Periodically Correlated Sequences with Rational Spectra and PARMA Systems

Andrzej Makagon

Abstract An acronym PARMA is used in different configurations, we talk about PARMA systems, PARMA sequences, or PARMA models. This paper is a result of the author search to understand this complex world of PARMAs.

1 Introduction

Periodically correlated sequences (PC) are sequences that are obtained by listing elements of a multi-variate stationary sequence in a linear order. There is no surprise therefore that both theories are strictly related. VARMA models are representations of multi-variate stationary sequences in a form of vector difference equations (VDE). PARMA models result from nonhomogeneous (periodic) VDE representations and under mild conditions yield periodically correlated sequences. PARMA models form a subset of VARMA models. Since only stationary sequences with rational densities admit VARMA models, it is natural that the study of PARMA models should involve an analysis of periodically correlated sequences with rational densities.

Sequences with rational densities play an important role in the theory of multi-variate stationary sequences. These are the only multi-variate stationary sequences for which the theoretical prediction problem has an explicit solution (under small additional assumptions), that is the only multi-variate stationary sequences for which it is possible to explicitly compute the coefficients of the innovation representation of the sequence from its density. This solution, however, is not fully satisfactory since it depends on infinitely many parameters (i.e. innovation coefficients). A VARMA model is an ingenious concept of reducing the number of parameters. The idea is to represent

Andrzej Makagon
Department of Mathematics, Hampton University, Hampton, VA 23668 e-mail: andrzej.makagon@hamptonu.edu

a sequence as a solution of a vector difference equations that involves only finitely many terms. In spectral domain this translates to a problem of writing a transfer function of the sequence (i.e. a square factor of its density) as a "quotient" of two polynomial matrices. The main purpose of this paper is to transfer these and other relevant results known for multi-variate stationary sequences to those periodically correlated sequences given by PARMA models.

A paper is organized as follows. In Section 2 we provide notation and definitions needed in the sequel. Section 3 contains an extensive review of the theory of periodically correlated sequences and multi-variate stationary sequences that sometimes goes beyond the needs of this paper. A focus is on the relation between PC and T-variate stationary sequences and on sequences with rational densities. In Section 4 we discuss a relationship between PC sequences with rational densities and PARMA systems. The next Section 5 contains few remarks about PARMA models. In that section we limit our attention to full rank sequences; a general case of any rank sequences seems to be still open even for multi-variate stationary sequences.

Most of the facts about multi-variate stationary sequences come from [11], while for the theory of VARMA systems from [4] and [5]. Up-to-date review of PARMA models and related topics, with emphasis on statistics, can be found in [2]. Sources of information about periodically correlated sequences will be cited as needed. To our best knowledge, periodically correlated sequences with rational densities have not been studied before.

2 Preliminaries

Sets and Matrices. In what follows $\mathcal{C}$ denotes the set of complex numbers, $D_{<r} = \{z \in \mathcal{C} : |z| < r\}$ is an open disk of radius r, and $D_r = \{z \in \mathcal{C} : |z| = r\}$ is a circle of radius r, $r > 0$. The interval $[0, 2\pi)$ will be understood as a group with addition modulo 2π and with standard Lebesgue measure structure. The abbreviation *a.e.* will mean *almost everywhere* with respect to the Lebesgue measure on $[0, 2\pi)$. A function on $[0, 2\pi)$ will be interpreted as a function of the unit circle D_1 and will be written as $f(e^{it})$ or $f(z)$, $z = e^{it}$, $t \in [0, 2\pi)$. This notation is convenient in analysis of sequences with rational densities and we use it all over the paper. The letter T will be always a fixed positive integer and $\mathcal{Z}$ will denote the set of integers. The symbols $\mathrm{q}(m)$ and $\langle m\rangle$ stand for the quotient and the nonnegative remainder in division of m by T, so that $m = \mathrm{q}(m)T + \langle m\rangle$. The entries of matrices and vectors in this paper are indexed from 0 instead from 1. The (i, j) entry of a matrix A will be denoted by $A^{i,j}$. For any matrix A, A^* is the complex conjugate of A, i.e. $(A^*)^{i,j} = \overline{A^{j,i}}$. If B is a square $n \times n$ matrix then the *adjugate* B^A of B is an $n \times n$ matrix whose (i, j) entry is given by $(B^A)^{ij} = (-1)^{i+j} \det(B_{[j,i]})$, where $B_{[j,i]}$ is obtained from B by deleting j-th row and i-th column. A square matrix B is invertible iff $\det(B) \neq 0$ and if it is then $B^{-1} = (\det(B))^{-1} B^A$. An $n \times n$ diagonal matrix A will be denoted *diag* $[A^{0,0}, A^{1,1}, \dots A^{n-1,n-1}]$. A *minor of degree* k, $k \leq \min(m,n)$, of an $m \times n$ matrix A is the determinant of a square $k \times k$ sub-matrix of A obtained by deleting $m-k$ rows and $n-k$ columns from A. A rank of a matrix A is the maximum number of linearly independent columns or rows of A; $\mathrm{rank}(A) = r$ iff there is a minor of order r which is nonzero and all minors of bigger order are zero. A *principal minor* of order k of a square $n \times n$ matrix A is a minor obtained by deleting last $n-k$ columns and the last $n-k$ rows. A square matrix A is non-negative ($A \geq 0$) if for every $a \in \mathcal{C}^T$, $aAa^* \geq 0$; $A \geq 0$ iff all principal minors are non-negative. A $1 \times n$ matrix will be called a *vector*.

A *rational* $m \times n$ matrix $R(z)$ is a matrix whose entries $R(z)^{j,k}$ are rational functions of a complex variable $z \in \mathcal{C}$; that is, each entry is a ratio of two polynomials. We will be *always* assuming that the entries of a rational matrix are written in the simplest form. Note that since minors of a rational matrix are rational functions, if a minor is non-zero for one z than it is nonzero for all $z \in \mathcal{C}$ except finitely many points. Therefore the rank of a rational matrix is constant except for finitely many z's. The *poles* of a rational matrix are numbers z such that $R(z)$ does not exist and the zeros of $R(z)$ are complex numbers z at which the matrix $R(z)$ drops its rank (e.g. [7], or [5], Section 3.2.). A *polynomial* $m \times n$ matrix is a matrix whose entries are polynomials of a complex variable z. A matrix function $R(e^{it})$ on $[0, 2\pi)$ is called rational (or polynomial) if there is a rational (polynomial) matrix $R(z)$ on $\mathcal{C}$ such that $R(e^{it})$ equals $R(z)$ a.e. on D_1. If $P(z)$ is a polynomial matrix and $p(z) = \det P(z)$ is not identically zero, then $p(z) \neq 0$ for all $z \in \mathcal{C}$ except finitely many, and $P(z)^{-1} = p(z)^{-1} P(z)^A$ is a rational matrix. The most important fact from the theory of rational matrices will be for us the following theorem proved by Rozanov.

Theorem 2.1 ([11], Thm. 10.1) *Each a.e. nonegative rational square matrix function $F(e^{it})$ on $[0, 2\pi)$ of rank r (i.e. rank of $F(z)$ is r) can be represented in the form $F(e^{it}) = G(e^{it}) G(e^{it})^*$ a.e. where $G(z)$ is rational, analytic on $D_{<1}$, and the rank of $G(z)$ is r for* all $z \in D_{<1}$.

In terms of zeros and poles the last statement means that all zeros and poles of $G(z)$ are outside of the open unit disk $D_{<1}$.

A square polynomial matrix is $U(z)$ is called *unimodular* if $\det(U(z))$ is constant (i.e. does not depend on z). An $m \times m$ matrix $L(z)$ is called a *left (common) divisor* of $m \times n$ polynomial matrices $A(z)$ and $B(z)$ if there are $m \times n$ polynomial matrices $A_1(z)$ and $B_1(z)$ such that $A(z) = L(z) A_1(z)$ and $B(z) = L(z) B_1(z)$. Note that if $\det A(z)$ is not identically zero, then also $\det L(z)$ is not. A left divisor $L(z)$ is called a *greatest left divisor* of $A(z)$ and $B(z)$ if for any other left divisor $L_1(z)$ of $A(z)$ and $B(z)$ there is a polynomial $m \times m$ matrix $T(z)$ such that $L(z) = L_1(z) T(z)$. Polynomial matrices $A(z)$ and $B(z)$ are called *left coprime* if the only left divisors of $A(z)$ and $B(z)$ are the unimodular once. For those and other interesting facts from the theory of polynomial and rational matrices we refer the reader to [7] or [5], Section 3.2.

Hilbert Spaces. A Hilbert space will be denoted $\mathcal{H}$ (or $\mathcal{K}$), and (x, y) will denote the inner product of $x, y \in \mathcal{H}$ (or $\mathcal{K}$). All Hilbert spaces are assumed to be complex and separable. If $\mathcal{H}$ is a Hilbert space and M is a closed subspace of $\mathcal{H}$, then $(x|M)$ denotes the orthogonal projection of $x \in \mathcal{H}$ onto M. If S is any subset of $\mathcal{H}$ then $\overline{\text{sp}}\{S\}$ denote the smallest closed linear subspace of $\mathcal{H}$ containing S. A linear mapping (operator) from $\mathcal{H}$ onto $\mathcal{K}$ is called unitary if $(Ux, Uy)_{\mathcal{K}} = (x, y)_{\mathcal{H}}$ for every $x, y \in \mathcal{H}$. A sequence (ξ_n), $n \in \mathcal{Z}$, of elements of $\mathcal{H}$ is *orthonormal* if $(\xi_n, \xi_m) = 1$ if $m = n$, and zero otherwise. Important for us will be the space $\mathcal{C}^T$ of all row vectors of the length T with entries in $\mathcal{C}$ and standard Euclidean inner product, the Hilbert space L^2 of all measurable complex functions (in fact equivalence classes of functions) on $[0, 2\pi)$ which are square integrable w.r.t the Lebesgue measure dt on $[0, 2\pi)$, and the Hilbert space $L^2(\mathcal{C}^T)$ of all $\mathcal{C}^T$-valued functions f on $[0, 2\pi)$ such that their entries are in L^2. The inner product in $L^2(\mathcal{C}^T)$ is $(f, g) = \int_0^{2\pi} f(e^{it})g(e^{it})^* dt$. The symbol L^2_+ will denote the subspace of L^2 consisting functions f whose Fourier coefficients with negative indices are zero, i.e. such that $f(e^{it}) = \sum_{k=0}^{\infty} f_k e^{itk}$; $L^2_+(\mathcal{C}^T)$ denotes the subspace of functions in $L^2_+(\mathcal{C}^T)$ such that their entries are in L^2_+. The *standard orthonormal basis* for $\mathcal{C}^T$ is e_k, $k = 0 \ldots, T-1$, where e_k is the row vector that has 1 at the k-th place and zero otherwise. The *standard orthonormal basis* for $L^2(\mathcal{C}^T)$ is the family of functions $\zeta_n(e^{it})$, $n \in \mathcal{Z}$, defined by $\zeta_{nT+r}(e^{it}) = (1/\sqrt{2\pi})e^{-int}e_r$, $n \in \mathcal{Z}$, $r = 0, \ldots, T-1$. If $G(e^{it})$ is a matrix function then we say that G is square integrable if all entries of G are in L^2.

