Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 97 (2006) 2101–2130
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Estimation of the memory parameter by fitting fractionally differenced autoregressive models

R.J. Bhansali[a], L. Giraitis[b,1], P.S. Kokoszka[c,*,2]

[a] *University of Liverpool, UK*
[b] *University of York, UK*
[c] *Utah State University, USA*

Received 15 December 2004
Available online 3 March 2006

Abstract

Estimation of the memory parameter, d, by fitting a fractionally differenced autoregression of order p, where p approaches infinity simultaneously with the observed series length, n, is examined. Under some conditions on growth of p with respect to n and on the short-memory component, which admits an infinite autoregressive representation with coefficients a_j, the estimator is shown to be $\sqrt{p/n}$ consistent and asymptotically normal, where p may be taken to be proportional to $\log n$. The joint asymptotic distribution of the estimators of d and of the a_j is also derived.

AMS 1991 subject classification: 62M10; 62F12

Keywords: Long memory; Autoregressive estimation

1. Introduction

This paper considers estimation of the memory parameter d in a covariance stationary time series $\{X_t, t =, \ldots, -1, 0, 1, \ldots\}$ whose spectral density is of the form

$$f(\lambda) = |2\sin(\lambda/2)|^{-2d} h(\lambda), \tag{1.1}$$

* Corresponding author.
E-mail address: piotr@stat.usu.edu (P.S. Kokoszka).

1 L. Giraitis was supported by the ESRC Grant R000239538.
2 P.S. Kokoszka was partially supported by NSF Grants DMS-0413653 and INT-0223262.

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2006.01.003

where $-\frac{1}{2} < d < \frac{1}{2}$ and h is a bounded continuous function, $h(0) \neq 0$. In the case $0 < d < \frac{1}{2}$, $\{X_t\}$ is said to exhibit long memory, the case $d = 0$ corresponds to the short memory, and the case $-\frac{1}{2} < d < 0$ to the negative memory or antipersistence.

A variety of different approaches to estimating d have been suggested. The currently available methods may be grouped under four broad headings: graphical, parametric, non-parametric and semi-parametric, we refer to Bhansali and Kokoszka [7] for a recent review.

In this paper, we develop theoretical foundations for a fractional autoregressive model fitting approach to the estimation of d. We say that a time series follows a fractionally differenced autoregressive model of order p, abbreviated FAR(p, d), if its dth fractional difference follows an autoregressive model of order p. We assume that the observed time series is a realization of an FAR(∞, d) process. Our estimator of d is obtained by fitting an FAR(p, d) model by a likelihood procedure, but we treat p as a function of the sample size n such that $p \to \infty$, as $n \to \infty$. A main difficulty in developing an asymptotic theory has lied in the lack of a suitable Central Limit Theorem for quadratic forms whose kernel depends on the sample size n. We established the required results in [6]. To be able to use them, we develop in the present paper a number of bounds on L_1, L_2 and spectral norms of several matrices related to the information matrix and their derivatives, inverses and products. The dimension of the matrices and vectors depends on the sample size, and no relevant results have been available prior to the present work in the context of long memory and antipersistent time series.

We show that $\sqrt{n/p}(\widehat{d}-d)$ is asymptotically standard normal. The joint asymptotic distribution of $\widehat{d}$ and an increasing number of the estimated autoregressive coefficients is also shown to be normal with the order of consistency $\sqrt{p/n}$. The factor $\sqrt{p}$, which slows down the convergence, does not appear in the parametric likelihood estimation, see e.g. Fox and Taqqu [11] and Dahlhaus [10], because there the model is assumed to be correctly specified and so the information matrix is constant.

The approach we take is akin to the parametric approach, but the differenced process is postulated to follow an *infinite* order autoregressive model, and since such a representation exists under mild conditions (see e.g. [9, p. 78]), the class of processes we consider is larger than in a parametric setting. Indeed, our approach is similar in spirit to the well-known practice of estimating the spectral density, the linear predictor and related parameters of a stationary process with bounded spectral density by autoregressive model fitting, see Parzen [18], Berk [2], Shibata [21,22] and Bhansali [3–5], among others. In our setting, however, the spectral density can be unbounded or vanish, so the theory developed by the cited authors does not apply.

Hurvich and Brodsky [13] and Moulines and Soulier [16,17] consider an FEXP approach to the estimation of d which bears some similarity to the approach studied in the present paper. In the FEXP approach, $\log f_Y(\lambda)$, the logarithm of the spectral density of the short-memory process Y_t, is postulated to possess an infinite Fourier series expansion and thus Y_t is specified to follow an exponential model of Bloomfield [8]. The estimator $\widehat{d}_{\mathrm{ER}}$, say, is obtained by first truncating the infinite Fourier expansion of $\log f_Y(\lambda)$ at some finite value k and then by using a linear least-squares regression procedure. On the assumption that $k \to \infty$, as $n \to \infty$, Moulines and Soulier [16] show that $\sqrt{n/k}(\widehat{d}_{\mathrm{ER}} - d)$ is asymptotically normal. If the Fourier coefficients converge to zero at an exponential rate, k may be taken to be proportional to $\log n$, and this is also true of the autoregressive order, p, used in our approach. In this sense both methods are comparable and yield estimators of d with convergence rates $(\log n/n)^{1/2}$. Our approach is, however, likelihood-based and it may be expected to be asymptotically more efficient than the regression-based approach of these authors. Also, unlike Moulines and Soulier [16], we do not assume that the observed series is Gaussian. Both methods are asymptotically superior to the nonparametric and semiparametric approaches for which the fastest possible rate of convergence is $n^{-0.4}$. Even though the rate of convergence for the parametric methods is $n^{-1/2}$, these methods are not robust to model misspecification. A comparison of finite sample performance of the various methods discussed above is presented in [7]. For a broad class of long-memory Gaussian processes, our method is shown to outperform commonly used non-parametric, semi-parametric and parametric methods, the latter if the parametric model is misspecified.

The paper is organized as follows: in Section 2, we state the assumptions and describe the estimation procedure. Section 3 contains the main results of the paper and their proofs. These proofs rely on a number of auxiliary results which are developed in Sections 4–6. Section 4 establishes bounds on the rate of growth of the information matrix and related matrices of an FAR(p) process, as $p \to \infty$, thus extending similar results for AR(p) models to fractionally differenced processes. Section 5 focuses on delicate bounds for the remainder term appearing in the proofs of our main theorems. These bounds use the bounds of Section 4 and the pre-estimation procedure described in Section 2. Section 6 applies the results of Bhansali et al. [6] to quadratic forms of FAR(p, d) processes and establishes a Central Limit Theorem which is used to find the asymptotic distribution of our estimator.

2. Assumptions, parameters and estimates

We suppose that the observed series $X_1, \ldots, X_n$ is a realization of a process $\{X_t\}$ satisfying the following assumption:

Assumption 1. The process $\{X_t\}$ is defined by

$$(1-B)^d X_t = Y_t, \quad -\tfrac{1}{2} < d < \tfrac{1}{2}, \tag{2.1}$$

where B is the usual backward shift operator and $\{Y_t\}$ is a weakly dependent autoregressive process defined by

$$\sum_{j=0}^{\infty} a_j Y_{t-j} = Z_t. \tag{2.2}$$

The random variables Z_t are independent identically distributed and satisfy

$$EZ_t = 0, \quad EZ_t^2 = \sigma^2, \quad EZ_t^4 < \infty. \tag{2.3}$$

The coefficients a_j are absolutely summable, $a_0 = 1$, and for some $\varepsilon > 0$

$$a(z) = \sum_{j=0}^{\infty} a_j z^j \neq 0, \quad |z| < 1 + \varepsilon. \tag{2.4}$$

In (2.1), $(1-B)^d$ is the fractional difference operator defined by

$$(1-B)^d = \sum_{j=0}^{\infty} b_j B^j \tag{2.5}$$

with $b_j = \Gamma(j-d)/[\Gamma(j+1)\Gamma(-d)]$, where $\Gamma(x)$ is the Gamma function.

The process $\{X_t\}$ has the spectral density

$$f(\lambda) = \frac{\sigma^2}{2\pi}|1 - e^{i\lambda}|^{-2d} \left| \sum_{j=0}^{\infty} a_j e^{i\lambda j} \right|^{-2}. \tag{2.6}$$

Denote by $\boldsymbol{\beta} = (d, a_1, a_2, \ldots)^T$ a generic element of the parameter space and by $\boldsymbol{\beta}^0 = (d^0, a_1^0, a_2^0, \ldots)^T$ the true value of the parameter. Assumption 1 pertains to the true value of the parameter, even though this is not explicitly indicated by using the superscript 0.

In the space l_2 of square summable sequences, introduce the norms

$$\|\boldsymbol{\beta}\|_2 = \left(d^2 + \sum_{j=1}^{\infty} a_j^2 \right)^{1/2}, \quad \|\boldsymbol{\beta}\|_1 = |d| + \sum_{j=1}^{\infty} |a_j|. \tag{2.7}$$

The norm $\| \cdot \|_1$ is used in Assumption 2.

Assumption 2. The parameter space E has the form $E = [-\frac{1}{2}, \frac{1}{2}] \times E_{\mathrm{a}}$, where $E_{\mathrm{a}} = I_1 \times I_2 \times \cdots$ and each I_j is a closed bounded interval. The set E is bounded with respect to the norm $\| \cdot \|_1$ and contains an open neighborhood of the true parameter $\boldsymbol{\beta}^0$.

To derive the limit distribution of the estimator of $\boldsymbol{\beta}$ studied in this paper, we need the following assumptions on the order $p = p_n$ of the fitted FAR(p) model and on the a_j:

Assumption 3. Suppose that, as $n \to \infty$,

$$p = p_n \to \infty, \quad p = o(n^{1/8}), \quad p = o(n^{1-2d^0} / \log^4 n) \tag{2.8}$$

and

$$\sum_{j=p}^{\infty} |a_j^0| = o(n^{-1/2}). \tag{2.9}$$

The upper bounds in condition (2.8) are technical assumptions which reflect the intuition that p cannot increase too fast with n, i.e. that the order of the fitted model should be much smaller than the sample size. As illustrated in Remark 2.3, in practical applications, p increases roughly like $\log n$. The bound $p = o(n^{1/8})$ is implied by condition (2.12) below and is included in Assumption 3 for ease of reference and to show the approximate maximal rate of growth of p which is needed in our proofs. For d close to $\frac{1}{2}$, it is also implied by the last bound in (2.8).

Condition (2.9) indicates that p must increase sufficiently fast to ensure that the "bias" due to the neglected autoregressive coefficients vanishes as $n \to \infty$. A similar assumption is made by Berk [2] for establishing the asymptotic normality of the autoregressive spectral estimator in the short-memory case.

We now describe the two-step estimation procedure. First, we pre-estimate the parameter $d^0 \in (-\frac{1}{2}, \frac{1}{2})$ by an estimator $\widetilde{d}$ such that

$$\widetilde{d} - d^0 = o_P(n^{-r}) \tag{2.10}$$

for some $0 < r < 1$. Then we define the interval

$$I_{n,d} = [\widetilde{d} - \Delta_n, \widetilde{d} + \Delta_n] \cap [-\tfrac{1}{2}, \tfrac{1}{2}], \tag{2.11}$$

where $\Delta_n = A/k_n$, $A > 0$ is a constant and the sequence $k_n \to \infty$ is chosen such that for some fixed $\gamma > 2$,

$$p^\gamma = O(k_n), \quad k_n = o(\min(n^{1/4}, n^r)) \quad \text{as } n \to \infty. \tag{2.12}$$

(Recall that $p = p_n$ is the order of the fitted FAR(p) model.)

The pre-estimation ensures that the remainder term $\boldsymbol{r}(p)$ appearing in the proofs of Theorems 3.1 and 3.2 is asymptotically negligible. In practice it allows to choose an initial value of d for the likelihood optimization described below. A specific estimator $\widetilde{d}$ satisfying (2.10) and the choice of the sequence k_n are discussed in Remark 2.3.

To formalize the procedure of fitting a FAR(p, d) model, define

$$G_p(\boldsymbol{\beta}) = \int_{-\pi}^{\pi} I_n(\lambda) g_p^{-1}(\lambda, \boldsymbol{\beta})\, d\lambda, \tag{2.13}$$

where

$$I_n(\lambda) = \frac{1}{2\pi n} \left| \sum_{j=1}^{n} X_j e^{i\lambda j} \right|^2 \tag{2.14}$$

is the periodogram of the observations and

$$g_p(\lambda, \boldsymbol{\beta}) = |1 - e^{i\lambda}|^{-2d} |A_p(\lambda)|^{-2}, \quad A_p(\lambda) = \sum_{j=0}^{p} a_j e^{i\lambda j} \tag{2.15}$$

is the power transfer function of the FAR(p, d) model.

