Bernoulli **24**(2), 2018, 1202–1232
DOI: 10.3150/16-BEJ897

Baxter's inequality for finite predictor coefficients of multivariate long-memory stationary processes

AKIHIKO INOUE[1], YUKIO KASAHARA[2] and MOHSEN POURAHMADI[3]

[1]*Department of Mathematics, Hiroshima University, Higashi-Hiroshima 739-8526, Japan.*
E-mail: inoue100@hiroshima-u.ac.jp
[2]*Department of Mathematics, Hokkaido University, Sapporo 060-0810, Japan.*
E-mail: y-kasa@math.sci.hokudai.ac.jp
[3]*Department of Statistics, Texas A&M University, College Station, TX 77843, USA.*
E-mail: pourahm@stat.tamu.edu

For a multivariate stationary process, we develop explicit representations for the finite predictor coefficient matrices, the finite prediction error covariance matrices and the partial autocorrelation function (PACF) in terms of the Fourier coefficients of its phase function in the spectral domain. The derivation is based on a novel alternating projection technique and the use of the forward and backward innovations corresponding to predictions based on the infinite past and future, respectively. We show that such representations are ideal for studying the rates of convergence of the finite predictor coefficients, prediction error covariances, and the PACF as well as for proving a multivariate version of Baxter's inequality for a multivariate FARIMA process with a common fractional differencing order for all components of the process.

Keywords: Baxter's inequality; long memory; multivariate stationary processes; partial autocorrelation functions; phase functions; predictor coefficients

1. Introduction

Baxter's inequality in [2] provides valuable information about the convergence of the finite predictor coefficients to their infinite past counterparts (autoregressive coefficients) of a short-memory univariate stationary process. It has been used by [3] in proving the consistency of the autoregressive model fitting process and the corresponding autoregressive spectral density estimator, and in proving the validity of autoregressive sieve bootstrap for a stationary time series in [9,10,31]. Due to the widespread applicability of Baxter's inequality in these areas and others, there has been a great deal of activities in extending it to the setups of multivariate stationary processes in [11,17], random fields in [33], and rectangular arrays in [34]. In these extensions, the *boundedness* of the spectral density function of the underlying process appears to be an absolutely essential and indispensable part of proving Baxter's inequality.

In [26], however, Baxter's inequality was established for univariate long-memory processes where the boundedness of the spectral density function is clearly violated. Unlike the classical proofs for short-memory processes involving the orthogonal polynomials or the Durbin–Levinson algorithm, the key ingredient of the proof in [26] was an explicit representation of the finite predictor coefficients in terms of the autoregressive (AR) and moving average (MA) coefficients. The derivation of the representation in turn was based on techniques that use von Neumann's alternating projections on the infinite past and future. These techniques were first used by [22] and have been developed to derive the needed representations for the finite prediction error variances [22–25], the partial autocorrelation functions [7,24,28], and the finite predictor coefficients [26]. Unfortunately, most of the details of the proofs in the univariate case do not carry over to the multivariate setup where, for example, all functions and the sequences of AR and MA coefficients are matrix-valued and hence in general do not commute with each other.

In this paper, for a multivariate stationary process, we prove the desired explicit representations for the finite predictor coefficients, the finite prediction error covariances and the partial autocorrelation function (PACF). See Theorems 5.2–5.4 in Section 5. The three new ingredients that enable us to obtain the results in the multivariate framework are:

(i) Use of the Fourier coefficients of the matrix-valued phase function of the process in the spectral domain, rather than the AR and MA coefficient matrices (see Section 4).
(ii) Development of an enhanced alternating projection technique tailored to the specific needs of the problem at hand (see Section 3).
(iii) Use of the forward and backward innovation processes corresponding to the predictions based on the infinite past and future, respectively (see Sections 2, 4 and 5).

Our representation theorems make it possible to extend Baxter's inequality and other univariate asymptotic results to the multivariate long-memory processes. Even when specialized to univariate processes, our method and results are more succinct, transparent and improve the known univariate results in several ways. For example, our representation theorem for the finite predictor coefficients, that is, Theorem 5.4 below, is stated under the minimality condition (see (M) in Section 5) only, which is weaker than the condition in the corresponding univariate result, i.e., Theorem 2.9 in [26].

In this paper, when applying the representation theorems, we restrict our attention to a class of q-variate long-memory processes, that is, the q-variate *FARIMA* (fractional autoregressive integrated moving-average) or *vector ARFIMA* processes with common fractional differencing order for all components. A process $\{X_k\}$ in this class has the spectral density w of the form

$$w(e^{i\theta}) = |1 - e^{i\theta}|^{-2d} g(e^{i\theta}) g(e^{i\theta})^*, \tag{1.1}$$

where $d \in (-1/2, 1/2) \setminus \{0\}$ and $g : \mathbb{T} \to \mathbb{C}^{q\times q}$ has rational entries satisfying some suitable conditions; see (F) in Section 6. The process $\{X_k\}$ is described by the equation

$$(1 - L)^d X_k = g(L)\xi_k, \qquad k \in \mathbb{Z}, \tag{1.2}$$

where L is the lag operator defined by $LX_m = X_{m-1}$ and $\{\xi_k\}$ is a q-variate white noise, that is, a q-variate, centered process such that $E[\xi_n \xi_m^*] = \delta_{nm} I_q$ with I_q being the $q \times q$ unit matrix. See, for example, [12]. We notice that the parameter d in (1.1) is the *fractional differencing degree* in (1.2). The q-variate FARIMA processes are multivariate analogues of univariate ones introduced independently by [16] and [19].

We present the following quick summary of the asymptotic results obtained by applying our representation theorems to a q-variate FARIMA process $\{X_k\}$ with (1.1):

(1) Baxter's inequality for $\{X_k\}$ with $d \in (0, 1/2)$ (see Theorem 6.9 below).
(2) The precise asymptotics for the finite prediction error covariances v_n and $\tilde{v}_n$ of $\{X_k\}$ with $d \in (-1/2, 1/2) \setminus \{0\}$ (see Theorem 6.5 below; see also Section 5 for the definitions of v_n and $\tilde{v}_n$).
(3) The precise asymptotic behavior for the PACF α_n of $\{X_k\}$ with $d \in (-1/2, 1/2) \setminus \{0\}$ (see Theorem 6.7 below; see also Section 5 for the definition of α_n).

First, Baxter's inequality for FARIMA processes is of the form

$$\sum_{j=1}^{n} \|\phi_{n,j} - \phi_j\| \leq K \sum_{j=n+1}^{\infty} \|\phi_j\|, \qquad n \in \mathbb{N}, \tag{1.3}$$

for some positive constant K, where, for $a \in \mathbb{C}^{q \times q}$, $\|a\|$ denotes the spectral norm of a (see Section 2), and ϕ_j and $\phi_{n,j}$ denote the forward infinite and finite predictor coefficients, respectively, of $\{X_k\}$ (see Sections 2 and 5, respectively, for their precise definitions). We also prove a backward analogue of (1.3); see Corollary 6.10 below. We refer to [26] for the corresponding result for univariate long-memory processes and [1,36,39] for its application; see also [21] for other applications of results in [26]. In [11], Baxter's inequality (1.3) was proved for a class of multivariate short-memory stationary processes. The original inequality (1.3) of Baxter [2] was an assertion for univariate short-memory processes. See also [3] and [37], Section 7.6.2.

Next, the asymptotic results in (2) above are of the form

$$v_n = v_\infty + \frac{d^2}{n} v_\infty + O\big(n^{-2}\big), \qquad n \to \infty, \tag{1.4}$$

$$\tilde{v}_n = \tilde{v}_\infty + \frac{d^2}{n} \tilde{v}_\infty + O\big(n^{-2}\big), \qquad n \to \infty, \tag{1.5}$$

where v_∞ (resp., $\tilde{v}_\infty$) is the forward (resp., backward) infinite prediction error covariance of $\{X_k\}$; see Section 6.3 for their precise definitions. We refer to [22–25] for the corresponding results for univariate long-memory processes. See also [15,20] for related work.

Finally, the result in (3) is of the form

$$\alpha_n = \frac{d}{n} V + O\big(n^{-2}\big), \qquad n \to \infty, \tag{1.6}$$

where V is a unitary matrix in $\mathbb{C}^{q \times q}$ which depends only on g (and not d). We refer to [7,22–25] for the corresponding results for univariate long-memory processes. In the theory of orthogonal polynomials on the unit circle, the PACF appears as the sequence of *Verblunsky coefficients* and plays a central role. See, for example, [5,13,28].

The above q-variate FARIMA process has a common fractional differencing order d for all components. The question arises of proving analogues of (1)–(3) above for more general q-variate FARIMA processes which have, in general, different order of differencing in each component, that is,

$$(1-L)^{\mathbf{d}} := \begin{pmatrix} (1-L)^{d_1} & & 0 \\ & \ddots & \\ 0 & & (1-L)^{d_q} \end{pmatrix}$$

with $\mathbf{d} = (d_1, \dots, d_q)$, instead of $(1-L)^d$ (see, e.g., [12]). We leave this question open here; the difficulty stems from the fact that, for such a general q-variate FARIMA process, the matrices $g(L)$ and $(1-L)^{\mathbf{d}}$ do not commute with each other.

This paper is organized as follows. In Section 2, we give preliminary definitions and basic facts. In Section 3, we prove the key projection theorem. In Section 4, we describe some basic facts about the Fourier coefficients of the phase function which is needed in Section 5. In Section 5, we prove the main results, i.e., the representation theorems for the finite prediction error covariances, the PACF and the finite predictor coefficients of multivariate stationary processes. In Section 6, we apply the main results to multivariate FARIMA processes with common fractional differencing order for all components, and establish the results (1)–(3) above for them.

2. Preliminaries

Let $\mathbb{C}^{m\times n}$ be the set of all complex $m \times n$ matrices; we write $\mathbb{C}^q$ for $\mathbb{C}^{q\times 1}$. We write I_n for the $n \times n$ unit matrix. For $a \in \mathbb{C}^{m\times n}$, a^{T} denotes the transpose of a, and $\bar{a}$ and a^* the complex and Hermitian conjugates of a, respectively; thus, in particular, $a^* := \bar{a}^{\mathrm{T}}$. For $a \in \mathbb{C}^{q\times q}$, we write $\|a\|$ for the spectral norm of a:

$$\|a\| := \sup_{u\in\mathbb{C}^q, |u|=1} |au|.$$

Here $|u| := (\sum_{i=1}^q |u^i|^2)^{1/2}$ denotes the Euclidean norm of $u = (u^1, \dots, u^q)^{\mathrm{T}} \in \mathbb{C}^q$. A Hermitian matrix $a \in \mathbb{C}^{q\times q}$ is said to be *positive*, denoted as $a \geq 0$, if $(au)^*u \geq 0$ for all $u \in \mathbb{C}^q$. When $a \geq 0$, we have $\|a\| = \sup_{u\in\mathbb{C}^q,|u|=1}(au)^*u$. For Hermitian matrices $a, b \in \mathbb{C}^{q\times q}$, we write $a \geq b$ if $a - b \geq 0$. If $a \geq b$, then we have $\|a\| \geq \|b\|$. For $p \in [1, \infty)$ and $K \subset \mathbb{Z}$, $\ell_p^{q\times q}(K)$ denotes the space of $\mathbb{C}^{q\times q}$-valued sequences $\{a_k\}_{k\in K}$ such that $\sum_{k\in K} \|a_k\|^p < \infty$. We write $\ell_{p+}^{q\times q}$ for $\ell_p^{q\times q}(\mathbb{N} \cup \{0\})$ and ℓ_{p+} for $\ell_{p+}^{1\times 1} = \ell_p^{1\times 1}(\mathbb{N} \cup \{0\})$.

Let $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ be the unit circle in $\mathbb{C}$. We write σ for the normalized Lebesgue measure $d\theta/(2\pi)$ on $([-\pi,\pi), \mathcal{B}([-\pi,\pi)))$, where $\mathcal{B}([-\pi,\pi))$ is the Borel σ-algebra of $[-\pi,\pi)$; thus we have $\sigma([-\pi,\pi)) = 1$. For $p \in [1,\infty)$, we write $L_p(\mathbb{T})$ for the Lebesgue space of measurable functions $f : \mathbb{T} \to \mathbb{C}$ such that $\|f\|_p < \infty$, where $\|f\|_p := \{\int_{-\pi}^{\pi} |f(e^{i\theta})|^p \sigma(d\theta)\}^{1/p}$. Let $L_p^{m\times n}(\mathbb{T})$ be the space of $\mathbb{C}^{m\times n}$-valued functions on $\mathbb{T}$ whose entries belong to $L_p(\mathbb{T})$.