Stochastic Sequences. We adopt a Hilbert space approach. A (univariate) *stochastic sequence* $(x(n))$ in a Hilbert space $\mathcal{H}$ is a sequence of elements of $\mathcal{H}$ indexed by the set of all integers $\mathcal{Z}$. The *correlation function* of $(x(n))$ is the function on $\mathcal{Z}^2$ defined by $R_x(m, n) = (x(m), x(n))$. If $(x(n))$ is a stochastic sequence then we denote $M_x = \overline{\text{sp}}\{x(m) : m \in \mathcal{Z}\}$. Two stochastic sequences $(x(n))$ in $\mathcal{H}$ and $(y(n))$ in possibly different Hilbert space $\mathcal{K}$ are said to be *equivalent* if $R_x(m, n) = R_y(m, n)$ for every $m, n \in \mathcal{Z}$ or, equivalently, if there is a unitary mapping Φ from M_x onto M_y such that $\Phi(x(n)) = y(n)$, $n \in \mathcal{Z}$. The concept of an equivalence of stochastic sequences makes the Hilbert space $\mathcal{H}$ appearing in the definition of a stochastic sequence irrelevant and we will stop writing it unless will be necessary. If T is a positive integer, then a *T-variate stochastic sequence in $\mathcal{H}$* is a family of T stochastic sequences $(x^k(n))$, $k = 0, \ldots, T-1$, in $\mathcal{H}$. It is convenient to write it as a sequence of column vectors $\boldsymbol{x}(n) = [x^k(n)]$ with entries in $\mathcal{H}$ and $x^0(n)$ being at the top. The *correlation function* of a T-variate stochastic sequence $\boldsymbol{x}(n) = [x^k(n)]$, $n \in \mathcal{Z}$, is $T \times T$ matrix valued function $R_{\boldsymbol{x}}$ on $\mathcal{Z}^2$ defined as $R_{\boldsymbol{x}}(m, n)^{j,k} = (x^j(n), x^k(m))$. A T-variate sequence $\boldsymbol{\xi}_n = [\xi^j_n]$, $n \in \mathcal{Z}$, is called *a T-variate orthonormal sequence* if $R_{\boldsymbol{x}}(n, n) = I$, the identity matrix, and $R_{\boldsymbol{x}}(m, n) = 0$ for all $m \neq n$.

If $(\boldsymbol{x}(n))$ is a T-variate stochastic sequence then we define $M_{\boldsymbol{x}}(n) = \overline{\mathrm{sp}}\{x^k(m) : k = 0, \dots, T-1, m \leq n\}$, $M_{\boldsymbol{x}} = M_{\boldsymbol{x}}(+\infty)$, and $N_{\boldsymbol{x}}(n) = M_{\boldsymbol{x}}(n) \ominus M_{\boldsymbol{x}}(n-1)$. A sequence $(\boldsymbol{x}(n)$ is called *regular* if $\bigcap_n M_{\boldsymbol{x}}(n) = \{0\}$. One of the main goals of prediction theory is to find an orthogonal projection $\hat{\boldsymbol{x}}(n) = (\boldsymbol{x}(n)|M_{\boldsymbol{x}}(n-1) := [(x^k(n)|M_{\boldsymbol{x}}(n-1)]$. In probabilistic terms $\hat{\boldsymbol{x}}(n)$ represents the best linear estimate (predictor) of a random vector $\boldsymbol{x}(n)$ from the past. All the above notation and definitions are valid for univariate sequences, that is when $T = 1$.

3 PCs and T-variate Stationary Sequences.

A T-variate stochastic sequence $(\boldsymbol{x}(n))$ is called *stationary* if for every $n \in \mathcal{Z}$, $R_{\boldsymbol{x}}(n+r, r)$ is constant in $r \in \mathcal{Z}$. If it is so, then the function $K_{\boldsymbol{x}}(n) := R_{\boldsymbol{x}}(n, 0)$ is called the *(matrix) correlation function of a stationary sequence* $(\boldsymbol{x}(n))$. If $(\boldsymbol{x}(n))$ is T-variate stationary then the mapping $Ux^k(n) = x^k(n+1)$, $n \in \mathcal{Z}$, $k = 0, \dots, T-1$, extends linearly to a unitary operator on $M_{\boldsymbol{x}}$ which is called the *shift* of $(\boldsymbol{x}(n))$. A T-variate stochastic sequence $(\boldsymbol{x}(n))$ is completely described by its shift and the vector $\boldsymbol{x}(0)$; namely $\boldsymbol{x}(n) = U^n \boldsymbol{x}(0) := [U^n x^k(0)]$. If we write $U^n = \int_0^{2\pi} e^{-iun} E(du)$, $n \in \mathcal{Z}$, (e.g. [3]) then we obtain that $K_{\boldsymbol{x}}(n) = \int_0^{2\pi} e^{-iun} F(du)$ where F is a $T \times T$ nonnegative matrix measure on $[0, 2\pi)$ defined by $F^{j,k}(\varDelta) = (E(\varDelta)x^k(0), x^j(0))$. The measure F is called the *spectrum* of a T-variate stationary sequence $(\boldsymbol{x}(n))$.

An important example is a sequence $(H(n))$ of $T \times T$ matrix valued functions with rows in $L^2(\mathcal{C}^T)$ defined as $H(n)(e^{it}) = e^{-int} H(e^{it})$, $n \in \mathcal{Z}$, $t \in [0, 2\pi)$. The sequence $(H(n))$ can be viewed as a T-variate stationary sequence in $\mathcal{H} = L^2(\mathcal{C}^T)$. The k-th coordinate $H^k(n)$ of $H(n)$ is the k-th row $H^{k\cdot}(e^{it})$ of the matrix $H(e^{it})$ multiplied by e^{-int}, the shift of $(H(n))$ is multiplication by e^{-it}, the correlation function of $(H(n))$ is $K_H(n) = \int_0^{2\pi} e^{-int} H(e^{it}) H(e^{it})^* dt$, so the spectrum F of $(H(n))$ is $F(dt) = H(e^{it}) H(e^{it})^* dt$.

A (univariate) sequence $(x(n))$ is called *periodically correlated with period* T (we will abbreviate it T-PC), if for every $n \in \mathcal{Z}$ the sequence $R_x(n+r, r) = (x(n+r), x(r))$ is T-periodic in $r \in \mathcal{Z}$. The discrete Fourier transform of $R_x(n+r, r)$ with respect r will be denoted $a_j(n)$; more precisely

$$a_j(n) := \sum_{r=0}^{T-1} e^{-2\pi i j r/T} R_x(n+r, r), \qquad j = 0, \dots, T-1. \tag{1}$$

If $(x(n))$ is T-PC then the mapping $Wx(n) = x(n+T)$, $n \in \mathcal{Z}$, extends linearly to a unitary operator in M_x which is called the *T-shift* of a T-PC sequence $(x(n))$. To describe a T-PC sequence we need $x(0)$ and two unitary operators; namely, a sequence $(x(n))$ is T-PC iff $(x(n))$ is of the form

$$x(n) = U^n \left[(1/T) \sum_{j=0}^{T-1} e^{-2\pi ijn/T} V^j x(0) \right], \qquad n \in \mathcal{Z}, \tag{2}$$

where U and V are unitary operators in some Hilbert space $\mathcal{K} \supseteq M_x$, $V^T = I$, and U, V satisfy a *canonical commutation relation* $VU = e^{-2\pi i/T}UV$. If we write operators U and V in terms of their spectral resolutions as follows $U^n = \int_0^{2\pi} e^{-iun}E(du)$, $n \in \mathcal{Z}$, and $V^j = \sum_{k=0}^{T-1} e^{2\pi ikj/T}P_k$, $j = 0, \ldots, T-1$, then the formula (2) takes the form $x(n) = U^n P_{\langle n \rangle} x(0)$ and we obtain that $R_x(n+r, r) = (U^n x(0), P_{\langle r \rangle} x(0))$, $n, r \in \mathcal{Z}$. Consequently $a_j(n) = \int_0^{2\pi} e^{-iun}\gamma^j(du)$, where $\gamma^j(\Delta) = (E(\Delta)x(0), V^j x(0))$, $j = 0, \ldots, T-1$ (see [9] for details). The vector measure $\gamma = (\gamma^0, \ldots, \gamma^{T-1})$ is called the *spectrum* of a T-PC sequence $(x(n))$. The existence of measures γ^j can be proved in a different way (e.g. [6], Section 6.2).

If the spectrum F or γ is absolutely continuous with respect to Lebesgue measure on $[0, 2\pi)$, then we call the respective sequence *absolutely continuous*, and abbreviate it *a.c.*. The Radon-Nikodym derivative of an a.c. spectrum with respect to the Lebesgue measure on $[0, 2\pi)$ will be called a *density* of the sequence. As indicated in Section 2 we will look at a density as a function on the unit circle rather than on $[0, 2\pi)$. Here are the precise definitions:

1. A *density of a T-variate stationary a.c. sequence* $(\boldsymbol{x}(n))$ is a $T \times T$ matrix function F on the unit circle D_1 with integrable entries such that

$$K_{\boldsymbol{x}}(n) = \int_0^{2\pi} e^{-itn} F\left(e^{it}\right) dt, \qquad n \in \mathcal{Z}. \tag{3}$$

2. A *density of an a.c. T-PC sequence* $(x(n))$ is a vector function $g = (g^0, \ldots, g^{T-1})$ on the unit circle D_1 with integrable entries such that

$$a_j(n) = \int_0^{2\pi} e^{-itn} g^j\left(e^{it}\right) dt, \qquad j = 0, \ldots, T-1, \ n \in \mathcal{Z}. \tag{4}$$

A *square factor of a density* F of a T-variate stationary sequence is a $T \times T$ matrix function H on the unit circle D_1 with entries in L^2 such that $F\left(e^{it}\right) = H\left(e^{it}\right) H\left(e^{it}\right)^*$ a.e. A *square factor of a density* g of a T-PC sequence is defined as a vector function h on the unit circle D_1 with values in $\mathcal{C}^T$ and with entries in L^2 such that for every $j = 0, \ldots, T-1$, $g^j(e^{it}) = h(e^{it})h(e^{i(t+2\pi j/T)})^*$ a.e. If g is a density of a T-PC sequence then g admits at least one square factor (see [8] or Lemma 3.1 below). Note that our definition of a square factor is slightly different from the one given in [8], where it was defined as $g^j(e^{it}) = (1/T)h(e^{it})h(e^{i(t+2\pi j/T)})^*$.

