Econometric Theory, **28**, 2012, 1313–1349.
doi:10.1017/S026646661200014X

A STATE SPACE CANONICAL FORM FOR UNIT ROOT PROCESSES

DIETMAR BAUER
Austrian Institute of Technology

MARTIN WAGNER
Institute for Advanced Studies and
Frisch Centre for Economic Research

In this paper we develop a canonical state space representation of autoregressive moving average (ARMA) processes with unit roots with integer integration orders at arbitrary unit root frequencies. The developed representation utilizes a state process with a particularly simple dynamic structure, which in turn renders this representation highly suitable for unit root, cointegration, and polynomial cointegration analysis. We also propose a new definition of polynomial cointegration that overcomes limitations of existing definitions and extends the definition of multicointegration for I(2) processes of Granger and Lee (1989a, *Journal of Applied Econometrics* 4, 145–159). A major purpose of the canonical representation for statistical analysis is the development of parameterizations of the sets of all state space systems of a given system order with specified unit root frequencies and integration orders. This is, e.g., useful for pseudo maximum likelihood estimation. In this respect an advantage of the state space representation, compared to ARMA representations, is that it easily allows one to put in place restrictions on the (co)integration properties. The results of the paper are exemplified for the cases of largest interest in applied work.

1. INTRODUCTION

Linear dynamic models incorporating unit roots and cointegration have become standard tools for modeling economic time series since the early 1980s. This is in particular true for the I(1) case, but also holds for more general settings including seasonal unit roots and seasonal cointegration (with early contributions by Hylleberg, Engle, Granger, and Yoo, 1990; Lee, 1992; Ahn and Reinsel, 1994)

The first author acknowledges support by the Austrian FWF via project P-14438-INF. Part of this work was done while the second author was visiting the Economics Departments of Princeton University and the European University Institute in Florence, whose hospitality is gratefully acknowledged. We thank Manfred Deistler, Søren Johansen, Peter C.B. Phillips, Benedikt M. Pötscher, and Thomas Ribarits for many useful discussions on the topic and for pointing out mistakes in earlier versions of the paper. We also thank seminar participants in Copenhagen, Dortmund, Florence, Konstanz, Princeton, Rochester, and Yale and conference participants in the EMM meeting in Varese and the Frontiers of Time Series Analysis conference in Olbia for comments. Furthermore the very helpful comments of the anonymous referees and the editor in charge, which are gratefully acknowledged, also led to substantial improvements. This work has received partial financial support from the ESF-EMM network. The usual disclaimer applies.
Address correspondence to Martin Wagner, Institute for Advanced Studies, Department of Economics and Finance, Stumpergasse 56, A-1060 Vienna, Austria; e-mail: Martin.Wagner@ihs.ac.at.

© Cambridge University Press 2012

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

or processes with higher integration orders (see, e.g., Stock and Watson, 1993; Johansen, 1997).

The concept of cointegration has been generalized by considering not only static transformations but also dynamic transformations of the form $\sum_{j=0}^{q} \beta_j' y_{t-j}$. The corresponding cointegration concept has been introduced first as multicointegration by Yoo (1986) and Granger and Lee (1989a, 1989b). Multicointegration prevails for an s-dimensional I(2) process $(y_t)_{t\in\mathbb{Z}}$, with $\mathbb{Z}$ denoting the integers, if there exists a linear combination of the two I(1) processes $(\beta' y_t)_{t\in\mathbb{Z}}$ (i.e., $\beta \in \mathbb{R}^s$ reduces the integration order from 2 to 1) and $(y_t - y_{t-1})_{t\in\mathbb{Z}}$, that is stationary. The concept of multicointegration has been generalized (by allowing for higher integration orders and higher-order linear dynamic transformations) to polynomial cointegration by Gregoir and Laroque (1994); see also Gregoir (1999a).