The estimator $\widehat{\boldsymbol{\beta}}_p = (\widehat{d}, \widehat{a}_1, \ldots, \widehat{a}_p)$ is obtained by minimizing the function $G_p(\boldsymbol{\beta})$ over the set

$$\bar{E}_n = I_{n,d} \times E_{a,p}, \quad E_{a,p} = I_1 \times \cdots \times I_p$$

with the I_j as in Assumption 2. It thus has only the first $p + 1$ nonvanishing components and can be defined formally as

$$\widehat{\boldsymbol{\beta}}_p = \operatorname{argmin}\{G_p(\boldsymbol{\beta});\ \boldsymbol{\beta} \in \bar{E}_n\}. \tag{2.16}$$

In Section 3, we show that the estimator $\widehat{\boldsymbol{\beta}}_p$ is an $(n/p)^{1/2}$ consistent estimator of d^0 and the parameters $a_1^0, \ldots, a_p^0$, and satisfies the central limit theorem.

The corresponding estimator of the white-noise variance σ^2 is given by

$$\widehat{\sigma}^2(p) = G_p(\widehat{\boldsymbol{\beta}}_p).$$

Remark 2.1. Condition (2.4) implies that $c(z) := 1/a(z)$ admits the expansion

$$c(z) = \sum_{j=0}^{\infty} c_j^0 z^j, \quad |z| \leqslant 1, \quad c_0^0 = 1$$

and that

$$a_j^0 = O(r^j) \quad \text{and} \quad c_j^0 = O(r^j) \quad \text{for some } 0 < r < 1. \tag{2.17}$$

We denote

$$A(\lambda) = \sum_{j=0}^{\infty} a_j^0 e^{i\lambda j}, \quad C(\lambda) = 1/A(\lambda) = \sum_{j=0}^{\infty} c_j^0 e^{i\lambda j}. \tag{2.18}$$

By (2.17), $\sum_{j=p}^{\infty} |a_j^0| = O(r^p),\ (n \to \infty)$, so for (2.9) to hold it is enough to ensure that $p \geqslant K \ln n$ with large enough K. Consequently, the rate of convergence $(p/n)^{1/2}$ in Theorems 3.1 and 3.2 may be as fast as $(\ln n/n)^{1/2}$.

Notice also that (2.17) implies the bound

$$\sum_{j=k}^{\infty} |a_j^0| = o(k^{-2}) \tag{2.19}$$

which we use in proofs below.

Remark 2.2. The bandwidth Δ_n in (2.11) has the following properties which will be used in the sequel:

$$\widetilde{d} - d^0 = o_P(\Delta_n), \quad \Delta_n \leqslant C p^{-\gamma}, \quad P\{d^0 \in I_{n,d}\} \to 1. \tag{2.20}$$

The first relation in (2.20), which implies the last one, follows from (2.10) and $k_n = o(n^r)$, see (2.12). The bound $\Delta_n \leqslant C p^{-\gamma}$ follows from $\Delta_n = A k_n^{-1}$ and $p^\gamma = O(k_n)$. Definition (2.11) of the minimization interval $I_{n,d}$ and (2.20) imply that our estimator of d defined via (2.16) is consistent: $\widehat{d} \xrightarrow{P} d^0$.

Remark 2.3. For pre-estimation of the parameter $d^0 \in (-\frac{1}{2}, \frac{1}{2})$ one can use, for example, the local Whittle estimator $\widetilde{d}$ (see [20]) defined as

$$\widetilde{d} = \operatorname{argmin}\{U_n(d),\ d \in [-1/2, 1/2]\}, \tag{2.21}$$

where $U_n(d)$ is a local contrast function

$$U_n(d) = \log\left(\frac{1}{m}\sum_{j=1}^{m} j^{2d} I_n(\lambda_j)\right) - \frac{2d}{m}\sum_{j=1}^{m} \log j.$$

Here $\lambda_k = \frac{2\pi k}{n}$, $k = 1, \ldots, [(n-1)/2]$ are the Fourier frequencies and $m = m_n \to \infty$ is the bandwidth parameter. Under our assumptions, when $m = O(n^{0.8})$,

$$\widetilde{d} - d^0 = O_P(m^{-1/2}).$$

Choosing $m = n^{0.7}$, for example, we see that (2.10) holds with $r = 0.3$. Since the coefficients a_j^0 decay exponentially fast, see Remark 2.1, we can set $p = (\log n)^\delta$ for some $\delta > 1$. For $m = n^{0.7}$, we can thus set, say, $k_n = n^{0.2}$ or $k_n = (\log n)^{3\delta}$.

3. Asymptotic properties of the estimator

To formulate our asymptotic results, we set

$$\tau_0 = 1, \quad \tau_i = \sum_{m=1}^{i} \frac{a_{i-m}}{m}, \quad 1 \leqslant i \leqslant p, \quad s_J = \left(\sum_{i=0}^{J} \tau_i^2 \right)^{1/2}. \tag{3.1}$$

Below $N(0,1)$ denotes a random variable with a standard normal distribution.

Unless stated otherwise, all asymptotic relations are as $n \to \infty$. Throughout the paper, $a_n \sim b_n$ means that $a_n/b_n \to C \neq 0$ as $n \to \infty$.

3.1. Main results

Theorem 3.1. *Suppose that Assumptions* 1–3 *hold. Then the estimator* $\widehat{\boldsymbol{\beta}}_p = (\widehat{d}, \widehat{a}_1, \ldots, \widehat{a}_p)$ *is* $(n/p)^{1/2}$ *consistent*:

$$(n/p)^{1/2} \| \widehat{d} - d^0, \widehat{a}_1 - a_1^0, \ldots, \widehat{a}_p - a_p^0 \|_2 = O_P(1). \tag{3.2}$$

Theorem 3.2. *Suppose that Assumptions* 1–3 *hold. Then for a fixed J and for increasing* $J = J_n = o(p)$,

$$(n/p)^{1/2} (\widehat{d} - d^0, \widehat{a}_1 - a_1^0, \ldots, \widehat{a}_J - a_J^0)^T = (\tau_0, \tau_1, \ldots, \tau_J)^T Z_n + \mathbf{r}_{n,J}, \tag{3.3}$$

where Z_n *is a scalar random variable such that*

$$Z_n \xrightarrow{d} N(0,1)$$

and the remainder term satisfies $\|\mathbf{r}_{n,J}\|_2 = o_P(1)$.

Corollary 3.1. *Under the assumptions of Theorem* 3.2,

$$(n/p)^{1/2} (\widehat{d} - d^0) \xrightarrow{d} N(0,1) \tag{3.4}$$

and for any fixed $k = 1, 2, \ldots$,

$$(n/p)^{1/2} (\widehat{a}_k - a_k^0) \xrightarrow{d} \tau_k N(0,1). \tag{3.5}$$

To obtain parameter free limits, set

$$\widehat{s}_J = \left(\sum_{i=0}^{J} \widehat{\tau}_i^2 \right)^{1/2}, \quad \widehat{\tau}_0 = 1, \quad \widehat{\tau}_i = \sum_{m=1}^{i} \frac{\widehat{a}_{i-m}}{m}, \quad 1 \leqslant i \leqslant J. \tag{3.6}$$

Theorem 3.3. *Suppose that Assumptions* 1–3 *hold and* $J = J_n = o(p)$. *Then*,

$$(n/p)^{1/2} s_J^{-1} \| (\widehat{d} - d^0, \widehat{a}_1 - a_1^0, \ldots, \widehat{a}_J - a_J^0) \|_2 \xrightarrow{d} |N(0,1)|, \tag{3.7}$$

$$\widehat{s}_J \xrightarrow{P} s_J \tag{3.8}$$

and

$$(n/p)^{1/2} \widehat{s}_J^{-1} \| \widehat{d} - d^0, \widehat{a}_1 - a_1^0, \ldots, \widehat{a}_J - a_J^0 \|_2 \xrightarrow{d} |N(0,1)|. \tag{3.9}$$

3.2. Idea of the proof and additional notation

Denote

$$\boldsymbol{\beta}_p = (d, a_1, \dots, a_p)^T = (\beta_0, \beta_1, \dots, \beta_p)^T$$

and define analogously $\widehat{\boldsymbol{\beta}}_p$ and $\boldsymbol{\beta}_p^0$. By the mean value theorem, for a vector $\boldsymbol{\beta}_p^*$ satisfying $\|\boldsymbol{\beta}_p^* - \boldsymbol{\beta}_p^0\|_2 \leqslant \|\widehat{\boldsymbol{\beta}_p} - \boldsymbol{\beta}_p^0\|_2$,

$$\nabla G_p(\widehat{\boldsymbol{\beta}}_p) - \nabla G_p(\boldsymbol{\beta}_p^0) = \widehat{W}_p(\boldsymbol{\beta}_p^*)(\widehat{\boldsymbol{\beta}}_p - \boldsymbol{\beta}_p^0), \tag{3.10}$$

where

$$\nabla G_p(\boldsymbol{\beta}_p) = \left(\frac{\partial}{\partial \beta_0} G_p(\boldsymbol{\beta}), \frac{\partial}{\partial \beta_1} G_p(\boldsymbol{\beta}), \dots, \frac{\partial}{\partial \beta_p} G_p(\boldsymbol{\beta})\right)^T \tag{3.11}$$

is the $(p+1)$-dimensional random vector of partial derivatives, and

$$\widehat{W}_p(\boldsymbol{\beta}_p^*) = \nabla^2 G_p(\boldsymbol{\beta}_p^*)$$

denotes the $(p+1) \times (p+1)$ random matrix with entries

$$\widehat{w}_{jk}(\boldsymbol{\beta}_p^*) = \frac{\partial^2}{\partial \beta_k \partial \beta_j} G_p(\boldsymbol{\beta}_p^*). \tag{3.12}$$

Set

$$\boldsymbol{\zeta}(p) = \nabla G_p(\boldsymbol{\beta}_p^0), \quad \boldsymbol{v}(p) = \left(\widehat{d}_0 - d_0^0, \widehat{a}_1 - a_1^0, \dots, \widehat{a}_p - a_p^0\right)^T. \tag{3.13}$$

If $\widehat{\boldsymbol{\beta}}_p$ is an inner point of $\bar{E}_n$, then $\nabla G_p(\widehat{\boldsymbol{\beta}}_p) = 0$ and the relation (3.10) can be rewritten as

$$\boldsymbol{v}(p) = -\widehat{W}_p(\boldsymbol{\beta}_p^*)^{-1} \boldsymbol{\zeta}(p). \tag{3.14}$$

Fox and Taqqu [11] and Giraitis and Surgailis [12], among others, considered correctly specified parametric models with finite fixed p. In their work the asymptotic distribution of $\boldsymbol{v}(p)$ may be found from the asymptotic distribution of $\boldsymbol{\zeta}(p)$ by showing that, as $n \to \infty$, $\widehat{W}_p(\boldsymbol{\beta}_p^*)$ has a deterministic limit in probability, which is an invertible matrix. In our setting, the main effort is to show that the inverse $\widehat{W}_p(\boldsymbol{\beta}_p^*)^{-1}$ can be replaced by a deterministic matrix $W_p(\boldsymbol{\beta}_p^0)^{-1} = (2\pi/\sigma^2) W(p)^{-1}$. We now proceed to define the matrix $W(p)$, providing first a heuristic argument why $(2\pi/\sigma^2) W(p)^{-1}$ is a reasonable replacement for $\widehat{W}_p(\boldsymbol{\beta}_p^*)^{-1}$.

The entries of the matrix $\widehat{W}_p(\boldsymbol{\beta}_p)$ can be written as

$$\widehat{w}_{jk}(\boldsymbol{\beta}_p) = \int_{-\pi}^{\pi} h_{jk}(\lambda, \boldsymbol{\beta}_p) I_n(\lambda)\, d\lambda, \quad j, k = 0, 1, \dots, p, \tag{3.15}$$

where

$$h_{jk}(\lambda, \boldsymbol{\beta}_p) = \frac{\partial^2}{\partial \beta_k \partial \beta_j} g_p^{-1}(\lambda, \boldsymbol{\beta}_p).$$

Note that

$$\frac{\partial^2}{\partial\beta_0\partial\beta_0} g_p^{-1}(\lambda, \boldsymbol{\beta}_p) = 4\ln^2|1-e^{i\lambda}||1-e^{i\lambda}|^{2d}|A_p(\lambda)|^2 \tag{3.16}$$

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

$$\frac{\partial^2}{\partial\beta_0\partial\beta_j} g_p^{-1}(\lambda, \boldsymbol{\beta}_p) = 2\ln|1-e^{i\lambda}||1-e^{i\lambda}|^{2d}(A_p(\lambda)e^{-i\lambda j} + A_p(-\lambda)e^{i\lambda j}); \tag{3.17}$$

$$\frac{\partial^2}{\partial\beta_j\partial\beta_k} g_p^{-1}(\lambda, \boldsymbol{\beta}_p) = |1-e^{i\lambda}|^{2d}(e^{i\lambda(k-j)} + e^{-i\lambda(k-j)}). \tag{3.18}$$

Define the matrix $W_p(\boldsymbol{\beta}_p)$ with the entries

$$w_{jk}(\boldsymbol{\beta}_p) = \frac{\sigma^2}{2\pi}\int_{-\pi}^{\pi} h_{jk}(\lambda, \boldsymbol{\beta}_p) g(\lambda, \boldsymbol{\beta})\, d\lambda, \quad j, k = 0, 1, \ldots, p, \tag{3.19}$$

which are defined analogously to (3.15) but with the periodogram replaced by the spectral density

$$f(\lambda) = \frac{\sigma^2}{2\pi}|1-e^{i\lambda}|^{-2d}|A(\lambda)|^{-2} =: \frac{\sigma^2}{2\pi} g(\lambda, \boldsymbol{\beta}).$$

Set

$$W(p) = \frac{2\pi}{\sigma^2} W_p(\boldsymbol{\beta}^0) \tag{3.20}$$

and note that $W(p)$ has the entries

$$w_{00} = 4\int_{-\pi}^{\pi} \ln^2|1-e^{i\lambda}|\, d\lambda = 8\int_0^{\pi} \ln^2(2\sin(\lambda/2))\, d\lambda = \frac{2}{3}\pi^3, \tag{3.21}$$

$$\begin{aligned} w_{0j} &= 4\int_{-\pi}^{\pi} e^{-i\lambda j} \ln|1-e^{i\lambda}||A(\lambda)|^{-2} A(\lambda)\, d\lambda \\ &= -4\pi \sum_{k=0}^{\infty} \frac{c_k^0}{k+j}, \quad j = 1, \ldots, p, \end{aligned} \tag{3.22}$$

$$w_{jk} = 2\int_{-\pi}^{\pi} e^{i\lambda(k-j)}|A(\lambda)|^{-2}\, d\lambda, \quad j, k = 1, \ldots, p. \tag{3.23}$$

It is convenient to write the matrix $W(p)$ as

$$W(p) = \begin{bmatrix} w_{00} & \boldsymbol{w}(p)^T \\ \boldsymbol{w}(p) & 2\Gamma(p) \end{bmatrix}, \tag{3.24}$$

where $\boldsymbol{w} = (w_{01}, \ldots, w_{0p})^T$ and $\Gamma(p) = (w_{jk})_{j,k=1,\ldots,p}$.