Remark 1 In the case of stationary sequences H is often called a *transfer function* (the nomenclature seems to come from signal processing). Suppose that $(\boldsymbol{x}(n))$ is T-variate stationary sequence with a density F and H is a square factor of F. We can think about a filter whose input is a T-variate "white noise" on T, that is, in our approach, a Hilbert space valued vector measure $\boldsymbol{z}(\Delta) = [z^k(\Delta)]$ on $[0, 2\pi)$ such that $(z^k(\Delta_1), z^j(\Delta_2)) = \ell(\Delta_1 \cap \Delta_2)$ if $k = j$ (where ℓ is the Lebesgue measure) and zero otherwise. The output is a T-variate sequence $\mathbf{y}(n) = \int_0^{2\pi} e^{-int} H\left(e^{it}\right) \boldsymbol{z}(dt)$ which is stationary and its spectral density is equal to F, that is $(\mathbf{y}(n))$ unitary equivalent to the sequence $(\boldsymbol{x}(n))$. A square factor h of a density g of a T-PC sequence $(x(n))$ has a similar interpretation. Now the input of a filter is a univariate "white noise" $z(\Delta)$. The output of the filter is

$$y(n) = \int_0^{2\pi} e^{-int} \left[\frac{1}{T} \sum_{k=0}^{T-1} e^{-i2\pi nk/T} h\left(e^{i(t+2\pi k/T)}\right) \right] z(dt). \tag{5}$$

Corollary 3.1 shows that $(y(n))$ is a T-PC sequence equivalent to $(x(n))$. □

There is a natural one-to-one correspondence between T-variate stationary sequences and T-PC sequences. Namely, if $\boldsymbol{x}(n) = [x^k(n)]$, $n \in \mathcal{Z}$, is a T-variate stationary sequence then by arranging $x^k(n)$'s in one sequence we obtain a T-PC sequence $x(n) = x^{\langle n \rangle}(\mathrm{q}(n))$, $n \in \mathcal{Z}$, and conversely, if $x(n)$ is T-PC and we define $x^k(n) = x(nT+k)$, $k = 0, \ldots, T-1$, $n \in \mathcal{Z}$, then $\boldsymbol{x}(n) = [x^k(n)]$, $n \in \mathcal{Z}$, is a T-variate stationary. Given a T-PC sequence $(x(n))$, the T-variate stationary sequence $(\boldsymbol{x}(n)) = [x(nT+k)]$ defined above will be called the *block sequence corresponding to* $(x(n))$; given a T-variate stationary sequence $(\boldsymbol{x}(n))$, the sequence $x(n) = x^{\langle n \rangle}(\mathrm{q}(n))$, $n \in \mathcal{Z}$, will be called the T-PC sequence *corresponding to* $(\boldsymbol{x}(n))$. Since $M_{\boldsymbol{x}}(n) = M_x(nT+T-1)$, both sequences are simultaneously regular or not. The following lemma describes the relation between the spectra and square factors of $(x(n))$ and $(\boldsymbol{x}(n))$.

Lemma 3.1 *Let $(x(n))$ be T-PC and $(\boldsymbol{x}(n))$ be the corresponding T-variate stationary block sequence.*

1. *Suppose that $(\boldsymbol{x}(n))$ is a.c. and $H\left(e^{it}\right)$ is a square factor of its density $F\left(e^{it}\right)$. Let $H^{k\cdot}$ denote the k-th row of H, and let*

$$h(e^{it}) := \sum_{k=0}^{T-1} e^{ikt} H^{k\cdot}(e^{iTt}), \quad t \in [0, 2\pi).$$

   Then $(x(n))$ is a.c. and h is a square factor of a density $g = (g^0, \ldots, g^{T-1})$ of $(x(n))$, that is $g^j(e^{it}) = h(e^{it}) h(e^{i(t+2\pi j/T)})^*$ a.e., $j = 0 \ldots, T-1$.
2. *Suppose that $(x(n))$ is a.c. and $h(e^{it})$ is a square factor of a density g of $(x(n))$. Define*

$$f_k(e^{it}) = (1/T) \sum_{j=0}^{T-1} e^{-ik(t+2\pi j/T)} h(e^{i(t+2\pi j/T)}), \quad k = 0, \ldots, T-1.$$

Then f_k is $2\pi/T$-periodic, and hence there is a function $h_k(e^{it})$ such that $f_k(e^{it}) = h_k(e^{iTt})$. Let $H(e^{it})$ be the $T \times T$ matrix function which k-th row $H^{k\cdot}$ is equal h_k. Then $(\boldsymbol{x}(n))$ is a.c. and H is a square factor of its density F, that is $F(e^{it}) = H(e^{it})H(e^{it})^*$ a.e.

The lemma is a consequence of Lemmas 3.2 and 3.3 from [8] (applied for μ to be the Lebsgue measure). For a convenience of the reader, and because the proof for a.c. sequences is much easier than in a general case, we will give a full proof in Appendix. The proof is not a replacement nor a simplification of the proof given in [8] since it works only for a.c. sequences.

We will express the above relations in a matrix form. To simplify the notation we denote $e^{it} = z$ and $e^{i2\pi/T} = d$. Note that $d^k = d^{\langle k\rangle}$ and that $\sum_{k=0}^{T-1} d^{jk} = 0$ unless $j = 0$ modulo T when it is equal to T. Let $H = [H^{j,k}]$ and $h = (h^0, \dots, h^{T-1})$ be the square factors defined in Lemma 3.1. Given h we define the *companion matrix functions* $H_d(z)$ by $H_d^{j,k}(z) = h^k(zd^j)$. Note that h is just a first row of H_d. Further let $D_z = \textit{diag}\,[1, z, z^2 \dots, z^{T-1}]$, and let D be the matrix with entries $D^{j,k} = d^{jk}$. Easy computation shows $D^{-1} = (1/T)[d^{-jk}]$. Under these notations the part 1 of Lemma 3.1 says that $h(z) = (1, z, \dots, z^{T-1})H(z^T)$, while part 2 says that

$$H^{k\cdot}(z^T) = (1/T)z^{-k}\sum_{j=0}^{T-1} d^{-jk}h(zd^j) = (D_z^{-1}D^{-1}H_d(z))^{k\cdot}, \tag{6}$$

which also shows that $H_d(z) = DD_zH(z^T)$. Summing up we have that

$$h(z) = (1, z, \dots, z^{T-1})H(z^T) \tag{7}$$

$$H(z^T) = D_z^{-1}D^{-1}H_d(z), \quad \text{where } H_d^{j,k}(z) = h^k(zd^j) \tag{8}$$

The operations (7) and (8) are inverse to each other, that is if we start with H, construct h as in (7), and then use (8), we will end up with the same H; indeed

$$H(z) \xrightarrow{(7)} h(z) = (1, z, \dots, z^{T-1})H(z^T) \xrightarrow{(8)} D_z^{-1}D^{-1}H_d(z) = H(z^T).$$

Therefore the correspondence $h \leftrightarrow H$ described in Lemma 3.1 is one-to-one and onto.

The substitution $z = e^{it}$ is not just symbolic convenience. In many cases it defines a concrete function of complex variable; for example if $R(e^{it})$ is a rational function of t then $R(z)$ is a rational function of complex variable that coincides with $R(e^{it})$ on the unit circle D_1.

From Lemma 3.1 we obtain the following functional model for a.c. T-PC sequences, which is a special a.c. case of [8], Theorem 3.2 (and also [9], Theorem 3.3).

Corollary 3.1 *Let $(x(n))$ be an a.c. T-PC sequence with density g and let h be a square factor of g. Let $(y(n))$ be a sequence in $L^2(\mathcal{C}^T)$ defined as*

$$y(n)(e^{it}) = (1/T) \sum_{j=0}^{T-1} e^{-in(t+2\pi j/T)} h(e^{i(t+2\pi j/T)}), \quad t \in [0, 2\pi).$$

Then $(y(n))$ and $(x(n))$ are unitarily equivalent.

Proof. From (6) it follows that, in terms of $z = e^{it}$, $d = e^{i2\pi/T}$ and $H(z)$, $y(n)(z) = (1/T) \sum_{j=0}^{T-1} z^{-n} d^{-jn} h(zd^j) = z^{-\mathrm{q}(n)T} H^{\langle n\rangle\cdot}(z^T)$. Hence $y(n)(e^{i\cdot})$, $n \in \mathcal{Z}$, is a T-PC sequence in $L^2(\mathcal{C}^T)$ that corresponds to a T-variate stationary sequence $(\mathbf{y}(n))$ in $L^2(\mathcal{C}^T)$ defined by $y^k(n)(e^{it}) = e^{-i\mathrm{q}(n)Tt} H^{\langle n\rangle\cdot}(e^{itT})$. Since for any integrable 2π periodic function ϕ, $\int_0^{2\pi} \phi(e^{iTt})dt = \int_0^{2\pi} \phi(e^{it})dt$, $(\mathbf{y}(n))$ has the same covariance as the T-variate block sequence corresponding to $(x(n))$, and hence $(y(n))$ and $(x(n))$ are unitarily equivalent. □

An immediate consequence of Lemma 3.1 is that the corresponding sequences $(\boldsymbol{x}(n))$ and $(x(n))$ are simultaneously a.c. or not. Since their densities are respectively $F(z) = H(z)H(z)^*$ and $g(z) = h(z)H_d(z)^* = e_0 H_d(z)H_d(z)^*$, $z = e^{it}$, where $e_0 = (1, 0, \ldots, 0)$, the relations (7) and (8) yield the following relations between the density g of a T-PC sequence $(x(n))$ and the density F of the corresponding T-variate stationary block sequence:

$$F(z^T) = (1/T) D_z^{-1} D^{-1} G(z) D D_z, \tag{9}$$

where $G(z) = H_d(z)H_d(z)^*$, that is $G(z)^{j,k} = h(zd^j)h(zd^k)^* = g^{\langle k-j\rangle}(zd^j)$; and

$$g(z) = e_0 D D_z F(z^T) D_z^* D^*. \tag{10}$$

Remember that in the above formulas $z = e^{it}$ and $d = e^{i2\pi/T}$. Therefore we have obtained the following corollary.

Corollary 3.2 *Let $(x(n))$ be T-PC and $(\boldsymbol{x}(n))$ be the corresponding T-variate stationary block sequence. Suppose that both are a.c., and let g and F be their densities, respectively. Then*

$$F^{j,k}(z^T) = (1/T^2) z^{k-j} \sum_{p=0}^{T-1} \sum_{q=0}^{T-1} d^{kq-jp} g^{\langle q-p\rangle}(zd^p),$$

$$g^k(z) = \sum_{p=0}^{T-1} \sum_{q=0}^{T-1} z^{p-q} d^{-qk} F^{p,q}(z^T),$$

where $z = e^{it}$ and $d = e^{i2\pi/T}$.