Cointegration analysis is performed both nonparametrically and parametrically. The difference between the two approaches is that in the former case only the cointegrating relationships are estimated, while other features of the data generating process are treated as nuisance parameters (see, e.g., Phillips and Hansen, 1990; Sims, Stock, and Watson, 1990; Stock and Watson, 1993), whereas in the latter a parametric model class characterizing the set of feasible data generating processes is considered. Among the parametric approaches the autoregressive (AR) framework, popularized by the work of Johansen and co-authors, is particularly prominent. The so-called error correction model (ECM) representation of autoregressive processes allows one to characterize AR models that correspond to I(1) processes (see the discussion in Johansen, 1995, Chap. 6), to seasonally integrated processes (see Johansen and Schaumburg, 1999), or to I(2) processes (see Johansen, 1997). With the exception of the I(1) case, the set of restrictions on the AR parameter set to impose certain integration properties becomes rather complicated, as is illustrated in Section 6. This in turn implies that the definition and in particular the analysis of, e.g., pseudo maximum likelihood estimators (based on maximizing the Gaussian likelihood) also become complicated. Based on the representation result developed in Gregoir (1999a, Thm. 3.2, p. 447), Gregoir (1999b) considers estimation for AR processes with unit roots at arbitrary frequencies with integer integration orders. However, due to the complicated nonlinear restrictions involved in the representation theorem, it appears to be highly complicated to define a parameterization.[1]

In this paper we show that the state space framework is highly suitable for unit root and cointegration analysis by developing a specific state space representation for unit root ARMA processes that highlights the unit root and cointegration properties. The results are thus similar in spirit to the results of Gregoir (1999a) in that a representation result for unit root processes, in this paper ARMA processes with unit roots at arbitrary frequencies with integer integration orders, is provided. A stochastic process $(y_t)_{t\in\mathbb{Z}},\ y_t \in \mathbb{R}^s$, is said to possess a state space representation if it is a solution of the equations

$$\begin{aligned} y_t &= Cx_t + \varepsilon_t, \\ x_{t+1} &= Ax_t + B\varepsilon_t, \end{aligned} \tag{1}$$

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

for a white noise process $(\varepsilon_t)_{t\in\mathbb{Z}}$. The unobserved process $(x_t)_{t\in\mathbb{Z}} \in \mathbb{R}^n$ is referred to as state process, and $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times s}$, and $C \in \mathbb{R}^{s\times n}$ are real matrices. As discussed (for the stationary case) in great detail in Hannan and Deistler (1988, Chaps. 1 and 2), henceforth referred to as HD, there are close connections between state space and ARMA models. Despite these connections, only a few early contributions use the state space framework for unit root and cointegration analysis. These include Aoki (1987) and Aoki and Havenner (1989, 1991), all of which deal with the I(1) case only.

So-called structural time series models (STSMs) as discussed in Harvey (1989), which are typically also cast in a state space format, allow one to model multiple time series with unit roots and cointegration. However, in that approach—as the name suggests—a large amount of structure is imposed a priori (like, e.g., trend and cycle decompositions with specific assumptions concerning the error processes driving these components). Our work, however, is geared towards characterizing the classes of state space models with certain unit root and cointegration properties, to allow for the development of estimation and inference procedures in which the dynamic properties, including the unit root and cointegration properties, of the modeled time series are not specified ex ante but are determined from data.

The form of the system equations (1) already indicates why the state space framework is useful for cointegration analysis. The dynamics of the state $(x_t)_{t\in\mathbb{Z}}$ allow for a simple description of its stochastic behavior, i.e., stationarity or integratedness of some order. The observed process $(y_t)_{t\in\mathbb{Z}}$ is simply given by a static transformation of the state (plus the white noise error process). This implies that the unit root and cointegration properties of $(y_t)_{t\in\mathbb{Z}}$ are easily understood once the respective properties of $(x_t)_{t\in\mathbb{Z}}$ are understood. The basis of the state can be freely chosen, as is discussed in Section 3. This degree of freedom is a major advantage of the state space framework compared to the ARMA framework, since it allows the usage of a state process with particularly simple unit root and cointegration properties, which are summarized in the newly defined state space unit root structure (see Definition 2 below). The main result of this paper is a specific *canonical form* (for a definition, see Section 3) based on a canonical state space representation of ARMA unit root processes that highlights the unit root, cointegration, and polynomial cointegration properties very conveniently.

The ARMA framework, on the contrary, does not offer the same level of flexibility with respect to representation as the state space framework. In an ARMA representation the unit root and cointegration properties depend upon polynomial matrices, which makes it comparably hard to impose restrictions on the (co)integration properties. For example, already for I(2) autoregressive models the imposition of properties like variables being integrated of certain orders or restricting the multicointegrating space to a given subspace necessitates complicated reparameterizations. It is exemplified in Section 6.3 how to straightforwardly impose such restrictions in the developed state space representation.