Direct verification shows that

$$W(p)^{-1} = \begin{bmatrix} \omega_0(p) & \boldsymbol{\omega}(p)^T \\ \boldsymbol{\omega}(p) & M(p) \end{bmatrix}, \tag{3.25}$$

where

$$\omega_0(p) = [w_{00} - \tfrac{1}{2}\boldsymbol{w}(p)^T\Gamma(p)^{-1}\boldsymbol{w}(p)]^{-1},$$
$$\boldsymbol{\omega}(p) = -\tfrac{1}{2}\omega_0(p)\Gamma(p)^{-1}\boldsymbol{w},$$
$$M(p) = \tfrac{1}{2}\Gamma(p)^{-1}[I + \tfrac{1}{2}\omega_0(p)\boldsymbol{w}(p)\boldsymbol{w}(p)^T\Gamma(p)^{-1}]. \tag{3.26}$$

If $\{X_t\}$ is an FAR(p, d) process of finite order p, and if this correct fixed order is used in estimation, then $4\pi W(p)^{-1}$ is the asymptotic covariance matrix of $\sqrt{n}(\widehat{\boldsymbol{\beta}}(p) - \boldsymbol{\beta}^0(p))$, see e.g. Fox and Taqqu [11], Giraitis and Surgailis [12], and $4\pi\omega_0(p)$ is the asymptotic variance of $\sqrt{n}(\widehat{d} - d^0)$. As shown in Lemma 4.1, in our setting $4\pi\omega_0(p) \sim p$, as $n \to \infty$.

We conclude this section by presenting a representation for the inverse of the matrix $\Gamma(p)$ appearing in (3.24). This representation is used extensively in the proofs.

Observe that

$$\Gamma(p) = (2\pi/\sigma^2)R(p), \tag{3.27}$$

where $R(p) = \{r(k-j)\}_{j,k=1,2,\ldots,p}$ denotes the covariance matrix of $\{Y_t\}$, c.f. (2.2). As is well known, see Kailath et al. [14], we may write

$$R(p)^{-1} = \sigma_p^{-2}[\tilde{S}(p)^T\tilde{S}(p) - \tilde{U}(p)^T\tilde{U}(p)], \tag{3.28}$$

where

$$\tilde{S}(p) = \begin{bmatrix} 1 & \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_{p-1} \\ & 1 & \alpha_1 & \alpha_2 & \cdots & \alpha_{p-2} \\ & & 1 & \alpha_1 & \cdots & \alpha_{p-3} \\ & & & & & \vdots \\ & & & & 1 & \alpha_1 \\ & & & & & 1 \end{bmatrix} \tag{3.29}$$

and

$$\tilde{U}(p) = \begin{bmatrix} \alpha_p & \alpha_{p-1} & \alpha_{p-2} & \alpha_{p-3} & \cdots & \alpha_1 \\ & \alpha_p & \alpha_{p-1} & \alpha_{p-2} & \cdots & \alpha_2 \\ & & \alpha_p & \alpha_{p-1} & \cdots & \alpha_3 \\ & & & & & \vdots \\ & & & & \alpha_p & \alpha_{p-1} \\ & & & & & \alpha_p \end{bmatrix}. \tag{3.30}$$

Here, the coefficients $\alpha_1, \alpha_2, \ldots, \alpha_p$, which depend on p, are the values which minimize

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |1 + u_1 e^{i\lambda} + \cdots + u_p e^{i\lambda p}|^2 \sigma^2 |A(\lambda)|^{-2}\, d\lambda \tag{3.31}$$

with respect to $u_1, \ldots, u_p$ and σ_p^2 is the minimum of (3.31).

We shall approximate $R^{-1}(p)$ by

$$H(p) = \sigma^{-2}[S(p)^T S(p) - U(p)^T U(p)], \tag{3.32}$$

where

$$S(p) = \begin{bmatrix} 1 & a_1 & a_2 & a_3 & \cdots & a_{p-1} \\ & 1 & a_1 & a_2 & \cdots & a_{p-2} \\ & & 1 & a_1 & \cdots & a_{p-3} \\ & & & & & \vdots \\ & & & & 1 & a_1 \\ & & & & & 1 \end{bmatrix}, \tag{3.33}$$

$$U(p) = \begin{bmatrix} a_p & a_{p-1} & a_{p-2} & a_{p-3} & \cdots & a_1 \\ & a_p & a_{p-1} & a_{p-2} & \cdots & a_2 \\ & & a_p & a_{p-1} & \cdots & a_3 \\ & & & & & \vdots \\ & & & & a_p & a_{p-1} \\ & & & & & a_p \end{bmatrix} \tag{3.34}$$

using the following inequality of Baxter [1]:

$$\sum_{j=1}^{p} |\alpha_j - a_j| = O\left(\sum_{j=p+1}^{\infty} |a_j|\right). \tag{3.35}$$

3.3. Proofs of Theorems 3.1 *and* 3.2

Proof of Theorem 3.1. By Lemma 5.1,

$$\boldsymbol{v}(p) = -(2\pi/\sigma^2) W(p)^{-1} \boldsymbol{\zeta}(p) + \boldsymbol{r}(p), \tag{3.36}$$

where $W(p)$ denotes matrix (3.20), and the remainder term $\boldsymbol{r}(p)$ has the property

$$\|\boldsymbol{r}(p)\|_2 = o_P((p/n)^{1/2}). \tag{3.37}$$

By Lemma 4.5,

$$\|W(p)^{-1}\boldsymbol{\zeta}(p)\|_2 = O((p/n)^{1/2}).$$

Thus

$$\|\boldsymbol{v}(p)\|_2 \leqslant C(\|W(p)^{-1}\boldsymbol{\zeta}(p)\|_2 + \|\boldsymbol{r}(p)\|_2) = O((p/n)^{1/2}),$$

to prove (3.2). □

Proof of Theorem 3.2. Denoting by $[\cdot]_i$ the ith element of the vector occurring in the brackets, and setting $\boldsymbol{\zeta}_+(p) = (\zeta_1, \ldots, \zeta_p)^T$ (recall that $\boldsymbol{\zeta}(p) = (\zeta_0, \zeta_1, \ldots, \zeta_p)^T$), by (3.25) we can write

$$[W(p)^{-1}\boldsymbol{\zeta}(p)]_0 \equiv -V, \quad V = \omega_0(p)\left(\frac{1}{2}\sum_{j=1}^{p}[\Gamma^{-1}\boldsymbol{w}]_j \zeta_j - \zeta_0\right), \tag{3.38}$$

$$[W(p)^{-1}\boldsymbol{\zeta}(p)]_i = \tfrac{1}{2}[\Gamma^{-1}\boldsymbol{w}]_i V + \tfrac{1}{2}[\Gamma^{-1}\boldsymbol{\zeta}_+(p)]_i, \quad i = 1, \ldots, p, \tag{3.39}$$

where to lighten the notation we denoted $\Gamma = \Gamma(p)$.

To prove (3.3) note that from (3.36) to (3.39) it follows that

$$\widehat{d} - d^0 = V + o_P((p/n)^{1/2}),$$
$$(\widehat{a}_1 - a_1^0, \dots, \widehat{a}_p - a_p^0)^T = -(2\pi/\sigma^2)\tfrac{1}{2}\Gamma^{-1}\boldsymbol{w}V - (2\pi/\sigma^2)\tfrac{1}{2}\Gamma^{-1}\zeta_+(p) + \boldsymbol{r}(p),$$

where $\|\boldsymbol{r}(p)\|_2 = o_P((p/n)^{1/2})$. In Proposition 6.1 below it is shown that

$$(2\pi/\sigma^2)(n/p)^{1/2}V \xrightarrow{d} N(0, 1). \tag{3.40}$$

By estimate (4.25) of Lemma 4.4 below, $\|\frac{1}{2}\Gamma^{-1}\boldsymbol{w} + \boldsymbol{\tau}\|_2 \leqslant Cp^{-1}$. Hence

$$(n/p)^{1/2}(\widehat{a}_1 - a_1^0, \dots, \widehat{a}_p - a_p^0)^T = (\tau_1, \dots, \tau_p)^T(2\pi/\sigma^2)(n/p)^{1/2}V$$
$$-\tfrac{1}{2}(2\pi/\sigma^2)(n/p)^{1/2}\Gamma^{-1}\boldsymbol{\zeta}_+(p) + \boldsymbol{r}_1(p),$$

where the remainder term $\boldsymbol{r}_1(p)$ can be bounded as

$$\|\boldsymbol{r}_1(p)\|_2 \leqslant (n/p)^{1/2}(\|\tfrac{1}{2}\Gamma^{-1}\boldsymbol{w} + \boldsymbol{\tau}\|_2\|\boldsymbol{\zeta}_+(p)\|_2 + \|q(p)\|_2)$$
$$\leqslant C(n/p)^{1/2}\left(p^{-1}O_P((p/n)^{1/2}) + o_P(1)\right) = o_P(1).$$

In Lemma 6.4 we show that

$$E|[\Gamma^{-1}\zeta_+(p)]_i| \leqslant Cn^{-1/2}$$

uniformly in $i = 1, \dots, p$. Hence, for $J = o(p)$,

$$(n/p)^{1/2}(\widehat{d}_0 - d^0, \widehat{a}_1 - a_1^0, \dots, \widehat{a}_J - a_J^0)^T$$
$$= (1, \tau_1, \dots, \tau_J)^T(2\pi/\sigma^2)(n/p)^{1/2}V + q''(p, J), \tag{3.41}$$

$$\|q''(p, J)\|_2 \leqslant (n/p)^{1/2}\|([\Gamma^{-1}\zeta(p)]_1, \dots, [\Gamma^{-1}\zeta(p)]_J)\|_2 + o_P(1)$$
$$= (n/p)^{1/2}(J/n)^{1/2}O_P(1) + o_P(1) = o_P(1)$$

to prove (3.3). □

4. L_1, L_2 and spectral norm bounds

In this section, we establish several bounds on L_1, L_2 and spectral norms of vectors and matrices appearing in the proofs of Theorems 3.1 and 3.2. All matrices and vectors and relations between them are defined in Section 3.2.

To lighten the notation, we denote throughout Section 4, $a_j = a_j^0$, $c_j = c_j^0$, $w_j = w_{0j}$ and, in the proofs, suppress the dependence of the matrices $\widetilde{S}$, S, $\widetilde{U}$, U on p.

Denote by $\|A\|_2 = (\sum_{i,j=1}^p a_{ij}^2)^{1/2}$ the Euclidean norm and by $\|A\|_{\text{sp}} = \sup_{\|x\|=1} \|Ax\|_2$ the spectral norm of the matrix $A = (a_{ij})_{i,j=1,\dots,p}$. Recall that

$$\|A\|_{\text{sp}} \leqslant \|A\|_2, \quad \|A\|_2 \leqslant \sqrt{p}\|A\|_{\text{sp}},$$

$$\|A + B\|_2 \leqslant \|A\|_2 + \|B\|_2, \quad \|A + B\|_{\text{sp}} \leqslant \|A\|_{\text{sp}} + \|B\|_{\text{sp}}$$

and

$$\|AB\|_2 \leqslant \|A\|_{\text{sp}}\|B\|_2, \quad \|Av\|_2 \leqslant \|A\|_{\text{sp}}\|v\|_2, \tag{4.1}$$

for any $p \times p$ matrix B and a vector $v = (v_1, \dots, v_p)^T$.