For not a.c. sequences, the relation between spectral measures of $(x(n))$ and $(\boldsymbol{x}(n))$ can be found in [8].

Definition 3.1 *A T-PC sequence $(x(n))$ is said to have a* rational density *if $(x(n))$ is a.c. and there is a rational vector function $g(z)$ of complex variable such that $g(e^{it})$ is a density of $(x(n))$. A T-variate stationary sequence $(\boldsymbol{x}(n))$*

is said to have a rational density *if* $(\boldsymbol{x}(n))$ *is a.c. and there is a rational* $T \times T$ *matrix function* $F(z)$ *of complex variable such that* $F(e^{it})$ *is a density of* $(\boldsymbol{x}(n))$.

The formulas (9) and (10) show, in particular, that F is rational iff g is rational. We summarize our discussion in the following theorem.

Theorem 3.1 *Let* $(x(n))$ *be an a.c.* T*-PC sequence and* $(\boldsymbol{x}(n))$ *be the corresponding* T*-variate stationary block sequence. Then* $(x(n))$ *has a rational density iff* $(\boldsymbol{x}(n))$ *has a rational density. Moreover, if* $F(z)$ *is a rational density of* $(\boldsymbol{x}(n))$ *and* $H(z)$ *is a rational square factor of* $F(z)$*, then* $h(z)$ *defined by (7) is a rational square factor of a density* $g(z)$ *of* $(x(n))$*; and vice verse, if* $g(z)$ *is a rational density of* $(x(n))$ *and* $h(z)$ *is a rational square factor of* $g(z)$*, then* $H(z)$ *defined by (8) is a rational square factor of a density* $F(z)$ *of* $(\boldsymbol{x}(n))$*. Everywhere above* $z = e^{it}$*,* $t \in [0, 2\pi)$*.*

Sequences with rational density may have square factors which are not rational.

A T-PC sequence $(x(n))$ is called a *moving average (MA)* if there exist an orthonormal system (ξ_n) in some Hilbert space $\mathcal{H} \supseteq M_x$ and a set of scalars $(c_k(n))$, $n, k \in \mathcal{Z}$, such that each $(c_k(n))$, $k \in \mathcal{Z}$, is T-periodic in n,

$$x(n) = \sum_{k=-\infty}^{\infty} c_k(n)\xi_{n-k}, \quad n \in \mathcal{Z}. \tag{11}$$

and $U^T x(n) = x(n+T)$, $n \in \mathcal{Z}$, where U is the shift of (ξ_n) defined by $U\xi_n = \xi_{n+1}$. Note that we allow some ξ_k's to be outside M_x; and we allow some $c_0(n)$'s to be zero. If two sequences $(x(n))$ and $(y(n))$ are unitarily equivalent and one has an MA representation, then the other also does. A sequence may have many different MA representations. We recognize an MA representation of $(x(n))$ (if it exists) by listing its coefficients $(c_k(n))$. A T-PC sequence has an MA representation iff it is a.c.. An MA representation $(c_k(n))$ of a T-PC sequence $(x(n))$ in called an *innovation representation of* $(x(n))$ if for every $n \in \mathcal{Z}$

$$M_x(n) = \overline{\mathrm{sp}}\{c_0(m)\xi_m : m \le n\} := M_{c\xi}(n). \tag{12}$$

For (12) to be true for every n it is enough that it is true for $n = 0, \ldots, T-1$. An MA representation $(c_k(n))$ of $(x(n))$ is an innovation representation iff $c_k(n) = 0$ for all $k < 0$ and $n \in \mathcal{Z}$, and $c_0(n)\xi_n$ is a one-step prediction error at n, that is $x(n) - (x(n)|M_x(n-1)) = c_0(n)\xi_n$, $n \in \mathcal{Z}$. A T-PC sequence $(x(n))$ has an innovation representation iff it is regular. If it does then the number of nonzero elements in the set $\{c_0(m) : m = 0, \ldots, T-1\}$, is called the *rank* of $(x(n))$.

A *moving average (MA)* representation of a T-variate stationary stochastic sequence $(\boldsymbol{x}(n))$ is a representation of $(\boldsymbol{x}(n))$ in the form

$$\boldsymbol{x}(n) = \sum_{k=-\infty}^{\infty} C_k \boldsymbol{\xi}_{n-k}, \qquad n \in \mathcal{Z}, \tag{13}$$

where C_k's are $T \times T$ matrices, and $\boldsymbol{\xi}_n = [\xi_n^j]$ is a T-variate orthonormal sequence in some space $\mathcal{H} \supseteq M_{\boldsymbol{x}}$ such that $W\boldsymbol{x}^k(n) = \boldsymbol{x}^k(n+1)$, $n \in \mathcal{Z}$, $k = 0, \ldots, T-1$, where W is the shift of $(\boldsymbol{\xi}_n)$ defined as $W\xi_n^k = \xi_{n+1}^k$. An MA representation of $(\boldsymbol{x}(n))$ (if exist) will be recognized by listing its matrix coefficients (C_k). A T-variate stationary sequence $(\boldsymbol{x}(n))$ has an MA representation iff $(\boldsymbol{x}(n))$ is a.c.. An MA representation (C_k) of a T-variate stationary sequence $(\boldsymbol{x}(n))$ is called an *innovation representation of* $(\boldsymbol{x}(n))$ iff for every $n \in \mathcal{Z}$,

$$M_{\boldsymbol{x}}(n) = \overline{\text{sp}}\{aC_0\boldsymbol{\xi}_m : m \leq n, a \in \mathcal{C}^T\}, \tag{14}$$

that is iff $C_k = 0$ for all $k < 0$ and for each n, $\boldsymbol{x}(n) - (\boldsymbol{x}(n)|M_{\boldsymbol{x}}(n-1)) = C_0(n)\boldsymbol{\xi}_n$, $n \in \mathcal{Z}$. Here $(\boldsymbol{x}(n)|M_{\boldsymbol{x}}(n-1)) = [(\boldsymbol{x}^k(n)|M_{\boldsymbol{x}}(n-1))]$. For (14) to be true it is enough that it is true for $n = 0$. A T-variate stationary sequence has an innovation representation iff it is regular. If it does then then dimension of the space $N_{\boldsymbol{x}}(0) = \overline{\text{sp}}\{aC_0\boldsymbol{\xi}_0 : a \in \mathcal{C}^T\}$ (which is equal to the rank of matrix C_0) is called the *rank* of the sequence. Note that since $N_{\boldsymbol{x}}(0) = N_x(0) \oplus \ldots N_x(T-1)$, the rank of a regular T-PC sequence $(x(n))$ is equal to the rank of the corresponding T-variate stationary block sequence $(\boldsymbol{x}(n))$ and is equal to the (matrix) rank the matrix

$$\Sigma := (\boldsymbol{x}(n) - (\boldsymbol{x}(n)|M_{\boldsymbol{x}}(n-1))(\boldsymbol{x}(n) - (\boldsymbol{x}(n)|M_{\boldsymbol{x}}(n-1))^*.$$

If (C_k) is an innovation representation of $(\boldsymbol{x}(n))$ then $\Sigma = C_0C_0^*$.

There is a one-to-one correspondence between MA representations $(c_k(n))$ of a T-PC sequence $(x(n))$ and MA representations (C_k) of its corresponding T-variate stationary block sequence $(\boldsymbol{x}(n))$ given by

$$C_k^{i,j} = c_{kT+i-j}(i) \quad \text{or} \quad c_k(n) = C_{-\text{q}(n-k)}^{\langle n\rangle, \langle n-k\rangle} \ . \tag{15}$$

Recall that $\text{q}(m)$ and $\langle m\rangle$ stand for the quotient and the remainder in division of m by T, respectively. To see (15) note that after substituting $x^j(n) = x(nT+j)$ and $\xi_{n-p}^k = \xi_{(n-p)T+k}$ into (13) we obtain that

$$\sum_{p=-\infty}^{\infty} \sum_{k=0}^{T-1} C_p^{j,k} \xi_{(n-p)T+k} = x(nT+j) = \sum_{r=-\infty}^{\infty} c_r(j)\xi_{nT+j-r},$$

because $c_r(nT+j) = c_r(j)$. Multiplying both sides by $\xi_{(n-p)T+k}$ we obtain that

$$C_p^{j,k} = \sum_{r=-\infty}^{\infty} c_r(j)(\xi_{nT+j-r}, \xi_{(n-p)T+k}) = c_{pT-k+j}(j)$$

On the other hand multiplying both sides by ξ_{nT+j-r} and noting that $(n-p)T+k = nT+j-r$ iff $-p = \mathrm{q}(j-r)$ and $k = \langle j-r \rangle$, we obtain that $c_r(j) = C^{j,\langle j-r\rangle}_{-\mathrm{q}(j-r)}$.

We will refer to the two MA representations described in (15) as *corresponding* to each other. The relation (15) is visualized in a matrix form below:

$$\left[\begin{array}{c|cccc|cccc|c} \cdots & c_T(0) & c_{T-1}(0) & \cdots & c_1(0) & c_0(0) & c_{-1}(0) & \cdots & c_{-T+1}(0) & \cdots \\ \cdots & c_{T+1}(1) & c_T(1) & \cdots & c_2(1) & c_1(1) & c_0(1) & \cdots & c_{-T+2}(1) & \cdots \\ \cdots & \cdots & \cdots & \mathbf{C_1} & \cdots & \cdots & \cdots & \mathbf{C_0} & \cdots & \cdots \\ \cdots & c_{2T-1}(T-1) & c_{2T-2}(T-1) & \cdots & c_T(T-1) & c_{T-1}(T-1) & c_{T-2}(T-1) & \cdots & c_0(T-1) & \cdots \end{array}\right]$$

If $(c_k(n))$ in (15) is an innovation representation of $(x(n))$ then (C_k) is an innovation representation of $(\boldsymbol{x}(n))$. The converse is not true even if C_0 is lower triangular, a counterexample is given in [10]. However, if additionally to being lower triangular C_0 is invertible (i.e. $(\boldsymbol{x}(n))$ is of full rank), then $(c_k(n))$ is an innovation representation of $(x(n))$.