The main purpose of the canonical representation for statistical analysis is its use in the development of parameterizations of model classes (for a formal

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

definition see, e.g., HD, p. 37) that allow one to take into account specified unit root and cointegration properties, such as integration orders of variables or knowledge about the cointegrating spaces. This in turn facilitates the definition and analysis of pseudo maximum likelihood estimators.[2]

With the developed state space representation, both cointegration and polynomial cointegration are characterized by simple orthogonality constraints. In this respect note that we provide a new definition of polynomial cointegration that overcomes some deficiencies of existing definitions and that generalizes the concept of multicointegration. A detailed reconsideration of polynomial cointegration, based on the developed state space representation, is provided elsewhere due to space limitations.

The paper is organized as follows: Section 2 describes the class of processes considered in this paper and gives some relevant definitions. Section 3 discusses state space representations for the considered class of processes, and in Section 4 the canonical form is presented and the state space unit root structure is defined. In Section 5 cointegration and polynomial cointegration are defined and the usefulness of the developed canonical form for cointegration and polynomial cointegration analysis is discussed. In Section 6 the ideas and results of this paper are exemplified with several illustrative examples covering the cases of main interest in economic applications. Finally, Section 7 summarizes and concludes the paper. All proofs are relegated to the Appendix.

The following notation is used: I_m denotes the $m \times m$ dimensional identity matrix and $0^{m\times n}$ the $m \times n$ dimensional null matrix. With x' we denote the complex conjugate transpose of a complex vector and with $\overline{x}$ its complex conjugate. Similar notation is used for matrices. Furthermore, $\lambda_{\max}(A)$ denotes an eigenvalue of maximum modulus of a square matrix A. With $\mathcal{R}(\cdot)$ and $\mathcal{I}(\cdot)$ we denote the real and imaginary parts of complex quantities and $\otimes$ denotes the Kronecker product.

2. UNIT ROOT PROCESSES

This paper deals with the analysis of multivariate ARMA processes $(y_t)_{t\in\mathbb{Z}}$, $y_t \in \mathbb{R}^s$, which are later on considered in the state space framework. A stochastic process $(y_t)_{t\in\mathbb{Z}}$ is referred to as an ARMA process in this paper, if there exist integers $p, q \geq 1$ and matrices $A_j \in \mathbb{R}^{s\times s}$, $j = 1, \ldots, p$ and $B_j \in \mathbb{R}^{s\times s}$, $j = 1, \ldots, q$ and a white noise process $(\varepsilon_t)_{t\in\mathbb{Z}}$, with $\mathbb{E}\varepsilon_t\varepsilon_t' = \Sigma > 0$, such that

$$y_t + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} = \varepsilon_t + B_1\varepsilon_{t-1} + \cdots + B_q\varepsilon_{t-q}, \quad t \in \mathbb{Z}. \quad \textbf{(2)}$$

This definition of an ARMA process is broader than usual, since we do not impose stationarity in the definition of ARMA processes, as is done, e.g., in HD (p. 25).

Defining the matrix polynomials $a(z) := I_s + A_1 z + \cdots + A_p z^p$ and $b(z) := I_s + B_1 z + \cdots + B_q z^q$ where $z \in \mathbb{C}$, the pair $(a(z), b(z))$ is called an ARMA system corresponding to the process $(y_t)_{t\in\mathbb{Z}}$.

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

It is well known (see HD, pp. 9-10) that equation (2) always has solutions and that any solution can be written as the sum of a particular solution and a solution of the homogenous equation

$$y_t + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} = 0, \quad t \in \mathbb{Z}. \tag{3}$$

Furthermore, if $\det a(z) \neq 0$ for $|z| \leq 1$, equation (2) has a unique causal and stationary solution given by $y_t = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j}$, $t \in \mathbb{Z}$. The coefficients C_j, $j \geq 0$, are given from the power series expansion at 0 of the rational transfer function $c(z) := a(z)^{-1} b(z) = \sum_{j=0}^{\infty} C_j z^j$. Clearly this unique solution depends only on the power series coefficients C_j and the noise $(\varepsilon_t)_{t \in \mathbb{Z}}$.

Below we consider *unit root* ARMA processes (see Definition 1 below). To this end the difference operator at frequency $0 \leq \omega \leq \pi$ is defined as

$$\Delta_\omega(L) := \begin{cases} 1 - e^{i\omega} L, & \omega \in \{0, \pi\} \\ (1 - e^{i\omega} L)(1 - e^{-i\omega} L), & \omega \in (0, \pi). \end{cases} \tag{4}$$

Here L denotes the backward shift operator, such that $L(y_t)_{t \in \mathbb{Z}} = (y_{t-1})_{t \in \mathbb{Z}}$. Keeping this definition in mind we also use the sloppy shorthand notation $L y_t = y_{t-1}$, with which we obtain, e.g., $\Delta_\omega(L) y_t = y_t - 2\cos(\omega) y_{t-1} + y_{t-2}$ for $0 < \omega < \pi$. Note that for $0 < \omega < \pi$ the difference operator $\Delta_\omega(L)$ filters the pair of complex conjugate unit roots $e^{\pm i\omega}$ jointly. This ensures that also filtered processes $\Delta_\omega(L)(y_t)_{t \in \mathbb{Z}}$ are real valued for real valued processes $(y_t)_{t \in \mathbb{Z}}$. With the necessary notation established we can now define the class of ARMA processes considered in this paper.