The Hardy class $H_2(\mathbb{T})$ on $\mathbb{T}$ is the closed subspace of $L_2(\mathbb{T})$ consisting of $f \in L_2(\mathbb{T})$ such that $\int_{-\pi}^{\pi} e^{im\theta} f(e^{i\theta})\sigma(d\theta) = 0$ for $m = 1, 2, \dots$. Let $H_2^{m\times n}(\mathbb{T})$ be the space of $\mathbb{C}^{m\times n}$-valued functions on $\mathbb{T}$ whose entries belong to $H_2(\mathbb{T})$. Let $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$ be the open unit disk in $\mathbb{C}$. We write $H_2(\mathbb{D})$ for the Hardy class on $\mathbb{D}$, consisting of holomorphic functions f on $\mathbb{D}$ such that $\sup_{r\in[0,1)} \int_{-\pi}^{\pi} |f(re^{i\theta})|^2 \sigma(d\theta) < \infty$. As usual, we identify each function

f in $H_2(\mathbb{D})$ with its boundary function $f(e^{i\theta}) := \lim_{r\uparrow 1} f(re^{i\theta})$, σ-a.e., in $H_2(\mathbb{T})$. A function h in $H_2^{n\times n}(\mathbb{T})$ is called *outer* if $\det h$ is a $\mathbb{C}$-valued outer function, that is, $\det h$ satisfies $\log|\det h(0)| = \int_{-\pi}^{\pi} \log|\det h(e^{i\theta})|\sigma(d\theta)$ (cf. [30], Definition 3.1).

For $q \in \mathbb{N}$, let $\{X_k\} = \{X_k : k \in \mathbb{Z}\}$ be a $\mathbb{C}^q$-valued, centered, weakly stationary process, defined on a probability space $(\Omega, \mathcal{F}, P)$, which we shall simply call a *q-variate stationary process*. Write $X_k = (X_k^1, \dots, X_k^q)^{\mathrm{T}}$, and let M be the complex Hilbert space spanned by all the entries $\{X_k^j : k \in \mathbb{Z}, j = 1, \dots, q\}$ in $L^2(\Omega, \mathcal{F}, P)$, which has inner product $(x, y)_M := E[x\overline{y}]$ and norm $\|x\|_M := (x, x)_M^{1/2}$. For $K \subset \mathbb{Z}$ such as $\{n\}$, $(-\infty, n] := \{n, n-1, \dots\}$, $[n, \infty) := \{n, n+1, \dots\}$, and $[m, n] := \{m, \dots, n\}$ with $m \le n$, we define the closed subspace M_K^X of M by

$$M_K^X := \overline{\mathrm{sp}}\{X_k^j : j = 1, \dots, q, k \in K\}.$$

We write $(M_K^X)^\perp$ for the orthogonal complement of M_K^X in M. Let P_K and $P_K^\perp$ be the orthogonal projection operators of M onto M_K^X and $(M_K^X)^\perp$, respectively.

Let M^q be the space of $\mathbb{C}^q$-valued random variables on $(\Omega, \mathcal{F}, P)$ whose entries belong to M. The norm $\|x\|_{M^q}$ of $x = (x^1, \dots, x^q)^{\mathrm{T}} \in M^q$ is given by $\|x\|_{M^q} := (\sum_{i=1}^q \|x^i\|_M^2)^{1/2}$. For $K \subset \mathbb{Z}$ and $x = (x^1, \dots, x^q)^{\mathrm{T}} \in M^q$, we write $P_K x$ for $(P_K x^1, \dots, P_K x^q)^{\mathrm{T}}$. We define $P_K^\perp x$ in a similar way. For $x = (x^1, \dots, x^q)^{\mathrm{T}}$ and $y = (y^1, \dots, y^q)^{\mathrm{T}}$ in M^q,

$$\langle x, y\rangle := E[xy^*] = \begin{pmatrix} (x^1, y^1)_M & (x^1, y^2)_M & \cdots & (x^1, y^q)_M \\ (x^2, y^1)_M & (x^2, y^2)_M & \cdots & (x^2, y^q)_M \\ \vdots & \vdots & \ddots & \vdots \\ (x^q, y^1)_M & (x^q, y^2)_M & \cdots & (x^q, y^q)_M \end{pmatrix} \in \mathbb{C}^{q\times q}$$

stands for the Gram matrix of x and y.

Let $\{X_k\}$ be a q-variate stationary process. If there exists a positive $q \times q$ Hermitian matrix-valued function w on $\mathbb{T}$, satisfying $w \in L_1^{q\times q}(\mathbb{T})$ and

$$\langle X_m, X_n\rangle = \int_{-\pi}^{\pi} e^{-i(m-n)\theta} w(e^{i\theta}) \frac{d\theta}{2\pi}, \qquad n, m \in \mathbb{Z},$$

then we call w the *spectral density* of $\{X_k\}$. We say that $\{X_k\}$ is *purely nondeterministic* (PND) if $\bigcap_{n\in\mathbb{Z}} M_{(-\infty,n]}^X = \{0\}$. Every PND process $\{X_k\}$ has spectral density (cf. Section 4 in [38], Chapter II). We consider the following condition:

$$\{X_k\} \text{ has spectral density } w \text{ such that } \log\det w \in L_1(\mathbb{T}). \tag{A}$$

A necessary and sufficient condition for (A) is that $\{X_k\}$ is PND and its spectral density w satisfies $\det w(e^{i\theta}) > 0$, σ-a.e. (see Theorem 6.1 in [38], Chapter II).

In what follows, we assume (A) for $\{X_k\}$. Let $\{\tilde{X}_k : k \in \mathbb{Z}\}$ be the time-reversed process of $\{X_k\}$:

$$\tilde{X}_k := X_{-k}, \qquad k \in \mathbb{Z}. \tag{2.1}$$

Then, since

$$\langle \tilde{X}_n, \tilde{X}_m \rangle = \langle X_{-n}, X_{-m} \rangle = \int_{-\pi}^{\pi} e^{-i(n-m)\theta} w(e^{-i\theta}) \frac{d\theta}{2\pi},$$

$\{\tilde{X}_k\}$ has the spectral density $\tilde{w}$ given by

$$\tilde{w}(e^{i\theta}) = w(e^{-i\theta}). \tag{2.2}$$

In particular, $\{\tilde{X}_k\}$ also satisfies (A). The spectral densities w and $\tilde{w}$ have the decompositions

$$w(e^{i\theta}) = h(e^{i\theta}) h(e^{i\theta})^*, \qquad \tilde{w}(e^{i\theta}) = \tilde{h}(e^{i\theta}) \tilde{h}(e^{i\theta})^*, \qquad \sigma\text{-a.e.}, \tag{2.3}$$

respectively, for some outer functions h and $\tilde{h}$ in $H_2^{q\times q}(\mathbb{T})$, and h and $\tilde{h}$ are unique up to constant unitary factors (see, e.g., [38], Chapter II, and [18], Theorem 11). We define the outer function $h_\sharp$ in $H_2^{q\times q}(\mathbb{T})$ by

$$h_\sharp(z) := \{\tilde{h}(\overline{z})\}^*. \tag{2.4}$$

Then, $h_\sharp$ satisfies

$$w(e^{i\theta}) = h_\sharp(e^{i\theta})^* h_\sharp(e^{i\theta}), \qquad \sigma\text{-a.e.} \tag{2.5}$$

We may take $h_\sharp = h$ for the univariate case $q = 1$ but there is no such simple relation between h and $h_\sharp$ for $q \geq 2$. We call $h^* h_\sharp^{-1}$ the *phase function* of $\{X_k\}$. Since

$$\{h(e^{i\theta})^* h_\sharp(e^{i\theta})^{-1}\}^* h(e^{i\theta})^* h_\sharp(e^{i\theta})^{-1} = \{h_\sharp(e^{i\theta})^*\}^{-1} w(e^{i\theta}) h_\sharp(e^{i\theta})^{-1} = I_q$$

holds σ-a.e., it is a unitary matrix valued function on $\mathbb{T}$. See Section 4 and [35], page 428.

Let

$$X_k = \int_{-\pi}^{\pi} e^{-ik\theta} \Lambda(d\theta), \qquad k \in \mathbb{Z},$$

be the spectral representation of $\{X_k\}$, where Λ is the $\mathbb{C}^q$-valued random spectral measure such that

$$\left(\int_{-\pi}^{\pi} \phi(e^{i\theta}) \Lambda(d\theta), \int_{-\pi}^{\pi} \psi(e^{i\theta}) \Lambda(d\theta) \right)_M = \int_{-\pi}^{\pi} \phi(e^{i\theta}) w(e^{i\theta}) \psi(e^{i\theta})^* \frac{d\theta}{2\pi}$$

for $\phi, \psi \in L(w)$ with $L(w)$ being the class of measurable $\phi : \mathbb{T} \to \mathbb{C}^{1\times q}$ satisfying $\int_{-\pi}^{\pi} \phi(e^{i\theta}) w(e^{i\theta}) \phi(e^{i\theta})^* \sigma(d\theta) < \infty$ (cf. [38], Chapter I). We define a q-variate stationary process $\{\xi_k : k \in \mathbb{Z}\}$, called the *forward innovation process* of $\{X_k\}$, by

$$\xi_k := \int_{-\pi}^{\pi} e^{-ik\theta} h(e^{i\theta})^{-1} \Lambda(d\theta), \qquad k \in \mathbb{Z}. \tag{2.6}$$

Then, $\{\xi_k\}$ satisfies $\langle \xi_n, \xi_m \rangle = \delta_{nm} I_q$ and

$$M^X_{(-\infty,n]} = M^\xi_{(-\infty,n]}, \qquad n \in \mathbb{Z} \tag{2.7}$$

(cf. Section 4 in [38], Chapter II), whence, for $n \in \mathbb{Z}$, $\{\xi_k^j : j = 1, \dots, q, k \geq n+1\}$ becomes a complete orthonormal basis of $(M^X_{(-\infty,n]})^{\perp}$.

On the other hand, the spectral representation of $\{\tilde{X}_k\}$ is given by

$$\tilde{X}_k = \int_{-\pi}^{\pi} e^{-ik\theta} \tilde{\Lambda}(d\theta), \qquad k \in \mathbb{Z}$$

with the $\mathbb{C}^q$-valued random measure $\tilde{\Lambda}$ defined by

$$\tilde{\Lambda}(E) := \Lambda(-E), \qquad E \in \mathcal{B}\big((-\pi, \pi)\big), \tag{2.8}$$

where $-E := \{-\theta : \theta \in E\}$. Let $\{\tilde{\xi}_k : k \in \mathbb{Z}\}$ be the forward innovation process of $\{\tilde{X}_k\}$ given by

$$\tilde{\xi}_k := \int_{-\pi}^{\pi} e^{-ik\theta} \tilde{h}\big(e^{i\theta}\big)^{-1} \tilde{\Lambda}(d\theta), \qquad k \in \mathbb{Z}. \tag{2.9}$$

Then, we easily see that $\{\tilde{\xi}_k\}$ satisfies $\langle \tilde{\xi}_n, \tilde{\xi}_m \rangle = \delta_{nm} I_q$ and

$$M^X_{[-n,\infty)} = M^{\tilde{\xi}}_{(-\infty,n]}, \qquad n \in \mathbb{Z}, \tag{2.10}$$

whence, for $n \in \mathbb{Z}$, $\{\tilde{\xi}_k^j : j = 1, \dots, q, k \geq n+1\}$ becomes a complete orthonormal basis of $(M^X_{[-n,\infty)})^{\perp}$. We also call $\{\tilde{\xi}_k\}$ the *backward innovation process* of $\{X_k\}$. Then, $\{\xi_k\}$ turns out to be the backward innovation process of $\{\tilde{X}_k\}$.

We define, respectively, the *forward MA and AR coefficients* c_k and a_k of $\{X_k\}$ by

$$h(z) = \sum_{k=0}^{\infty} z^k c_k, \qquad -h(z)^{-1} = \sum_{k=0}^{\infty} z^k a_k, \qquad z \in \mathbb{D}, \tag{2.11}$$

and the *backward MA and AR coefficients* $\tilde{c}_k$ and $\tilde{a}_k$ of $\{X_k\}$ by

$$\tilde{h}(z) = \sum_{k=0}^{\infty} z^k \tilde{c}_k, \qquad -\tilde{h}(z)^{-1} = \sum_{k=0}^{\infty} z^k \tilde{a}_k, \qquad z \in \mathbb{D}. \tag{2.12}$$

It should be noticed that c_k and a_k (resp., $\tilde{c}_k$ and $\tilde{a}_k$) are the backward (resp., forward) MA and AR coefficients of the time-reversed process $\{\tilde{X}_k\}$, respectively. All of $\{c_k\}$, $\{a_k\}$, $\{\tilde{c}_k\}$ and $\{\tilde{a}_k\}$ are $\mathbb{C}^{q\times q}$-valued sequences, and we have $\{c_k\}, \{\tilde{c}_k\} \in \ell_{2+}^{q\times q}$ and $c_0 a_0 = \tilde{c}_0 \tilde{a}_0 = -I_q$. We have the following forward and backward MA representations of $\{X_k\}$, respectively:

$$X_n = \sum_{k=-\infty}^{n} c_{n-k} \xi_k, \qquad X_{-n} = \sum_{k=-\infty}^{n} \tilde{c}_{n-k} \tilde{\xi}_k, \qquad n \in \mathbb{Z} \tag{2.13}$$

(cf. Section 4 in [38], Chapter II). If we further assume

$$\{a_k\}, \{\tilde{a}_k\} \in \ell_{1+}^{q\times q}, \tag{2.14}$$

then the following forward and backward AR representations of $\{X_k\}$, respectively, also hold:

$$\sum_{k=-\infty}^{n} a_{n-k} X_k + \xi_n = 0, \qquad \sum_{k=-\infty}^{n} \tilde{a}_{n-k} X_{-k} + \tilde{\xi}_n = 0, \qquad n \in \mathbb{Z} \tag{2.15}$$

(see, e.g., the proof of [22], Theorem 4.4). From (2.15), we obtain the following forward and backward infinite prediction formulas, respectively, for $\{X_k\}$:

$$P_{(-\infty,-1]} X_0 = \sum_{k=1}^{\infty} \phi_k X_{-k}, \qquad P_{[1,\infty)} X_0 = \sum_{k=1}^{\infty} \tilde{\phi}_k X_k.$$

Here

$$\phi_k := c_0 a_k, \qquad \tilde{\phi}_k := \tilde{c}_0 \tilde{a}_k, \qquad k \in \mathbb{N}. \tag{2.16}$$

We call ϕ_k (resp., $\tilde{\phi}_k$) the *forward* (resp., *backward*) *infinite predictor coefficients* of $\{X_k\}$. It should be noticed that ϕ_k (resp., $\tilde{\phi}_k$) are the backward (resp., forward) infinite predictor coefficients of $\{\tilde{X}_k\}$.