Set $\boldsymbol{t} = (t_1, \dots, t_p)^T$ and $\boldsymbol{\tau} = (\tau_1, \dots, \tau_p)^T$ where

$$t_i = -\frac{4\pi}{i}, \quad \tau_i = \sum_{j=1}^{i} \frac{a_{i-j}}{j}. \tag{4.2}$$

Lemma 4.1. *Under Assumptions* 1–3,

$$\omega_0(p) \sim (4\pi)^{-1} p, \tag{4.3}$$

$$\|W(p)^{-1}\|_{\mathrm{sp}} \leqslant Cp \tag{4.4}$$

Proof of Lemma 4.1. To show (4.3) recall that

$$\omega_0(p) = [w_{00} - \tfrac{1}{2}\boldsymbol{w}(p)^T \Gamma(p)^{-1} \boldsymbol{w}(p)]^{-1}.$$

We have $\Gamma(p)^{-1} = \sigma^2 (2\pi)^{-1} R(p)^{-1}$ where

$$R(p)^{-1} = H(p) + (R(p)^{-1} - H(p)) = \sigma^{-2}(S^T S - U^T U) + (R(p)^{-1} - H(p)). \tag{4.5}$$

Thus

$$\boldsymbol{w}(p)^T \Gamma(p)^{-1} \boldsymbol{w}(p) = (2\pi)^{-1} \|S\boldsymbol{w}\|_2^2 + (2\pi)^{-1} R_p,$$

where

$$\begin{aligned} |R_p| &\leqslant C(\|U\boldsymbol{w}\|_2^2 + \|w^T(R(p)^{-1} - H(p))w\|_2) \\ &\leqslant C(\|U\boldsymbol{w}\|_2^2 + \|R(p)^{-1} - H(p)\|_{\mathrm{sp}} \|\boldsymbol{w}\|_2^2) \leqslant Cp^{-2} \end{aligned}$$

since $\|\boldsymbol{w}\|_2^2 \leqslant C$ by (4.8) and $\|Uw\|_2^2 \leqslant Cp^{-2}$ by (4.21) of Lemma 4.4 below, and $\|R(p)^{-1} - H(p)\|_{\mathrm{sp}} \leqslant Cp^{-2}$ by (4.17) of Lemma 4.3 below. Moreover,

$$\|S\boldsymbol{w}\|_2^2 = \|\boldsymbol{t}\|_2^2 + 2\sum_{i=1}^{p} t_i [S\boldsymbol{w} - \boldsymbol{t}]_i + \|S\boldsymbol{w} - \boldsymbol{t}\|_2^2 = \|\boldsymbol{t}\|_2^2 + O(p^{-2})$$

by (4.22) of Lemma 4.4 which implies

$$\boldsymbol{w}(p)^T \Gamma(p)^{-1} \boldsymbol{w}(p) = (2\pi)^{-1} \|\boldsymbol{t}\|_2^2 + O(p^{-2}).$$

On the other hand, by definition (3.21),

$$w_{00} = 4\int_{-\pi}^{\pi} \ln^2 |2\sin(\lambda/2)|\, d\lambda.$$

It is known that

$$\int_0^{\pi} \cos(j\lambda) \ln(2\sin(\lambda/2))\, d\lambda = -\frac{\pi}{2j}, \quad j = 1, 2, \dots \tag{4.6}$$

and

$$\ln(2\sin(\lambda/2)) = -\sum_{j=1}^{\infty} \frac{\cos j\lambda}{j}. \tag{4.7}$$

So we can write $\ln(2\sin(\lambda/2)) = \sum_{j=-\infty}^{\infty} e^{i\lambda j} v_j$ where $v_0 = 0$ and for $j = \pm 1, \pm 2, \ldots$,

$$v_j = \int_{-\pi}^{\pi} e^{i\lambda j} \ln|2\sin(\lambda/2)|\, d\lambda = 2\int_0^{\pi} \cos(j\lambda)\ln(2\sin(\lambda/2))\, d\lambda = -\frac{\pi}{j}.$$

Thus, by Parseval equality

$$\int_{-\pi}^{\pi} (\ln|2\sin(\lambda/2)|)^2\, d\lambda = (2\pi)^{-1}\sum_j v_j^2 = 2(2\pi)^{-1}\sum_{j=1}^{\infty}(\pi/j)^2 = \pi\sum_{j=1}^{\infty} j^{-2}.$$

Hence

$$\begin{aligned}\omega_0(p)^{-1} &= 4\pi\sum_{j=1}^{\infty} j^{-2} - \frac{1}{2}(2\pi)^{-1}\|\boldsymbol{t}\|_2^2 + o(p^{-1})\\ &= 4\pi\sum_{j=1}^{\infty} j^{-2} - (4\pi)^{-1}\sum_{j=1}^{p}(4\pi/j)^2 + o(p^{-1})\\ &= 4\pi\sum_{j=p+1}^{\infty} j^{-2} + o(p^{-1}) \sim 4\pi p^{-1},\end{aligned}$$

to prove (4.3).

Using (3.25), we can estimate

$$\|W(p)^{-1}\|_{\mathrm{sp}} \leqslant |\boldsymbol{\omega}_0(p)| + 2\|\boldsymbol{\omega}(p)\|_2 + \|M(p)\|_{\mathrm{sp}}.$$

We have

$$\|\boldsymbol{\omega}(p)\|_2 = |\omega_0(p)|\,\|\Gamma(p)^{-1}\boldsymbol{w}\|_2 \leqslant C\|\Gamma(p)^{-1}\|_{\mathrm{sp}}\|\boldsymbol{w}\|\,|\omega_0(p)| \leqslant C|\omega_0(p)|$$

since $\|\boldsymbol{w}\|_2 \leqslant C$, by (4.8) and $\|\Gamma(p)^{-1}\|_{\mathrm{sp}} \leqslant C$ by (4.13). On the other hand,

$$\begin{aligned}\|M(p)\|_{\mathrm{sp}} &= \|\tfrac{1}{2}\Gamma(p)^{-1}[I + \tfrac{1}{2}\omega_0(p)\boldsymbol{w}(p)\boldsymbol{w}(p)^T\Gamma(p)^{-1}]\|_{\mathrm{sp}}\\ &\leqslant \tfrac{1}{2}\|\Gamma(p)^{-1}\|_{\mathrm{sp}} + \tfrac{1}{4}|\omega_0(p)|\,\|\Gamma(p)^{-1}\boldsymbol{w}(p)\|_{\mathrm{sp}}^2\\ &\leqslant C + |\omega_0(p)|\,\|\Gamma(p)^{-1}\|_{\mathrm{sp}}^2\|\boldsymbol{w}\|_2^2 \leqslant Cp,\end{aligned}$$

by (4.3). Thus $\|W(p)^{-1}\|_{\mathrm{sp}} \leqslant Cp$, to prove (4.4). □

Lemma 4.2. *Under Assumptions* 1–3, *uniformly in p*,

$$\|\boldsymbol{w}\|_2 \leqslant C, \tag{4.8}$$

$$\|S(p)\|_{\mathrm{sp}} \leqslant C, \quad \|\widetilde{S}(p)\|_{\mathrm{sp}} \leqslant C, \tag{4.9}$$

$$\|U(p)\|_{\mathrm{sp}} \leqslant C, \quad \|\widetilde{U}(p)\|_{\mathrm{sp}} \leqslant C, \tag{4.10}$$

$$\|S(p)\|_2 \leqslant Cp^{1/2}, \quad \|\widetilde{S}(p)\|_2 \leqslant Cp^{1/2}, \tag{4.11}$$

$$\|U(p)\|_2 \leqslant Cp^{1/2}, \quad \|\widetilde{U}(p)\|_2 \leqslant Cp^{1/2}, \tag{4.12}$$

$$\|\Gamma^{-1}(p)\|_{\mathrm{sp}} \leqslant C. \tag{4.13}$$

Proof of Lemma 4.2. By (3.22),

$$|w_{0j}| \leqslant \sum_{k=0}^{\infty} \left| \frac{c_k}{k+j} \right| \leqslant j^{-1} \sum_{k=0}^{\infty} |c_k| \leqslant C j^{-1}, \quad j = 1, \ldots, p, \tag{4.14}$$

since $\sum_{k=0}^{\infty} |c_k| < \infty$ (see Remark 2.1). Thus

$$\|\boldsymbol{w}\|_2 \leqslant \left(\sum_{i=1}^{p} w_i^2 \right)^{1/2} \leqslant \left(C \sum_{i=1}^{\infty} i^{-2} \right)^{1/2} \leqslant C.$$

To show (4.9) note that

$$\sum_{i=0}^{p} |a_i| \leqslant \sum_{i=0}^{\infty} |a_i| =: A < \infty \tag{4.15}$$

and

$$\sum_{i=0}^{p} |\alpha_i| \leqslant \sum_{i=0}^{p} |\alpha_i - a_i| + \sum_{i=0}^{p} |a_i| \leqslant 2 \sum_{i=0}^{\infty} |a_i| =: 2A < \infty \tag{4.16}$$

by (3.35). Hence the elements of the matrix $\widetilde{S} = \{\widetilde{s}_{ij}\}_{i,j=1,\ldots,p}$ have the property

$$\sum_{i=1}^{p} |\widetilde{s}_{ij}| \leqslant A, \quad \sum_{j=1}^{p} |\widetilde{s}_{ij}| \leqslant A$$

and so

$$\|\widetilde{S}\|_{\text{sp}}^2 = \sup_{\|x\|_2=1} \sum_{i,k,j=1}^{p} x_i \widetilde{s}_{ki} \widetilde{s}_{kj} x_j \leqslant \sup_{\|x\|_2=1} \sum_{i,k,j=1}^{p} |\widetilde{s}_{ki} \widetilde{s}_{kj}| (x_i^2 + x_j^2) \leqslant 2A^2 < \infty.$$

This also implies that $\|\widetilde{S}\|_2 \leqslant \sqrt{p} \|\widetilde{S}\|_{\text{sp}} \leqslant C\sqrt{p}$. The bounds for $S, U, \widetilde{U}$ follow by the same argument using (4.15)–(4.16).

Finally, it is well known, see e.g. Berk [2], inequalities (2.14), that $\inf_{|\lambda| \leqslant \pi} |A(\lambda)|^{-2} > 0$ implies $\|R(p)^{-1}\|_{\text{sp}} \leqslant C$. Therefore (4.13) follows from (3.27). □

Lemma 4.3. *Under Assumptions* 1–3,

$$\|R(p)^{-1} - H(p)\|_{\text{sp}} = o(p^{-2}). \tag{4.17}$$

Proof of Lemma 4.3. We show that

$$|\sigma_p^{-2} - \sigma^{-2}| = o(p^{-2}), \tag{4.18}$$

$$\|\widetilde{S}^T \widetilde{S} - S^T S\|_{\text{sp}} = o(p^{-2}), \tag{4.19}$$

$$\|\widetilde{U}^T \widetilde{U} - U^T U\|_{\text{sp}} = o(p^{-2}). \tag{4.20}$$

Then, together with (4.9) and (4.10), the above relations imply

$$\begin{aligned} \|R(p)^{-1} - H(p)\|_{\text{sp}} &\leqslant \sigma_p^{-2} (\|\widetilde{S}^T \widetilde{S} - S^T S\|_{\text{sp}} + \|\widetilde{U}^T \widetilde{U} - U^T U\|_{\text{sp}}) \\ &\quad + |\sigma^{-2} - \sigma_p^{-2}| (\|S^T\|_{\text{sp}} \|S\|_{\text{sp}} + \|U^T\|_{\text{sp}} \|U\|_{\text{sp}}) = o(p^{-2}). \end{aligned}$$

To prove (4.18) note that

$$2\pi(\sigma_p^2 - \sigma^2) = \sigma^2 \int_{-\pi}^{\pi} [|1 + \alpha_1 e^{i\lambda} + \cdots + \alpha_p e^{i\lambda p}|^2 - |A(\lambda)|^2]|A(\lambda)|^{-2} d\lambda$$

and that by (3.35) and (2.19)

$$|(1 + \alpha_1 e^{i\lambda} + \cdots + \alpha_p e^{i\lambda p}) - A(\lambda)| \leqslant C \sum_{j=p+1}^{\infty} |a_j| = o(p^{-2}).$$

To show (4.19), note that $\widetilde{S} - S = \{s'_{kj}\}_{k,j=1,\ldots,p}$ where $\sum_{k=1}^{p} |s'_{kj}| \leqslant \sum_{i=0}^{p} |\alpha_i - a_i| \leqslant C \sum_{i=p+1}^{\infty} |a_i| =: A_p = o(p^{-2})$ by (3.35) and (2.19). Hence, by the same argument as in the proof of Lemma 4.2, it follows that

$$\|\widetilde{S} - S\|_{\mathrm{sp}} \leqslant \sqrt{2} A_p = o(p^{-2}), \quad \|\widetilde{U} - U\|_{\mathrm{sp}} = o(p^{-2})$$

which together with (4.9) imply (4.19). Verification of (4.20) is similar. □

Lemma 4.4. *Under Assumptions* 1–3, *there exists* $C > 0$ *such that*

$$\|U(p)\boldsymbol{w}\|_2 \leqslant Cp^{-1}, \tag{4.21}$$

$$\|S(p)\boldsymbol{w} - \boldsymbol{t}\|_2 \leqslant Cp^{-1} \tag{4.22}$$

and

$$|[S(p)\boldsymbol{w} - \boldsymbol{t}]_i| \leqslant C(p-i)^{-2}p^{-1}, \quad i = 1, \ldots, p,$$

$$\|(4\pi)^{-1} S(p)^T S(p)\boldsymbol{w} + \boldsymbol{\tau}\|_2 \leqslant Cp^{-1}, \tag{4.23}$$

$$\sum_{i=1}^{p} \left| \left[\frac{1}{2}\Gamma(p)^{-1}\boldsymbol{w} + \boldsymbol{\tau} \right]_i \right| \leqslant Cp^{-1}, \tag{4.24}$$

$$\left\| \frac{1}{2}\Gamma(p)^{-1}\boldsymbol{w} + \boldsymbol{\tau} \right\|_2 \leqslant Cp^{-1}, \tag{4.25}$$

where $\boldsymbol{t}$ *and* $\boldsymbol{\tau}$ *are defined in* (4.2).

