leq \mathrm{P}\left(a_n^{-1}T\sum_{t=1}^{n-p}\left\|\left[\frac{z_{t+i}(\phi_0)}{\phi_0(B^{-1})} - \frac{z_{t+i}}{\phi_0(B^{-1})}\right]_{i=1}^p\right\| > \frac{\eta}{2}\right) \\
&\leq \mathrm{P}\left(a_n^{-1}T\sum_{i=1}^{p}\sum_{t=1}^{n-p}\left|\frac{z_{t+i}(\phi_0)}{\phi_0(B^{-1})} - \frac{z_{t+i}}{\phi_0(B^{-1})}\right| > \frac{\eta}{2}\right) \\
&\leq \left(\frac{2}{\eta}\right)^\beta \mathrm{E}\left\{a_n^{-1}T\sum_{i=1}^{p}\sum_{t=1}^{n-p}\left|\frac{z_{t+i}(\phi_0)}{\phi_0(B^{-1})} - \frac{z_{t+i}}{\phi_0(B^{-1})}\right|\right\}^\beta \\
&\leq \left(\frac{2}{\eta}\right)^\beta T^\beta a_n^{-\beta}\sum_{i=1}^{p}\sum_{t=1}^{n-p}\mathrm{E}\left|\frac{z_{t+i}(\phi_0)}{\phi_0(B^{-1})} - \frac{z_{t+i}}{\phi_0(B^{-1})}\right|^\beta. \qquad (4.43)
\end{aligned}$$

By (4.37),

$$\sum_{i=1}^{p}\sum_{t=1}^{n-p}\mathrm{E}\left|\frac{z_{t+i}(\phi_0)}{\phi_0(B^{-1})} - \frac{z_{t+i}}{\phi_0(B^{-1})}\right|^\beta = O(1),$$

and so, since $a_n \to \infty$, (4.43) is $o(1)$. Because

$$\|\phi^*_{t,n}(\mathbf{u}) - \phi_0\| \leq a_n^{-1}\|\mathbf{u}\| \leq a_n^{-1}T,$$

from (4.42), there exist constants $\tilde{C} > 0$ and $0 < \tilde{D} < 1$ such that

$$\sup_{\|\mathbf{u}\|\leq T}\left|\mathbf{u}'\frac{\partial^2 z_t(\phi^*_{t,n}(\mathbf{u}))}{\partial\phi\partial\phi'}\mathbf{u}\right| \leq \sum_{j=-\infty}^{\infty}\tilde{C}\tilde{D}^{|j|}|z_{t-j}|$$

for all n sufficiently large and all $t \in \{1, \ldots, n-p\}$. If follows that, for all n sufficiently large,

$$\begin{aligned} \mathrm{P}\left(\sup_{\|\mathbf{u}\|\le T} a_n^{-2}\sum_{t=1}^{n-p}\left|\mathbf{u}'\frac{\partial^2 z_t(\boldsymbol{\phi}^*_{t,n}(\mathbf{u}))}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}\mathbf{u}\right| > \eta\right) &\le \mathrm{P}\left(a_n^{-2}\sum_{t=1}^{n-p}\sum_{j=-\infty}^{\infty}\tilde{C}\tilde{D}^{|j|}|z_{t-j}| > \eta\right) \\ &\le \eta^{-\beta}a_n^{-2\beta}n\mathrm{E}|z_1|^{\beta}\sum_{j=-\infty}^{\infty}\tilde{C}\left(\tilde{D}^{\beta}\right)^{|j|}. \end{aligned} \tag{4.44}$$

By (4.9), $a_n^{-2\beta}n \to 0$, and so (4.44) is $o(1)$. Thus, since $T > 0$ was arbitrarily chosen, $V_n(\cdot) - V_n^\dagger(\cdot) = o_p(1)$ on $C(K)$ for every compact set $K \subset \mathbb{R}^p$, and hence the proof is complete. □

Lemma 10 *Under the conditions of Theorem 2,*

$$W_n^*(\cdot,\cdot) - W_n^\dagger(\cdot,\cdot) = o_p(1)$$

on $C(\mathbb{R}^{p+2})$, where $W_n^*(\cdot,\cdot)$ and $W_n^\dagger(\cdot,\cdot)$ are defined in (4.21) and (4.22) respectively.

Proof: Let $T > 0$. We will first show that $W_n^*(\cdot,\cdot) - W_n^\dagger(\cdot,\cdot) = o_p(1)$ on $C([-T,T]^{p+2})$. In particular, for $\epsilon > 0$ and $\eta > 0$, we show that

$$\mathrm{P}\left(\sup_{\mathbf{u}\in[-T,T]^p,\,\mathbf{v}\in[-T,T]^2}|W_n^*(\mathbf{u},\mathbf{v}) - W_n^\dagger(\mathbf{u},\mathbf{v})| > \eta\right) < \epsilon \tag{4.45}$$

for all n sufficiently large. For any fixed integer $m < n-p$, any $\mathbf{u} \in \mathbb{R}^p$, and any $\mathbf{v} = (v_1, v_2)' \in \mathbb{R}^2$,

$$\begin{aligned} &|W_n^*(\mathbf{u},\mathbf{v}) - W_n^\dagger(\mathbf{u},\mathbf{v})| \\ &\le \sum_{t=1}^{n-p-m}\left|\ln f\left(\frac{z_t + n^{-1/\alpha}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}}{\sigma/|\phi_{0r}| + v_1/n^{1/2}};\alpha + \frac{v_2}{n^{1/2}}\right)\right. \\ &\qquad\left. - \ln f\left(\frac{z_t(\boldsymbol{\phi}_0) + n^{-1/\alpha}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}}{\sigma/|\phi_{0r}| + v_1/n^{1/2}};\alpha + \frac{v_2}{n^{1/2}}\right)\right| &(4.46) \\ &\quad + \sum_{t=1}^{n-p-m}\left|\ln f\left(\frac{z_t}{\sigma/|\phi_{0r}|};\alpha\right) - \ln f\left(\frac{z_t(\boldsymbol{\phi}_0)}{\sigma/|\phi_{0r}|};\alpha\right)\right| &(4.47) \\ &\quad + \sum_{t=n-p-m+1}^{n-p}\left|\ln\left[\left(\frac{\sigma}{|\phi_{0r}|} + \frac{v_1}{n^{1/2}}\right)^{-1} f\left(\frac{z_t + n^{-1/\alpha}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}}{\sigma/|\phi_{0r}| + v_1/n^{1/2}};\alpha + \frac{v_2}{n^{1/2}}\right)\right]\right. \\ &\qquad\left. - \ln\left[\left(\frac{\sigma}{|\phi_{0r}|}\right)^{-1} f\left(\frac{z_t}{\sigma/|\phi_{0r}|};\alpha\right)\right]\right| &(4.48) \\ &\quad + \sum_{t=n-p-m+1}^{n-p}\left|\ln\left[\left(\frac{\sigma}{|\phi_{0r}|} + \frac{v_1}{n^{1/2}}\right)^{-1} f\left(\frac{z_t(\boldsymbol{\phi}_0) + n^{-1/\alpha}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}}{\sigma/|\phi_{0r}| + v_1/n^{1/2}};\alpha + \frac{v_2}{n^{1/2}}\right)\right]\right. \\ &\qquad\left. - \ln\left[\left(\frac{\sigma}{|\phi_{0r}|}\right)^{-1} f\left(\frac{z_t(\boldsymbol{\phi}_0)}{\sigma/|\phi_{0r}|};\alpha\right)\right]\right|. &(4.49) \end{aligned}$$