3. A projection theorem

In this section, we present a projection theorem which facilitates finding explicit representations of the finite predictor coefficients, the finite prediction error covariances and the PACF of a q-variate stationary process $\{X_k\}$, in terms of the Fourier coefficients of the phase function.

Let H be a Hilbert space with inner product $(\cdot,\cdot)$. Let $I : H \to H$ be the identity map. For a closed subspace A of H, we write P_A for the orthogonal projection operator of H onto A and $P_A^{\perp}$ for that onto the orthogonal complement $A^{\perp}$ of A, that is, $P_A^{\perp} = I - P_A$. For closed subspaces A and B of H, von Neumann's Alternating Projection Theorem (cf. [37], Section 9.6.3) states that $(P_A P_B)^n$ converges to $P_{A\cap B}$ as $n \to \infty$ in the strong operator topology. From this, we have the following projection theorem.

Theorem 3.1 ([22,24]). *Let A and B be closed subspaces of H. Then, we have, for $x, y \in H$,*

$$P_{A\cap B}^{\perp} x = \sum_{k=0}^{\infty} \bigl\{ P_B^{\perp} (P_A P_B)^k x + P_A^{\perp} P_B (P_A P_B)^k x \bigr\}, \tag{3.1}$$

$$\begin{aligned}\bigl(P_{A\cap B}^{\perp} x, P_{A\cap B}^{\perp} y\bigr) = \sum_{k=0}^{\infty} \bigl\{ &\bigl(P_B^{\perp} (P_A P_B)^k x, P_B^{\perp} (P_A P_B)^k y\bigr) \\ &+ \bigl(P_A^{\perp} P_B (P_A P_B)^k x, P_A^{\perp} P_B (P_A P_B)^k y\bigr) \bigr\},\end{aligned} \tag{3.2}$$

the sum in (3.1) converging strongly.

The assertion (3.2) (resp., (3.1)) is an abstract form of [22], Theorem 4.1, and [24], Theorem 3.1 (resp., Remarks to [24], Theorem 3.1), and can be proved in a similar way.

For our applications in this paper, we need the next variant.

Theorem 3.2. *Let A and B be closed subspaces of H. Then, we have*

$$P_{A\cap B}^{\perp} a = \sum_{k=0}^{\infty} \{ P_B^{\perp} (P_A^{\perp} P_B^{\perp})^k a - (P_A^{\perp} P_B^{\perp})^{k+1} a \}, \qquad a \in A, \tag{3.3}$$

$$(P_{A\cap B}^{\perp} a_1, P_{A\cap B}^{\perp} a_2) = \sum_{k=0}^{\infty} (P_B^{\perp} (P_A^{\perp} P_B^{\perp})^k a_1, a_2), \qquad a_1, a_2 \in A, \tag{3.4}$$

$$(P_{A\cap B}^{\perp} a, P_{A\cap B}^{\perp} b) = -\sum_{k=0}^{\infty} ((P_A^{\perp} P_B^{\perp})^{k+1} a, b), \qquad a \in A, b \in B, \tag{3.5}$$

the sum in (3.3) *converging strongly.*

Proof. If $a \in A$, then

$$\begin{aligned} P_B^{\perp} P_A P_B a &= P_B^{\perp} (I - P_A^{\perp}) P_B a = -P_B^{\perp} P_A^{\perp} P_B a = -P_B^{\perp} P_A^{\perp} (I - P_B^{\perp}) a \\ &= P_B^{\perp} P_A^{\perp} P_B^{\perp} a. \end{aligned}$$

Hence, we have, for $k = 1, 2, \ldots,$

$$P_B^{\perp} (P_A P_B)^k a = P_B^{\perp} (P_A^{\perp} P_B^{\perp}) (P_A P_B)^{k-1} a = \cdots = P_B^{\perp} (P_A^{\perp} P_B^{\perp})^k a,$$

and, for $k = 0, 1, \ldots,$

$$\begin{aligned} P_A^{\perp} P_B (P_A P_B)^k a &= P_A^{\perp} (I - P_B^{\perp}) (P_A P_B)^k a = -P_A^{\perp} P_B^{\perp} (P_A P_B)^k a \\ &= -(P_A^{\perp} P_B^{\perp})^{k+1} a. \end{aligned}$$

Therefore, (3.3) and

$$\begin{aligned} (P_{A\cap B}^{\perp} a_1, P_{A\cap B}^{\perp} a_2) = \sum_{m=0}^{\infty} \{ &(P_B^{\perp} (P_A^{\perp} P_B^{\perp})^m a_1, P_B^{\perp} (P_A^{\perp} P_B^{\perp})^m a_2) \\ &+ ((P_A^{\perp} P_B^{\perp})^{m+1} a_1, (P_A^{\perp} P_B^{\perp})^{m+1} a_2) \}, \qquad a_1, a_2 \in A \end{aligned} \tag{3.6}$$

follow from (3.1) and (3.2), respectively. However, we have, for $a_1, a_2 \in A$ and $m = 0, 1, \ldots,$

$$\begin{aligned} (P_B^{\perp} (P_A^{\perp} P_B^{\perp})^m a_1, P_B^{\perp} (P_A^{\perp} P_B^{\perp})^m a_2) &= (P_B^{\perp} (P_A^{\perp} P_B^{\perp})^{2m} a_1, a_2), \\ ((P_A^{\perp} P_B^{\perp})^{m+1} a_1, (P_A^{\perp} P_B^{\perp})^{m+1} a_2) &= (P_B^{\perp} (P_A^{\perp} P_B^{\perp})^{2m+1} a_1, a_2). \end{aligned}$$

Thus, (3.4) follows from (3.6).

Let $a \in A$ and $b \in B$. Then, $(P_B^\perp a, P_B^\perp b) = 0$. For $m = 1, 2, \ldots$, we have

$$P_B^\perp (P_A P_B)^m b = P_B^\perp P_A (P_B P_A)^{m-1} b = -\big(P_B^\perp P_A^\perp\big)^m b,$$

whence

$$\begin{aligned}\big(P_B^\perp (P_A P_B)^m a, P_B^\perp (P_A P_B)^m b\big) &= -\big(P_B^\perp \big(P_A^\perp P_B^\perp\big)^m a, \big(P_B^\perp P_A^\perp\big)^m b\big) \\ &= -\big(\big(P_A^\perp P_B^\perp\big)^{2m} a, b\big).\end{aligned}$$

Similarly, we have, for $m = 0, 1, \ldots$,

$$P_A^\perp P_B (P_A P_B)^m b = P_A^\perp (P_B P_A)^m b = P_A^\perp \big(P_B^\perp P_A^\perp\big)^m b,$$

whence

$$\begin{aligned}\big(P_A^\perp P_B (P_A P_B)^m a, P_A^\perp P_B (P_A P_B)^m b\big) &= -\big(\big(P_A^\perp P_B^\perp\big)^{m+1} a, P_A^\perp \big(P_B^\perp P_A^\perp\big)^m b\big) \\ &= -\big(\big(P_A^\perp P_B^\perp\big)^{2m+1} a, b\big).\end{aligned}$$

Thus, (3.5) follows from (3.2). □

In the applications of this paper, A and B correspond to the infinite past and future of a multivariate stationary process.

4. Fourier coefficients of the phase function

Let $\{X_k\}$ be a q-variate stationary process satisfying the condition (A), with spectral density w. Let $\{\tilde{X}_k\}$, h and $h_\sharp$ be as in Section 2. We define a sequence $\{\beta_k\}_{k=-\infty}^{\infty}$ as the (minus of the) Fourier coefficients of the phase function $h^* h_\sharp^{-1}$:

$$\beta_k := -\int_{-\pi}^{\pi} e^{-ik\theta} h\big(e^{i\theta}\big)^* h_\sharp\big(e^{i\theta}\big)^{-1} \frac{d\theta}{2\pi}, \qquad k \in \mathbb{Z}. \tag{4.1}$$

Since $h^* h_\sharp^{-1}$ is unitary matrix valued (see Section 2), we see that $\{\beta_k\} \in \ell_2^{q\times q}(\mathbb{Z})$. The sequence $\{\beta_k\}$ plays a central role in our representation theorems.

Recall the forward and backward innovation processes $\{\xi_k\}$ and $\{\tilde{\xi}_k\}$, respectively, of $\{X_k\}$ from Section 2.

Lemma 4.1. *We assume* (A). *Then we have*

$$\langle \xi_j, \tilde{\xi}_k \rangle = -\beta_{j+k}, \qquad \langle \tilde{\xi}_k, \xi_j \rangle = -\beta_{k+j}^*, \qquad j, k \in \mathbb{Z}.$$

Proof. From (2.4), (2.8) and (2.9), we see that

$$\tilde{\xi}_k := \int_{-\pi}^{\pi} e^{ik\theta}\{h_\sharp(e^{i\theta})^*\}^{-1}\Lambda(d\theta), \qquad k \in \mathbb{Z}.$$

Combining this with (2.3) and (2.6), we obtain

$$\begin{aligned}\langle \xi_j, \tilde{\xi}_k\rangle &= \int_{-\pi}^{\pi} e^{-i(j+k)\theta} h(e^{i\theta})^{-1} h(e^{i\theta}) h(e^{i\theta})^* h_\sharp(e^{i\theta})^{-1} \frac{d\theta}{2\pi} \\ &= \int_{-\pi}^{\pi} e^{-i(j+k)\theta} h(e^{i\theta})^* h_\sharp(e^{i\theta})^{-1} \frac{d\theta}{2\pi} = -\beta_{j+k},\end{aligned}$$

which also implies the second equality. □

Remark 1. By Lemma 4.1, we have the following mutual representations between $\{\xi_k\}$ and $\{\tilde{\xi}_k\}$:

$$\xi_j = -\sum_{k=-\infty}^{\infty} \beta_{j+k}\tilde{\xi}_{jk}, \qquad \tilde{\xi}_k = -\sum_{j=-\infty}^{\infty} \beta^*_{k+j}\xi_j.$$

Lemma 4.2. *We assume* (A). *Then, for* $\{s_l\} \in \ell_{2+}^{q\times q}$ *and* $n \in \mathbb{Z}$, *we have*

$$P^\perp_{[-n,\infty)}\left(\sum_{l=0}^{\infty} s_l \xi_l\right) = -\sum_{j=0}^{\infty}\left(\sum_{l=0}^{\infty} s_l \beta_{n+j+l+1}\right)\tilde{\xi}_{n+j+1}, \tag{4.2}$$

$$P^\perp_{(-\infty,-1]}\left(\sum_{l=0}^{\infty} s_l \tilde{\xi}_{n+l+1}\right) = -\sum_{j=0}^{\infty}\left(\sum_{l=0}^{\infty} s_l \beta^*_{n+j+l+1}\right)\xi_j. \tag{4.3}$$

In particular, $\{\sum_{l=0}^{\infty} s_l\beta_{n+j+l+1}\}_{j=0}^{\infty}, \{\sum_{l=0}^{\infty} s_l\beta^*_{n+j+l+1}\}_{j=0}^{\infty} \in \ell_{2+}^{q\times q}$.