DEFINITION 1. *The s-dimensional zero mean real ARMA process $(y_t)_{t \in \mathbb{Z}}$ has unit root structure $\Omega := ((\omega_1, h_1), \ldots, (\omega_l, h_l))$ with $0 \leq \omega_1 < \omega_2 < \cdots < \omega_l \leq \pi, h_k \in \mathbb{N}, k = 1, \ldots, l$, if it is a solution of the difference equation*

$$D(L) y_t := \Delta_{\omega_1}^{h_1}(L) \cdots \Delta_{\omega_l}^{h_l}(L) y_t = v_t, \quad t \in \mathbb{Z}, \tag{5}$$

where $(v_t)_{t \in \mathbb{Z}}$ is a stationary ARMA process, thus there exists a pair of left coprime matrix polynomials $(a(z), b(z)), \det a(z) \neq 0, |z| \leq 1, \det b(z) \neq 0, |z| < 1$ such that $(v_t)_{t \in \mathbb{Z}} = a(L)^{-1} b(L)(\varepsilon_t)_{t \in \mathbb{Z}} = c(L)(\varepsilon_t)_{t \in \mathbb{Z}}$ with $c(z) \neq 0$ for $z = e^{i\omega_k}, k = 1, \ldots, l$, for a white noise process $(\varepsilon_t)_{t \in \mathbb{Z}}$ with $\mathbb{E}\varepsilon_t \varepsilon_t' = \Sigma > 0$. Henceforth, the degree of the polynomial $D(z) = \Delta_{\omega_1}^{h_1}(z) \cdots \Delta_{\omega_l}^{h_l}(z)$ is denoted by H.

Such a zero mean real ARMA process is called unit root process. *The set $\{\omega_1, \ldots, \omega_l\}$ is referred to as set of* unit root frequencies *and the integers h_k, $k = 1, \ldots, l$ are called* integration orders.

A unit root process with unit root structure $((0, m))$, $m \in \mathbb{N}$, is called I(m) *process and a unit root process with unit root structure $((\omega_1, 1), \ldots, (\omega_l, 1))$ is called* multiple frequency I(1)*, in short* MFI(1)*, process.*

A zero mean stationary ARMA process $(v_t)_{t \in \mathbb{Z}} = c(L)(\varepsilon_t)_{t \in \mathbb{Z}}$ is said to have empty unit root structure $\Omega_0 := \{\}$.

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

Equation (5) implies that

$$a(L)\Delta_{\omega_1}^{h_1}(L)\cdots\Delta_{\omega_l}^{h_l}(L)y_t = a(L)v_t = b(L)\varepsilon_t, \quad t \in \mathbb{Z}. \tag{6}$$

Hence the unit root process $(y_t)_{t\in\mathbb{Z}}$ is a solution of the vector difference equation (VDE)

$$y_t + \sum_{j=1}^{p+H} \tilde{A}_j y_{t-j} = \varepsilon_t + \sum_{j=1}^{q} B_j \varepsilon_{t-j}, \quad t \in \mathbb{Z}; \tag{7}$$

i.e., $(y_t)_{t\in\mathbb{Z}}$ is a solution of the ARMA system $(\tilde{a}(z), b(z))$ with $\tilde{a}(z) = D(z)a(z)$ and $D(z)$ as given in Definition 1.