From Calder (1998), as $z \to \infty$,

$$\frac{\partial \ln f(z;\theta)}{\partial z} = -(\theta+1)z^{-1} + O(z^{-\theta-1}). \tag{4.50}$$

Hence, for all n sufficiently large and all $|v_1|, |v_2| \le T$, there exist constants $K_1, K_2 > 0$ such that

$$\left|\frac{\partial \ln f(z;\alpha + v_2/n^{1/2})}{\partial z}\right| \le K_1 \quad \forall z \in \mathbb{R} \qquad \text{and} \qquad \frac{\sigma}{|\phi_{0r}|} + \frac{v_1}{n^{1/2}} \ge \frac{\sigma}{|\phi_{0r}|} - K_2 > 0,$$

and therefore (4.46) and (4.47) are both bounded above by

$$\frac{K_1}{\sigma/|\phi_{0r}| - K_2} \sum_{t=1}^{n-p-m} \left| z_t - z_t(\phi_0) \right|.$$

Also from Calder (1998), as $z \to \infty$,

$$\frac{\partial \ln f(z;\theta)}{\partial \theta} = \frac{(\partial h(\theta)/\partial\theta)}{h(\theta)} + \ln z + O(z^{-\theta}\ln z),$$

where

$$h(\theta) := \Gamma(\theta+1)\sin\left(\frac{\pi}{2}\theta\right).$$

Thus, for all n sufficiently large and all $|v_2| \le T$, there exists a $K_3 > 0$ such that

$$\left| \left.\frac{\partial \ln f(z;\theta)}{\partial \theta}\right|_{\theta = \alpha + v_2/n^{1/2}} \right| \le K_3 + |\ln|z|| \quad \forall z \in \mathbb{R},$$

and so, when $|v_1| \le T$, (4.48) is bounded above by

$$\begin{aligned}
&\sum_{t=n-p-m+1}^{n-p} \left| \ln\left[\left(\frac{\sigma}{|\phi_{0r}|} + \frac{v_1}{n^{1/2}}\right)^{-1} f\left(\frac{z_t + n^{-1/\alpha}\sum_{j\neq 0} c_j(\mathbf{u}) z_{t-j}}{\sigma/|\phi_{0r}| + v_1/n^{1/2}}; \alpha + \frac{v_2}{n^{1/2}}\right)\right]\right. \\
&\qquad \left. - \ln\left[\left(\frac{\sigma}{|\phi_{0r}|}\right)^{-1} f\left(\frac{z_t + n^{-1/\alpha}\sum_{j\neq 0} c_j(\mathbf{u}) z_{t-j}}{\sigma/|\phi_{0r}| + v_1/n^{1/2}}; \alpha + \frac{v_2}{n^{1/2}}\right)\right]\right| \\
&\quad + \sum_{t=n-p-m+1}^{n-p} \left| \ln\left[\left(\frac{\sigma}{|\phi_{0r}|}\right)^{-1} f\left(\frac{z_t + n^{-1/\alpha}\sum_{j\neq 0} c_j(\mathbf{u}) z_{t-j}}{\sigma/|\phi_{0r}| + v_1/n^{1/2}}; \alpha + \frac{v_2}{n^{1/2}}\right)\right]\right. \\
&\qquad \left. - \ln\left[\left(\frac{\sigma}{|\phi_{0r}|}\right)^{-1} f\left(\frac{z_t + n^{-1/\alpha}\sum_{j\neq 0} c_j(\mathbf{u}) z_{t-j}}{\sigma/|\phi_{0r}| + v_1/n^{1/2}}; \alpha\right)\right]\right| \\
&\quad + \sum_{t=n-p-m+1}^{n-p} \left| \ln\left[\left(\frac{\sigma}{|\phi_{0r}|}\right)^{-1} f\left(\frac{z_t + n^{-1/\alpha}\sum_{j\neq 0} c_j(\mathbf{u}) z_{t-j}}{\sigma/|\phi_{0r}| + v_1/n^{1/2}}; \alpha\right)\right]\right. \\
&\qquad \left. - \ln\left[\left(\frac{\sigma}{|\phi_{0r}|}\right)^{-1} f\left(\frac{z_t}{\sigma/|\phi_{0r}|}; \alpha\right)\right]\right|
\end{aligned}$$

$$
\begin{aligned}
\leq\ & m\left|\ln\left(\frac{\sigma}{|\phi_{0r}|}+\frac{v_1}{n^{1/2}}\right)^{-1}-\ln\left(\frac{\sigma}{|\phi_{0r}|}\right)^{-1}\right|+n^{-1/2}mK_3T \\
& +n^{-1/2}T\sum_{t=n-p-m+1}^{n-p}\left|\ln\left|\frac{z_t+n^{-1/\alpha}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}}{\sigma/|\phi_{0r}|+v_1/n^{1/2}}\right|\right| \\
& +K_1\sum_{t=n-p-m+1}^{n-p}\left|\frac{z_t+n^{-1/\alpha}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}}{\sigma/|\phi_{0r}|+v_1/n^{1/2}}-\frac{z_t}{\sigma/|\phi_{0r}|}\right| \\
\leq\ & n^{-1/2}\frac{mT}{\sigma/|\phi_{0r}|-K_2}+n^{-1/2}mK_3T+n^{-1/2}T\sum_{t=n-p-m+1}^{n-p}\left|\ln\left|\frac{z_t+n^{-1/\alpha}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}}{\sigma/|\phi_{0r}|+v_1/n^{1/2}}\right|\right| \\
& +n^{-1/\alpha}\frac{K_1}{\sigma/|\phi_{0r}|-K_2}\sum_{t=n-p-m+1}^{n-p}\sum_{j\neq 0}|c_j(\mathbf{u})z_{t-j}|+n^{-1/2}\frac{K_1T}{(\sigma/|\phi_{0r}|-K_2)^2}\sum_{t=n-p-m+1}^{n-p}|z_t|.
\end{aligned}
$$