Proof. By Lemma 4.1, we have $\langle \sum_{l=0}^{\infty} s_l\xi_l, \tilde{\xi}_{n+j+1}\rangle = -\sum_{l=0}^{\infty} s_l\beta_{n+j+l+1}$. On the other hand, $\{\tilde{\xi}^k_{n+j+1} : k = 1,\dots,q,\ j \ge 0\}$ is a complete orthonormal basis of $(M^X_{[-n,\infty)})^\perp$. Thus (4.2) follows. We can prove (4.3) in a similar way. □

Remark 2. In Lemma 4.2, the map $\{s_l\}_{l=0}^{\infty} \mapsto \{\sum_{l=0}^{\infty} s_l\beta_{n+j+l+1}\}_{j=0}^{\infty}$ defines a bounded Hankel operator $\Gamma_n : \ell_{2+}^{q\times q} \to \ell_{2+}^{q\times q}$ with block Hankel matrix

$$\begin{pmatrix} \beta_{n+1} & \beta_{n+2} & \beta_{n+3} & \cdots \\ \beta_{n+2} & \beta_{n+3} & \beta_{n+4} & \cdots \\ \beta_{n+3} & \beta_{n+4} & \beta_{n+5} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

(cf. [35]), and similarly for $\{s_l\}_{l=0}^{\infty} \mapsto \{\sum_{l=0}^{\infty} s_l\beta^*_{n+j+l+1}\}_{j=0}^{\infty}$.

Lemma 4.2 allows one to define, for $n \in \mathbb{N}$ and $k \in \mathbb{N} \cup \{0\}$, the sequences $\{b_{n,j}^k\}_{j=0}^\infty \in \ell_{2+}^{q\times q}$ by the recursion

$$b_{n,j}^0 = \delta_{0j} I_q, \qquad b_{n,j}^{2k+1} = \sum_{l=0}^\infty b_{n,l}^{2k} \beta_{n+j+l+1}, \qquad b_{n,j}^{2k+2} = \sum_{l=0}^\infty b_{n,l}^{2k+1} \beta_{n+j+l+1}^*. \tag{4.4}$$

For $n \in \mathbb{N}$, we define the sequence $\{W_n^k\}_{k=0}^\infty$ in M^q by

$$W_n^{2k} = P_{(-\infty,-1]}^\perp \bigl(P_{[-n,\infty)}^\perp P_{(-\infty,-1]}^\perp\bigr)^k X_0, \qquad k = 0, 1, \dots, \tag{4.5}$$

$$W_n^{2k+1} = -\bigl(P_{[-n,\infty)}^\perp P_{(-\infty,-1]}^\perp\bigr)^{k+1} X_0, \qquad k = 0, 1, \dots. \tag{4.6}$$

Proposition 4.3. *We assume* (A). *Then, for $n \in \mathbb{N}$ and $k \in \mathbb{N} \cup \{0\}$, we have*

$$W_n^{2k} = c_0 \sum_{j=0}^\infty b_{n,j}^{2k} \xi_j, \qquad W_n^{2k+1} = c_0 \sum_{j=0}^\infty b_{n,j}^{2k+1} \tilde{\xi}_{n+j+1} \tag{4.7}$$

and

$$\langle W_n^{2k}, X_0 \rangle = c_0 b_{n,0}^{2k} c_0^*, \qquad \langle W_n^{2k+1}, X_{-(n+1)} \rangle = c_0 b_{n,j}^{2k+1} \tilde{c}_0^*. \tag{4.8}$$

Proof. Note that, from the definition of W_n^k,

$$W_n^{2k+1} = -P_{[-n,\infty)}^\perp W_n^{2k}, \qquad W_n^{2k+2} = -P_{(-\infty,-1]}^\perp W_n^{2k+1}.$$

We prove (4.7) by induction. First, from (2.7) and (2.13), we have

$$W_n^0 = P_{(-\infty,-1]}^\perp X_0 = c_0 \xi_0 = c_0 \sum_{j=0}^\infty b_{n,j}^0 \xi_j.$$

For $k = 0, 1, \dots$, assume that $W_n^{2k} = c_0 \sum_{j=0}^\infty b_{n,j}^{2k} \xi_j$. Then, by (4.2),

$$\begin{aligned} W_n^{2k+1} &= -P_{[-n,\infty)}^\perp \left(c_0 \sum_{j=0}^\infty b_{n,j}^{2k} \xi_j \right) = c_0 \sum_{j=0}^\infty \left(\sum_{l=0}^\infty b_{n,l}^{2k} \beta_{n+j+l+1} \right) \tilde{\xi}_{n+j+1} \\ &= c_0 \sum_{j=0}^\infty b_{n,j}^{2k+1} \tilde{\xi}_{n+j+1}, \end{aligned}$$

and, by (4.3),

$$W_n^{2k+2} = -P^{\perp}_{(-\infty,-1]}\left(c_0 \sum_{j=0}^{\infty} b_{n,j}^{2k+1}\tilde{\xi}_{n+j+1}\right)$$
$$= c_0 \sum_{j=0}^{\infty}\left(\sum_{l=0}^{\infty} b_{n,l}^{2k+1}\beta^*_{n+j+l+1}\right)\xi_j = c_0 \sum_{j=0}^{\infty} b_{n,j}^{2k+2}\xi_j.$$

Thus (4.7) follows. We obtain the first (resp., second) equality in (4.8) from the first (resp., second) equalities in (4.7) and (2.13). □

Lemma 4.2 also allows one to define, for $n \in \mathbb{N}$ and $k \in \mathbb{N}\cup\{0\}$, the sequences $\{\tilde{b}^k_{n,j}\}_{j=0}^{\infty} \in \ell_{2+}^{q\times q}$ by the recursion

$$\tilde{b}^0_{n,j} = \delta_{0j} I_q, \qquad \tilde{b}^{2k+1}_{n,j} = \sum_{l=0}^{\infty} \tilde{b}^{2k}_{n,l}\beta^*_{n+j+l+1}, \qquad \tilde{b}^{2k+2}_{n,j} = \sum_{l=0}^{\infty} \tilde{b}^{2k+1}_{n,l}\beta_{n+j+l+1}. \tag{4.9}$$

Proposition 4.4. *We assume* (A). *Then, for $n \in \mathbb{N}$ and $k \in \mathbb{N}\cup\{0\}$, we have $b^{2k}_{n,0} \geq 0$ and $\tilde{b}^{2k}_{n,0} \geq 0$.*

Proof. Let $A = M^X_{[-n,\infty)}$ and $B = M^X_{(-\infty,-1]}$. Then, in the same way as the proof of (3.4) in Theorem 3.2, we have

$$\langle W_n^{2k}, X_0\rangle = \begin{cases} \langle P_B^{\perp}(P_A^{\perp}P_B^{\perp})^m X_0, P_B^{\perp}(P_A^{\perp}P_B^{\perp})^m X_0\rangle, & k = 2m\text{: even}, \\ \langle (P_A^{\perp}P_B^{\perp})^{m+1} X_0, (P_A^{\perp}P_B^{\perp})^{m+1} X_0\rangle, & k = 2m+1\text{: odd}. \end{cases}$$

This and the first equality in (4.8) give $c_0 b^{2k}_{n,0} c_0^* \geq 0$ or $b^{2k}_{n,0} \geq 0$. The second equality follows from the first one applied to $\{\tilde{X}_k\}$. □

5. Representation theorems

In this section, we develop explicit representations for the finite predictor coefficients, the finite prediction error covariances and the PACF of a q-variate stationary process $\{X_k\}$, in terms of the sequence $\{\beta_j\}$ defined in Section 4. We focus on the one-step ahead predictions to keep the notation simple.

In deriving the representation theorems for the finite predictors of a q-variate stationary process $\{X_k\}$, the following *intersection of past and future* property of $\{X_k\}$ plays a key role:

$$M^X_{(-\infty,-1]} \cap M^X_{[-n,\infty)} = M^X_{[-n,-1]}, \qquad n = 1, 2, \ldots. \tag{IPF}$$

A useful sufficient condition for (IPF) is the following *minimality* condition:

$$\begin{aligned}&\{X_k\} \text{ has spectral density } w \text{ satisfying } \det w(e^{i\theta}) > 0,\ \sigma\text{-a.e.,}\\ &\text{and } w^{-1} \in L_1^{q\times q}(\mathbb{T}).\end{aligned} \tag{M}$$

In fact, by [27], Corollary 3.6, (M) implies (IPF). The condition (M) also implies (A) by [32], Lemma 2.5 and Theorem 2.8, or more directly by

$$\begin{aligned}|\log\det w| &= q\left|\log(\det w)^{1/q}\right| \le q\left\{(\det w)^{1/q} + (\det w)^{-1/q}\right\}\\ &= q\left\{(\lambda_1\cdots\lambda_q)^{1/q} + \left(\lambda_1^{-1}\cdots\lambda_q^{-1}\right)^{1/q}\right\}\\ &\le q\left\{\frac{\lambda_1+\cdots+\lambda_q}{q} + \frac{\lambda_1^{-1}+\cdots+\lambda_q^{-1}}{q}\right\} = \operatorname{Tr} w + \operatorname{Tr} w^{-1},\end{aligned}$$

where $\lambda_1, \dots, \lambda_q$ denote the eigenvalues of w and we have used the inequality $|\log y| \le y + (1/y)$ for $y > 0$.

The property (IPF) is closely related to the property

$$M^X_{(-\infty,-1]} \cap M^X_{[0,\infty)} = \{0\} \tag{CND}$$

called *complete nondeterminacy* by [40]. In fact, by [27], Theorem 3.5, (IPF) and (CND) are equivalent under (A). The condition (CND) is also closely related to the *rigidity* for matrix-valued Hardy functions (see [29]). It should be noticed that if $\{X_k\}$ satisfies (IPF), then so does the time-reversed process $\{\tilde{X}_k\}$, and that the same holds for (M) and (CND).

Recall W^k_n from (4.5) and (4.6). The next proposition is a direct consequence of (3.3) in Theorem 3.2.

Proposition 5.1. *We assume* (IPF). *Then, for* $n \in \mathbb{N}$, *we have*

$$P^{\perp}_{[-n,-1]} X_0 = \sum_{k=0}^{\infty} W^k_n,$$

the sum converging strongly in M^q.

Proof. The equality follows from (IPF) and (3.3) in Theorem 3.2 applied to $A = M^X_{[-n,\infty)}$, $B = M^X_{(-\infty,-1]}$ and $a = X^j_0$, $j = 1, \dots, q$. □

Under (A), and for $n \in \mathbb{N}$ and $k = 1, \dots, n$, the *forward and backward finite predictor coefficients* $\phi_{n,k} \in \mathbb{C}^{q\times q}$ and $\tilde{\phi}_{n,k} \in \mathbb{C}^{q\times q}$, respectively, of a q-variate stationary process $\{X_k\}$ are defined by

$$P_{[-n,-1]} X_0 = \phi_{n,1} X_{-1} + \cdots + \phi_{n,n} X_{-n}, \tag{5.1}$$

$$P_{[-n,-1]} X_{-(n+1)} = \tilde{\phi}_{n,1} X_{-n} + \cdots + \tilde{\phi}_{n,n} X_{-1}. \tag{5.2}$$

Recall c_0, $\tilde{c}_0$, β_j, $b^{2k}_{n,j}$ and $\tilde{b}^{2k}_{n,j}$ from (2.11), (2.12), (4.1), (4.4) and (4.9), respectively. Here is the representation theorem for $\phi_{n,n}$ and $\tilde{\phi}_{n,n}$, which are closely related to the PACF of $\{X_k\}$.

Theorem 5.2. *We assume* (A) *and* (IPF). *Then, for* $n \in \mathbb{N}$,

$$\phi_{n,n} = c_0 \sum_{k=0}^{\infty} \left(\sum_{j=0}^{\infty} b_{n,j}^{2k} \beta_{n+j} \right) \tilde{c}_0^{-1}, \qquad \tilde{\phi}_{n,n} = \tilde{c}_0 \sum_{k=0}^{\infty} \left(\sum_{j=0}^{\infty} \tilde{b}_{n,j}^{2k} \beta_{n+j}^* \right) c_0^{-1}.$$

Proof. Since $P_{[-n,-1]}^{\perp} X_0 \equiv -\phi_{n,n} X_{-n} \equiv -\phi_{n,n} \tilde{c}_0 \tilde{\xi}_n \bmod M_{[-n+1,\infty)}^X$, we have

$$\langle P_{[-n,-1]}^{\perp} X_0, \tilde{\xi}_n \rangle = -\phi_{n,n} \tilde{c}_0 \langle \tilde{\xi}_n, \tilde{\xi}_n \rangle = -\phi_{n,n} \tilde{c}_0.$$

On the other hand, from Propositions 5.1 and 4.3 and Lemma 4.1, we get

$$\langle P_{[-n,-1]}^{\perp} X_0, \tilde{\xi}_n \rangle = \sum_{k=0}^{\infty} \langle W_n^{2k}, \tilde{\xi}_n \rangle = -c_0 \sum_{k=0}^{\infty} \left(\sum_{j=0}^{\infty} b_{n,j}^{2k} \beta_{n+j} \right).$$

Thus the first formula follows. We obtain the second formula by applying the first one to the time-reversed process $\{\tilde{X}_k\}$, □

For $n = 0, 1, \ldots$, we define the *forward and backward finite prediction error covariances* v_n and $\tilde{v}_n$, respectively, of a q-variate stationary process $\{X_k\}$ by $v_0 = \tilde{v}_0 = \langle X_0, X_0 \rangle$ and

$$v_n := \langle P_{[-n,-1]}^{\perp} X_0, P_{[-n,-1]}^{\perp} X_0 \rangle, \qquad n = 1, 2, \ldots, \tag{5.3}$$

$$\tilde{v}_n := \langle P_{[-n,-1]}^{\perp} X_{-(n+1)}, P_{[-n,-1]}^{\perp} X_{-(n+1)} \rangle, \qquad n = 1, 2, \ldots. \tag{5.4}$$

Notice that $\tilde{v}_n$ (resp., v_n) is the forward (resp., backward) finite prediction error covariance of the time-reversed process $\{\tilde{X}_k\}$. In this paper, under (A), we fix the definition of the *partial autocorrelation function* (PACF) α_n of $\{X_k\}$ by

$$\alpha_n := \begin{cases} (v_0)^{-1/2} \langle X_0, X_{-1} \rangle (\tilde{v}_0)^{-1/2}, & n = 1, \\ (v_{n-1})^{-1/2} \langle P_{[-n+1,-1]}^{\perp} X_0, P_{[-n+1,-1]}^{\perp} X_{-n} \rangle (\tilde{v}_{n-1})^{-1/2}, & n = 2, 3, \ldots \end{cases}$$

(cf. [14]).