Lemma 3.2 *Let $(x(n))$ be a regular T-PC sequence and $(\boldsymbol{x}(n))$ be the corresponding T-variate stationary block sequence. Let (C_k) be an innovation representation of $(\boldsymbol{x}(n))$ and $(c_k(n))$ be the corresponding MA representation of $(x(n))$ defined in (15). If C_0 is lower triangular and invertible, then $(c_k(n))$ is an innovation representation of $(x(n))$.*

Proof. Recall that $x^k(0) = x(k)$, $\boldsymbol{\xi}^k_0 = \xi_k$, $k = 0,\dots,T-1$, $M_{\boldsymbol{x}}(0) = M_x(T-1)$, and $M_{\boldsymbol{x}}(-1) = M_x(-1)$. Also note that since C_0 is invertible and C_0 is lower triangular, then all $C_0^{k.k}$ are different than zero. By assumption $C_0\boldsymbol{\xi}_0$ is equal to the orthogonal projection of $\boldsymbol{x}(0)$ onto $N_{\boldsymbol{x}}(0) = M_{\boldsymbol{x}}(0) \ominus M_{\boldsymbol{x}}(-1)$. Since C_0 is lower triangular, we have that $x(k) - (x(k)|M_{\boldsymbol{x}}(-1)) = C_0^{k,0}\xi_0 + \dots C_0^{k,k}\xi_k$, $k = 0,\dots,T-1$. Suppose first that $k=0$. Then from $M_{\boldsymbol{x}}(-1) = M_x(-1)$ it follows that $x(0) - (x(0)|M_x(-1)) = x(0) - (x(0)|M_{\boldsymbol{x}}(-1)) = C_0^{0,0}\xi_0$. Since $C_0^{0,0} \neq 1$, we conclude that $N_x(0) = M_x(0) - M_x(-1) = \overline{\mathrm{sp}}\{\xi_0\}$. Assume that we have already shown that $x(j) = (x(j)|M_x(j-1)) = C_0^{j,j}\xi_j$ for $j = 0,\dots,k-1$, $0 < k < T-1$. Then $x(k) - x(k)|M_{\boldsymbol{x}}(-1) = C_0^{k,0}\xi_0 + \dots + C_0^{k-1,k-1}\xi_{k-1} + C_0^{k,k}\xi_k$, and hence

$$x(k) - C_0^{k,k}\xi_k = (x(k)|M_{\boldsymbol{x}}(-1)) + C_0^{k,0}\xi_0 + \dots C_0^{k-1,k-1}\xi_{k-1} = (x(k)|M_x(k-1)),$$

i.e. $x(k) - (x(k)|M_x(k-1)) = C_0^{k,k}\xi_k$. Note that the proof will not work if $C_0^{0,0} = 0$ but both $C_0^{1,0}$ and $C_0^{1,1}$ are different than zero. □

There is an obvious one-to-one correspondence between MA representations (C_n) of a T-variate stationary sequence $(\boldsymbol{x}(n))$ and square factors H of a density F of $(\boldsymbol{x}(n))$ given by

$$\boldsymbol{x}(n) = \sum_{k=-\infty}^{\infty} C_k \boldsymbol{\xi}_{n-k} \longleftrightarrow H\left(e^{it}\right) = (1/\sqrt{2\pi}) \sum_{k=-\infty}^{\infty} C_k e^{ikt}. \qquad (16)$$

This relation is easily seen if we choose $\boldsymbol{\xi}_n^j = (1/\sqrt{2\pi})e^{-int}e_j$. Square factors of F that correspond to innovation representations of $(\boldsymbol{x}(n))$ are called *maximal factors* (Rozanov [11]). The corresponding notion for PC sequences was introduced in [10] under the name an *i-factor*. Maximal factors can be characterized in terms of subspaces of L^2_+ spanned by their coordinates (see for example [11] for T-variate stationary case, and [10] for the PC case). Finding a maximal factor is equivalent to finding coefficients of an innovation representation of a sequence. The latter constitutes a solution to so called *prediction problem*. So far the prediction problem has been solved only for full rank stationary T-variate sequences having rational densities ([11, 4, 5]). Theorem 3.1 and Lemma 3.2 allow us to obtain the solution for full rank PC sequences with rational densities. A procedure is following: given a full rank T-PC sequence $(x(n))$ with rational density g compute the rational density F of the corresponding T-variate block sequence $(\boldsymbol{x}(n))$ using formula (9), find a maximal rational square factor G of F using a construction given in [5] or [11], multiply G by a proper unitary matrix Q so that the "zero" term of Fourier series of $H(e^{it}) = G(e^{it})Q$ is a lower triangular matrix, and then use Lemma 3.2 to recover innovation coefficients of $(x(n))$.

4 PARMA Systems

A VARMA system of dimension $T \geq 1$ is a system of vector difference equations (VDE)

$$\sum_{j=0}^{l} A_j \boldsymbol{x}(n-j) = \sum_{j=0}^{r} B_j \boldsymbol{\xi}_{n-j}, \qquad n \in \mathcal{Z}, \tag{17}$$

where A_j's, and B_j's are complex $T \times T$ matrices, A_0 is invertible, A_l, B_0, B_r are nonzero, and $\boldsymbol{\xi}_n$, $n \in \mathcal{Z}$, is a given T-variate orthonormal sequence in some Hilbert space $\mathcal{H}$. A *proper stationary solution* to a VARMA system (17) is a T-variate stationary sequence $(\boldsymbol{x}(n))$ which satisfies the system and such that $x^j(n) \in M_{\boldsymbol{\xi}}$ and $Wx^j(n) = x^j(n+1)$, $n \in \mathcal{Z}$, $j = 0, \ldots, T-1$, where W denotes the shift of $(\boldsymbol{\xi}_n)$. In many publications and books the last requirement is replaced by some additional assumptions about the coefficients of the system or by a requirement that the solution has an MA representation (see for example in [5]). Without any additional assumptions the system (17) may have multiple or not regular stationary solutions ([5], p. 13). A PARMA system is an infinite system of difference equations

$$\sum_{j=0}^{l} a_j(n) x(n-j) = \sum_{j=0}^{r} b_j(n) \xi_{n-j}, \qquad n \in \mathcal{Z}, \tag{18}$$

where $l, r \geq 0$, $a_j(n)$ and $b_j(n)$ are T-periodic (in n) sequences of complex numbers, $a_0(n) = 1$ for every $n \in Z$, and none of the sequences $(b_0(n))$, $(a_l(n))$, and $(b_r(n))$ is identically zero. Let U be the shift of (ξ_n), that is $U\xi_n = \xi_{n+1}$, $n \in Z$. We will be interested only in T-PC solutions $(x(n))$ to the system whose T-shift coincides with U^T, that is such that $x(n) \in M_\xi$ and $U^T x(n) = x(n+T)$, $n \in \mathcal{Z}$. We label them *proper PC solutions*. The assumption $U^T x(n) = x(n+T)$ allows us to avoid having multiple or not regular PC solutions. If we arrange the coefficients $a_j(n)$ into an $T \times (L+1)T$ matrix $[A_L \ \dots \ A_1 \ A_0]$ where L is such that the matrix contains all nonzero $a_j(n)$'s as shown below

$$\left[\begin{array}{c|cccc|cccc} \cdots & a_T(0) & \cdots & a_2(0) & a_1(0) & a_0(0) & 0 & \cdots & 0 \\ \cdots & a_{T+1}(1) & \cdots & a_3(1) & a_2(1) & a_1(1) & a_0(1) & \cdots & 0 \\ \cdots & \cdots & \mathbf{A(1)} & \cdots & \cdots & \cdots & \mathbf{A(0)} & \cdots & \cdots \\ \cdots\cdots & a_{2T-1}(T-1) & \cdots & a_{T+1}(T-1) & a_T(T-1) & a_{T-1}(T-1) & a_{T-1}(T-1) & \cdots & a_0(T-1) \end{array}\right]$$

and do the same for the $b_j(n)$'s creating $T \times (R+1)T$ matrix $[B_R \ \dots \ B_0]$, then, using matrices A_j and B_j, the system (18) can be written as a VARMA system

$$\sum_{j=0}^{L} A_j \boldsymbol{x}(n-j) = \sum_{j=0}^{R} B_j \xi_{n-j}, \qquad n \in \mathcal{Z}, \tag{19}$$

where A_0 and B_0 are lower triangular, and $(\boldsymbol{x}(n))$ and $(\boldsymbol{\xi}_n)$ are T-variate block sequences corresponding to $(x(n))$ and (ξ_n) respectively, that is $\boldsymbol{x}(n) = [x^k(n)]$ with $x^k(n) = x(nT+k)$ and $\boldsymbol{\xi}_n = [\xi_n^k]$ with $\xi_n^k = \xi_{nT+k}$. The system (17), and hence the system (18), is completely described by a pair of polynomial matrices $(A(z),\, B(z))$ defined as

$$A(z) = \sum_{k=0}^{L} A(k) z^k \quad \text{and} \quad B(z) = \sum_{k=0}^{R} B(k) z^k. \tag{20}$$

In the sequel we will identify both (17) and (18) by giving the corresponding pair $(A(z),\, B(z))$. The only difference between PARMA and general VARMA systems is that in PARMA systems A_0 and B_0 are lower triangular, and $A_0^{i,i} = 1$ for each $i = 0, \dots, T-1$, so PARMA systems form a subset of the family of VARMA systems. Note that since $a(z) = \det(A(z))$ is a polynomial and by assumption $a(0) = \det(A_0) \neq 0$, $a(z) \neq 0$ for all $z \in \mathcal{C}$ except finitely many points, and consequently $A(z)^{-1}$ exists for all $z \in \mathcal{C}$ except finitely many points.

Theorem 4.1 *A PARMA system $(A(z), B(z))$ has a proper PC solution iff the rational matrix function $A(z)^{-1}B(z)$ has no poles of modulus 1. If a proper PC solution $(x(n))$ exists, then it is unique, absolutely continuous, and its density $g = (g^0, \dots, g^{T-1})$ is given by*

$$g^j(e^{it}) = h(e^{it}) h(e^{i(t+2\pi j/T)})^*, \qquad a.e. \tag{21}$$

where $h(z) = (1, z, \ldots, z^{T-1})H(z^T)$ *and* $H(z) = (1/\sqrt{2\pi})A(z)^{-1}B(z)$, $z = e^{it}$.

The theorem remains true when in the above formulation we replace *PARMA* by *VARMA*, *PC* by *T-variate stationary*, $(x(n))$ by $(\boldsymbol{x}(n))$, g by F, and the formula (21) by $F(e^{it}) = H(e^{it})H(e^{it})^*$, where $H(z) = (1/\sqrt{2\pi})A(z)^{-1}B(z)$.