Proof of Lemma 4.4. We first verify (4.21). By (4.14), $|w_j| \leqslant Cj^{-1}$, so by (2.19)

$$\begin{aligned} |[U\boldsymbol{w}]_i| &\leqslant \sum_{j=i}^{p} |a_{p-j+i} w_j| \leqslant \sum_{j=i:j\leqslant p/2}^{p} |a_{p-j+i}|C + \sum_{j=i:j>p/2}^{p} |a_{p-j+i}|Cp^{-1} \\ &\leqslant C \left(\sum_{j=p/2}^{\infty} |a_j| + p^{-1} \sum_{j=i}^{\infty} |a_j| \right) \leqslant C(p^{-2} + i^{-2}p^{-1}) \leqslant Cp^{-1}i^{-1} \end{aligned} \tag{4.26}$$

which shows that $\|U\boldsymbol{w}\|_2 = (\sum_{i=1}^{p} [U\boldsymbol{w}]_i^2)^{1/2} \leqslant Cp^{-1}$.

Now we establish (4.22). Since for every $t \geqslant 1$, $\sum_{u=0}^{t} a_u c_{t-u} = 0$, (3.22) and (4.7) imply that

$$t_i = -\frac{4\pi}{i} = \sum_{j=0}^{\infty} a_j w_{i+j}. \tag{4.27}$$

Consequently,

$$[S\boldsymbol{w}]_i = \sum_{j=0}^{p-i} a_j w_{i+j} = t_i - q_i \tag{4.28}$$

with

$$t_i = \sum_{j=0}^{\infty} a_j w_{i+j}, \quad q_i = \sum_{j=p-i+1}^{\infty} a_j w_{i+j}.$$

For $j \geqslant p - i + 1$, $|w_{i+j}| \leqslant C(i+j)^{-1} \leqslant Cp^{-1}$, and therefore by (2.19)

$$|[S\boldsymbol{w}]_i - t_i| = |q_i| \leqslant \sum_{j=p-i+1}^{\infty} |a_j| C p^{-1} \leqslant C(p-i)^{-2} p^{-1}, \quad i = 1, \ldots, p. \tag{4.29}$$

Hence

$$\|S\boldsymbol{w} - \boldsymbol{t}\|_2^2 = \sum_{i=1}^{p} ([S\boldsymbol{w}]_i - t_i)^2 = \sum_{i=1}^{p} q_i^2 \leqslant Cp^{-2} \tag{4.30}$$

to prove (4.22).

To prove (4.23), observe that

$$(4\pi)^{-1}[S^T \boldsymbol{t}]_i = -\sum_{j=1}^{i} \frac{a_{i-j}}{j} = -\tau_i, \quad i = 1, \ldots, p$$

so that

$$(4\pi)^{-1}[S^T S\boldsymbol{w}]_i = (4\pi)^{-1}[S^T(S\boldsymbol{w} - \boldsymbol{t})]_i - \tau_i \tag{4.31}$$

and by (4.30) and (4.9)

$$\begin{aligned} \|(4\pi)^{-1} S^T S\boldsymbol{w} + \boldsymbol{\tau}\|_2 &\leqslant (4\pi)^{-1} \|S^T(S\boldsymbol{w} - \boldsymbol{t})\|_2 \\ &\leqslant (4\pi)^{-1} \|S\|_{\mathrm{sp}} \|S\boldsymbol{w} - \boldsymbol{t}\|_2 \leqslant Cp^{-1}. \end{aligned}$$

To show (4.24) note that by (3.28)

$$\begin{aligned} &\sum_{i=1}^{p} \left| \left[\frac{1}{2} \Gamma(p)^{-1} \boldsymbol{w} + \boldsymbol{\tau} \right]_i \right| \\ &\quad \leqslant C \sum_{i=1}^{p} |[(4\pi)^{-1} S^T S\boldsymbol{w} + \boldsymbol{\tau}]_i| + \sum_{i=1}^{p} |[U^T U\boldsymbol{w}]_i| + \sum_{i=1}^{p} |[(R(p)^{-1} - H(p))\boldsymbol{w}]_i| \\ &\quad =: s_{n,1} + s_{n,2} + s_{n,3}. \end{aligned}$$

By (4.28) and (4.31),

$$s_{n,1} \leqslant C \sum_{i=1}^{p} |[(4\pi)^{-1} S^T (S\boldsymbol{w} - \boldsymbol{t})]_i| \leqslant C \sum_{i,j=1}^{p} |[S^T]_{ij} q_j| \leqslant C \sum_{i} |a_i| \sum_{j=1}^{p} |q_j| \leqslant Cp^{-1},$$

in view of (4.29). Using (4.26), (4.31) and (2.19), we can estimate

$$s_{n,2} \leqslant C \sum_{i=1}^{p} \sum_{j=1}^{p} |[U^T]_{ij} [U\boldsymbol{w}]_j| \leqslant C \sum_{i=1}^{p} |a_i| \sum_{j=1}^{p} |[U\boldsymbol{w}]_j| \leqslant C \sum_{j=1}^{p} (p^{-2} + j^{-2} p^{-1}) \leqslant C p^{-1}.$$

On the other hand,

$$\begin{aligned} |s_{n,3}| &\leqslant C p^{1/2} \|(R(p)^{-1} - H(p))\boldsymbol{w}\|_2 \\ &\leqslant C p^{1/2} \|R(p)^{-1} - H(p)\|_{\mathrm{sp}} \|\boldsymbol{w}\|_{\mathrm{sp}} \leqslant C p^{1/2} p^{-2} \leqslant C p^{-1} \end{aligned}$$

by (4.17) and (4.8).

Finally, (4.25) follows from (4.24), since

$$\begin{aligned} \left\| \frac{1}{2} \Gamma(p)^{-1} \boldsymbol{w} + \boldsymbol{\tau} \right\|_2^2 &= \sum_{j=1}^{p} \left[\frac{1}{2} \Gamma(p)^{-1} \boldsymbol{w} + \boldsymbol{\tau} \right]_j^2 \\ &\leqslant \left(\sum_{j=1}^{p} \left| \left[\frac{1}{2} \Gamma(p)^{-1} \boldsymbol{w} + \boldsymbol{\tau} \right]_j \right| \right)^2 \leqslant C p^{-2}. \qquad \square \end{aligned}$$

The next lemma and Lemma 5.1 form the proof of Theorem 3.1.

Lemma 4.5. *Under Assumptions* 1–3,

$$E \|W(p)^{-1} \boldsymbol{\zeta}(p)\|_2 = O((p/n)^{1/2}). \tag{4.32}$$

Proof of Lemma 4.5. Using relations (3.38) and (3.39), we obtain

$$\begin{aligned} \|W(p)^{-1} \boldsymbol{\zeta}(p)\|_2 &\leqslant C(|V| + \|\Gamma^{-1} \boldsymbol{w}\|_2 |V| + \|\Gamma^{-1} \boldsymbol{\zeta}_+(p)\|_2) \\ &\leqslant C(|V| + \|\Gamma^{-1}\|_{\mathrm{sp}} \|\boldsymbol{w}\|_2 |V| + \|\Gamma^{-1}\|_{\mathrm{sp}} \|\boldsymbol{\zeta}_+(p)\|_2). \end{aligned}$$

By Lemma 4.2 we have that $\|\Gamma^{-1}\|_{\mathrm{sp}} \leqslant C$, $\|\boldsymbol{w}\|_2 \leqslant C$. By Lemma 6.3 below, $E\|\boldsymbol{\zeta}(p)\|_2 = O((p/n)^{1/2})$ and in Proposition 6.1 below it is shown that

$$E|V| = O((p/n)^{1/2}). \tag{4.33}$$

Thus $E\|W(p)^{-1}\boldsymbol{\zeta}(p)\|_2 = O((p/n)^{1/2})$, to prove (4.32). $\square$

5. The remainder term

In this section, we prove two lemmas which establish relation (3.37) on which, together with Lemma 4.5, the proof of Theorem 3.1 rests.

Lemma 5.1. *Under assumptions of Theorem* 3.1,

$$\boldsymbol{v}(p) = -\frac{2\pi}{\sigma^2} W(p)^{-1} \boldsymbol{\zeta}(p) + \boldsymbol{r}(p) \tag{5.1}$$

where

$$\|\boldsymbol{r}(p)\|_2 = o_P((p/n)^{1/2}). \tag{5.2}$$

Proof of Lemma 5.1. We show first that

$$\|\boldsymbol{v}(p)\|_2 = O_P((p^{3/2}/n^{1/2})). \tag{5.3}$$

By the mean value theorem,

$$0 \geqslant G_p(\widehat{\boldsymbol{\beta}}_p) - G_p(\boldsymbol{\beta}_p^0) = \boldsymbol{v}(p)^T \nabla G_p(\boldsymbol{\beta}_p^0) + \boldsymbol{v}(p)^T \widehat{W}_p(\boldsymbol{\beta}_p^0)\boldsymbol{v}(p) + R_n(\boldsymbol{\beta}_p^*),$$

where

$$R_n(\boldsymbol{\beta}_p^*) = \sum_{i,j,k=0}^{p} v_i v_j v_k \widehat{w}_{ijk}(\boldsymbol{\beta}_p^*), \tag{5.4}$$

$(v_0, \ldots, v_p) = (\widehat{\boldsymbol{\beta}}_p - \boldsymbol{\beta}_p^0) = (\widehat{d} - d^0, \widehat{a}_1 - a_1^0, \ldots, \widehat{a}_p - a_p^0)$, and $\|\boldsymbol{\beta}_p^* - \boldsymbol{\beta}_p^0\|_2 \leqslant \|\widehat{\boldsymbol{\beta}}_p - \boldsymbol{\beta}_p^0\|_2$. Recall that (see (3.12)),

$$\widehat{w}_{ijk}(\boldsymbol{\beta}_p^*) = \frac{\partial^3}{\partial\beta_i \partial\beta_j \partial\beta_k} G_p(\boldsymbol{\beta}_p^*), \tag{5.5}$$

where G_p is given by (2.13).

We show below that

$$|\boldsymbol{v}(p)^T \nabla G_p(\boldsymbol{\beta}_p^0)| = \|\boldsymbol{v}(p)\|_2 O_P((p/n)^{1/2}), \tag{5.6}$$

$$\boldsymbol{v}(p)^T \widehat{W}_p(\boldsymbol{\beta}_p^0)\boldsymbol{v}(p) \geqslant cp^{-1}\|\boldsymbol{v}(p)\|_2^2(1 + o_P(1)) \tag{5.7}$$

for some $c > 0$. By (5.14) of Lemma 5.2 below, $|R_n(\boldsymbol{\beta}_p^*)| = \|\boldsymbol{v}(p)\|_2^2 o_P(p^{-1})$ which together with (5.6) and (5.7) yields

$$cp^{-1}\|\boldsymbol{v}(p)\|_2^2(1 + o_P(1)) \leqslant \|\boldsymbol{v}(p)\|_2 O_P((p/n)^{1/2}),$$

to prove (5.3).

By Lemma 6.3 below,

$$\|\boldsymbol{\zeta}(p)\|_2 \equiv \|\nabla G_p(\boldsymbol{\beta}_p^0)\|_2 = O_P((p/n)^{1/2}).$$

Therefore

$$\|\boldsymbol{v}(p)^T \boldsymbol{\zeta}(p)\|_2 \leqslant \|\boldsymbol{v}(p)\|_2 \|\boldsymbol{\zeta}(p)\|_2 = \|\boldsymbol{v}(p)\|_2 O_P((p/n)^{1/2}),$$

to prove (5.6).