Every zero mean solution $(y_t)_{t\in\mathbb{Z}}$ of (7) such that the filtered process $D(L)(y_t)_{t\in\mathbb{Z}}$ is stationary is a unit root process. The difference between any two such solutions is a solution of the homogenous equation $\tilde{a}(L)y_{t,h} = D(L)a(L)y_{t,h} = 0$, $t \in \mathbb{Z}$. These are linear combinations of terms of the form $t^a\rho^t\cos(\omega(t-m))$ and $t^a\rho^t\sin(\omega(t-m))$; see, e.g., HD (p. 10). Here ρ and ω depend on the roots $\rho^{-1}e^{i\omega}$ of $\det\tilde{a}(z)$ and a depends on the corresponding multiplicities.[3] For stable roots it holds that $0 < \rho < 1$ and thus the corresponding solution components decay exponentially (for increasing t). These components are, similar to the focus on the stationary solution only in the stable case, excluded in Definition 1. For unit roots it holds that $\rho = 1$ and thus the corresponding homogenous solution components can be accommodated by including regressors of the form $t^a\cos(\omega(t-m)), t^a\sin(\omega(t-m))$ in the model. Consequently, the paper focuses on one specific particular solution of (7), defined below in (8), that is convenient for unit root and cointegration analysis.

We consider only zero mean processes to simplify the notation and exposition, whereas for actual modeling purposes one would include appropriate deterministic components (which may, as just discussed, even be necessary for zero mean processes). As is well known the presence or absence of deterministic components will in general influence the asymptotic distributions of the parameter estimators, but considering them has no influence on the unit root and cointegration properties. Note in this respect that in the literature definitions of integrated processes $(z_t)_{t\in\mathbb{Z}}$ are often formulated for demeaned processes, i.e., for processes $(z_t - \mathbb{E}(z_t))_{t\in\mathbb{Z}}$, compare, e.g., Johansen (1995, Def. 3.3, p. 35) and in this paper we also consider this latter zero mean process only.

Remark 1. Clearly the restriction to ARMA processes renders the above definition of unit root processes narrow. Such a narrow definition, however, suffices for the present paper in which a convenient state space representation for unit root ARMA processes is developed. The definition can of course be broadened by not constraining $(v_t)_{t\in\mathbb{Z}}$ in Definition 1 to be an ARMA process; e.g., Bauer and Wagner (2007) consider MFI(1) processes defined without the restriction to ARMA processes.

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

Concerning Definition 1 some further comments are in order. First, positing in the definition only the constraint that $c(e^{i\omega_k}) \neq 0$ for $k = 1, \ldots, l$, i.e., only at the unit root frequencies $z_k = e^{i\omega_k}$ and not for all points $z = e^{i\omega}$ with $0 \leq \omega \leq \pi$, allows one to classify also processes (with respect to unit root structure) that have been over-differenced at non-unit root frequencies. Second, the definition of the unit root structure and hence of the integration orders is given for the multivariate process and not for its individual coordinates. Since we only assume $c(e^{i\omega_k}) \neq 0$ and not $\det c(e^{i\omega_k}) \neq 0$ for $k = 1, \ldots, l$, the integration orders h_k are the maximal integration orders of the coordinates $(y_{j,t})_{t\in\mathbb{Z}}$, $j = 1, \ldots, s$ of $(y_t)_{t\in\mathbb{Z}}$ at the unit root frequency ω_k. Alternatively one might consider basing the definition of the unit root structure in a univariate fashion on the integration properties of the individual coordinates. Based on these, the unit root structure of the multivariate process could then be defined by including all unit root frequencies occurring in the different coordinates with the corresponding integration orders being given by the maximal integration orders at these frequencies over all coordinates. Such a unit root structure constructed from the individual coordinates corresponds with the unit root structure Ω from Definition 1 if and only if the correspondingly filtered multivariate process is stationary. This, however, need not be the case in general, since stacking stationary processes does not necessarily lead to a jointly stationary stacked process. Starting from a multivariate perspective rules out such problems. Note also that in the literature most definitions are formulated in a multivariate setting; see, e.g., Johansen (1995, 1997), Johansen and Schaumburg (1999), or Stock and Watson (1993).

Third, our definition of the empty unit root structure, which only requires the process to be stationary, implies, e.g., that processes that are solutions of $y_t = y_{t-1} + v_t$, $t \in \mathbb{Z}$ with $(v_t)_{t\in\mathbb{Z}}$ an ARMA process with empty unit root structure may have different unit root structures. In case that $v_t = \varepsilon_t, t \in \mathbb{Z}$ with $(\varepsilon_t)_{t\in\mathbb{Z}}$ white noise, the process $(y_t)_{t\in\mathbb{Z}}$ has unit root structure $((0, 1))$, if we consider a solution that has a unit root structure. If $v_t = \varepsilon_t - \varepsilon_{t-1}, t \in \mathbb{Z}$, then the process $(y_t)_{t\in\mathbb{Z}}$ has empty unit root structure, considering again a solution with a unit root structure. This could be overcome by introducing negative integration orders; compare Gregoir (1999a). For the representation results developed in this paper this issue is of no importance.