The next theorem gives explicit representations for v_n, $\tilde{v}_n$ and α_n.

Theorem 5.3. *We assume* (A) *and* (IPF). *Then, for* $n \in \mathbb{N}$, *we have*

$$v_n = c_0 \left(\sum_{k=0}^{\infty} b_{n,0}^{2k} \right) c_0^*, \qquad \tilde{v}_n = \tilde{c}_0 \left(\sum_{k=0}^{\infty} \tilde{b}_{n,0}^{2k} \right) \tilde{c}_0^*, \tag{5.5}$$

$$\langle P_{[-n+1,-1]}^{\perp} X_0, P_{[-n+1,-1]}^{\perp} X_{-n} \rangle = c_0 \left(\sum_{k=0}^{\infty} b_{n,0}^{2k+1} \right) \tilde{c}_0^*. \tag{5.6}$$

Proof. First, by (IPF) and (3.4) in Theorem 3.2 applied to $A = M^X_{[-n,\infty)}$, $B = M^X_{(-\infty,-1]}$ and $a_1 = X^i_0$, $a_2 = X^j_0 (i, j = 1, \dots, q)$, we have $v_n = \sum_{k=0}^{\infty} \langle W^{2k}_n, X_0 \rangle$. This and (4.8) give the first equality in (5.5). Next, we obtain the second equality in (5.5) by applying the first one to the time-reversed process $\{\tilde{X}_k\}$. Finally, by (IPF) and (3.5) in Theorem 3.2 applied to $A = M^X_{[-n,\infty)}$, $B = M^X_{(-\infty,-1]}$ and $a = X^i_0$, $b = X^j_{-(n+1)} (i, j = 1, \dots, q)$, we have

$$\langle P^{\perp}_{[-n,-1]} X_0, P^{\perp}_{[-n,-1]} X_{-(n+1)} \rangle = \sum_{k=0}^{\infty} \langle W^{2k+1}_n, X_{-(n+1)} \rangle.$$

This and (4.8) give (5.6). □

We can prove

$$\langle P^{\perp}_{[-n+1,-1]} X_0, P^{\perp}_{[-n+1,-1]} X_{-n} \rangle = \phi_{n,n} \tilde{v}_{n-1}, \qquad n = 2, 3, \dots,$$

in the same way as in the univariate case (cf. Corollary 5.2.1 in [8]). From this, we have

$$\alpha_n = (v_{n-1})^{-1/2} \phi_{n,n} (\tilde{v}_{n-1})^{1/2}, \qquad n = 1, 2, \dots, \tag{5.7}$$

and so Theorem 5.2 with (5.5) in Theorem 5.3 gives another explicit representation of α_n.

We turn to the representation of all the finite predictor coefficients $\phi_{n,j}$ and $\tilde{\phi}_{n,j}$. It turns out that, to deal with this problem, we need to assume the minimality (M) which is more stringent than (IPF) or (CND). A q-variate stationary process $\{X_k\}$ satisfying (M) has a dual process $\{X'_k : k \in \mathbb{Z}\}$, characterized by the biorthogonality relation $\langle X_j, X'_k \rangle = \delta_{jk} I_q$; see [32] for more information. Recall a_k and $\tilde{a}_k$ from (2.11) and (2.12), respectively. The dual process $\{X'_k\}$ admits the following two MA representations:

$$X'_n = -\sum_{k=0}^{\infty} a^*_k \xi_{n+k}, \qquad X'_{-n} = -\sum_{k=0}^{\infty} \tilde{a}^*_k \tilde{\xi}_{n+k}, \qquad n \in \mathbb{Z}. \tag{5.8}$$

Here notice that (M) implies

$$\{a_k\}, \{\tilde{a}_k\} \in \ell^{q \times q}_{2+}. \tag{5.9}$$

By (5.9), we can also define, for $n \in \mathbb{N}$ and $k, j \in \mathbb{N} \cup \{0\}$,

$$\phi^{2k}_{n,j} := c_0 \sum_{l=0}^{\infty} b^{2k}_{n,l} a_{j+l}. \qquad \phi^{2k+1}_{n,j} := c_0 \sum_{l=0}^{\infty} b^{2k+1}_{n,l} \tilde{a}_{j+l},$$

$$\tilde{\phi}^{2k}_{n,j} := \tilde{c}_0 \sum_{l=0}^{\infty} \tilde{b}^{2k}_{n,l} \tilde{a}_{j+l}, \qquad \tilde{\phi}^{2k+1}_{n,j} := \tilde{c}_0 \sum_{l=0}^{\infty} \tilde{b}^{2k+1}_{n,l} a_{j+l}.$$

Here is the representation theorem for the finite predictor coefficients.

Theorem 5.4. *We assume* (M). *Then, for* $n = 1, 2, \dots$ *and* $j = 1, \dots, n$,

$$\phi_{n,j} = \sum_{k=0}^{\infty} \{\phi_{n,j}^{2k} + \phi_{n,n-j+1}^{2k+1}\}, \qquad \tilde{\phi}_{n,j} = \sum_{k=0}^{\infty} \{\tilde{\phi}_{n,j}^{2k} + \tilde{\phi}_{n,n-j+1}^{2k+1}\}. \tag{5.10}$$

Proof. From $\langle X_k, X'_j \rangle = \delta_{kj} I_q$, we have $\langle P_{[-n,-1]}^{\perp} X_0, X'_{-j} \rangle = -\phi_{n,j}$ for $j = 1, \dots, n$, and, from Proposition 5.1, we find that

$$\langle P_{[-n,-1]}^{\perp} X_0, X'_{-j} \rangle = \sum_{k=0}^{\infty} \{\langle W_n^{2k}, X'_{-j} \rangle + \langle W_n^{2k+1}, X'_{-j} \rangle\}.$$

Moreover, from Proposition 4.3 and (5.8) rewritten as

$$X'_{-j} = -\sum_{l=-j}^{\infty} a_{j+l}^{*} \xi_l, \qquad X'_{-j} = -\sum_{l=-(n-j+1)}^{\infty} \tilde{a}_{n-j+l+1}^{*} \tilde{\xi}_{n+l+1},$$

we have

$$\langle W_n^{2k}, X'_{-j} \rangle = -c_0 \sum_{l=0}^{\infty} b_{n,l}^{2k} a_{j+l} = -\phi_{n,j}^{2k},$$

$$\langle W_n^{2k+1}, X'_{-j} \rangle = -c_0 \sum_{l=0}^{\infty} b_{n,l}^{2k+1} \tilde{a}_{n-j+l+1} = -\phi_{n,n-j+1}^{2k+1}.$$

Combining, we obtain the first equality in (5.10). Its second equality follows from the first one applied to the time-reversed process $\{\tilde{X}_k\}$. □

6. Applications to long-memory processes

In this section, we apply the representation theorems in Section 5 to a q-variate FARIMA process with common fractional differencing order for all components and derive the asymptotics of the finite prediction error covariances and the PACF as well as that of the finite predictor coefficients, and establish Baxter's inequality.

6.1. Univariate FARIMA processes

We start with some properties of univariate FARIMA$(0, d, 0)$ processes which we need in our perturbation technique below. This technique reduces the study of asymptotic properties of multivariate FARIMA processes to that of the corresponding problems for univariate FARIMA$(0, d, 0)$ processes.

For $d \in (-1/2, 1/2) \setminus \{0\}$, let $\{Y_k : k \in \mathbb{Z}\}$ be a univariate FARIMA$(0, d, 0)$ process with spectral density

$$w_Y(e^{i\theta}) = \left|1 - e^{i\theta}\right|^{-2d}, \qquad \theta \in (-\pi, \pi) \tag{6.1}$$

(see [16,19]; see also [8], Section 13.2). Then $u_0 := E[|Y_0|^2]$ is equal to $\Gamma(1-2d)/\Gamma(1-d)^2$ and the nth finite predictor coefficients $\psi_{n,n}$ of $\{Y_k\}$ (see (5.1)) are given by

$$\psi_{n,n} = \frac{d}{n-d}, \qquad n \in \mathbb{N}. \tag{6.2}$$

Let u_n be the finite prediction error variance of $\{Y_k\}$ defined by (5.3) with $\{X_k\}$ replaced by $\{Y_k\}$, for which we use the notation u_n rather than v_n. Then the Durbin–Levinson algorithm implies $u_n = u_0 \prod_{k=1}^{n}\{1 - (\psi_{k,k})^2\}$, whence

$$u_n = \frac{\Gamma(n+1-2d)\Gamma(n+1)}{\Gamma(n+1-d)^2}, \qquad n = 0, 1, \ldots. \tag{6.3}$$

For u_n, we present next its precise asymptotic behavior.

Proposition 6.1. *For $d \in (-1/2, 1/2) \setminus \{0\}$, we have $u_n = 1 + (d^2/n) + O(n^{-2})$ as $n \to \infty$.*

Proof. By Stirling's formula $\Gamma(x) = \sqrt{2\pi} e^{-x} x^{x+(1/2)}\{1 + (1/12x) + O(x^{-2})\}$ as $x \to \infty$ and

$$\left(\frac{n+1-2d}{n+1}\right)^{-d} = 1 + \frac{2d^2}{n} + O(n^{-2}), \qquad \frac{\sqrt{(n+1-2d)(n+1)}}{(n+1-d)} = 1 + O(n^{-2})$$

as $n \to \infty$, we have, as $n \to \infty$,

$$\begin{aligned} u_n &= \frac{\Gamma(n+1-2d)\Gamma(n+1)}{\Gamma(n+1-d)^2} \\ &= \left(1 - \frac{d}{n+1-d}\right)^{n+1-d}\left(1 + \frac{d}{n+1-d}\right)^{n+1-d}\left\{1 + \frac{2d^2}{n} + O(n^{-2})\right\}. \end{aligned}$$

On the other hand, by l'Hopital's rule, we have, for $a \in \mathbb{R}$,

$$\left(1 + \frac{a}{x}\right)^x = e^a - \frac{a^2 e^a}{2x} + O(x^{-2}), \qquad x \to \infty,$$

whence, as $n \to \infty$,

$$\left(1 - \frac{d}{n+1-d}\right)^{n+1-d}\left(1 + \frac{d}{n+1-d}\right)^{n+1-d} = \left\{1 - \frac{d^2}{n} + O(n^{-2})\right\}.$$

Combining, we obtain the proposition. □

Since

$$1 - e^{i\theta} = \begin{cases} \left|1 - e^{i\theta}\right| e^{(i/2)(\theta-\pi)} & \text{if } 0 < \theta < \pi, \\ \left|1 - e^{i\theta}\right| e^{(i/2)(\theta+\pi)} & \text{if } -\pi < \theta < 0, \end{cases}$$

the phase function

$$\Omega(e^{i\theta}) := \overline{(1 - e^{i\theta})^{-d}} / (1 - e^{i\theta})^{-d}$$

of the univariate FARIMA$(0, d, 0)$ process $\{Y_k\}$ above is given by

$$\Omega(e^{i\theta}) = \begin{cases} e^{id(\theta-\pi)} & \text{if } 0 < \theta < \pi, \\ e^{id(\theta+\pi)} & \text{if } -\pi < \theta < 0. \end{cases} \tag{6.4}$$

Therefore, the minus of the Fourier coefficients of the phase function $\Omega(e^{i\theta})$ for $\{Y_k\}$, which we write as ρ_n rather than β_n, are given by

$$\rho_n = -\int_{-\pi}^{\pi} e^{-in\theta} \Omega(e^{i\theta}) \frac{d\theta}{2\pi} = \frac{\sin(\pi d)}{\pi(n-d)}, \qquad n \in \mathbb{Z}. \tag{6.5}$$

One can also obtain (6.5) using [7], Remark 1 and Lemma 4.4.