Similarly, for all n sufficiently large and all $|v_1|, |v_2| \leq T$, (4.49) is bounded above by

$$
\begin{aligned}
& n^{-1/2}\frac{mT}{\sigma/|\phi_{0r}|-K_2}+n^{-1/2}mK_3T+n^{-1/2}T\sum_{t=n-p-m+1}^{n-p}\left|\ln\left|\frac{z_t(\phi_0)+n^{-1/\alpha}\sum_{j\neq 0}c_j(\mathbf{u})z_{t-j}}{\sigma/|\phi_{0r}|+v_1/n^{1/2}}\right|\right| \\
+\ & n^{-1/\alpha}\frac{K_1}{\sigma/|\phi_{0r}|-K_2}\sum_{t=n-p-m+1}^{n-p}\sum_{j\neq 0}|c_j(\mathbf{u})z_{t-j}|+n^{-1/2}\frac{K_1T}{(\sigma/|\phi_{0r}|-K_2)^2}\sum_{t=n-p-m+1}^{n-p}|z_t(\phi_0)|.
\end{aligned}
$$

Therefore, using an argument similar to one used in proof of Lemma 8, it can be shown that (4.45) holds for all n sufficiently large. Consequently, $W_n^*(\cdot,\cdot) - W_n^\dagger(\cdot,\cdot) = o_p(1)$ on $C(K)$ for any compact set $K \subset \mathbb{R}^{p+2}$, and so the result holds. □

Lemma 11 *Under the conditions of Theorem 2,*

$$W_n(\cdot,\cdot) - W_n^\dagger(\cdot,\cdot) = o_p(1)$$

on $C(\mathbb{R}^{p+2})$, where $W_n(\cdot,\cdot)$ and $W_n^\dagger(\cdot,\cdot)$ are defined in (4.20) and (4.22) respectively.

Proof: Fix $T > 0$. From (4.50), for all n sufficiently large and all $|v_2| \leq T$, there exists a $K_1 > 0$ such that

$$\left|\frac{\partial \ln f(z;\alpha+v_2/n^{1/2})}{\partial z}\right| \leq K_1 \quad \forall z \in \mathbb{R}.$$

Also, for all n sufficiently large and all $|v_1| \leq T$, there exists a $K_2 > 0$ such that

$$\frac{\sigma}{|\phi_{0r}|}+\frac{v_1}{n^{1/2}} \geq \frac{\sigma}{|\phi_{0r}|} - K_2 > 0.$$

Thus, for all n sufficiently large, all $\mathbf{u} \in \mathbb{R}^p$, and all $\mathbf{v} = (v_1, v_2)' \in \mathbb{R}^2$ such that $|v_1|, |v_2| \leq T$,

$$
\begin{aligned}
&|W_n(\mathbf{u}, \mathbf{v}) - W_n^{\dagger}(\mathbf{u}, \mathbf{v})| \\
&\leq \sum_{t=1}^{n-p} \left| \ln f\left(\frac{z_t(\boldsymbol{\phi}_0 + n^{-1/\alpha}\mathbf{u})}{\sigma/|\phi_{0r}| + v_1/n^{1/2}}; \alpha + \frac{v_2}{n^{1/2}} \right) - \ln f\left(\frac{z_t(\boldsymbol{\phi}_0) + n^{-1/\alpha} \sum_{j \neq 0} c_j(\mathbf{u}) z_{t-j}}{\sigma/|\phi_{0r}| + v_1/n^{1/2}}; \alpha + \frac{v_2}{n^{1/2}} \right) \right| \\
&\leq \frac{K_1}{\sigma/|\phi_{0r}| - K_2} \sum_{t=1}^{n-p} \left| z_t(\boldsymbol{\phi}_0 + n^{-1/\alpha}\mathbf{u}) - z_t(\boldsymbol{\phi}_0) - n^{-1/\alpha} \sum_{j \neq 0} c_j(\mathbf{u}) z_{t-j} \right|,
\end{aligned}
$$

and so the result holds by the proof of Lemma 9. □

Appendix

In this section, we correct some of the results in Breidt, Davis, and Trindade (2001). We begin by changing the two lemmas in the Appendix of the paper so that Theorem 1 and some of the order selection results in Section 3 can be fixed. The notation we use corresponds to notation in Breidt, Davis, and Trindade (2001).

Lemma 1 *Suppose* $\{Y_t\}$ *and* $\{V_t\}$ *are the linear processes*

$$Y_t = \sum_{j=-\infty}^{\infty} c_j z_{t-j} \qquad \textit{and} \qquad V_t = \sum_{j=-\infty}^{\infty} d_j z_{t-j},$$

where $c_0 = 0$, $\sum_{j=-\infty}^{\infty} |c_j| < \infty$, $\sum_{j=-\infty}^{\infty} |d_j| < \infty$, *and* $\{z_t\}$ *is iid with mean* 0, *finite variance, and common distribution function* G. *If* G *has median* 0 *and is continuously differentiable in a neighborhood of* 0, *then*

$$S_n := \sum_{t=1}^{n-p} \left(|z_t - n^{-1/2} Y_t - n^{-1} V_t| - |z_t| \right) \xrightarrow{d} Var(Y_1) g(0) - d_0 E|z_1| + N,$$

where

$$N \sim N\left(0, \gamma^*(0) + 2\sum_{h=1}^{\infty} \gamma^*(h)\right),$$
$$\gamma^*(h) = E[Y_1 sgn(z_1) Y_{1+h} sgn(z_{1+h})],$$

and $g(z)$ *is the density corresponding to* G.