The unit root structure of unit root processes is uniquely defined, which can be shown by contradiction. Collecting the set of all unit root frequencies occurring in two unit root structures in $\{\omega_1, \ldots, \omega_m\}$, we can write the filters corresponding to the unit root structures as $D_1(L) := \Delta_{\omega_1}^{h_1}(L) \cdots \Delta_{\omega_m}^{h_m}(L)$ and $D_2(L) := \Delta_{\omega_1}^{\tilde{h}_1}(L) \cdots \Delta_{\omega_m}^{\tilde{h}_m}(L)$. In these filters $h_j = 0$ or $\tilde{h}_j = 0$ if a particular frequency ω_j is not contained in the respective unit root structure. Define $D^*(L) := \Delta_{\omega_1}^{\max(h_1,\tilde{h}_1)}(L) \cdots \Delta_{\omega_m}^{\max(h_m,\tilde{h}_m)}(L)$ and let $d_i(L) := D^*(L)/D_i(L)$, $i = 1, 2$. Consider a unit root frequency ω_k contained in the set of unit root frequencies corresponding to $D_1(L)$ but not contained in the set corresponding to $D_2(L)$. Then $d_1(e^{-i\omega_k}) \neq 0$ and hence the spectrum of $D^*(L)(y_t)_{t\in\mathbb{Z}} =$

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

$d_1(L)D_1(L)(y_t)_{t\in\mathbb{Z}} = d_2(L)D_2(L)(y_t)_{t\in\mathbb{Z}}$ is not equal to zero at frequency ω_k. However, $d_2(e^{-i\omega_k}) = 0$, which shows that the sets of unit root frequencies have to coincide. Analogously it follows that the integration orders have to coincide.

3. STATE SPACE SYSTEMS

It is well known that for stable matrix polynomials $a(z)$, i.e., $\det a(z) \neq 0$ for all $|z| \leq 1$, any ARMA system $(a(z), b(z))$ has a unique stationary solution on $\mathbb{Z}$. This solution is given by $y_t = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j}$, with the impulse response coefficients $C_j \in \mathbb{R}^{s\times s}$, $j \geq 0$, corresponding to the transfer function $c(z) = a(z)^{-1}b(z) = \sum_{j=0}^{\infty} C_j z_j$. Hence, this unique stationary solution only depends upon the transfer function $c(z)$ and the process $(\varepsilon_t)_{t\in\mathbb{Z}}$.

It is also well known that the unique stationary solution $(y_t)_{t\in\mathbb{Z}}$ just considered can be represented as a solution of a state space system (A, B, C), with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times s}$, and $C \in \mathbb{R}^{s\times n}$ as given in (1). The matrix triple (A, B, C) is referred to as a state space realization of the transfer function $c(z) = a(z)^{-1}b(z) = I_s + \sum_{j=1}^{\infty} CA^{j-1}Bz^j$. Similarly to the ARMA case, the set of all solutions of the state space equations is given by the sum of the stationary solution and a solution $(y_{t,h})_{t\in\mathbb{Z}}$ of the corresponding homogenous equations, i.e., $y_{t,h} = Cx_{t,h}$ and $x_{t+1,h} = Ax_{t,h}$.

Like ARMA systems, state space systems also are not unique, i.e., there exist classes of *observationally equivalent* state space systems that all realize the same transfer function. The state space equivalent to left coprimeness is minimality. A state space system is called minimal if there exists no other state space system realizing the same transfer function with a smaller state dimension. This minimal state dimension is also referred to as the order of the transfer function. Two minimal state space systems (A, B, C) and $(\tilde{A}, \tilde{B}, \tilde{C})$ are observationally equivalent if and only if there exists a nonsingular transformation matrix $T \in \mathbb{R}^{n\times n}$ such that $A = T\tilde{A}T^{-1}$, $B = T\tilde{B}$, and $C = \tilde{C}T^{-1}$. Such a transformation implies a corresponding basis change of the associated state vectors, i.e., $x_t = T\tilde{x}_t$.

Consider a given rational transfer function $c(z)$ and let (A, B, C) be a minimal state space realization and $(a(z), b(z))$ a left coprime ARMA realization. Then according to HD (Thm. 1.2.2, p. 19) the nonzero eigenvalues of A equal the inverses of the roots of $\det a(z)$. Hence, for minimal state space systems the stability condition $|\lambda_{\max}(A)| < 1$ corresponds to the stability assumption $\det a(z) \neq 0$ for $|z| \leq 1$ and eigenvalues of A of unit modulus correspond to unit roots. Considering the inverse transfer function $c^{-1}(z) = b^{-1}(z)a(z) = I_s - zC(I_n - z(A - BC))^{-1}B$ it follows analogously that the miniphase assumption $\det b(z) \neq 0$ for $|z| < 1$ (HD, p. 25) imposed in Definition 1 is equivalent to the condition $|\lambda_{\max}(A - BC)| \leq 1$.