Lemma 6.2. *Let $\{s_k\}_{k=-\infty}^{\infty}$ be a complex sequence such that $\sum_{k=-\infty}^{\infty} k^2 |s_k| < \infty$. Then, we have*

$$\lim_{n\to\infty} n \left(\rho_n^{-1} \sum_{k=-\infty}^{\infty} \rho_{n-k} s_k - \sum_{k=-\infty}^{\infty} s_k \right) = \sum_{k=-\infty}^{\infty} k s_k.$$

Proof. Since $\rho_{n-k}/\rho_n = (n-d)/(n-k-d)$, we have

$$n \left(\rho_n^{-1} \sum_{k=-\infty}^{\infty} \rho_{n-k} s_k - \sum_{k=-\infty}^{\infty} s_k \right) = \sum_{k=-\infty}^{\infty} \frac{n k s_k}{n-k-d}. \tag{6.6}$$

For $k \in \mathbb{Z}$, the function $f_{k,d} : \mathbb{Z} \to [0, \infty)$ defined by

$$f_{k,d}(n) := \left| \frac{nk}{n-k-d} \right| = \left| k + \frac{k(k+d)}{n-(k+d)} \right|$$

takes the maximum value at either $n = k-1, k$, or $k+1$, whence

$$\max_{n\in\mathbb{N}} f_{k,d}(n) \le \max\left\{ \frac{k(k-1)}{1+d}, \frac{k^2}{|d|}, \frac{k(k+1)}{1-d} \right\} \le ck^2$$

for some $c \in (0, \infty)$. Therefore, we have dominated convergence, as $n \to \infty$, on the right of (6.6), and the sum converges to $\sum_{k=-\infty}^{\infty} k s_k$, as desired. □

6.2. Multivariate FARIMA processes

Let $\overline{\mathbb{D}} := \{z \in \mathbb{C} : |z| \leq 1\}$ be the closed unit disk in $\mathbb{C}$. We consider the following condition for $g : \mathbb{T} \to \mathbb{C}^{q\times q}$:

$$\text{the entries of } g(z) \text{ are rational functions in } z \text{ that have no poles on } \overline{\mathbb{D}}, \text{ and } \det g \text{ has no zeros on } \overline{\mathbb{D}}. \tag{C}$$

The condition (C) implies that g is an outer function in $H_2^{q\times q}(\mathbb{T})$.

Lemma 6.3. *For $g : \mathbb{T} \to \mathbb{C}^{q\times q}$ with (C), there exists $\tilde{g} : \mathbb{T} \to \mathbb{C}^{q\times q}$ that satisfies (C) and*

$$g(e^{-i\theta})g(e^{-i\theta})^* = \tilde{g}(e^{i\theta})\tilde{g}(e^{i\theta})^*. \tag{6.7}$$

The function $\tilde{g}$ is uniquely determined from g up to a constant unitary factor.

Proof. Since the entries of $g(1/z)$ are rational, the lemma follows from the proof of Theorem 10.1 in [38], Chapter I. □

Let g and $\tilde{g}$ be as in Lemma 6.3. As in (2.4), we define the outer function $g_\sharp$ in $H_2^{q\times q}(\mathbb{T})$ by

$$g_\sharp(z) := \{\tilde{g}(\overline{z})\}^*. \tag{6.8}$$

Then, $g_\sharp$ satisfies both (C) and

$$g(e^{i\theta})g(e^{i\theta})^* = g_\sharp(e^{i\theta})^* g_\sharp(e^{i\theta}), \qquad \theta \in [-\pi, \pi). \tag{6.9}$$

It should be noticed that the proof of Theorem 10.1 in [38], Chapter I, is constructive, whence so is the above proof of the existence of $\tilde{g}$ and $g_\sharp$.

Example 3. For $c \in \mathbb{D}$, let

$$g(z) = \begin{pmatrix} 1 & 0 \\ 1/(1-cz) & 1 \end{pmatrix}.$$

Then g satisfies (C). From the proof of Lemma 6.3 and (6.8), we obtain

$$g_\sharp(z) = \frac{1}{\sqrt{1-|c|^2+|c|^4}} \begin{pmatrix} 1-|c|^2 & 1 \\ -1 + \dfrac{1-|c|^2}{1-cz} & -|c|^2 + \dfrac{1}{1-cz} \end{pmatrix}.$$

One can also directly check that $g_\sharp$ satisfies both (C) and (6.9).

Let $d \in (-1/2, 1/2) \setminus \{0\}$, and let $\{X_k\}$ be a q-variate stationary process which has spectral density w of the form

$$w(e^{i\theta}) = |1 - e^{i\theta}|^{-2d} g(e^{i\theta})g(e^{i\theta})^*, \qquad \text{where } g : \mathbb{T} \to \mathbb{C}^{q\times q} \text{ satisfies (C)}. \tag{F}$$

We call the process $\{X_k\}$ a *q-variate FARIMA process*. We easily find that $\{X_k\}$ satisfies (M), whence (A) and (IPF) (see Section 5). Let $\tilde{g}$ and $g_\sharp$ be as in Lemma 6.3 and (6.8), respectively. In what follows, as the outer functions h and $\tilde{h}$ for $\{X_k\}$ in Section 2, we take

$$h(z) = (1-z)^{-d} g(z), \qquad \tilde{h} = (1-z)^{-d} \tilde{g}(z). \tag{6.10}$$

Then, $h_\sharp$ defined by (2.4) is given by

$$h_\sharp(z) = (1-z)^{-d} g_\sharp(z). \tag{6.11}$$

From the second equality in (6.10), we see that the time-reversed process $\{\tilde{X}_k\}$ of $\{X_k\}$ is also a q-variate FARIMA process satisfying (F) with the same differencing order d and $\tilde{g}$ as g.

Let $\{c_n\}$ and $\{\tilde{c}_n\}$ be the forward and backward MA coefficients of $\{X_k\}$, respectively (see (2.11) and (2.12)). Then

$$c_0 = h(0) = g(0), \qquad \tilde{c}_0 = \tilde{h}(0) = h_\sharp(0)^* = g_\sharp(0)^* = \tilde{g}(0). \tag{6.12}$$

The sequence $\{\beta_n\}$ for $\{X_k\}$, which is defined by (4.1), is given by

$$\beta_n = -\int_{-\pi}^{\pi} e^{-in\theta} \Omega\big(e^{i\theta}\big) g\big(e^{i\theta}\big)^* g_\sharp\big(e^{i\theta}\big)^{-1} \frac{d\theta}{2\pi}, \qquad n \in \mathbb{Z},$$

with $\Omega(e^{i\theta})$ in (6.4).

We define a $q \times q$ unitary matrix U by

$$U := g(1)^* g_\sharp(1)^{-1}. \tag{6.13}$$

Recall the spectral norm $\|a\|$ of $a \in \mathbb{C}^{q\times q}$ from Section 2. The next proposition may be viewed as an improvement of Proposition 4.5 in [7].

Proposition 6.4. *For $d \in (-1/2, 1/2) \setminus \{0\}$, let $\{X_k\}$ be a q-variate FARIMA process with* (F). *For $n \in \mathbb{N}$, define $\Delta_n, \Delta_n' \in \mathbb{C}^{q\times q}$ by*

$$\beta_n = \rho_n (I_q + \Delta_n) U = \rho_n U \big(I_q + \Delta_n'\big),$$

respectively. Then there exists a positive constant M satisfying the two conditions

$$\|\Delta_n\| \le M n^{-1}, \qquad n \in \mathbb{N}, \tag{6.14}$$

$$\big\|\Delta_n'\big\| \le M n^{-1}, \qquad n \in \mathbb{N}. \tag{6.15}$$

Proof. We put $G(z) := \{g(1/\overline{z})\}^*$. Then $g(e^{i\theta})^* g_\sharp(e^{i\theta})^{-1} = G(e^{i\theta}) g_\sharp(e^{i\theta})^{-1}$ holds. By the property (C) for g and $g_\sharp$, there exists an open annulus A containing the unit circle $\mathbb{T}$ such that both $G(z)$ and $g_\sharp(z)^{-1}$ are holomorphic in A, whence $G(z)g_\sharp(z)^{-1}$ has the Laurent series expansion

$$G(z) g_\sharp(z)^{-1} = \sum_{k=-\infty}^{\infty} s_k z^k, \qquad z \in A.$$

Since $A \supset \mathbb{T}$, the entries of s_k decay exponentially as $k \to \pm\infty$. Moreover, since $\beta_n = \sum_{k=-\infty}^{\infty} \rho_{n-k} s_k$ and $U = \sum_{k=-\infty}^{\infty} s_k$, we have

$$\Delta_n = \left(\rho_n^{-1} \sum_{k=-\infty}^{\infty} \rho_{n-k} s_k - \sum_{k=-\infty}^{\infty} s_k\right) U^{-1},$$

$$\Delta_n' = U^{-1}\left(\rho_n^{-1} \sum_{k=-\infty}^{\infty} \rho_{n-k} s_k - \sum_{k=-\infty}^{\infty} s_k\right).$$

Therefore, the proposition follows from Lemma 6.2. □

6.3. Asymptotics of the finite prediction error covariances

In this section, we derive the precise asymptotics of the finite prediction error covariance matrices for q-variate FARIMA processes with (F).

For $d \in (-1/2, 1/2) \setminus \{0\}$, let $\{X_k\}$ be a q-variate FARIMA process with (F). Let v_n and $\tilde{v}_n$ be the forward and backward finite prediction error covariances of $\{X_k\}$ defined by (5.3) and (5.4), respectively. We define the *forward and backward infinite prediction error covariances* $v_\infty \in \mathbb{C}^{q\times q}$ and $\tilde{v}_\infty \in \mathbb{C}^{q\times q}$, respectively, of $\{X_k\}$ by

$$v_\infty := \langle P^\perp_{(-\infty,-1]} X_0, P^\perp_{(-\infty,-1]} X_0\rangle = c_0 c_0^*, \tag{6.16}$$

$$\tilde{v}_\infty := \langle P^\perp_{[1,\infty)} X_0, P^\perp_{[1,\infty)} X_0\rangle = \tilde{c}_0 \tilde{c}_0^*, \tag{6.17}$$

where $\{c_n\}$ and $\{\tilde{c}_n\}$ are the forward and backward MA coefficients of $\{X_k\}$, respectively (see (2.11) and (2.12)). It should be noticed that $\tilde{v}_\infty$ (resp., v_∞) is the forward (resp., backward) infinite prediction error covariance of the time-reversed process $\{\tilde{X}_k\}$.

Theorem 6.5. *For $d \in (-1/2, 1/2) \setminus \{0\}$, let $\{X_k\}$ be a q-variate FARIMA process with* (F). *Then* (1.4) *and* (1.5) *hold.*

Proof. Let u_n be as in (6.3); it is the nth finite prediction error variance for a univariate fractional ARIMA$(0, d, 0)$ process $\{Y_k\}$ with spectral density (6.1). We prove the assertion (1.4) by comparing v_n with u_n.

From the representation of v_n in (5.5), we have

$$v_n - v_\infty = c_0\left(\sum_{k=1}^{\infty} b_{n,0}^{2k}\right) c_0^*.$$

Similarly, u_n can be expressed, in terms of $\{\rho_j\}$ in (6.5) only, as

$$u_n - 1 = \sum_{k=1}^{\infty} r_{n,0}^{2k},$$

where, for $n \in \mathbb{N}$ and $k \in \mathbb{N} \cup \{0\}$, $\{r^k_{n,j}\}_{k=0}^{\infty} \in \ell_{2+}$ is the analogue of $\{b^k_{n,l}\}_{k=0}^{\infty}$ for $\{Y_k\}$, defined by the recursion

$$r^0_{n,j} = \delta_{0j}, \qquad r^{k+1}_{n,j} = \sum_{l=0}^{\infty} r^k_{n,l}\rho_{n+j+l+1}. \tag{6.18}$$

Let Δ_n and M be as in Proposition 6.4. Recall U from (6.13).