To prove (5.7) note that by Lemma 6.3 below, $\|(\widehat{W}_p(\boldsymbol{\beta}_p^0) - W_p(\boldsymbol{\beta}_p^0))\|_2 = O_P(p/n^{1/2})$. Hence,

$$\begin{aligned}|\boldsymbol{v}(p)^T (\widehat{W}_p(\boldsymbol{\beta}_p^0) - W_p(\boldsymbol{\beta}_p^0))\boldsymbol{v}(p)| &\leqslant \|\boldsymbol{v}(p)\|_2^2 \|\widehat{W}_p(\boldsymbol{\beta}_p^0) - W_p(\boldsymbol{\beta}_p^0)\|_2 \\ &= \|\boldsymbol{v}(p)\|_2^2 O_P(p/n^{1/2}) = \|\boldsymbol{v}(p)\|_2^2 o_P(p^{-1})\end{aligned}$$

because $p = o(n^{1/8})$ by Assumption 3, which implies that

$$\boldsymbol{v}(p)^T \widehat{W}_p(\boldsymbol{\beta}_p^0)\boldsymbol{v}(p) = \boldsymbol{v}(p)^T W_p(\boldsymbol{\beta}_p^0)\boldsymbol{v}(p) + \|\boldsymbol{v}(p)\|_2^2 o_P(p^{-1}). \tag{5.8}$$

By Lemma 4.1,

$$\|W_p(\boldsymbol{\beta}_p^0)^{-1}\|_{\text{sp}} = \frac{2\pi}{\sigma^2}\|W(p)^{-1}\|_{\text{sp}} \leqslant cp$$

with $0 < c < \infty$. Therefore $\|W_p(\boldsymbol{\beta}_p^0)\|_{\mathrm{sp}} \geqslant c/p$, see e.g. Rao [19], p. 33, and so $\boldsymbol{v}(p)^T W_p(\boldsymbol{\beta}_p^0)\boldsymbol{v}(p) \geqslant c\|\boldsymbol{v}(p)\|_2^2/p$ which together with (5.8) implies (5.7). This completes the verification of (5.3).

We now establish (5.1) and (5.2). Note that, as $n \to \infty$, (2.20) implies that, with probability tending to 1, the estimate $\widehat{d}$ of $d^0 \in (-\frac{1}{2}, \frac{1}{2})$ is an inner point of the interval $I_{n,d}$. Indeed, since $p = o(n^{1/8})$ by Assumption 3, then by (5.3), $|\widehat{d} - d^0| = O_P((p^3/n)^{1/2}) = o_P(n^{-1/4})$, and therefore

$$|\widehat{d} - \widetilde{d}| \leqslant |\widehat{d} - d^0| + |d^0 - \widetilde{d}| = o(n^{-1/4}) + o_P(\Delta_n) = o_P(\Delta_n),$$

by (2.12) and (2.20). Hence, by Taylor expansion,

$$\nabla G_p(\widehat{\boldsymbol{\beta}}_p) - \nabla G_p(\boldsymbol{\beta}_p^0) = \widehat{W}_p(\boldsymbol{\beta}_p^0)\boldsymbol{v}(p) + \tfrac{1}{2}\boldsymbol{v}(p)^T \nabla^3 G_p(\boldsymbol{\beta}_p^*)\boldsymbol{v}(p), \tag{5.9}$$

where $\nabla G_p(\widehat{\boldsymbol{\beta}}_p) = 0$. Setting $\boldsymbol{\zeta}(p) = \nabla G_p(\boldsymbol{\beta}_p^0)$, we obtain

$$-\boldsymbol{\zeta}(p) = W_p(\boldsymbol{\beta}_p^0)\boldsymbol{v}(p) + \boldsymbol{R}_n^*, \tag{5.10}$$

where

$$\boldsymbol{R}_n^* = (\widehat{W}_p(\boldsymbol{\beta}_p^0) - W_p(\boldsymbol{\beta}_p^0))\boldsymbol{v}(p) + \tfrac{1}{2}\boldsymbol{v}(p)^T \nabla^3 G_p(\boldsymbol{\beta}_p^*)\boldsymbol{v}(p).$$

Recall that $W_p(\boldsymbol{\beta}_p^0) = \frac{\sigma^2}{2\pi} W(p)$. Hence

$$\boldsymbol{v}(p) = -\frac{2\pi}{\sigma^2} W(p)^{-1}\boldsymbol{\zeta}(p) - \frac{2\pi}{\sigma^2} W(p)^{-1}\boldsymbol{R}_n^* \tag{5.11}$$

and (5.2) follows if we show that

$$\|W(p)^{-1}\boldsymbol{R}_n^*\|_2 = o_P((p/n)^{1/2}). \tag{5.12}$$

We will show that

$$\|W(p)^{-1}\boldsymbol{R}_n^*\|_2 = o_P(\|\boldsymbol{v}(p)\|_2). \tag{5.13}$$

Since by Lemma 4.5 $\|W(p)^{-1}\boldsymbol{\zeta}(p)\|_2 = O_P((p/n)^{1/2})$, then (5.11) and (5.13) imply that $\|\boldsymbol{v}(p)\|_2 = O_P((p/n)^{1/2})$ which in view of (5.13) proves (5.12).

To show (5.13), observe that by Lemma 4.1, $\|W(p)^{-1}\|_{\mathrm{sp}} \leqslant Cp$. Therefore

$$\begin{aligned}
\|W(p)^{-1}\boldsymbol{R}_n^*\|_2 &\leqslant \|W(p)^{-1}\|_{\mathrm{sp}}\|\boldsymbol{R}_n^*\|_2 \\
&\leqslant Cp(\|\widehat{W}_p(\boldsymbol{\beta}_p^0) - W_p(\boldsymbol{\beta}_p^0)\|_2\|\boldsymbol{v}(p)\|_2 + \|\nabla^3 G_p(\boldsymbol{\beta}_p^*)\|_2\,\|\boldsymbol{v}(p)\|_2^2) \\
&\leqslant C\|\boldsymbol{v}(p)\|_2(p\|\widehat{W}_p(\boldsymbol{\beta}_p^0) - W_p(\boldsymbol{\beta}_p^0)\|_2 + p\|\nabla^3 G_p(\boldsymbol{\beta}_p^*)\|_2\,\|\boldsymbol{v}(p)\|_2) \\
&= o_P(\|\boldsymbol{v}(p)\|_2)
\end{aligned}$$

since by Lemma 6.3 below, $p\|(\widehat{W}_p(\boldsymbol{\beta}_p^0) - W_p(\boldsymbol{\beta}_p^0))\|_2 = O_P(p^2/n^{1/2}) = o_P(1)$, because $p = o(n^{1/8})$, whereas (5.3) and (5.15) of Lemma 5.2 below imply that

$$p\|\nabla^3 G_p(\boldsymbol{\beta}_p^*)\|_2\,\|\boldsymbol{v}(p)\|_2 = o_P(p^4 n^{-1/2}) = o_P(1),$$

in view of assumption (2.8). □

Lemma 5.2. *Suppose that assumptions of Lemma* 5.1 *are satisfied. Then* $R_n(\boldsymbol{\beta}_p^*)$ *in* (5.4) *satisfies the bound*

$$|R_n(\boldsymbol{\beta}_p^*)| = \|\boldsymbol{v}(p)\|_2^2 o_P(p^{-1}), \tag{5.14}$$

and $\nabla^3 G_p(\boldsymbol{\beta}_p^*)$ *in* (5.9) *has the property*

$$\|\nabla^3 G_p(\boldsymbol{\beta}_p^*)\|_2 = o_P(p^{3/2}). \tag{5.15}$$

Proof of Lemma 5.2. Since, by (3.18),

$$\widehat{w}_{ijk}(\boldsymbol{\beta}_p^*) = 0 \quad \text{if } 1 \leqslant i, j, k \leqslant p, \tag{5.16}$$

we can write

$$R_n(\boldsymbol{\beta}_p^*) = 3v_0 \sum_{j,k=1}^{p} v_j v_k \widehat{w}_{0jk}(\boldsymbol{\beta}_p^*) + 3v_0^2 \sum_{k=1}^{p} v_k \widehat{w}_{00k}(\boldsymbol{\beta}_p^*) + v_0^3 \widehat{w}_{000}(\boldsymbol{\beta}_p^*). \tag{5.17}$$

Because $\sup_{\beta \in E} \|\boldsymbol{\beta}\|_1 < \infty$ by Assumption 2, then

$$|\widehat{w}_{ijk}(\boldsymbol{\beta}_p^*)| \leqslant C \int_{-\pi}^{\pi} (|\ln|1 - e^{i\lambda}\|+1)^3 |1 - e^{i\lambda}|^{2d^*} I_n(\lambda)\, d\lambda =: C j_n(d^*) \tag{5.18}$$

uniformly in $0 \leqslant i, j, k \leqslant n$. Hence

$$|R_n(\boldsymbol{\beta}_p^*)| \leqslant C j_n(d^*) \upsilon_n, \tag{5.19}$$

where

$$\begin{aligned} \upsilon_n &= |v_0| \sum_{j,k=1}^{p} |v_j v_k| + v_0^2 \sum_{k=1}^{p} |v_k| + |v_0|^3 \\ &\leqslant C \left(p|v_0| \sum_{j=1}^{p} v_j^2 + p^{1/2} v_0^2 \left(\sum_{j=1}^{p} v_j^2 \right)^{1/2} + |v_0|^3 \right) \\ &\leqslant C \|\boldsymbol{v}(p)\|_2^2 (p|v_0| + p^{1/2}|v_0| + |v_0|), = \|\boldsymbol{v}(p)\|_2^2 o_P(p^{-1}) \end{aligned}$$

since by (2.20), $|v_0| = |\widehat{d} - d^0| = o_P(p^{-\gamma})$ with $\gamma > 2$. To complete proof of (5.14) it remains to show that

$$j_n(d^*) = O_P(1). \tag{5.20}$$

We have that for any $\varepsilon \in (0, 1/4)$,

$$E[j_n^2(d^*) 1_{\{d^* \geqslant d^0 - \varepsilon\}}] \leqslant CE \left[\int_{-\pi}^{\pi} |(|\ln|1 - e^{i\lambda}| + 1)^3 |1 - e^{i\lambda}|^{2d^0 - 2\varepsilon} I_n(\lambda)\, d\lambda \right]^2 = O(1)$$

by (6.6) of Lemma 6.2, which shows that $j_n(d^*) 1_{\{d^* \geqslant d^0 - \varepsilon\}} = O_P(1)$. On the other hand, $P(d^* < d^0 - \varepsilon) \leqslant P(|d^* - d^0| \leqslant \varepsilon) \to 0$, since $|d^* - d^0| \leqslant |\widehat{d} - d^0| \xrightarrow{P} 0$, to prove (5.20).

Relations (5.18) and (5.20) imply that

$$\|\nabla^3 G_p(\boldsymbol{\beta}_p^*)\|^2 = \sum_{i,j,k=0}^{p} \widehat{w}_{ijk}^2(\boldsymbol{\beta}_p^*) \leqslant C j_n^2(d^*) \sum_{i,j,k=0}^{p} 1 = O_P(p^3)$$

to prove (5.15). □

6. Central limit theorem and approximation bounds

The main result of this section is Proposition 6.1 on which the proof of Theorem 3.2 rests. Its proof relies on Lemma 6.2 which collects several results established by Bhansali et al. [6]. These results include L_2 approximations and a CLT for a form of the integrated periodogram needed to prove Proposition 6.1. Lemma 6.2 holds for linear processes (6.1) satisfying (6.2), (6.3). In Lemma 6.1 we verify that Assumption 1 implies these conditions. Lemmas 6.3 and 6.4 are used in the proofs of Theorems 3.1 and 3.2.

As in Section 4, in this section we denote $a_j = a_j^0$, $c_j = c_j^0$.

Lemma 6.1. *If the process* $\{X_t\}$ *satisfies Assumption* 1, *then it admits the representation*

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j} \tag{6.1}$$

with the ψ_j *satisfying*

$$\psi_j = cj^{-1+d}(1 + O(j^{-1})) \quad \text{if } d \neq 0,\ |d| < 1/2, \tag{6.2}$$

where $c = [a(1)\Gamma(d)]^{-1}$. *Moreover, if* $d \in (-1/2, 0)$ *then* $\sum_{j=0}^{\infty} \psi_j = 0$, *and*

$$\sum_{j=n}^{\infty} |\psi_j| = O(n^{-2}) \quad \text{if } d = 0. \tag{6.3}$$

Proof of Lemma 6.1. In the case $d = 0$, $\psi_j = c_j^0$, so (6.3) follows from (2.17).

The case $d \neq 0$ can be handled as in Sections 2 and 3 of Kokoszka and Taqqu [15] who considered the special case of $c(z) = \theta_p(z)/\phi_q(z)$, where $\theta_p(z)$ and $\phi_q(z)$ are the usual moving average and autoregressive polynomials and $\phi_q(z)$ has no roots in the closed unit disk. The latter assumption implies that the coefficients of $c(z)$ decay exponentially fast. In our setting, the c_j also decay exponentially fast, so (6.2) can be established for the ψ_j defined by

$$\sum_{j=0}^{\infty} \psi_j z^j = c(z)(1-z)^{-d}, \quad |z| < 1, \tag{6.4}$$

by following exactly the proof of Lemma 3.2 of Kokoszka and Taqqu [15]. The proof is exactly the same for positive and negative d because it uses the asymptotic order of the ratio $\Gamma(j+d)/\Gamma(j+1)$, as $j \to \infty$.