Proof. From the preceding discussion it follows that a PARMA system (18) has a proper T-PC solution $(x(n))$ iff the associated VARMA system (19) has a proper T-variate stationary solutions $(\boldsymbol{x}(n))$, and if it does then $(\boldsymbol{x}(n))$ is the T-variate stationary block sequence corresponding to $(x(n))$. Because $(\boldsymbol{x}(n))$ is to be proper, it is enough to find $\boldsymbol{x}(0)$ since then $\boldsymbol{x}(n) = W^n\boldsymbol{x}(0)$. Therefore it is enough to solve (19) just for $n = 0$, that is solve the equation $\sum_{j=0}^{L} A_j\boldsymbol{x}(-j) = \sum_{j=0}^{R} B_j\boldsymbol{\xi}_{-j}$. Substituting $\boldsymbol{x}(-j) = W^{-j}\boldsymbol{x}(0)$ and $\boldsymbol{\xi}_j = W^{-j}\boldsymbol{\xi}_0$, we can write the above equation as

$$\sum_{j=0}^{L} A_j W^{-j}\boldsymbol{x}(0) = \sum_{j=0}^{R} B_j W^{-j}\boldsymbol{\xi}_0, \tag{22}$$

or symbolically, using polynomials $A(z)$ and $B(z)$ defined in (20), as $A(W^{-1})\boldsymbol{x}(0) = B(W^{-1})\boldsymbol{\xi}(0)$. To solve (22) let us consider a T-variate orthonormal system $\boldsymbol{\zeta}_n = [\zeta_n^k]$ in $L^2(\mathcal{C}^T)$ defined as $\zeta_n^k = (1/\sqrt{2\pi})e^{-int}e_k$, where (e_k) is the standard basis in $\mathcal{C}^T$, and define the unitary operator $\Phi : M_{\boldsymbol{\xi}} \to L^2(\mathcal{C}^T)$ by $\Phi(\xi_n^k) = \zeta_n^k$, $n \in \mathcal{Z}$, $k = 0, \ldots, T-1$. Note that the shift of $(\boldsymbol{\zeta}_n)$ is the operator of multiplication by e^{-it} and that $\boldsymbol{\zeta}_0 = [\zeta_0^k] = (1/\sqrt{2\pi})I$, where I is the $T \times T$ identity matrix. The mapping Φ transfers the equation (22) into matrix equation

$$A(e^{it})H(e^{it}) = (1/\sqrt{2\pi})B(e^{it}), \tag{23}$$

where $H(e^{it})$ is a $T \times T$ matrix function with rows $H^{k\cdot}(e^{it}) = \Phi(x^k(0))$. Summing up, (22) has a solution $\boldsymbol{x}(0) \in M_{\boldsymbol{\xi}}$ iff there is a $T \times T$ matrix function H with rows in $L^2(\mathcal{C}^T)$ that satisfies (23). Since $A(e^{it})^{-1}$ exists a.e., the only candidate for H is $H(e^{it}) = (1/\sqrt{2\pi})A^{-1}(e^{it})B(e^{it})$. Hence (22) has a solution iff all entries of $A^{-1}(e^{it})B(e^{it})$ belong to L^2. Since $(e^{it} - c)^{-1}$ is square integrable always except when $|c| = 1$, we conclude the system (19) has a proper T-variate stationary solution iff $A^{-1}(z)B(z)$ has no poles of modulus 1, assuming as always that all entries of the rational matrix $A^{-1}(z)B(z)$, are written in the simplest forms. If this condition is satisfied then the solution $(\boldsymbol{x}(n))$ to (19) is given by $x^k(n) = \Phi^{-1}(e^{-int}H^{k\cdot}(e^{it}))$, where $H(e^{it}) = (1/\sqrt{2\pi})A^{-1}(e^{it})B(e^{it})$. The uniqueness follows from the fact that, because of a.e. invertibility of $A(e^{it})$, $H(e^{it})$ defined above is the only matrix function satisfying (23). The covariance of $(\boldsymbol{x}(n))$ is

$$K_{\boldsymbol{x}}^{j,k} = (\Phi^{-1}(e^{-int}H^{j\cdot}(e^{it})), \Phi^{-1}(H^{k\cdot}(e^{it}))) = \int_0^{2\pi} e^{-int}H^{j\cdot}(e^{it})H^{k\cdot}(e^{it})^*dt,$$

and hence $(\boldsymbol{x}(n))$ is a.c. and its density is $F(e^{it}) = H(e^{it})H(e^{it})^*$. A proper T-PC solution the original PARMA system $(A(z), B(z))$ is therefore the T-PC sequence that corresponds to $(\boldsymbol{x}(n))$, that is $x(n) = x^{\langle n \rangle}(\mathrm{q}(n))$, $n \in \mathcal{Z}$. From Theorem 3.1 we conclude that $(x(n))$ is a.c. and that $h(z) = (1, z, \dots, z^{T-1})H(z^T)$, $z = e^{it}$, is a square factor of the density g of $(x(n))$, which proves the formula (21). □

The first part (existence) of Theorem 4.1 is well known but difficult to attribute to a particular name. In fact more is known. From a description of all solutions to (19) given for example in [4], p. 11, it follows that if additionally $\det A(z) \neq 0$ for all $|z| = 1$, then the system has only one bounded solution which therefore must be a proper T-PC solution. Regarding computing a density of a proper T-PC solution, two other different procedures were given in [12] and [13]. Our formula seems similar to [13].

The formula (21) shows that a density of a proper PC solution to any PARMA system is a rational function. From the next theorem it follows that the opposite is also true.

Theorem 4.2 *Let $(x(n))$ be a T-PC sequence. Then the following conditions are equivalent:*

1. *$(x(n))$ is a proper T-PC solution to some PARMA system.*
2. *$(x(n))$ has a rational density.*
3. *there exists a PARMA system $(A(z), B(z))$ such that:*
   a. *polynomial matrices $A(z)$ and $B(z)$ are left co-prime,*
   b. *$A(z)$ has no zeros in an open disk $D_{<r}$ of a radius $r > 1$, and $B(z)$ has no zeros in the open disk $D_{<1}$ of radius 1,*
   c. *$(x(n))$ is the only T-PC solution to the system $(A(z), B(z))$.*

The theorem remains valid for T-variate stationary sequences, that is when in the above formulation we replace *T-PC* by *T-variate stationary*, $(x(n))$ by $(\boldsymbol{x}(n))$, and the word *PARMA* by *VARMA*.

Proof. Let $(\boldsymbol{x}(n))$ denote the T-variate stationary block sequence corresponding to $(x(n))$ and F be its spectrum.

(1. ⇒ 2.) From the proof of Theorem 4.1 it follows that if $(x(n))$ is a proper PC solution to some PARMA system (18), then the corresponding block sequence $(\boldsymbol{x}(n))$ satisfies the associated VARMA system (19) and that $(\boldsymbol{x}(n))$ is unitary equivalent to the T-variate stationary sequence $(H(n))$ in $\mathcal{H} = L^2(\mathcal{C}^T)$ defined by $H(n) = e^{-int}H(e^{it})$, where $H(e^{it}) = (1/\sqrt{2\pi})A^{-1}(e^{it})B(e^{it})$, $n \in \mathcal{Z}$. Note that $H(z) = (1/\sqrt{2\pi})A(z)^{-1}B(z)$ is rational. The correlation function of $(H(n))$ is $K_H(n) = \int_0^{2\pi} e^{-int}H(e^{it})H(e^{it})^* dt$. Hence $(\boldsymbol{x}(n))$ is a.c. and its spectral density is $F(e^{it}) = H(e^{it})H(e^{it})^*$ a.e., that is $H(z)$ is a rational square factor of $F(z)$. From Lemma 3.1 part 1. we conclude that the spectrum of $(x(n))$ is a.c. and $h(z) = \sum_{k=0}^{T-1} z^k H^{k\cdot}(z^T)$ is a square factor of the density g of $(x(n))$. Since $H(z)$ is a rational matrix, $h(z)$ is also, and consequently all $g^j(z) = h(z)h(d^j z)^*$, $d = e^{2\pi/T}$, are rational.

(2. $\Rightarrow$ 3.) Suppose now that $(x(n))$ is a.c. and its density g is rational. Then the corresponding T-variate stationary block sequence $(\boldsymbol{x}(n)$ is a.c. and, by Theorem 3.1, its density $F(z)$ is rational. The Rozanov's Theorem 2.1 implies that $F(z)$ has a factorization $F(z) = G(z)G(z)^*$ where G is rational, analytic in the the open disk $D_{<1}$ and $G(z)$ has no zeros in $D_{<1}$. Since entries of $F(e^{it})$ are integrable, the function $G(z)$ has no poles of modulus 1. Therefore $G(z)$ is analytic in some open disk $D_{<r}$, $r > 1$. The matrix $G(0)$ is different than 0 since otherwise $z = 0$ would be a zero of $G(z)$. If we factor the least common multiple, say $q(z)$, of all denominators of entries of $G(z)$, then we can write $G(z) = P(z)/q(z)$ where $P(z)$ is $T \times T$ polynomial matrix with no zeros in $D_{<1}$ and $q(z)$ is a scalar polynomial with all zeros outside $D_{\leq 1}$, and hence outside of a certain disk $D_{<r}$ of radius $r > 0$. In particular $z = 0$ is not a zero of $q(z)$, so $q(0) \neq 0$. Let $A_0(z) = (q(z)/q(0))I$ and $B_0(z) = (1/q(0))P(z)$. Then both are analytic polynomial matrices, the constant term of $A_0(z)$ is the identity matrix I, $A_0(z)^{-1} = (q(0)/q(z))I$ and $A_0(z)^{-1}B_0(z) = (q(0)/q(z))(1/q(0))P(z) = G(z)$. Factoring out the left greatest common divisor $L(z)$ of $A_0(z)$ and $B_0(z)$ we obtain that $A_0(z) = L(z)C(z)$ and $B_0(z) = L(z)D(z)$. Since $\det A_0(z) \neq 0$ on some $D_{<r}$, $r > 1$, both $\det L(z) \neq 0$ and $\det C(z) \neq 0$ on $D_{<r}$, $r > 1$. Hence $L(z)^{-1}$ exists and we conclude that $C(z)^{-1}D(z) = A_0(z)^{-1}B_0(z) = G(z)$. Let S be an invertible matrix such that $SC(0)$ is lower triangular, and let Q be a unitary matrix such that $SD(0)Q$ is lower triangular. Define $A(z) = SC(z)$ and $B(z) = (\sqrt{2\pi})SD(z)Q$. Since $A(0)$ and $B(0)$ are lower triangular, $(A(z), B(z))$ is a PARMA system. Polynomial matrices $A(z)$ and $B(z)$ satisfy conditions a. and b. of 3., and $(1/\sqrt{2\pi})A(z)^{-1}B(z) = C(z)^{-1}D(z)Q$, $z \in \mathcal{C}$. Define $H(z) = (\sqrt{2\pi})A(z)^{-1}B(z)$. Then $H(z) = G(z)Q$, and hence $H(z)H(z)^* = G(z)G(z)^* = F(z)$. Hence $(\boldsymbol{x}(n))$ is a proper T-variate stationary solution to the system $(A(z), B(z))$. Since $(\boldsymbol{x}(n))$ is a block sequence corresponding to $(x(n))$, the sequence $(x(n))$ is a proper T-PC solution to the system $(A(z), B(z))$. Since $\det A(z) = (\det S)(\det A_0(z))/(\det L(z)) \neq 0$ on the circle $|z| = 1$, the system $(A(z), B(z))$ has only one bounded solution, and hence only one T-PC solution.