From the definitions, we have

$$b^2_{n,0} = \sum_{l=0}^{\infty} \beta_{n+l+1}\beta^*_{n+l+1}, \qquad r^2_{n,0} = \sum_{l=0}^{\infty} \rho_{n+l+1}\rho_{n+l+1}.$$

Since U is unitary, we have, for $j, k \geq n$,

$$\beta_j \beta^*_k = \rho_j \rho_k (I_q + \Delta_j)\big(I_q + \Delta^*_k\big).$$

By Proposition 6.4 and the inequality $(1+x)^2 - 1 \leq 2x(1+x)^2$ for $x \geq 0$, we have

$$\begin{aligned}\big\|(I_q + \Delta_j)\big(I_q + \Delta^*_k\big) - I_q\big\| &= \big\|\Delta_j + \Delta^*_k + \Delta_j\Delta^*_k\big\| \\ &\leq \big(1 + \|\Delta_j\|\big)\big(1 + \|\Delta_k\|\big) - 1 \leq \big(1 + Mn^{-1}\big)^2 - 1 \\ &\leq 2Mn^{-1}\big(1 + Mn^{-1}\big)^2\end{aligned}$$

for $j, k \geq n$. Thus,

$$\big\|b^2_{n,0} - r^2_{n,0} I_q\big\| \leq 2Mn^{-1}\big(1 + Mn^{-1}\big)^2 r^2_{n,0}, \qquad n \in \mathbb{N}.$$

In the same way, we have, for $k = 1, \ldots,$

$$\big\|b^{2k}_{n,0} - r^{2k}_{n,0} I_q\big\| \leq 2kMn^{-1}\big(1 + Mn^{-1}\big)^{2k} r^{2k}_{n,0}, \qquad n \in \mathbb{N}.$$

Take $t > 1$ such that $t^2 \sin(\pi|d|) < 1$. Define $\tau_{2k} \in (0, \infty)$ by

$$\big(\pi^{-1}\arcsin x\big)^2 = \sum_{k=1}^{\infty} \tau_{2k} x^{2k}, \qquad |x| < 1 \tag{6.19}$$

(cf. Lemma 3.1 in [26]). Then, as in the proof of Proposition 3.2 in [26], there exists an $N \in \mathbb{N}$ such that

$$1 + Mn^{-1} \leq t, \qquad r^{2k}_{n,0} \leq n^{-1}\big\{t \sin\big(\pi|d|\big)\big\}^{2k}\tau_{2k} \qquad (k \in \mathbb{N}, n \geq N).$$

Combining, we have, for $n \geq N$,

$$\begin{aligned}\left\|n(v_n - v_\infty) - n(u_n - 1)v_\infty\right\| &\leq \|c_0\|^2 \sum_{k=1}^{\infty} n\left\|b_{n,0}^{2k} - r_{n,0}^{2k} I_q\right\| \\ &\leq n^{-1} M \|c_0\|^2 \sum_{k=1}^{\infty} 2k\tau_{2k}\left\{t^2 \sin\bigl(\pi|d|\bigr)\right\}^{2k},\end{aligned}$$

whence $\|n(v_n - v_\infty) - n(u_n - 1)v_\infty\| = O(n^{-1})$ as $n \to \infty$. This and Proposition 6.1 yield (1.4). We obtain (1.5) by applying (1.4) to the time-reversed process $\{\tilde{X}_k\}$. □

6.4. Asymptotics of the PACF

In this section, we derive the precise asymptotics of the PACF for a q-variate FARIMA process $\{X_k\}$ with (F). Recall U from (6.13). As above, $\{c_n\}$ and $\{\tilde{c}_n\}$ denote the forward and backward MA coefficients of $\{X_k\}$, respectively (see (2.11) and (2.12)).

First, we consider the asymptotics of $\phi_{n,n}$ in (5.1).

Theorem 6.6. *Let $d \in (-1/2, 1/2) \setminus \{0\}$, and let $\{X_k\}$ be a q-variate FARIMA process with* (F). *Then*

$$\phi_{n,n} = \frac{d}{n} c_0 U \tilde{c}_0^{-1} + O\bigl(n^{-2}\bigr), \qquad n \to \infty.$$

Proof. The proof is similar to that of Theorem 6.5. From the representation of $\phi_{n,n}$ in Theorem 5.2, we have

$$\phi_{n,n} = c_0\left(\sum_{k=0}^{\infty} \phi_n^k\right)\tilde{c}_0^{-1} \qquad \text{with } \phi_n^k := \sum_{j=0}^{\infty} b_{n,j}^{2k} \beta_{n+j}.$$

Similarly, the scalar coefficient $\psi_{n,n}$ for a univariate FARIMA$(0, d, 0)$ process $\{Y_k\}$, which is given by (6.2), can be expressed, in terms of $\{\rho_j\}$ in (6.5) only, as

$$\psi_{n,n} = \sum_{k=0}^{\infty} \psi_n^k \qquad \text{with } \psi_n^k := \sum_{j=0}^{\infty} r_{n,j}^{2k} \rho_{n+j},$$

where $r_{n,j}^k$ are defined by the recursion (6.18). We define $\varepsilon := d/|d|$ so that $|\rho_n| = \varepsilon\rho_n$. Let Δ_n and M be as in Proposition 6.4.

First, since

$$\phi_n^0 = \beta_n = \rho_n(I_q + \Delta_n)U, \qquad \psi_n^0 = \rho_n,$$

it follows from Proposition 6.4 that

$$\left\|\phi_n^0 - \psi_n^0 U\right\| \leq M n^{-1} \varepsilon\rho_n = M n^{-1} \varepsilon\psi_n^0.$$

Next, we have

$$\phi_n^1 = \sum_{j=0}^{\infty} \left(\sum_{l=0}^{\infty} \beta_{n+l+1} \beta_{n+j+l+1}^* \right) \beta_{n+j},$$

$$\psi_n^1 = \sum_{j=0}^{\infty} \left(\sum_{l=0}^{\infty} \rho_{n+l+1} \rho_{n+j+l+1} \right) \rho_{n+j}.$$

Then, since U is unitary, we have, for $j, k, l \geq n$,

$$\beta_j \beta_k^* \beta_l = \rho_j \rho_k \rho_l (I_q + \Delta_j)\big(I_q + \Delta_k^*\big)(I_q + \Delta_l) U.$$

By Proposition 6.4 and the inequality $(1+x)^3 - 1 \leq 3x(1+x)^3$ for $x \geq 0$, we have

$$\begin{aligned}
&\big\| (I_q + \Delta_j)\big(I_q + \Delta_k^*\big)(I_q + \Delta_l) - I_q \big\| \\
&\quad = \big\| \Delta_j + \Delta_k^* + \Delta_l + \Delta_j \Delta_k^* + \Delta_j \Delta_l + \Delta_k^* \Delta_l + \Delta_j \Delta_k^* \Delta_l \big\| \\
&\quad \leq \big(1 + \|\Delta_j\|\big)\big(1 + \|\Delta_k\|\big)\big(1 + \|\Delta_l\|\big) - 1 \leq \big(1 + Mn^{-1}\big)^3 - 1 \\
&\quad \leq 3Mn^{-1}\big(1 + Mn^{-1}\big)^3
\end{aligned}$$

for $j, k, l \geq n$. Thus,

$$\big\| \phi_n^1 - \psi_n^1 U \big\| \leq 3Mn^{-1}\big(1 + Mn^{-1}\big)^3 \varepsilon \psi_n^1, \qquad n \in \mathbb{N}.$$

In the same way, we have, for $k = 0, 1, \ldots,$

$$\big\| \phi_n^k - \psi_n^k U \big\| \leq (2k+1) Mn^{-1}\big(1 + Mn^{-1}\big)^{2k+1} \varepsilon \psi_n^k, \qquad n \in \mathbb{N}.$$

Take $t > 1$ such that $t^2 \sin(\pi |d|) < 1$. Define $\tau_{2k+1} \in (0, \infty)$ by

$$\pi^{-1} \arcsin x = \sum_{k=0}^{\infty} \tau_{2k+1} x^{2k+1}, \qquad |x| < 1 \tag{6.20}$$

(cf. Lemma 3.1 in [26]). Then, as in the proof of Proposition 3.2 in [26], there exists an $N \in \mathbb{N}$ such that

$$1 + Mn^{-1} \leq t, \qquad \varepsilon \psi_n^k \leq n^{-1} \big\{ t \sin\big(\pi |d|\big) \big\}^{2k+1} \tau_{2k+1} \qquad \big(k \in \mathbb{N} \cup \{0\}, n \geq N\big).$$

Combining, we have, for $n \geq N$,

$$\begin{aligned}\left\| n\phi_{n,n} - \frac{n}{n-d} dc_0 U \tilde{c}_0^{-1} \right\| &= n \| \phi_{n,n} - \psi_{n,n} c_0 U \tilde{c}_0^{-1} \| \leq \|c_0\| \| \tilde{c}_0^{-1} \| \sum_{k=0}^{\infty} n \| \phi_n^k - \psi_n^k U \| \\ &\leq n^{-1} \|c_0\| \| \tilde{c}_0^{-1} \| M \sum_{k=0}^{\infty} (2k+1) \tau_{2k+1} \{ t^2 \sin(\pi |d|) \}^{2k+1},\end{aligned}$$

whence $\|n\phi_{n,n} - dc_0 U \tilde{c}_0^{-1}\| = O(n^{-1})$ as $n \to \infty$. Thus, the theorem follows. □

Recall v_∞ and $\tilde{v}_\infty$ from (6.16) and (6.17), respectively. Notice that $v_\infty^{-1/2} c_0$ (resp., $\tilde{v}_\infty^{-1/2} \tilde{c}_0$) is the polar part of c_0 (resp., $\tilde{c}_0$). Recall the PACF α_n of $\{X_k\}$ from Section 5. The above theorem gives the following rate of convergence for α_n as $n \to \infty$.

Theorem 6.7. *Let $d \in (-1/2, 1/2) \setminus \{0\}$, and let $\{X_k\}$ be a q-variate FARIMA process with (F). Then (1.6) holds with the unitary matrix $V \in \mathbb{C}^{q\times q}$ given by*

$$V := v_\infty^{-1/2} c_0 \cdot U \cdot \left(\tilde{v}_\infty^{-1/2} \tilde{c}_0 \right)^*.$$

Proof. From the first equality in (5.5) in Theorem 5.3 and Proposition 4.4, we have $v_n \geq v_\infty$. Therefore, we see from Theorem 6.5 and [4], Theorem X.3.7, that $\|v_n^{1/2} - v_\infty^{1/2}\| = O(n^{-1})$ as $n \to \infty$. Similarly, we have $\tilde{v}_n^{1/2} = \tilde{v}_\infty^{1/2} + O(n^{-1})$ as $n \to \infty$.

From $v_n \geq v_\infty$ and [4], Propositions V.1.6 and V.1.8, we have $v_n^{-1/2} \leq v_\infty^{-1/2}$, so that $\|v_n^{-1/2}\| \leq \|v_\infty^{-1/2}\|$. Hence, as $n \to \infty$,

$$\begin{aligned}\| v_n^{-1/2} - v_\infty^{-1/2} \| &= \| v_n^{-1/2} (v_\infty^{1/2} - v_n^{1/2}) v_\infty^{-1/2} \| \\ &\leq \| v_\infty^{-1/2} \|^2 \| v_n^{1/2} - v_\infty^{1/2} \| = O(n^{-1}).\end{aligned}$$

Combining these with (5.7) and Theorem 6.6, we have

$$\begin{aligned}n\alpha_n &= v_{n-1}^{-1/2} \cdot n\phi_{n,n} \cdot \tilde{v}_{n-1}^{1/2} \\ &= \{ v_\infty^{-1/2} + O(n^{-1}) \} \{ dc_0 U \tilde{c}_0^{-1} + O(n^{-1}) \} \{ \tilde{v}_\infty^{1/2} + O(n^{-1}) \} \\ &= d v_\infty^{-1/2} c_0 \cdot U \cdot \left(\tilde{v}_\infty^{-1/2} \tilde{c}_0 \right)^* + O(n^{-1})\end{aligned}$$

as $n \to \infty$. Thus, the theorem follows. □

Remark 4. If we choose g and $\tilde{g}$ so that both $g(0) \geq 0$ and $\tilde{g}(0) \geq 0$ hold, then we see from (6.12), (6.16) and (6.17) that $c_0 = v_\infty^{1/2}$ and $\tilde{c}_0 = \tilde{v}_\infty^{1/2}$, whence $V = U$.

6.5. Baxter's inequality

In this section, we present Baxter's inequality for multivariate FARIMA processes with $0 < d < 1/2$. It extends the corresponding univariate result in [26].

For $d \in (-1/2, 1/2) \setminus \{0\}$, let $\{X_k\}$ be a q-variate FARIMA process with (F). Recall the forward and backward AR coefficients a_n and $\tilde{a}_n$ of $\{X_k\}$ from (2.11) and (2.12), respectively. They satisfy

$$\left\| n^{1+d} a_n + \frac{1}{\Gamma(-d)} g(1)^{-1} \right\| = O\left(n^{-1}\right), \qquad n \to \infty, \tag{6.21}$$

$$\left\| n^{1+d} \tilde{a}_n + \frac{1}{\Gamma(-d)} \left\{g_\sharp(1)^*\right\}^{-1} \right\| = O\left(n^{-1}\right), \qquad n \to \infty \tag{6.22}$$

(cf. [23], Lemma 2.2). In particular, we have

$$\lim_{n\to\infty} n^{1+d} \|a_n\| = \frac{\|g(1)^{-1}\|}{|\Gamma(-d)|}, \tag{6.23}$$

$$\lim_{n\to\infty} n^{1+d} \|\tilde{a}_n\| = \frac{\|\{g_\sharp(1)^*\}^{-1}\|}{|\Gamma(-d)|}. \tag{6.24}$$

We see from (6.23) and (6.24) that (2.14) holds if $0 < d < 1/2$.

Recall $\phi_{n,k}$ and ϕ_k from (5.1) and (2.16), respectively.