For the ψ_j defined by (6.4), it can then be verified, e.g. as in the proof of Theorem 2.1 of Kokoszka and Taqqu [15], that (6.1) is the unique causal moving average solution to Eqs. (2.1) and (2.2). □

Set

$$\eta_n(\lambda) = 2\pi b_n(\lambda) f(\lambda),$$

where b_n an even real function and f is the spectral density of (X_t). Define also the matrix

$$E_n = (e_{t-s})_{t,s=1}^n, \quad e_t = \int_{-\pi}^{\pi} e^{it\lambda}\eta_n(\lambda)\,d\lambda. \tag{6.5}$$

Lemma 6.2. *Assume* $\{X_t\}$, (6.1), *satisfies either* (6.2) *or* (6.3), *and* $b_n(\lambda)$ *is a real even function such that*

$$|b_n(\lambda)| \leqslant K_n|\lambda|^{2d-\delta}, \quad |\lambda| \leqslant \pi,$$

where $0 \leqslant \delta < \frac{1}{4}$. *Then, as* $n \to \infty$,

$$E\left|\int_{-\pi}^{\pi} b_n(\lambda)I_n(\lambda)\,d\lambda - \int_{-\pi}^{\pi} b_n(\lambda)f(\lambda)\,d\lambda\right|^2 \leqslant CK_n^2 n^{-1}, \tag{6.6}$$

where C does not depends on n and K_n. *If in addition*

$$\frac{K_n n^{\max(\delta,d,0)}\log n}{\|E_n\|_2} \to 0 \tag{6.7}$$

and

$$\int_{-\pi}^{\pi} \eta_n(\lambda)\,d\lambda = o(n^{-1/2}\|E_n\|_2), \tag{6.8}$$

then

$$\frac{\sqrt{2}\pi n}{\|E_n\|_2}\left(\int_{-\pi}^{\pi} b_n(\lambda)I_n(\lambda)\,d\lambda - \int_{-\pi}^{\pi} b_n(\lambda)f(\lambda)\,d\lambda\right) \xrightarrow{d} N(0, 1) \tag{6.9}$$

and

$$\frac{2(\pi n)^2}{\|E_n\|_2^2} E\left(\int_{-\pi}^{\pi} b_n(\lambda)I_n(\lambda)\,d\lambda - \int_{-\pi}^{\pi} b_n(\lambda)f(\lambda)\,d\lambda\right)^2 \to 1. \tag{6.10}$$

Lemma 6.3. *Under Assumptions* 1–3,

$$\|\widehat{W}_p(\boldsymbol{\beta}_p^0) - W_p(\boldsymbol{\beta}_p^0)\|_2 = O_P(p/n^{1/2}) \tag{6.11}$$

and

$$E\|\boldsymbol{\zeta}(p)\|_2 = O((p/n)^{1/2}). \tag{6.12}$$

Proof of Lemma 6.3. By definition, $\boldsymbol{\zeta}(p) = (\zeta_0, \zeta_1, \ldots, \zeta_p)$ and $\widehat{W}_p(\boldsymbol{\beta}_p^0) - W_p(\boldsymbol{\beta}_p^0) = (\widetilde{w}_{jk})_{j,k=0,\ldots,p}$ where

$$\zeta_j = \int_{-\pi}^{\pi} h_j(\lambda, \boldsymbol{\beta}_p^0)I_n(\lambda)\,d\lambda, \quad \tilde{w}_{jk} = \int_{-\pi}^{\pi} h_{jk}(\lambda, \boldsymbol{\beta}_p^0)(I_n(\lambda) - f(\lambda))\,d\lambda,$$
$$j, k = 0, 1, \ldots, p,$$

and

$$h_j(\lambda, \boldsymbol{\beta}_p) = \frac{\partial}{\partial \beta_j} g_p^{-1}(\lambda, \boldsymbol{\beta}_p), \quad h_{jk}(\lambda, \boldsymbol{\beta}_p) = \frac{\partial^2}{\partial \beta_k \partial \beta_j} g_p^{-1}(\lambda, \boldsymbol{\beta}_p).$$

Since for any $\varepsilon > 0$, uniformly in $|\lambda| \leqslant \pi$ and $0 \leqslant j, k \leqslant p$,

$$|h_j(\lambda, \boldsymbol{\beta}_p^0)| \leqslant C|1 - e^{i\lambda}|^{2d^0 - \varepsilon}, \quad |h_{jk}(\lambda, \boldsymbol{\beta}_p^0)| \leqslant C|1 - e^{i\lambda}|^{2d^0 - \varepsilon},$$

then by (6.6), $E|\tilde{w}_{jk}|^2 \leqslant Cn^{-1}$ uniformly in $0 \leqslant j, k \leqslant p$, and therefore

$$E\|(\widehat{W}_p(\boldsymbol{\beta}_p^0) - W_p(\boldsymbol{\beta}_p^0))\|_2^2 \leqslant \sum_{j,k=0}^{p} E|\widetilde{w}_{jk}|^2 \leqslant Cp^2 n^{-1},$$

to prove (6.11).

Next, we show that

$$E|\zeta_j|^2 \leqslant Cn^{-1} \tag{6.13}$$

uniformly in $1 \leqslant j \leqslant p$ which implies (6.12). We have that

$$\begin{aligned} E|\zeta_j|^2 &\leqslant 2E\left|\int_{-\pi}^{\pi} h_j(\lambda, \boldsymbol{\beta}_p^0)(I_n(\lambda) - f(\lambda))d\lambda\right|^2 \\ &\quad +2\left|\int_{-\pi}^{\pi} h_j(\lambda, \boldsymbol{\beta}_p^0) f(\lambda)\, d\lambda\right|^2 = q_{n,1}(j) + q_{n,2}(j), \end{aligned}$$

where $q_{n,1}(j) = O(n^{-1})$ uniformly in $0 \leqslant j \leqslant p$, by (6.6). On the other hand, it is easy to see that under Assumption 3,

$$\begin{aligned} \int_{-\pi}^{\pi} h_j(\lambda, \boldsymbol{\beta}_p^0) f(\lambda)\, d\lambda &= \int_{-\pi}^{\pi} \left[\frac{\partial}{\partial \beta_j} g_p^{-1}(\lambda, \boldsymbol{\beta}_p^0)\right] f(\lambda)\, d\lambda \\ &= \int_{-\pi}^{\pi} \left[\frac{\partial}{\partial \beta_j} f^{-1}(\lambda)\right] f(\lambda)\, d\lambda + O(n^{-1/2}) = O(n^{-1/2}), \\ &\quad j = 0, 1, \ldots, p, \end{aligned}$$

since

$$\int_{-\pi}^{\pi} \left[\frac{\partial}{\partial \beta_j} f^{-1}(\lambda)\right] f(\lambda)\, d\lambda = 0, \quad j = 0, 1, \ldots, p,$$

which completes the proof of (6.13).

The last equality is well-known as the normalization condition and is satisfied in our case because

$$\int_{-\pi}^{\pi} \left[\frac{\partial}{\partial \beta_0} f^{-1}(\lambda)\right] f(\lambda)\, d\lambda = \int_{-\pi}^{\pi} |1 - e^{i\lambda}|^2\, d\lambda = 0$$

and

$$\int_{-\pi}^{\pi} \left[\frac{\partial}{\partial \beta_j} f^{-1}(\lambda)\right] f(\lambda)\, d\lambda = 2\int_{-\pi}^{\pi} \sum_{u=0}^{\infty} c_u \cos((u + j)\lambda)\, d\lambda = 0,$$
$$j = 1, \ldots, p. \qquad \square$$

Lemma 6.4. *Under Assumptions* 1–3, *uniformly in* $i = 1, \dots, p$,

$$E|[\Gamma(p)^{-1}\zeta_+(p)]_i| = O(n^{-1/2}), \tag{6.14}$$

where $\zeta_+(p) = (\zeta_1, \dots, \zeta_p)$.

Proof of Lemma 6.4. To lighten the notation, we suppress the dependence of the matrices $\Gamma(p)$, $R(p)$, $H(p)$, $S(p)$ and $U(p)$ on p.

By (3.27), we have that

$$E|[\Gamma^{-1}\zeta_+(p)]_i| \leqslant CE|[R^{-1}\zeta_+(p)]_i| \leqslant C\sum_{j=1}^{p}|[R^{-1}]_{ij}|E|\zeta_j| \leqslant Cn^{-1/2}\sum_{j=1}^{p}|[R^{-1}]_{ij}|$$

by (6.13). Now,

$$\sum_{j=1}^{p}|[R^{-1}]_{ij}| \leqslant \sum_{j=1}^{p}|[R^{-1}]_{ij} - [H]_{ij}| + \sum_{j=1}^{p}|[H]_{ij}| =: T_1(i) + T_2(i),$$

where H is defined by (3.32). It remains to show that $|T_u(i)| \leqslant C$, $u = 1, 2$, uniformly in $1 \leqslant i \leqslant p$. We have that

$$T_1(i) \leqslant p^{1/2}\left(\sum_{j=1}^{p}|[R^{-1}]_{ij} - [H]_{ij}|^2\right)^{1/2} \leqslant p^{1/2}\|R^{-1} - H\|_2 = O(1)$$

by Lemma 4.3. On the other hand, by definition $H = \sigma^{-2}[S^TS - U^TU]$. Denote by a_{ij} and b_{ij}, $i, j = 1, \dots, p$ the elements of matrices S and U. Then, in view of (3.33) and (3.34), it is easy to see that

$$\sum_{j=1}^{p}|[S^TS]_{ij}| \leqslant C\sum_{k,j=1}^{p}|a_{ki}a_{kj}| \leqslant \left(\sum_{j=0}^{\infty}|a_j|\right)^2 < \infty$$

and

$$\sum_{j=1}^{p}|[U^TU]_{ij}| \leqslant C\sum_{k,j=1}^{p}|b_{ki}b_{kj}| \leqslant \left(\sum_{j=0}^{\infty}|a_j|\right)^2 < \infty$$

which implies that $|T_2(i)| \leqslant C < \infty$, uniformly in $1 \leqslant i \leqslant p$. □

Proposition 6.1. *Under Assumptions* 1–3, *the random variable V defined in* (3.38) *satisfies*

$$(2\pi/\sigma^2)(n/p)^{1/2}V \xrightarrow{d} N(0, 1) \tag{6.15}$$

and

$$(4\pi^2/\sigma^4)(n/p)EV^2 \to 1. \tag{6.16}$$

Proof of Proposition 6.1. By (3.38)

$$V = \omega_0(p)\left(\frac{1}{2}\sum_{j=1}^{p}[\Gamma^{-1}\boldsymbol{w}]_j\zeta_j - \zeta_0\right),$$

where in view of (4.3), $\omega_0(p) \sim (4\pi)^{-1}p$. Therefore to prove (6.15) and (6.16), it suffices to show that

$$Q_n := \frac{(pn)^{1/2}}{2\sigma^2}\left(\frac{1}{2}\sum_{j=1}^{p}[\Gamma^{-1}\boldsymbol{w}]_j\zeta_j - \zeta_0\right) \xrightarrow{d} N(0,1) \tag{6.17}$$

and

$$EQ_n^2 \to 1. \tag{6.18}$$

By definition,

$$\zeta_0 = 2\int_{-\pi}^{\pi} I_n(\lambda)|1-e^{i\lambda}|^{2d^0}|A_p(\lambda)|^2 \ln|1-e^{i\lambda}|\,d\lambda \tag{6.19}$$

and for $j \geqslant 1$

$$\zeta_j = 2\int_{-\pi}^{\pi} I_n(\lambda)|1-e^{i\lambda}|^{2d^0}e^{-i\lambda j}A_p(\lambda)\,d\lambda. \tag{6.20}$$

Hence we can write

$$Q_n = \int_{-\pi}^{\pi} I_n(\lambda)\widetilde{b}_n(\lambda)d = \int_{-\pi}^{\pi} I_n(\lambda)Re(\widetilde{b}_n(\lambda))\,d\lambda, \tag{6.21}$$

where

$$\widetilde{b}_n(\lambda) = \frac{(pn)^{1/2}}{\sigma^2}A_p(\lambda)|1-e^{i\lambda}|^{2d^0}\left(\sum_{j=1}^{p}e^{-i\lambda j}\left[\frac{1}{2}\Gamma^{-1}\boldsymbol{w}\right]_j - A_p(-\lambda)\ln|1-e^{i\lambda}|\right)$$

is a complex valued function with the property $\widetilde{b}_n(\lambda) = \overline{\widetilde{b}_n(-\lambda)}$. To prove (6.17) and (6.18) we shall show that conditions (6.7) and (6.8) of Lemma 6.2 are satisfied with

$$\eta_n(\lambda) = 2\pi b_n(\lambda)f(\lambda), \quad b_n(\lambda) = \mathrm{Re}(\widetilde{b}_n(\lambda)),$$

$$K_n = C(pn)^{1/2}\log n.$$

We focus on the verification of (6.7), condition (6.8) will the follow easily and will be verified towards the end of the proof.