(3. $\Rightarrow$ 1.) In view of Theorem 4.1, the condition b. itself implies that that system $(A(z), B(z))$ has a proper T-PC solution. Moreover it implies that $\det A(z) \neq 0$ on the circle $|z| = 1$, and hence that the system has only one bounded solution. Therefore the sequence $(x(n))$ must be a proper T-PC solution to $(A(z), B(z))$. □

An immediate consequence of the above theorem is regularity of every T-PC sequence with rational density.

Corollary 4.1 *Every T-PC (or T-variate stationary) sequence with rational density is regular. Consequently, if a PARMA (or VARMA) system has a proper PC (or proper T-variate stationary) solution, then this solution is regular.*

Proof. Suppose that $(x(n))$ is a T-PC sequence with rational density g, and let $(\boldsymbol{x}(n))$ be the corresponding T-variate stationary block sequence. From Theorem 4.2, part 3., we conclude that there is a PARMA system $(A(z), B(z))$ such that $(\boldsymbol{x}(n))$ is a proper T-variate stationary solution of the system $(A(z), B(z))$ and the rational function $H(z) = (\sqrt{2\pi})A(z)^{-1}B(z)$ is analytic in some open disk $D_{<r}$ of radius $r > 1$. Moreover $H(e^{it})$ is a square factor of the density of $(\boldsymbol{x}(n))$. Being analytic, H has an expansion $H(z) = \sum_{k=0}^{\infty} C_k z^k$, $|z| < r$, $r > 1$. Therefore the corresponding MA representation of $(\boldsymbol{x}(n))$ is one-sided and hence $(\boldsymbol{x}(n))$, and also $(x(n))$, are regular. □

5 PARMA Models

We failed to find a precise definition of a PARMA (or VARMA) model, so we have assumed the following.

Definition 5.1 *A PARMA (or VARMA) system* $(A(z), B(z))$ *is called a* PARMA (or VARMA) model *if the polynomial matrices* $A(z)$ *and* $B(z)$ *satisfy the conditions a. and b. of part 3 of Theorem 4.2*

Theorems 4.1 and 4.2 show that every PARMA (or VARMA) model has a unique T-PC (or T-variate stationary) solution and this solution has rational density, and vice verse, that every T-PC (or T-variate stationary) sequence with rational density admits a PARMA (respectively VARMA) model. The sole reason that we added the condition a. saying that $A(z)$ and $B(z)$ are left co-prime is to reduce the set of allowed models. Theorem 4.2 remains valid if we remove this condition from part 3.

Why do we like to have a model? Because then the sequence $(b_0(n)\xi_n)$ in (18) or $(B_0\boldsymbol{\xi}_n)$ in (17) are innovation sequences for $(x(n))$ and $(\boldsymbol{x}(n))$, that is $\overline{\mathrm{sp}}\{b_0(n)\xi_n\} = N_x(n)$ and $\overline{\mathrm{sp}}\{aB_0\boldsymbol{\xi}_n : a \in \mathcal{C}^T\} = N_{\boldsymbol{x}}(n)$, respectively. At least we believe so, since so far we can prove it only under some additional assumptions: a *miniphase assumption* about $B(z)$, or a *full density rank* assumption about the sequence. If $(b_0(n)\xi_n)$ (or $(B_0\boldsymbol{\xi}_n)$) are innovations then the one step prediction of an element $x(n)$ (or $\boldsymbol{x}(n)$) based on the immediate past is obtained by simply setting $\xi_n = 0$ in the model equation (18) (or $\boldsymbol{\xi}_n = 0$ in (17), respectively).

The *miniphase assumption* is the assumption that $\det B(z)$ is not identically zero ([5], page 25). In the case of VARMA systems, the miniphase assumption implies that a proper T-variate stationary solution $(\boldsymbol{x}(n))$ of the system is a full rank sequence and, hence, its density $F(e^{it})$ is a.e. invertible. It is easy to see that inverse also holds true, that is if a T-variate stationary sequence $(\boldsymbol{x}(n))$ has a rational density F with $\det F(e^{it}) \neq 0$ a.e., and if $(A(z), B(z))$ is a VARMA model for $(\boldsymbol{x}(n))$, then $B(z)$ must satisfy the miniphase assumption and $(\boldsymbol{x}(n))$ is full rank. Below we discuss a consequence of the *miniphase assumption* for PARMA models. Note that until this moment we have not assumed anything about the rank of a sequence or the matrix rank of a polynomial matrix $B(z)$.

Definition 5.2 *For any a.c. T-PC sequence $(x(n))$ with density g let $G(e^{it})$ be the $T\times T$ matrix function defined by $G^{j,k}(e^{it}) = g^{\langle k-j\rangle}(e^{i(t+2\pi j/T)})$, $t\in[0,2\pi)$, $k,j=0,\dots,T-1$. If $\det G(e^{it})\neq 0$ a.e. then we say that $(x(n))$ is is of a* full density rank.

Note that if h is a square factor of g, then $G(z) = H_d(z)H_d(z)^*$, where $z=e^{it}$ and $H_d(z)$ is the matrix whose k-th row is equal $h(zd^k)$ as in (8). Also recall that the *rank r of a T-PC sequence* $(x(n))$ is the number of nonzero elements in a sequence $x(k)-(x(k)|M_x(k-1))$, $k=0,\dots,T-1$. A T-PC sequence $(x(n))$ is said to be of *full rank* if $r=T$.

Theorem 5.1 *Suppose that a T-PC sequence $(x(n))$ has a rational density g and that $(A(z),B(z))$ is a PARMA model (18) for $(x(n))$. Assume additionally that* $\det B(z)$ *is not identically zero (a miniphase assumption). Then $(x(n))$ has full density rank and the sequence (ξ_n) in (18) is an innovation sequence for $(x(n))$. Consequently $x(n)$ is of full rank.*

Proof. Let $(\boldsymbol{x}(n))$ be a T-stationary block sequence corresponding to $(x(n))$ and F be its density. Let $(A(z),B(z))$ be a PARMA model for $(x(n))$. Then $H(z) = (1/\sqrt{2\pi})A(z)^{-1}B(z)$ is a square factor of F. Since $A(z)$ has no zeros in an open disk $D_{<r}$, $r>1$, $\det A(z)\neq 0$ for all $|z|<r$. By assumption $b(z)=\det B(z)$ is not identically zero and hence, because $b(z)$ is a polynomial, $b(z)\neq 0$ everywhere except finitely many z's. Since $B(z)$ has no zeros in the open disk $D_{<1}$, $\det B(z)\neq 0$ everywhere on $D_{<1}$. Summing up, $H(z)$ is analytic on $D_{<1}$ and $\det H(z) = (1/\sqrt{2\pi})(\det B(z))/(\det A(z))\neq 0$ on $D_{<1}$. From [11], page 76, we conclude that $H(z)$ is a maximal factor of F, which means that if we write $H(z)$ as a power series $H(z) = (1/\sqrt{2\pi})\sum_{k=0}^{\infty} C_k z^k$, $|z|<r$, then

$$\boldsymbol{x}(n) = \sum_{k=0}^{\infty} C_k \boldsymbol{\xi}_{n-k}, \tag{24}$$

is an innovation representation of $(\boldsymbol{x}(n))$. Note that, because $B(z)=\sqrt{2\pi}A(z)H(z)$ and all three are analytic in $D_{<r}$, $B(0)=\sqrt{2\pi}A(0)C_0$. Also $A(0)$ is invertible because otherwise $z=0$ would be a zero of $A(z)$. This and the fact that by definition both $B(0)$ and $A(0)$ are lower triangular imply that $C_0 = (1/\sqrt{2\pi})A(0)^{-1}B(0)$ is lower triangular and invertible. From Lemma 3.2, we conclude that the MA representation of $(x(n))$ generated by (24)

$$x(n) = \sum_{k=0}^{\infty} c_k(n)\xi_{n-k}, \qquad c_k(n) = C_{-\mathrm{q}(n-k)}^{\langle n\rangle,\langle n-k\rangle},$$

is an innovation representation of $(x(n))$, that is $(c_0(n)\xi_n)$ is an innovation for $(x(n))$. Since by (15), $c_0(mT+j)=c_0(j)=C_0^{j,j}\neq 0$ for all $j=0,\dots,T-1$, and $m \in \mathcal{Z}$, (ξ_n) is also an innovation for $(x(n))$. This implies that the rank of $(x(n))$ is T. Moreover, since $\det H(e^{it}) \neq 0$ a.e., the matrix function $H_d(e^{it})$ appearing in (8) is also invertible a.e., and hence $G(e^{it}) = H_d(e^{it})H_d(e^{it})^*$ is also invertible a.e., that is $(x(n))$ is a full density rank. □

The next theorem is sort of inverse and shows that if $(x(n))$ is a T-PC sequence with full rank rational density, then any PARMA model for $(x(n))$ must satisfy the miniphase assumption.

Theorem 5.2 *Suppose that $(x(n))$ is T-PC with a rational density g of full density rank. Let $(A(z), B(z))$ be a PARMA model for $(x(n))$. Then $\det B(z)$ is not identically zero, and (ξ_n) in (18) is an innovation sequence for $(x(n))$. Moreover $(x(n))$ is of full rank.*

Proof. Full density rank means that $\det G(e^{it}) = |\det H_d(e^{it}))|^2 \neq 0$ a.e. From the relation (8) it follows that also $|\det H(e^{itT})| = (1/T)|\det H_d(e^{it})| \neq 0$ a.e. Hence $H(z) = (1/\sqrt{2\pi})A(z)^{-1}B(z)$ is invertible except finitely many z's. Consequently $(A(z), B(z))$ satisfies a miniphase assumption and the rest follows from Theorem 5.1. □

We do not know whether Theorems 5.1 or 5.2 are true for sequences of not full density rank r or without a miniphase assumption.

An immediate consequence of the two theorems is that if a T-PC sequence $(x(n))$ has a rational density, then $(x(n))$ is of full rank iff it has a full density rank (we already know that every T-PC sequence $(x(n))$ with a rational density is regular).

In some publications a VARMA (as well as a PARMA) model is defined as a system $(A(z), B(z))$ which additionally to the condition b. of part 3 of Theorem 4.2, satisfies the so called *invertibility assumption* which says that $\det B(z) \neq 0$ for all $|z| \leq 1$ (e.g. [1], p. 409). The invertibility assumption (together with b.) immediately gives that a T-variate stationary solution $(\boldsymbol{x}(n))$ to the system is proper, full rank, and $(\boldsymbol{\xi}_n)$ an innovation sequence for $(\boldsymbol{x}(n))$. The invertibility assumption is much stronger that a miniphase assumption which allows $\det B(z) = 0$ for finitely many z's of modulus 1, and significantly reduces a number of T-variate stationary sequences with rational densities that can be modeled in that way. For example a pair $A(z) = 1$ and $B(z) = 1 - z$ is a model for a univariate stationary sequence $x(n) = \xi_n - \xi_{n-1}$, $n \in \mathcal{Z}$, but $B(z)$ does not satisfy an invertibility assumption.