Proof: Using the identity for $z \neq 0$

$$|z - y| - |z| = -y\,\mathrm{sgn}(z) + 2(y - z)\left\{\mathbf{1}_{\{0<z<y\}} - \mathbf{1}_{\{y<z<0\}}\right\},$$

we have

$$\begin{aligned} S_n &= -n^{-1/2}\sum_{t=1}^{n-p} Y_t \operatorname{sgn}(z_t) - n^{-1}\sum_{t=1}^{n-p} V_t \operatorname{sgn}(z_t) \\ &\quad +2\sum_{t=1}^{n-p}\left(n^{-1/2}Y_t + n^{-1}V_t - z_t\right)\left\{\mathbf{1}_{\{0<z_t<n^{-1/2}Y_t+n^{-1}V_t\}} - \mathbf{1}_{\{n^{-1/2}Y_t+n^{-1}V_t<z_t<0\}}\right\} \\ &=: A_n + B_n + C_n. \end{aligned}$$

From the proof of the original version of this Lemma, we know that $A_n \xrightarrow{d} N$ and

$$2\sum_{t=1}^{n-p}\left(n^{-1/2}Y_t - z_t\right)\left\{\mathbf{1}_{\{0<z_t<n^{-1/2}Y_t\}} - \mathbf{1}_{\{n^{-1/2}Y_t<z_t<0\}}\right\} \xrightarrow{P} \operatorname{Var}(Y_1)g(0).$$

By the ergodic theorem,

$$B_n \xrightarrow{P} -\mathrm{E}\left[V_1 \operatorname{sgn}(z_1))\right] = -d_0 \mathrm{E}\,|z_1|,$$

and so it suffices to show

$$\begin{aligned} &\sum_{t=1}^{n-p}\left(n^{-1/2}Y_t + n^{-1}V_t - z_t\right)\left\{\mathbf{1}_{\{0<z_t<n^{-1/2}Y_t+n^{-1}V_t\}} - \mathbf{1}_{\{n^{-1/2}Y_t+n^{-1}V_t<z_t<0\}}\right\} \\ &\qquad -\sum_{t=1}^{n-p}\left(n^{-1/2}Y_t - z_t\right)\left\{\mathbf{1}_{\{0<z_t<n^{-1/2}Y_t\}} - \mathbf{1}_{\{n^{-1/2}Y_t<z_t<0\}}\right\} \\ &\xrightarrow{P} 0. \end{aligned}$$

Since $\mathrm{E}\left|V_1\left\{\mathbf{1}_{\{0<z_1<n^{-1/2}Y_1+n^{-1}V_1\}} - \mathbf{1}_{\{n^{-1/2}Y_1+n^{-1}V_1<z_1<0\}}\right\}\right| \to 0$ as $n \to \infty$, we have

$$n^{-1}\sum_{t=1}^{n-p} V_t\left\{\mathbf{1}_{\{0<z_t<n^{-1/2}Y_t+n^{-1}V_t\}} - \mathbf{1}_{\{n^{-1/2}Y_t+n^{-1}V_t<z_t<0\}}\right\} \xrightarrow{P} 0.$$

Now consider

$$\begin{aligned} &\mathrm{E}\left|\sum_{t=1}^{n-p}\left(n^{-1/2}Y_t - z_t\right)\left\{\mathbf{1}_{\{0<z_t<n^{-1/2}Y_t+n^{-1}V_t\}} - \mathbf{1}_{\{0<z_t<n^{-1/2}Y_t\}}\right\}\right| \\ &\leq \mathrm{E}\sum_{t=1}^{n-p}\left|n^{-1/2}Y_t - z_t\right|\left\{\mathbf{1}_{\{0<n^{-1/2}Y_t\leq z_t<n^{-1/2}Y_t+n^{-1}V_t\}} + \mathbf{1}_{\{n^{-1/2}Y_t\leq 0<z_t<n^{-1/2}Y_t+n^{-1}V_t\}}\right\} \\ &\quad +\mathrm{E}\sum_{t=1}^{n-p}\left|n^{-1/2}Y_t - z_t\right|\left\{\mathbf{1}_{\{0<n^{-1/2}Y_t+n^{-1}V_t\leq z_t<n^{-1/2}Y_t\}} + \mathbf{1}_{\{n^{-1/2}Y_t+n^{-1}V_t\leq 0<z_t<n^{-1/2}Y_t\}}\right\}. \end{aligned}$$

For any $\epsilon > 0$, the right-hand side is bounded by

$$\mathrm{E}\sum_{t=1}^{n-p}\left|n^{-1/2}Y_t - z_t\right|\left\{\mathbf{1}_{\{0<n^{-1/2}Y_t\leq z_t<n^{-1/2}Y_t+n^{-1}V_t<\epsilon\}} + \mathbf{1}_{\{-\epsilon<n^{-1/2}Y_t\leq 0<z_t<n^{-1/2}Y_t+n^{-1}V_t<\epsilon\}}\right\}$$

$$+\mathrm{E}\sum_{t=1}^{n-p}\left|n^{-1/2}Y_t - z_t\right|\left\{\mathbf{1}_{\{0<n^{-1/2}Y_t+n^{-1}V_t\le z_t<n^{-1/2}Y_t<\epsilon\}} + \mathbf{1}_{\{-\epsilon<n^{-1/2}Y_t+n^{-1}V_t\le 0<z_t<n^{-1/2}Y_t<\epsilon\}}\right\}$$

$$+\mathrm{E}\sum_{t=1}^{n-p}\left(2n^{-1/2}|Y_t| + n^{-1}|V_t|\right)3\left\{\mathbf{1}_{\{|n^{-1/2}Y_t+n^{-1}V_t|>\epsilon\}} + \mathbf{1}_{\{|n^{-1/2}Y_t|>\epsilon\}}\right\}$$

$$\le 3\epsilon(n-s)\mathrm{P}\left(-\epsilon < n^{-1/2}Y_1 \le z_1 < n^{-1/2}Y_1 + n^{-1}V_1 < \epsilon\right)$$

$$+2\epsilon(n-s)\mathrm{P}\left(-\epsilon < n^{-1/2}Y_1 + n^{-1}V_1 \le z_1 < n^{-1/2}Y_1 < \epsilon\right)$$

$$+6n^{1/2}\left\{\mathrm{E}\left(|Y_1| + |V_1|\right)^2\right\}^{1/2}\left(\left\{\mathrm{P}\left(|n^{-1/2}Y_1 + n^{-1}V_1| > \epsilon\right)\right\}^{1/2} + \left\{\mathrm{P}\left(|n^{-1/2}Y_1| > \epsilon\right)\right\}^{1/2}\right).$$