Consider the matrix $E_n = (e_{t-s})$ defined by (6.5). We shall show below that for any $\delta > 0$,

$$|b_n(\lambda)| \leqslant CK_n|\lambda|^{2d^0-\delta}, \quad |\lambda| \leqslant \pi, \tag{6.22}$$

where $C > 0$ does not depend on n and λ, and

$$\|E_n\|_2^2 = 2(\pi n)^2(1+o(1)). \tag{6.23}$$

Relations (6.22) and (6.23) imply condition (6.7) since, with $d = d^0$, bf and small enough δ

$$\frac{K_n n^{\max(\delta,d,0)} \log n}{\|E_n\|_2} \leqslant C(p/n)^{1/2}(\log n)^2 n^{\max(\delta,d,0)} \to 0,$$

by Assumption 3.

To show (6.22), note that $|\sum_{j=1}^{p} [\frac{1}{2}\Gamma^{-1}\boldsymbol{w} + \boldsymbol{\tau}]_j| \leqslant Cp^{-1}$, by (4.24), which implies that

$$\widetilde{b}_n(\lambda) = -\frac{(pn)^{1/2}}{\sigma^2} A_p(\lambda)|1 - e^{i\lambda}|^{2d^0} \times \left(\sum_{j=1}^{p} \tau_j e^{-i\lambda j} + A_p(-\lambda) \ln|1 - e^{i\lambda}| + O(p^{-1})\right). \tag{6.24}$$

Observing that $|A_p(\lambda)| \leqslant \sum_{j=0}^{\infty} |a_j| < \infty$, and using the bound

$$|\tau_j| = \left|\sum_{k=1}^{j} \frac{a_{j-k}}{k}\right| \leqslant \sum_{1\leqslant k\leqslant j/2} |a_{j-k}| + \sum_{j/2<k\leqslant j} \frac{|a_{j-k}|}{j/2} \leqslant Cj^{-1}, \tag{6.25}$$

which follows from (2.17), we obtain that for any $\delta > 0$,

$$\begin{aligned}|b_n(\lambda)| &\leqslant C(pn)^{1/2}|1 - e^{i\lambda}|^{2d^0} \left(\sum_{j=1}^{p} j^{-1} + |\ln|1 - e^{i\lambda}||\right) \\ &\leqslant C(pn)^{1/2}(\log n)|1 - e^{i\lambda}|^{2d^0-\delta} = CK_n|\lambda|^{2d^0-\delta},\end{aligned}$$

to prove (6.22).

We now establish (6.23). Denote

$$\widetilde{\eta}_n(\lambda) = 2\pi\widetilde{b}_n(\lambda) f(\lambda).$$

Then $\eta_n(\lambda) = \mathrm{Re}(\widetilde{\eta}_n(\lambda)) = 2\pi\mathrm{Re}(\widetilde{b}_n(\lambda)) f(\lambda)$. First, observe that $\sum_{j=1}^{p} e^{-i\lambda j}\tau_j = \sum_{j=1}^{p+1} e^{-i\lambda j} \tau_j + O(p^{-1})$ since $|\tau_{p+1}| \leqslant Cp^{-1}$, by (6.25). We can write

$$\begin{aligned}\sum_{j=1}^{p+1} \tau_j e^{-i\lambda j} &\equiv \sum_{j=1}^{p+1} e^{-i\lambda j} \sum_{k=1}^{j} a_{j-k}k^{-1} = \sum_{s=0}^{p} e^{-i\lambda s} a_s \sum_{k=1}^{p+1-s} e^{-i\lambda k}k^{-1} \\ &= A_p(-\lambda) \sum_{k=1}^{p} e^{-i\lambda k}k^{-1} + o(p^{-1/2})\end{aligned}$$

since

$$\begin{aligned}\left|\sum_{s=0}^{p} e^{-i\lambda s} a_s \sum_{k=p-s+2}^{p} e^{-i\lambda k}k^{-1}\right| &\leqslant \sum_{s=0}^{p} |a_s| \sum_{k=p-s+2}^{p} k^{-1} \\ &\leqslant \sum_{0\leqslant s\leqslant p^{3/8}} |a_s| \sum_{k=p-s+2}^{p} k^{-1} + \sum_{p^{3/8}\leqslant s\leqslant p} |a_s| \sum_{k=1}^{p} k^{-1} \\ &= o(p^{-1/2})\end{aligned}$$

using the estimate $\sum_{k=p-s+2}^{p} k^{-1} \leqslant Cp^{-1}s = o(p^{-1/2})$ for $1 \leqslant s \leqslant p^{3/8}$, and the estimate $\sum_{p^{3/8}\leqslant s\leqslant p} |a_s| = o(p^{-6/8})$ which holds in view of (2.19).

This, together with (6.24) and (2.6) implies that, for $0 \leqslant \lambda \leqslant \pi$,

$$\widetilde{\eta}_n(\lambda) = -(pn)^{1/2}|A_p(\lambda)|^2|A(\lambda)|^{-2}\left(\sum_{j=1}^{p} e^{-i\lambda j} j^{-1} + \ln(2\sin(\lambda/2)) + o(p^{-1/2})\right).$$

Since $A_p(\lambda) = A(\lambda) + o(n^{-1/2})$, by Assumption 3, we obtain that

$$\widetilde{\eta}_n(\lambda) = -(pn)^{1/2}\left(\sum_{j=1}^{p} e^{-i\lambda j} j^{-1} + \ln(2\sin(\lambda/2)\right) + o(n^{1/2})(|\ln(2\sin(\lambda/2))| + 1).$$

This and equality $\ln(2\sin(\lambda/2)) = -\sum_{j=1}^{\infty} j^{-1}\cos(j\lambda)$ yields that

$$\eta_n(\lambda) := \mathrm{Re}(\widetilde{\eta}_n(\lambda)) = \eta_{n,1}(\lambda) + \eta_{n,2}(\lambda), \tag{6.26}$$

where

$$\eta_{n,1}(\lambda) = (pn)^{1/2} \sum_{j=p+1}^{\infty} j^{-1}\cos(j\lambda), \quad \eta_{n,2}(\lambda) = o(n^{1/2})(|\ln|\lambda\|| + 1).$$

Write

$$\|E_n\|_2^2 = \sum_{t,s=1}^{n} e_{t-s}^2 = \int_{-\pi}^{\pi}\left|\sum_{t=1}^{n} e^{it(x+y)}\right|^2 \eta_n(x)\eta_n(y)\,dx\,dy = v_{n,1} + v_{n,2},$$

where

$$v_{n,1} = \int_{-\pi}^{\pi}\left|\sum_{t=1}^{n} e^{it(x+y)}\right|^2 \eta_{n,1}(x)\eta_{n,1}(y)\,dx\,dy$$

and

$$|v_{n,2}| \leqslant \int_{-\pi}^{\pi}\left|\sum_{t=1}^{n} e^{it(x+y)}\right|^2 (2|\eta_{n,1}(x)\eta_{n,2}(y)| + |\eta_{n,2}(x)\eta_{n,2}(y)|)\,dx\,dy.$$

To estimate $v_{n,1}$, note

$$\begin{aligned} e_{n,1}(t) &:= \int_{-\pi}^{\pi} e^{it\lambda}\eta_{n,1}(\lambda)\,d\lambda = (pn)^{1/2}\int_{-\pi}^{\pi} e^{i\lambda t}\sum_{j=p+1}^{\infty} j^{-1}\cos(j\lambda)\,d\lambda \\ &= (pn)^{1/2}\frac{\pi}{t}, \quad t \geqslant p+1 \end{aligned}$$

and $e_{n,1}(t) = 0$, for $0 \leqslant t \leqslant p$. Then

$$\begin{aligned} v_{n,1} &= \sum_{t,s=1}^{n} e_{n,1}^2(t-s) = \sum_{t=-n}^{n}(n-|t|)e_{n,1}^2(|t|) \\ &= 2\pi^2 np \sum_{t=p+1}^{n}(n-t)t^{-2} = 2\pi^2 n^2(1+o(1)). \end{aligned}$$

On the other hand, by Cauchy inequality,

$$
\begin{aligned}
v_{n,2} \leqslant\, & C \left(\int_{-\pi}^{\pi} \left| \sum_{t=1}^{n} e^{it(x+y)} \right|^2 |\eta_{n,1}(x)|^2 \, dx \, dy \right)^{1/2} \\
& \times \left(\int_{-\pi}^{\pi} \left| \sum_{t=1}^{n} e^{it(x+y)} \right|^2 |\eta_{n,2}(y)|^2 \, dx \, dy \right)^{1/2} \\
& + C \int_{-\pi}^{\pi} \left| \sum_{t=1}^{n} e^{it(x+y)} \right|^2 |\eta_{n,2}(x)|^2 \, dx \, dy \\
\leqslant\, & Cn \left(\int_{-\pi}^{\pi} |\eta_{n,1}(x)|^2 \, dx \int_{-\pi}^{\pi} |\eta_{n,2}(y)|^2 \, dy \right)^{1/2} + Cn \int_{-\pi}^{\pi} |\eta_{n,2}(x)|^2 \, dx = o(n^2),
\end{aligned}
$$

since by Parseval equality,

$$
\int_{-\pi}^{\pi} |\eta_{n,1}(x)|^2 \, dx \leqslant Cpn \sum_{t=p+1}^{\infty} t^{-2} \leqslant Cn
$$

and

$$
\int_{-\pi}^{\pi} |\eta_{n,2}(y)|^2 \, dy = o(n),
$$

to complete proof of (6.23).

Finally, from (6.26) we obtain that

$$
\int_{-\pi}^{\pi} \eta_n(\lambda) \, d\lambda = \int_{-\pi}^{\pi} \sum_{j=p+1}^{\infty} \frac{\cos j\lambda}{j} \, d\lambda + \int_{-\pi}^{\pi} o(n^{1/2}) |\ln |\lambda|| \, d\lambda = o(n^{1/2})
$$

which in view of (6.23) implies (6.8). □

References

[1] G. Baxter, An asymptotic result for the finite predictor, Math. Scand. 10 (1962) 137–144.

[2] K.N. Berk, Consistent autoregressive spectral estimates, Ann. Statist. 2 (1974) 489–502.

[3] R.J. Bhansali, Linear prediction by autoregressive model fitting in the time domain, Ann. Statist. 6 (1978) 224–231.

[4] R.J. Bhansali, Autoregressive and window estimates of the inverse correlation function, Biometrika 67 (1980) 551–566.

[5] R.J. Bhansali, Order selection for linear time series models: a review, in: T. Subba Rao (Ed.), Developments in Time Series Analysis, Chapman & Hall, London, 1993, pp. 50–56.

[6] R.J. Bhansali, L. Giraitis, P.S. Kokoszka, Approximations and limit theory for quadratic forms of linear processes, Technical Report, Utah State University, 2005.

[7] R.J. Bhansali, P.S. Kokoszka, Estimation of the long-memory parameter: a review and an extension, in: I.V. Basawa, C.C. Heyde, R.L. Taylor (Eds.), Proceedings of the Symposium on Inference for Stochastic Processes, IMS Lecture Notes, 2001, pp. 125–150.

[8] P. Bloomfield, An exponential model for the spectrum of a scalar time series, Biometrika 60 (1973) 217–226.

[9] D.R. Brillinger, Time Series: Data Analysis and Theory, Holt, New York, 1975.

[10] R. Dahlhaus, Efficient parameter estimation for self similar processes, Ann. Statist. 17 (1989) 1749–1766.

[11] R. Fox, M.S. Taqqu, Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series, Ann. Statist. 14 (1986) 517–532.

[12] L. Giraitis, D. Surgailis, CLT for quadratic forms in strongly dependent linear variables and application to asymptotical normality of Whittle's estimate, Probab. Theory Related Fields 86 (1990) 87–104.

[13] C.M. Hurvich, J. Brodsky, Broadband semiparametric estimation of the memory parameter of a long-memory time series using fractional exponential models, J. Time Ser. Anal. 23 (2001) 221–249.

[14] T. Kailath, A. Vieira, M. Morf, Inverses of Toeplitz operators, innovations, and orthogonal polynomials, SIAM Rev. 20 (1978) 106–119.

[15] P.S. Kokoszka, M.S. Taqqu, Fractional ARIMA with stable innovations, Stochastic Process. Appl. 60 (1995) 19–47.

[16] E. Moulines, P. Soulier, Broad band log-periodogram regression of time series with long range dependence, Ann. Statist. 27 (1999) 1415–1439.

[17] E. Moulines, P. Soulier, Data-driven order selection for projection estimator of the spectral density of time series with long range dependence, J. Time Ser. Anal. 21 (2000) 193–218.

[18] E. Parzen, Multiple time series modelling, in: P.R. Krishnaiah (Ed.), Multivariate Analysis II, Academic Press, New York, 1969.

[19] C.R. Rao, Linear Statistical Inference, Wiley, New York, 1973.

[20] P.M. Robinson, Gaussian semiparametric estimation of long range dependence, Ann. Statist. 23 (1995) 1630–1661.

[21] R. Shibata, Asymptotically efficient selection of the order of the model for estimating parameters of a linear process, Ann. Statist. 8 (1980) 147–164.

[22] R. Shibata, An optimal autoregressive spectral estimate, Ann. Statist. 9 (1981) 300–306.