Note is that a construction of the rational matrix $G(z)$ in Rozanov's Theorem 2.1 is explicit, as well as all constructions presented in this paper are explicit. Therefore given a PARMA system we can explicitly compute a density of its T-PC solution via Theorem 4.1, while a construction given in the proof of Theorem 4.2 allows us to construct a PARMA model for any T-PC sequence with rational density given its density g. A needed procedure for finding a left greatest common divisor of polynomial matrices $A_0(z)$ and $B_0(z)$ can be found in [5], p. 38, or in [7], sec 1.15.2. Moreover, if g is if full density rank then the proofs of Theorems 5.1 and 5.2 show us how to find coefficients of an innovation representation of $(x(n))$.

Remark 2 For each T-PC sequence (or T-variate stationary sequence) with rational density one can find many different PARMA (respectively VARMA) models, even if we assume the miniphase assumption. This is because there are many pairs $(A_k(z), B_k(z))$ satisfying the conditions of Definition 5.1 and such that that $H_k(z)H_k(z)^*$, where $H_k(z) = (1/\sqrt{2\pi})A_k(z)^{-1}B_k(z)$, coincide a.e. on the unit circle $|z| = 1$. Although each model serves its prediction purpose, this lack of uniqueness is not convenient in model identification for it would be nice to have a unique, preferably minimal, set of model coefficient to be estimated. Because of this Hannan introduced the notion of *identifiability*. The idea is to impose some constraints on allowable models such that each sequence with rational density would have one and only one model satisfying these constraints. Identifiability problem for VARMA models is discussed in [5]. We do not address this question in our paper. ☐

Appendix: Proof of Lemma 3.1

Let $(x(n))$ be T-PC, $(\boldsymbol{x}(n) = [x^k(n)])$ be the corresponding T-variate stationary block sequence, and γ and F be respectively their spectral measures.

1. First we prove that: *if* $(\boldsymbol{x}(n))$ *is a.c.,* $H\left(e^{it}\right)$ *is a square factor of its density* $F\left(e^{it}\right)$ *of* $(\boldsymbol{x}(n))$*,* $H^{k\cdot}$ *denotes the k-th row of H, and we define* $h(e^{it}) = \sum_{k=0}^{T-1} e^{ikt}H^{k\cdot}(e^{iTt})$*, then* $(x(n))$ *is a.c. and h is a square factor of a density* $g = (g^0, \ldots, g^{T-1})$ *of* $(x(n))$. Write $H^{k\cdot}$ as a Fourier series $H^{k\cdot}(e^{it}) = \sum_{n=-\infty}^{\infty} H_n^k e^{int}$. Then

$$h(e^{it}) = \left[\sum_{n=-\infty}^{\infty}\sum_{k=0}^{T-1} H_n^k e^{i(nT+k)t}\right] = \sum_{p=-\infty}^{\infty} h_p e^{ipt}, \quad \text{where } h_p := H_{\mathrm{q}(p)}^{\langle p\rangle}.$$

Denoting $d = e^{2\pi/T}$, we obtain that

$$\begin{aligned} b_j(n) &:= \int_0^{2\pi} e^{-int}h(e^{it})h(e^{i(t+2\pi j/T)})^* dt \\ &= \sum_{p=-\infty}^{\infty}\sum_{q=-\infty}^{\infty} h_p h_q^* d^{-jq}\int_0^{2\pi} e^{i(p-q-n)t}dt = 2\pi\sum_{q=-\infty}^{\infty} h_{q+n}h_q^* d^{-jq}. \end{aligned}$$

Because $\sum_{j=0}^{T-1} d^{j(r-q)}$ is nonzero only if $q = mT + r$ for some $m \in \mathcal{Z}$, the inverse discrete Fourier transform of $b_j(n)$ is

$$(1/T)\sum_{j=0}^{T-1} d^{jr} b_j(r) = (2\pi/T)\sum_{q=-\infty}^{\infty} h_{q+n} h_q^* \sum_{j=0}^{T-1} d^{j(r-q)}$$
$$= 2\pi \sum_{m=-\infty}^{\infty} h_{mT+r+n} h_{mT+r}^* = 2\pi \sum_{m=-\infty}^{\infty} H_{m+\mathrm{q}(r+n)}^{\langle r+n\rangle} (H_m^r)^*,$$

$r = 0, \dots, T-1$. Recall that the spectrum of $(x(n))$ is a vector measure $\gamma = (\gamma^0, \dots, \gamma^{T-1})$ such that $a_j(n) = \int_0^{2\pi} e^{-int}\gamma^j(dt)$, and that $R_x(n+r, r) = (1/T)\sum_{j=0}^{T-1} e^{2\pi ijr/T} a_j(r)$. Since for $r = 0, \dots, T-1$,

$$R_x(n+r, r) = K_{\boldsymbol{x}}^{\langle n+r\rangle, r}(\mathrm{q}(n+r)) = \int_0^{2\pi} e^{-i\mathrm{q}(n+r)t} H^{\langle n+r\rangle \cdot}(e^{it}) H^{r\cdot}(e^{it})^* dt$$
$$= \sum_{p=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} H_p^{\langle n+r\rangle} H_m^r \int_0^{2\pi} e^{i(p-m-\mathrm{q}(n+r))t} dt$$
$$= 2\pi \sum_{m=-\infty}^{\infty} H_{m+\mathrm{q}(n+r)}^{\langle n+r\rangle} H_m^r, \tag{25}$$

comparing this with the inverse discrete Fourier transform of $(b_j(n))$, we conclude that $b_j(n) = a_j(n)$, $n \in \mathcal{Z}$, $j = 0, \dots, T-1$, and hence γ is a.c. and the density of γ^j is equal $g^j(e^{it}) = h(e^{it})h(e^{i(t+2\pi j/T)})^*$ a.e.

2. We will show opposite, that is *assuming that* $(x(n))$ *is a.c., from a factor* $h(e^{it})$ *of a density* g *of* $(x(n))$ *we will construct a square factor* H *of the density of* $F(e^{it})$ *of* $(\boldsymbol{x}(n))$*, showing at the same time that* $(\boldsymbol{x}(n))$ *is a.c.*. Write $h(e^{it}) = \sum_{p=-\infty}^{\infty} h_p e^{ipt}$, so that $e^{-ikt}h(e^{it}) = \sum_{q=-\infty}^{\infty} h_{q+k} e^{iqt}$. We want to construct a function whose Fourier coefficients are equal h_{q+k} if $q = mT$, and other are zero. It is easy to see that the function $f_k(e^{iu})) = \frac{1}{T}\sum_{j=0}^{T-1} e^{-ik(u+2\pi j/T)} h(e^{i(u+2\pi j/T)})$ has this property. Indeed, denoting as previously $d = e^{2\pi i/T}$ we obtain that

$$f_k(e^{it}) = \sum_{p=-\infty}^{\infty}\left(\frac{1}{T}\sum_{j=0}^{T-1} d^{(p-k)j}\right) h_p e^{i(p-k)t} = \sum_{m=-\infty}^{\infty} h_{mT+k} e^{imTt}.$$

Clearly each function f_k is a function of tT, hence there is $h_k(e^{it})$ such that $h_k(e^{iTt}) = f_k(e^{it})$. The Fourier series of h_k is $h_k(e^{it}) = \sum_{m=-\infty}^{\infty} h_{mT+k} e^{imt}$. Let $H(e^{it})$ be the $T \times T$ matrix function which k-th row $H^{k\cdot}$ is equal h_k, i.e.

$$H^{k\cdot}(e^{it}) = \sum_{n=-\infty}^{\infty} H_n^k e^{int}, \quad \text{where} \quad H_n^k = h_{nT+k},$$

and let $(\mathbf{y}(n))$ be a T-variate stationary sequence with the density $H(e^{it})H(e^{it})^*$. Repeating computation (25) we conclude that the covariance of the T-PC sequence $(y(n))$ corresponding to $(\mathbf{y}(n))$ equals

$$R_y(n+r,r) = 2\pi \sum_{m=-\infty}^{\infty} H_{m+\mathrm{q}(n+r)}^{\langle n+r\rangle} (H_m^r)^* = 2\pi \sum_{m=-\infty}^{\infty} h_{Tm+n+r} h_{mT+r}^*.$$

On the other hand, as it was computed in part 1.,

$$R_x(n+r,r) = (1/T) \sum_{j=0}^{T-1} d^{jr} \int_0^{2\pi} e^{-int} h(e^{it}) h(e^{i(t+2\pi j/T)})^* dt$$
$$= 2\pi \sum_{m=-\infty}^{\infty} h_{mT+r+n} h_{mT+r}^*.$$

Comparing this with the previous formula we see that $R_x = R_y$, and hence they have the same spectrum. We therefore conclude that $(\boldsymbol{x}(n))$ is a.c. and its density $F(e^{it}) = H(e^{it}) H(e^{it})^*$ a.e.

References

1. Brockwell, P. J., Davis, R. A.: Time Series: Theory and Methods. Springer Series in Statistics, Springer (1987)
2. Dudek, A., Hurd, H., Wójtowicz, W.: PARMA Methods Based on Fourier Representation of Periodic Coefficients, preprint, (2015)
3. Dunford, N., Schwartz, J. T.: Linear Opperators Part II: Spectral Theory. John Wiley & Sons, Inc., (1988)
4. Hannan, E. J.: Multiple Time Series. John Wiley & Sons, Inc., (1970)
5. Hannan, E. J., Deistler, M.; The Statistical Theory of Linear Systems. Society for Industrial and Applied Mathematics (SIAM) (2012)
6. Hurd, H. L., Miamee, A., Periodically Corerlated Random Sequences; Spectral Theory and Practice. John Wiley & Sons, Inc., (2007)
7. Kaczorek, T.: Polynomial and Rational Matrices. Springer (2007)
8. Makagon, A., and Miamee, A.G.: Spectral Representation of Periodically Correlated Sequences. Probability Math. Stat. **33** no. 1, 175 – 188, (2013)
9. Makagon, A., and Miamee, A.G.: Structure of PC Sequences and the 3rd Prediction Problem. in: Cyclostationarity: Theory and Methods, Eds. F. Chaari, J. Leskow, A. Neapolitano, A. Sanchez-Ramirea. Lecture Notes in Mechanical Engineering. Springer, 53 – 72, (2014)
10. Makagon, A.: Innovation and Factorization of the Density of a Regular PC Sequence. Probability Math. Stat. **35**, no. 1, 15–30 (2015)
11. Rozanov, Yu. A.: Stationary random Processes. Holden-Day Series in Time Series Analysis, Holsden-Day, (1967)
12. Sakai,H: On The Spectral Density Matrix of a Periodic ARMA Process. J. Time Ser. Anal. **12**, 73–82, (1991)
13. Wylomanska, A.: Spectral Measures of PARMA Sequences. J. Time Series Anal. **29** no. 1, 1-13, (2008)