Note that for ϵ sufficiently small there exists a $\delta > 0$ such that

$$\mathrm{P}\left(-\epsilon < n^{-1/2}Y_1 \le z_1 < n^{-1/2}Y_1 + n^{-1}V_1 < \epsilon\right)$$

$$\le \mathrm{P}\left(|z_1| < \epsilon, n^{-1/2}Y_1 \le z_1 < n^{-1/2}Y_1 + n^{-1}\left\{|d_0|\epsilon + \sum_{\substack{j=-\infty\\ j\neq 0}}^{\infty}|d_j z_{1-j}|\right\}\right)$$

$$\le (g(0)+\delta)n^{-1}\left(|d_0|\epsilon + \sum_{\substack{j=-\infty\\ j\neq 0}}^{\infty}|d_j|\mathrm{E}|z_1|\right)$$

and

$$\mathrm{P}\left(-\epsilon < n^{-1/2}Y_1 + n^{-1}V_1 \le z_1 < n^{-1/2}Y_1 < \epsilon\right) \le (g(0)+\delta)n^{-1}\left(|d_0|\epsilon + \sum_{\substack{j=-\infty\\ j\neq 0}}^{\infty}|d_j|\mathrm{E}|z_1|\right).$$

Since $\mathrm{E}(|Y_1| + |V_1|)^2 < \infty$,

$$n\mathrm{P}\left(|n^{-1/2}Y_t + n^{-1}V_t| > \epsilon\right) \to 0 \quad \text{and} \quad n\mathrm{P}\left(|n^{-1/2}Y_t| > \epsilon\right) \to 0,$$

and it follows that

$$\mathrm{E}\left|\sum_{t=1}^{n-p}\left(n^{-1/2}Y_t - z_t\right)\left\{\mathbf{1}_{\{0<z_t<n^{-1/2}Y_t+n^{-1}V_t\}} - \mathbf{1}_{\{0<z_t<n^{-1/2}Y_t\}}\right\}\right| \to 0.$$

□

If we approximate $\varphi(B^{-1})z_t(\phi)$ by

$$\varphi_0(B^{-1})z_t(\phi_0) + \sum_{j=1}^{s}\frac{\partial}{\partial\phi_j}\left\{\varphi(B^{-1})z_t(\phi)\right\}\Big|_{\phi=\phi_0}(\phi_j - \phi_{0j})$$

$$+\frac{1}{2}\sum_{j=1}^{s}\sum_{k=1}^{s}\frac{\partial^2}{\partial\phi_j\partial\phi_k}\left\{\varphi(B^{-1})z_t(\phi)\right\}\Big|_{\phi=\phi_0}(\phi_j - \phi_{0j})(\phi_k - \phi_{0k}),$$

the criterion function can be written as

$$\begin{aligned}
m_n &= \sum_{t=1}^{n-p} |\varphi(B^{-1})z_t(\phi) - \phi(B)X_{t+s}| \\
&= \sum_{t=1}^{n-p} |\varphi(B^{-1})B^s z_{t+s}(\phi) - \phi_0(B)X_{t+s} + (\phi_0(B) - \phi(B))X_{t+s}| \\
&\simeq \sum_{t=1}^{n-p} \Big|\varphi_0(B^{-1})B^s z_{t+s}(\phi_0) - B^s z_{t+s}(\phi_0) + z_t(\phi_0) \\
&\quad + \sum_{j=1}^{s} \frac{\partial}{\partial \phi_j}\{\varphi(B^{-1})z_t(\phi)\}\Big|_{\phi=\phi_0}(\phi_j - \phi_{0j}) \\
&\quad + \frac{1}{2}\sum_{j=1}^{s}\sum_{k=1}^{s} \frac{\partial^2}{\partial\phi_j\partial\phi_k}\{\varphi(B^{-1})z_t(\phi)\}\Big|_{\phi=\phi_0}(\phi_j - \phi_{0j})(\phi_k - \phi_{0k}) \\
&\quad - \phi_0(B)X_{t+s} + n^{1/2}(\phi - \phi_0)' n^{-1/2}(X_{t+s-1}, \ldots, X_t)'\Big| \\
&= \sum_{t=1}^{n-p} \Bigg| z_t(\phi_0) + n^{-1/2}\mathbf{u}' \left[\frac{\partial}{\partial\phi_j}\{\varphi(B^{-1})z_t(\phi)\}\Big|_{\phi=\phi_0} + X_{t+s-j}\right]_{j=1}^{s} \\
&\quad + \frac{1}{2}n^{-1}\mathbf{u}'\left[\frac{\partial^2}{\partial\phi_j\partial\phi_k}\{\varphi(B^{-1})z_t(\phi)\}\Big|_{\phi=\phi_0}\right]_{j,k=1}^{s}\mathbf{u}\Bigg|, \qquad (5.1)
\end{aligned}$$

where $\mathbf{u} = n^{1/2}(\phi - \phi_0)$. The coefficient of $n^{-1/2}$ in (5.1) was discussed in the original paper, so we now focus on the coefficient of n^{-1}. Because

$$\frac{\partial^2}{\partial\phi_j\partial\phi_k}\{\varphi(B^{-1})z_t(\phi)\} = \frac{1}{\phi(B^{-1})}\left\{\frac{\partial}{\partial\phi_j}z_{t+k}(\phi) + \frac{\partial}{\partial\phi_k}z_{t+j}(\phi)\right\},$$