The following Theorem 1 clarifies the relations between unit root processes and the solutions to corresponding state space representations. In light of the discussion in the previous section, a central element in this respect is one particular solution of a minimal state space representation that only depends on (A, B, C) and $(\varepsilon_t)_{t\in\mathbb{Z}}$. This solution is built from a specific state space representation with the

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

A matrix in the format $A = \mathrm{diag}(A_u, A_\bullet)$, where all eigenvalues of A_u have unit modulus (hence A_u is invertible) and $|\lambda_{\max}(A_\bullet)| < 1$. Considering such a state space realization is clearly not a restriction of generality, since an observationally equivalent realization of this form always exists (e.g., by using the real Jordan normal form of A). We also partition the corresponding matrices B, C and the state vector x_t accordingly, i.e., $B = [B_u', B_\bullet']'$, $C = [C_u, C_\bullet]$, and $x_t = [x_{t,u}', x_{t,\bullet}']'$. Then the considered particular solution $(y_{t,p})_{t\in\mathbb{Z}}$ is given as

$$
\begin{aligned}
y_{t,p} &= C_u x_{t,u} + C_\bullet x_{t,\bullet} + \varepsilon_t,\\
x_{t,u} &= \begin{cases} \sum_{j=1}^{t-1} A_u^{j-1} B_u \varepsilon_{t-j}, & t \geq 1 \\ -\sum_{j=t}^{0} A_u^{j-1} B_u \varepsilon_{t-j}, & t < 1, \end{cases}\\
x_{t,\bullet} &= \sum_{j=1}^{\infty} A_\bullet^{j-1} B_\bullet \varepsilon_{t-j}. \qquad \textbf{(8)}
\end{aligned}
$$

Clearly, since we consider only minimal state space representations, $(x_{t,\bullet})_{t\in\mathbb{Z}}$ is the unique stationary solution of the state equation $x_{t+1,\bullet} = A_\bullet x_{t,\bullet} + B_\bullet \varepsilon_t$. With respect to $x_{t,u}$ the particular solution chosen corresponds to $x_{1,u} = 0$. These choices imply that $(y_{t,p})_{t\in\mathbb{Z}}$ only depends on the transfer function corresponding to the state space system (A, B, C) and the noise $(\varepsilon_t)_{t\in\mathbb{Z}}$. The relation between such state space systems and unit root processes is clarified in the following theorem.

THEOREM 1. *Let $(y_t)_{t\in\mathbb{Z}}$ be a unit root process as given in Definition 1. Then there exists a solution $(y_{t,h})_{t\in\mathbb{Z}}$ of the homogenous equation $D(L)(y_{t,h})_{t\in\mathbb{Z}} = 0$, such that the process $y_{t,p} := y_t - y_{t,h}$, $t \in \mathbb{Z}$ is the particular solution given in (8) to a minimal state space system (A, B, C) with $A = \mathrm{diag}(A_u, A_\bullet)$, where all eigenvalues of A_u have unit modulus and $|\lambda_{\max}(A_\bullet)| < 1$.*

The process $(y_{t,p})_{t\in\mathbb{Z}}$ given in (8) is a zero mean unit root process and a particular solution of a minimal state space system (A, B, C) such that $|\lambda_{\max}(A)| = 1$ and $|\lambda_{\max}(A - BC)| \leq 1$. Denoting with $((\omega_1, h_1), \ldots, (\omega_l, h_l))$ the unit root structure of $(y_{t,p})_{t\in\mathbb{Z}}$, it holds that $z_k = e^{i\omega_k}$, $k = 1, \ldots, l$ are the eigenvalues of A with unit modulus and non-negative imaginary part and the integration orders $h_k > 0$ equal the dimensions of the largest Jordan blocks corresponding to z_k for $k = 1, \ldots, l$.

The theorem shows that every unit root process can, up to a component in the kernel of $D(L)$, be represented as a specific particular solution of a minimal state space model. This particular solution only depends on the transfer function corresponding to the state space system as well as the error process $(\varepsilon_t)_{t\in\mathbb{Z}}$. Conversely, all solutions of state space systems where the corresponding matrix A has eigenvalues on the unit circle are up to a solution of the homogenous equation unit root processes.