Theorem 6.8. *For $d \in (0, 1/2)$, let $\{X_k\}$ be a q-variate FARIMA process with* (F). *Then the forward finite and infinite predictor coefficients $\phi_{n,k}$ and ϕ_k, respectively, of $\{X_k\}$ satisfy*

$$\sum_{j=1}^{n} \|\phi_{n,j} - \phi_j\| = O\left(n^{-d}\right), \qquad n \to \infty.$$

Proof. For $k = 0, 1, \ldots,$ we show by induction on k that

$$\left\|b^k_{n,l}\right\| \le \left(1 + Mn^{-1}\right)^k r^k_{n,l}, \qquad n \in \mathbb{N}, l \in \mathbb{N} \cup \{0\}, \tag{6.25}$$

where M is a positive constant satisfying (6.14) and $r^k_{n,l}$ are defined by (6.18). Indeed, the case $k = 0$ is evident by the definitions $b^0_{n,l} = \delta_{0l} I_q$ and $r^0_{n,l} = \delta_{0l}$. Assuming (6.25) for $k \ge 0$, we see from Proposition 6.4 that

$$\begin{aligned}\left\|b^{k+1}_{n,l}\right\| &\le \sum_{m=0}^{\infty} \left\|b^k_{n,m}\right\| \|\beta_{n+l+m+1}\| \\ &\le \left(1 + Mn^{-1}\right)^{k+1} \sum_{m=0}^{\infty} r^k_{n,m} \rho_{n+l+m+1} = \left(1 + Mn^{-1}\right)^{k+1} r^{k+1}_{n,l}.\end{aligned}$$

Thus (6.25) also holds for $k+1$.

Define $\tau_k \in (0,\infty)$ by (6.19) and (6.20). Then we see from Proposition 3.2 in [26] that, for any $t>1$, there exits an $N \in \mathbb{N}$ such that

$$nr_{n,l}^k \le \tau_k \big\{t \sin(\pi d)\big\}^k, \qquad 1 + Mn^{-1} \le t \qquad \big(l \in \mathbb{N} \cup \{0\}, k \in \mathbb{N}, n \ge N\big).$$

Here we take $t>1$ such that $t^2 \sin(\pi d) < 1$. Then, from (6.25),

$$n \sum_{k=1}^{\infty} \|b_{n,l}^k\| \le \sum_{k=1}^{\infty} \tau_k \big\{t^2 \sin(\pi d)\big\}^k < \infty, \qquad l \in \mathbb{N} \cup \{0\}, n \ge N. \tag{6.26}$$

From $\phi_j = c_0 a_j = c_0 \sum_{l=0}^{\infty} b_{n,l}^0 a_{j+l}$ and Theorem 5.4, we have

$$\phi_{n,j} - \phi_j = c_0 \sum_{k=1}^{\infty} \sum_{l=0}^{\infty} b_{n,l}^{2k} a_{j+l} + c_0 \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} b_{n,l}^{2k+1} \tilde{a}_{n-j+l+1},$$

whence

$$\sum_{j=1}^{n} \|\phi_{n,j} - \phi_j\| \le \sum_{j=1}^{n} \sum_{l=0}^{\infty} R_{j+l} \sum_{k=1}^{\infty} \|b_{n,l}^k\|,$$

where $R_j = \max\{\|\phi_j\|, \|\tilde{\phi}_j\|\}$. Since $n^{1+d} R_n$ is bounded by (6.23) and (6.24), we have, for $n \in \mathbb{N}$,

$$n^{-1+d} \sum_{j=1}^{n} \sum_{l=j}^{\infty} R_l \le \Big\{\sup_{l \in \mathbb{N}} l^{1+d} R_l\Big\} \left\{\sup_{m \in \mathbb{N}} m^{-1+d} \sum_{j=1}^{m} \sum_{l=j}^{\infty} l^{-1-d}\right\} < \infty.$$

Hence we see from (6.26) that, for $n \ge N$,

$$n^d \sum_{j=1}^{n} \|\phi_{n,j} - \phi_j\| \le \left\{\sum_{k=1}^{\infty} \tau_k \big\{r^2 \sin(\pi d)\big\}^k\right\} \left\{\sup_{m \in \mathbb{N}} m^{-1+d} \sum_{j=1}^{m} \sum_{l=j}^{\infty} R_l\right\} < \infty.$$

The desired result follows from this. □

Since $\phi_n = c_0 a_n$, we see from (6.21) that

$$\left\| n^{1+d} \phi_n + \frac{1}{\Gamma(-d)} c_0 g(1)^{-1} \right\| = O\big(n^{-1}\big), \qquad n \to \infty.$$

In particular,

$$\lim_{n \to \infty} n^{1+d} \|\phi_n\| = \frac{\|c_0 g(1)^{-1}\|}{|\Gamma(-d)|}.$$

From this and [6], Proposition 1.5.8, we obtain the following asymptotic behavior of $\sum_{j=n+1}^{\infty} \|\phi_j\|$ as $n \to \infty$:

$$\lim_{n\to\infty} n^d \sum_{j=n+1}^{\infty} \|\phi_j\| = \frac{\|c_0 g(1)^{-1}\|}{\Gamma(1-d)}. \tag{6.27}$$

Here is Baxter's inequality for multivariate FARIMA processes with $0 < d < 1/2$.

Theorem 6.9. *For $d \in (0, 1/2)$, let $\{X_k\}$ be a q-variate FARIMA process with (F), and let $\phi_{n,k}$ and ϕ_n be as in Theorem 6.8. Then, there exists a positive constant K such that (1.3) holds.*

Proof. In view of (6.27), Theorem 6.8 gives the desired assertion. □

By applying Theorem 6.9 to the time-reversed process $\{\tilde{X}_k\}$, we immediately obtain the following backward Baxter inequality.

Corollary 6.10. *For $d \in (0, 1/2)$, let $\{X_k\}$ be a q-variate FARIMA process with (F), and let $\tilde{\phi}_{n,k}$ and $\tilde{\phi}_k$ be the backward finite and infinite predictor coefficients, respectively, of $\{X_k\}$. Then, there exists a positive constant $\tilde{K}$ such that*

$$\sum_{j=1}^{n} \|\tilde{\phi}_{n,j} - \tilde{\phi}_j\| \le \tilde{K} \sum_{j=n+1}^{\infty} \|\tilde{\phi}_j\|, \qquad n \in \mathbb{N}. \tag{6.28}$$

Acknowledgements

We would like to thank the anonymous referees for their helpful comments. M. Pourahmadi was supported by the NFS Grant DMS-13-09586.

References

[1] Baillie, R.T. and Kapetanios, G. (2013). Estimation and inference for impulse response functions from univariate strongly persistent processes. *Econom. J.* **16** 373–399. MR3146771
[2] Baxter, G. (1962). An asymptotic result for the finite predictor. *Math. Scand.* **10** 137–144. MR0149584
[3] Berk, K.N. (1974). Consistent autoregressive spectral estimates. *Ann. Statist.* **2** 489–502. MR0421010
[4] Bhatia, R. (1997). *Matrix Analysis*. New York: Springer. MR1477662
[5] Bingham, N.H. (2012). Multivariate prediction and matrix Szegö theory. *Probab. Surv.* **9** 325–339. MR2956574
[6] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989). *Regular Variation*. Cambridge: Cambridge Univ. Press. MR1015093
[7] Bingham, N.H., Inoue, A. and Kasahara, Y. (2012). An explicit representation of Verblunsky coefficients. *Statist. Probab. Lett.* **82** 403–410. MR2875229
[8] Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd ed. New York: Springer. MR1093459

[9] Bühlmann, P. (1995). Moving-average representation of autoregressive approximations. *Stochastic Process*. *Appl*. **60** 331–342. MR1376807

[10] Bühlmann, P. (1997). Sieve bootstrap for time series. *Bernoulli* **3** 123–148. MR1466304

[11] Cheng, R. and Pourahmadi, M. (1993). Baxter's inequality and convergence of finite predictors of multivariate stochastic processes. *Probab. Theory Related Fields* **95** 115–124. MR1207310

[12] Chung, C.-F. (2001). Calculating and analyzing impulse responses for the vector ARFIMA model. *Econom*. *Lett*. **71** 17–25. MR1821722

[13] Damanik, D., Pushnitski, A. and Simon, B. (2008). The analytic theory of matrix orthogonal polynomials. *Surv*. *Approx*. *Theory* **4** 1–85. MR2379691

[14] Dègerine, S. (1990). Canonical partial autocorrelation function of a multivariate time series. *Ann*. *Statist*. **18** 961–971. MR1056346

[15] Ginovian, M.S. (1999). Asymptotic behavior of the prediction error for stationary random sequences. *Izv*. *Nats*. *Akad*. *Nauk Armenii Mat*. **34** 18–36. MR1854056

[16] Granger, C.W.J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *J*. *Time Series Anal*. **1** 15–29. MR0605572

[17] Hannan, E.J. and Deistler, M. (1988). *The Statistical Theory of Linear Systems*. New York: Wiley. MR0940698

[18] Helson, H. and Lowdenslager, D. (1961). Prediction theory and Fourier series in several variables. II. *Acta Math*. **106** 175–213. MR0176287

[19] Hosking, J.R.M. (1981). Fractional differencing. *Biometrika* **68** 165–176. MR0614953

[20] Ibragimov, I.A. and Solev, V.N. (1968). Asymptotic behavoir of the prediction error of a stationary sequence with a spectral density of special form. *Theory Probab. Appl*. **13** 703–707.

[21] Ing, C.-K., Chiou, H.-T. and Guo, M. (2016). Estimation of inverse autocovariance matrices for long memory processes. *Bernoulli* **22** 1301–1330. MR3474817

[22] Inoue, A. (2000). Asymptotics for the partial autocorrelation function of a stationary process. *J*. *Anal*. *Math*. **81** 65–109. MR1785278

[23] Inoue, A. (2002). Asymptotic behavior for partial autocorrelation functions of fractional ARIMA processes. *Ann*. *Appl*. *Probab.* **12** 1471–1491. MR1936600

[24] Inoue, A. (2008). AR and MA representation of partial autocorrelation functions, with applications. *Probab. Theory Related Fields* **140** 523–551. MR2365483

[25] Inoue, A. and Kasahara, Y. (2004). Partial autocorrelation functions of the fractional ARIMA processes with negative degree of differencing. *J*. *Multivariate Anal*. **89** 135–147. MR2041213

[26] Inoue, A. and Kasahara, Y. (2006). Explicit representation of finite predictor coefficients and its applications. *Ann*. *Statist*. **34** 973–993. MR2283400

[27] Inoue, A., Kasahara, Y. and Pourahmadi, M. (2016). The intersection of past and future for multivariate stationary processes. *Proc*. *Amer. Math*. *Soc*. **144** 1779–1786. MR3451253

[28] Kasahara, Y. and Bingham, N.H. (2014). Verblunsky coefficients and Nehari sequences. *Trans*. *Amer. Math*. *Soc*. **366** 1363–1378. MR3145734

[29] Kasahara, Y., Inoue, A. and Pourahmadi, M. (2016). Rigidity for matrix-valued Hardy functions. *Integral Equations Operator Theory* **84** 289–300. MR3456943

[30] Katsnelson, V.E. and Kirstein, B. (1997). On the theory of matrix-valued functions belonging to the Smirnov class. In *Topics in Interpolation Theory* (*Leipzig*, 1994). *Oper. Theory Adv. Appl*. **95** 299–350. Basel: Birkhäuser. MR1473261

[31] Kreiss, J.-P., Paparoditis, E. and Politis, D.N. (2011). On the range of validity of the autoregressive sieve bootstrap. *Ann*. *Statist*. **39** 2103–2130. MR2893863

[32] Masani, P. (1960). The prediction theory of multivariate stochastic processes. III. Unbounded spectral densities. *Acta Math*. **104** 141–162. MR0121952

[33] Meyer, M., Jentsch, C. and Kreiss, J.-P. (2017). Baxter's inequality and sieve bootstrap for random fields. *Bernoulli* **23** 2988–3020. MR3654797
[34] Meyer, M., McMurry, T. and Politis, D. (2015). Baxter's inequality for triangular arrays. *Math. Methods Statist.* **24** 135–146. MR3366950
[35] Peller, V.V. (2003). *Hankel Operators and Their Applications*. New York: Springer. MR1949210
[36] Poskitt, D.S., Grose, S.D. and Martin, G.M. (2015). Higher-order improvements of the sieve bootstrap for fractionally integrated processes. *J. Econometrics* **188** 94–110. MR3371662
[37] Pourahmadi, M. (2001). *Foundations of Time Series Analysis and Prediction Theory*. New York: Wiley. MR1849562
[38] Rozanov, Yu.A. (1967). *Stationary Random Processes*. San Francisco: Holden-Day. MR0214134
[39] Rupasinghe, M. and Samaranayake, V.A. (2012). Asymptotic properties of sieve bootstrap prediction intervals for *FARIMA* processes. *Statist. Probab. Lett.* **82** 2108–2114. MR2979746
[40] Sarason, D. (1978). *Function Theory on the Unit Circle*. Blacksburg, VA: Virginia Polytechnic Institute and State Univ. Notes for lectures given at a Conference at Virginia Polytechnic Institute and State University, Blacksburg, Va., June 19–23, 1978. MR0521811

Received January 2016 and revised May 2016