the coefficient of n^{-1} is

$$\begin{aligned}
&\frac{1}{2}\mathbf{u}'\left[\frac{\partial^2}{\partial\phi_j\partial\phi_k}\{\varphi(B^{-1})z_t(\phi)\}\Big|_{\phi=\phi_0}\right]_{j,k=1}^{s}\mathbf{u} \\
&= \frac{1}{2}\mathbf{u}'\left[\frac{1}{\phi(B^{-1})}\left\{\frac{\partial}{\partial\phi_j}z_{t+k}(\phi) + \frac{\partial}{\partial\phi_k}z_{t+j}(\phi)\right\}\Big|_{\phi=\phi_0}\right]_{j,k=1}^{s}\mathbf{u} \\
&= \frac{1}{2}\mathbf{u}'\left[\frac{1}{\phi_0^2(B^{-1})}\{X_{t+s+j-k} + X_{t+s+k-j} + 2z_{t+j+k}(\phi_0)\}\right]_{j,k=1}^{s}\mathbf{u} \\
&\simeq \frac{1}{2}\mathbf{u}'\left[\frac{-z_{t+j-k} - z_{t+k-j}}{\phi_0(B^{-1})\phi_0(B)} + \frac{2z_{t+j+k}}{\phi_0^2(B^{-1})}\right]_{j,k=1}^{s}\mathbf{u} \\
&= \frac{1}{2}\mathbf{u}'\left[-\sum_{m=0}^{\infty}\sum_{l=0}^{\infty}\psi_m\psi_l(z_{t+j-k-l+m} + z_{t+k-j-l+m}) + \frac{2z_{t+j+k}}{\phi_0^2(B^{-1})}\right]_{j,k=1}^{s}\mathbf{u} \\
&=: -V_t = -\sum_{j=-\infty}^{\infty} d_j z_{t-j}.
\end{aligned}$$

Note that

$$\begin{aligned} d_0 \mathrm{E}|z_1| &= \frac{1}{2}\frac{\mathrm{E}|Z_1|}{|\phi_{0r}|}\mathbf{u}'\left[\sum_{l=0}^{\infty}(\psi_l\psi_{l-j+k}+\psi_l\psi_{l+j-k})\right]_{j,k=1}^{s}\mathbf{u} \\ &= \frac{\mathrm{E}|Z_1|}{\sigma^2|\phi_{0r}|}\mathbf{u}'\mathbf{\Gamma}_s\mathbf{u}. \end{aligned}$$

The modified version of Lemma 2 follows.

Lemma 2 *For* $\mathbf{u} \in I\!R^s$, *let*

$$S_n(\mathbf{u}) = m_n(\boldsymbol{\phi}_0 + n^{-1/2}\mathbf{u}) - \sum_{t=1}^{n-p}|z_t(\boldsymbol{\phi}_0)|$$

and define

$$\begin{aligned} S_n^*(\mathbf{u}) &= \sum_{t=1}^{n-p}\Bigg| z_t(\boldsymbol{\phi}_0) + n^{-1/2}\mathbf{u}'\left[\frac{\partial}{\partial\phi_j}\{\varphi(B^{-1})z_t(\boldsymbol{\phi})\}\Big|_{\boldsymbol{\phi}=\boldsymbol{\phi}_0} + X_{t+s-j}\right]_{j=1}^{s} \\ &\qquad + \frac{1}{2}n^{-1}\mathbf{u}'\left[\frac{\partial^2}{\partial\phi_j\partial\phi_k}\{\varphi(B^{-1})z_t(\boldsymbol{\phi})\}\Big|_{\boldsymbol{\phi}=\boldsymbol{\phi}_0}\right]_{j,k=1}^{s}\mathbf{u}\Bigg| \\ &\quad - \sum_{t=1}^{n-p}|z_t(\boldsymbol{\phi}_0)|. \end{aligned}$$

Then

1. $S_n^* \xrightarrow{d} S$ *on* $C(I\!R^s)$ *where*

$$S(\mathbf{u}) = \left(\frac{2f_\sigma(0)}{|\phi_{0r}|} - \frac{E|Z_1|}{\sigma^2|\phi_{0r}|}\right)\mathbf{u}'\mathbf{\Gamma}_s\mathbf{u} + \mathbf{u}'\mathbf{N}$$

and

$$\mathbf{N} \sim N\left(\mathbf{0}, \frac{2\,Var|Z_1|}{\phi_{0r}^2\sigma^2}\mathbf{\Gamma}_s\right).$$

2. $S_n \xrightarrow{d} S.$

Proof: The proof is essentially the same as the original, except

$$\begin{aligned} S_n^\dagger(\mathbf{u}) &= \sum_{t=1}^{n-p}\left\{|z_t - n^{-1/2}Y_t - n^{-1}V_t| - |z_t|\right\} \\ &= -n^{-1/2}\sum_{t=1}^{n-p}Y_t\mathrm{sgn}(z_t) + \left(\frac{2f_\sigma(0)}{|\phi_{0r}|} - \frac{\mathrm{E}|Z_1|}{\sigma^2|\phi_{0r}|}\right)\mathbf{u}'\mathbf{\Gamma}_s\mathbf{u} + o_p(1). \end{aligned}$$

□

Therefore, Theorem 1 now reads

Theorem 1 *Assume $2\sigma^2 f_\sigma(0) > E|Z_1|$ and A1–A3 hold. Then there exists a sequence of local minimizers $\hat{\boldsymbol{\phi}}_{LAD}$ of $m_n(\boldsymbol{\phi})$ such that*

$$n^{1/2}(\hat{\boldsymbol{\phi}}_{LAD} - \boldsymbol{\phi}_0) \xrightarrow{d} \frac{\sigma^2|\phi_{0r}|}{2E|Z_1| - 4\sigma^2 f_\sigma(0)}\boldsymbol{\Gamma}_s^{-1}\mathbf{N} \sim N\left(\mathbf{0}, \frac{Var|Z_1|}{2\left(2\sigma^2 f_\sigma(0) - E|Z_1|\right)^2}\sigma^2\boldsymbol{\Gamma}_s^{-1}\right),$$

where $\boldsymbol{\Gamma}_s = [\gamma(j-k)]_{j,k=1}^s$ and $\gamma(\cdot)$ is the autocovariance function of the causal AR(r) $\{Z_t/\phi_0(B)\}$.

For the Laplace density,

$$\frac{\mathrm{Var}\,|Z_1|}{2\left(2\sigma^2 f_\sigma(0) - \mathrm{E}\,|Z_1|\right)^2} = \frac{1}{2},$$

and, for the Students' t-distribution with $\nu > 2$ degrees of freedom,

$$\frac{\mathrm{Var}\,|Z_1|}{2\left(2\sigma^2 f_\sigma(0) - \mathrm{E}\,|Z_1|\right)^2} = \frac{\nu-2}{8\Gamma^2((\nu+1)/2)}\left(\pi(\nu-1)^2\Gamma^2(\nu/2) - 4(\nu-2)\Gamma^2((\nu+1)/2)\right). \qquad (5.2)$$

When $\nu = 3$, (5.2) is 0.7337.

Corollary 1 states

Corollary 1 *Assume the conditions of Theorem 1. If the true all-pass model order is r and the fitted model order is $p > r$ then*

$$n^{1/2}\hat{\phi}_{p,LAD} \xrightarrow{d} N\left(0, \frac{Var|Z_1|}{2\left(2\sigma^2 f_\sigma(0) - E|Z_1|\right)^2}\right),$$

where $\hat{\phi}_{p,LAD}$ is the pth element of $\hat{\boldsymbol{\phi}}_{LAD}$.

So, the term $\hat{\theta}^2$ used in the order determination procedure becomes

$$\hat{\theta}^2 := \frac{\hat{v}_1}{2\left(2\hat{v}_2\hat{d} - \hat{e}_1\right)^2} \xrightarrow{P} \frac{\mathrm{Var}\,|Z_1|}{2\left(2\sigma^2 f_\sigma(0) - \mathrm{E}\,|Z_1|\right)^2},$$

where $\hat{e}_1$ and $\hat{v}_1$ are the empirical mean and variance respectively of $\{|z_t(\hat{\boldsymbol{\phi}})|\}$, $\hat{v}_2$ is the empirical variance of $\{z_t(\hat{\boldsymbol{\phi}})\}$, and $\hat{d}$ is a kernel estimator of the density at zero based on $\{z_t(\hat{\boldsymbol{\phi}})\}$.

We now make a few changes to the derivation of AIC(·). Again, $X_1, \ldots, X_n$ and $X_1^*, \ldots, X_n^*$ represent two independent realizations of the model $(\boldsymbol{\phi}_0', \kappa_0)'$, where we assume $\boldsymbol{\phi}_0 \in \mathbb{R}^p$ and $p \leq s$. If we obtain $\hat{\boldsymbol{\phi}}$ and $\hat{\kappa}$ using $X_1, \ldots, X_n$,

$$-2\mathcal{L}_{X^*}(\hat{\boldsymbol{\phi}}, \hat{\kappa}) = -2\mathcal{L}_X(\hat{\boldsymbol{\phi}}, \hat{\kappa}) - 2\frac{\sqrt{2}\sum_{t=1}^{n-p}|z_t(\hat{\boldsymbol{\phi}})|}{\hat{\kappa}} + 2\frac{\sqrt{2}\sum_{t=1}^{n-p}|z_t^*(\hat{\boldsymbol{\phi}})|}{\hat{\kappa}}$$

$$= \quad -2\mathcal{L}_X(\hat{\phi},\hat{\kappa}) - 2(n-p) + 2\sqrt{2}\frac{\sum_{t=1}^{n-p}|z_t^*(\hat{\phi})| - \sum_{t=1}^{n-p}|z_t^*(\phi_0)|}{\hat{\kappa}}$$
$$+2\sqrt{2}\frac{\sum_{t=1}^{n-p}|z_t^*(\phi_0)|}{\hat{\kappa}}.$$

Note that

$$\frac{\sum_{t=1}^{n-p}|z_t^*(\hat{\phi})| - \sum_{t=1}^{n-p}|z_t^*(\phi_0)|}{\hat{\kappa}} \xrightarrow{d} \frac{\mathbf{u}'\mathbf{N}^*}{\sqrt{2}\mathrm{E}|Z_1||\phi_{0r}|^{-1}} + \left(\frac{2f_\sigma(0)}{|\phi_{0r}|} - \frac{\mathrm{E}|Z_1|}{\sigma^2|\phi_{0r}|}\right)\frac{\mathbf{u}'\mathbf{\Gamma}_p\mathbf{u}}{\sqrt{2}\mathrm{E}|Z_1||\phi_{0r}|^{-1}},$$

where

$$\mathbf{u} = \frac{\sigma^2|\phi_{0r}|}{2\mathrm{E}|Z_1| - 4\sigma^2 f_\sigma(0)}\mathbf{\Gamma}_p^{-1}\mathbf{N}$$

and $\mathbf{N}, \mathbf{N}^*$ are iid $N(\mathbf{0}, 2\mathrm{Var}|Z_1|\phi_{0r}^{-2}\sigma^{-2}\mathbf{\Gamma}_p)$.

$$\mathrm{E}\left[\frac{\sum_{t=1}^{n-p}|z_t^*(\hat{\phi})| - \sum_{t=1}^{n-p}|z_t^*(\phi_0)|}{\hat{\kappa}}\right] \simeq \left(\frac{\sqrt{2}f_\sigma(0)}{\mathrm{E}|Z_1|} - \frac{1}{\sqrt{2}\sigma^2}\right)\mathrm{trace}(\mathbf{\Gamma}_p\mathrm{E}[\mathbf{u}\mathbf{u}'])$$
$$= \frac{\mathrm{Var}|Z_1|\sigma^2}{2\left(2\sigma^2 f_\sigma(0) - \mathrm{E}|Z_1|\right)^2}\left(\frac{\sqrt{2}f_\sigma(0)}{\mathrm{E}|Z_1|} - \frac{1}{\sqrt{2}\sigma^2}\right)p.$$

Therefore,

$$\mathrm{AIC}(p) := -2\mathcal{L}_X(\hat{\phi},\hat{\kappa}) + \frac{\mathrm{Var}|Z_1|}{(2\sigma^2 f_\sigma(0) - \mathrm{E}|Z_1|)^2}\left(\frac{2\sigma^2 f_\sigma(0)}{\mathrm{E}|Z_1|} - 1\right)p.$$

The penalty term can be estimated consistently with

$$\frac{\hat{v}_1}{(2\hat{v}_2\hat{d} - \hat{e}_1)^2}\left(\frac{2\hat{v}_2\hat{d}}{\hat{e}_1} - 1\right).$$

References

Allal, J., Kaaouachi A., and Paindaveine, D. (2001). R-estimation for ARMA models. *Journal of Nonparametric Statistics* **13**, 815–831.

Benveniste, A., Goursat, M., and Roget, G. (1980). Robust identification of a nonminimum phase system: blind adjustment of linear equalizer in data communications. *IEEE Transactions on Automatic Control* **25**, 385–398.

Billingsley, P. (1999). *Convergence of Probability Measures,* 2nd ed. Wiley, New York.

Bingham, N.H., Goldie, C.M., and Teugels, J.L. (1987). *Regular Variation.* Cambridge University Press, Cambridge.

Blass, W.E. and Halsey, G.W. (1981). *Deconvolution of Absorption Spectra.* Academic Press, New York.

Boldin, M.V. (1982). Estimation of the distribution of noise in an autoregression scheme. *Theory of Probability and its Applications* **27**, 866–871.

Breidt, F.J. and Davis, R.A. (1991). Time-reversibility, identifiability and independence of innovations for stationary time series. *Journal of Time Series Analysis* **13**, 377–390.

Breidt, F.J., Davis, R.A., Lii, K.-S., and Rosenblatt, M. (1991). Maximum likelihood estimation for noncausal autoregressive processes. *Journal of Multivariate Analysis* **36**, 175–198.

Breidt, F.J., Davis, R.A., and Trindade, A.A. (2001). Least absolute deviation estimation for all-pass time series models. *Annals of Statistics* **29**, 919–946.

Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods,* 2nd ed. Springer-Verlag, New York.

Brockwell, P.J. and Davis, R.A. (1996). *Introduction to Time Series and Forecasting.* Springer-Verlag, New York.

Calder, M. (1998). Parameter Estimation for Noncausal and Heavy Tailed Autoregressive Processes. Ph.D. dissertation, Department of Statistics, Colorado State University, Fort Collins, CO.

Chi, C.-Y. and Kung, J.-Y. (1995). A new identification algorithm for allpass systems by higher-order statistics. *Signal Processing* **41**, 239–256.

Chien, H.-M., Yang, H.-L., and Chi, C.-Y. (1997). Parametric cumulant based phase estimation of 1-d and 2-d nonminimum phase systems by allpass filtering. *IEEE Transactions on Signal Processing* **45**, 1742–1762.

Cline, D.B.H. (1983). Estimation and Linear Prediction for Regression, Autoregression and ARMA with Infinite Variance Data. Ph.D. dissertation, Department of Statistics, Colorado State University, Fort Collins, CO.

Davis, R.A. (1996). Gauss-Newton and M-estimation for ARMA processes with infinite variance. *Stochastic Processes and their Applications* **63**, 75–95.

Davis, R.A., Knight, K., and Liu, J. (1992). M-estimation for autoregressions with infinite variance. *Stochastic Processes and their Applications* **40**, 145–180.

Davis, R.A. and Resnick S. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Annals of Probability* **13**, 179–195.

Davis, R.A. and Wu, W. (1997). Bootstrapping M-estimates in regression and autoregression with infinite variance. *Statistica Sinica* **7**, 1135–1154.

Donoho, D. (1981). On minimum entropy deconvolution. In *Applied Time Series Analysis II* (D.F. Findley, ed.) 565–608. Academic Press, New York.

DuMouchel, W.H. (1973). On the asymptotic normality of the maximum likelihood estimate when sampling from a stable distribution. *Annals of Statistics* **1**, 948–957.

Feigin, P.D. and Resnick, S.I. (1994). Limit distributions for linear programming time series estimators. *Stochastic Processes and their Applications* **51**, 135–165.

Feller, W. (1971). *An Introduction to Probability Theory and its Applications,* Vol. 2, 2nd ed. Wiley, New York.

Giannakis, G.B. and Swami, A. (1990). On estimating noncausal nonminimum phase ARMA models of non-Gaussian processes. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **38**, 478–495.

Gnedenko, B.V. and Kolmogorov, A.N. (1968). *Limit Distributions for Sums of Independent Random Variables.* Addison-Wesley, Reading, MA.

Godfrey, R. and Rocca, F. (1981). Zero memory nonlinear deconvolution. *Geophysical Prospecting* **29**, 189–228.

Hooke, R. and Jeeves, T. (1961). A direct search solution of numerical and statistical problems. *Journal of Association for Computing Machinery* **8**, 212–229.

Hsueh, A.-C. and Mendel, J.M. (1985). Minimum-variance and maximum-likelihood deconvolution for noncausal channel models. *IEEE Trans. Geoscience and Remote Sensing* **23**, 797–808.

Huang, J. and Pawitan, Y. (2000). Quasi-likelihood estimation of non-invertible moving average processes. *Scandinavian Journal of Statistics* **27**, 689–702.

Jaeckel, L.A. (1972). Estimating regression coefficients by minimizing the dispersion of the residuals. *Annals of Mathematical Statistics* **43**, 1559–1458.

Koul, H.L. and Saleh, A.K.Md.E. (1993). R-estimation of the parameters of autoregressive [Ar(p)] models. *Annals of Statistics* **21**, 534–551.

Koul, H.L., Sievers, G.L., and McKean, J.W. (1987). An estimator of the scale parameter for the rank analysis of linear models under general score functions. *Scandinavian Journal of Statistics* **14**, 131–141.

Lii, K.-S. and Rosenblatt, M. (1988). Nonminimum phase non-Gaussian deconvolution. *Journal of Multivariate Analysis* **27**, 359–374.

Lii, K.-S. and Rosenblatt, M. (1992). An approximate maximum likelihood estimation for non-Gaussian non-minimum phase moving average processes. *Journal of Multivariate Analysis* **43**, 272–299.

Lii, K.-S. and Rosenblatt, M. (1996). Maximum likelihood estimation for nonGaussian nonminimum phase ARMA sequences. *Statistica Sinica* **6**, 1–22.

Mukherjee, K. and Bai, Z.D. (2002). R-estimation in autoregression with square-integrable score function. *Journal of Multivariate Analysis* **81**, 167–186.

Nolan, J.P. (1997). Numerical calculation of stable densities and distribution functions. *Communications in Statistics, Stochastic Models* **13**, 759–774.

Ooe, M. and Ulrych, T.J. (1979). Minimum entropy deconvolution with an exponential transformation. *Geophysical Prospecting* **27**, 458–473.

Scargle, J.D. (1981). Phase-sensitive deconvolution to model random processes, with special reference to astronomical data. In *Applied Time Series Analysis II* (D.F. Findley, ed.) 549–564. Academic Press, New York.

Silverman, B.W. (1978). Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *Annals of Statistics* **6**, 177-184.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, New York.

Terpstra, J.T., McKean, J.W., and Naranjo, J.D. (2001). Weighted wilcoxon estimates for autoregression. *Australia and New Zealand Journal of Statistics* **43**, 399–419.

Wiggins, R.A. (1978). Minimum entropy deconvolution. *Geoexploration* **16**, 21–35.

Zolotarev, V.M. (1986). *One-dimensional Stable Distributions.* American Mathematical Society, Providence, RI.