Partitioning $x_{t,h} = [x_{t,u,h}', x_{t,\bullet,h}']'$ according to the partitioning of $A = \mathrm{diag}(A_u, A_\bullet)$ it follows from the proof of Theorem 1 that a solution $(y_t)_{t\in\mathbb{Z}}$ to

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

the state space system is a unit root process, when the corresponding $(x_{t,h})_{t\in\mathbb{Z}}$ fulfills $x_{t,\bullet,h}=0,\ t\in\mathbb{Z}$ and $\mathbb{E}(x_{t,u,h})=0,\ t\in\mathbb{Z}$. Note, however, that the set of unit root processes given as solutions to the ARMA representation as considered in Definition 1 and the solution processes to corresponding minimal state space systems as discussed in the theorem are not identical. This stems from the fact that in the theorem only minimal state space systems are considered, whereas the ARMA representation $(\tilde{a}(z),b(z))$ is not necessarily left coprime, with potential rank reductions in $(\tilde{a}(z),b(z))$ on the unit circle due to roots of $\det b(z)$ at unit root frequencies. Rather than providing a full discussion, let us just consider a simple example here, with $D(L)=1-L$, $a(L)=I_2$ and $b(L)=\mathrm{diag}(1,1-L)$. The solution set, where for brevity we only consider $t\in\mathbb{N}$, to the ARMA system is given by $y_t=(\sum_{j=1}^{t}\varepsilon_{j,1}+c_1,\varepsilon_{t,2}+c_2)'$, for $c_1,c_2\in\mathbb{R}$ zero mean random variables. A minimal state space representation corresponding to the above ARMA system is given by $A=1$, $B=(1,\ 0)$, and $C=(1,\ 0)'$. Thus, the corresponding solution set is given only by $y_t=(\sum_{j=1}^{t}\varepsilon_{j,1}+c_1,\varepsilon_{t,2})'$, i.e., $c_2=0$.

In the remainder of the paper, based on the particular solution (8), we exploit the flexibility of the state space framework to develop a specific, unique state space representation for unit root processes that highlights the unit root and cointegration properties. In this respect, as can already be guessed from Theorem 1, the eigenvalue structure of A will be of key importance since the unit modulus eigenvalues of A correspond to the unit roots. Formally, we construct a so-called *canonical form*, i.e., a mapping that attaches unique state space representations to transfer functions $c(z)$.

4. THE CANONICAL FORM

In constructing the canonical representation first more structure (as already used in the proof of Theorem 1) is imposed on the block-diagonal A matrix used so far. Henceforth, we consider the A matrix in the format $A=\mathrm{diag}(A_1,\ldots,A_l,A_\bullet)$, where the matrices A_k, $k=1,\ldots,l$ have as their only eigenvalues the pair $z_k,\overline{z}_k=e^{\pm i\omega_k}$ (including here for brevity the cases $z_k=\pm1$) for $0\le\omega_k<\omega_{k+1}\le\pi$, $k=1,\ldots,l-1$ and where as before $|\lambda_{\max}(A_\bullet)|<1$. As argued before, such a realization exists, the A-matrix can, e.g., be taken in real Jordan normal form with the eigenvalues ordered accordingly. Also, as before we consider the matrices B, C, and the state vector x_t partitioned accordingly, i.e., $B=[B_1',\ldots,B_l',B_\bullet']'$, $C=[C_1,\ldots,C_l,C_\bullet]$, and $x_t=[x_{t,1}',\ldots,x_{t,l}',x_{t,\bullet}']'$. Note that for block-diagonal systems with $A=\mathrm{diag}(A_1,\ldots,A_m)$, $B=[B_1',\ldots,B_m']'$, and $C=[C_1,\ldots,C_m]$ partitioned accordingly, it follows that the power series coefficients of $c(z)=I_s+\sum_{j=1}^{\infty}CA^{j-1}Bz^j$ fulfill $CA^jB=\sum_{h=1}^{m}C_hA_h^jB_h$ for $j\ge0$. This implies that the subsystems (A_h,B_h,C_h) can be analyzed separately when developing a canonical form.

The block-diagonality constraints are not sufficient to select a unique state space realization (A,B,C) of the transfer function $c(z)$. The additionally necessary constraints are imposed in two steps. In the first step restrictions that detail

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

the form of the matrix A are imposed. In the second step a set of constraints on the matrices B and C to impose uniqueness is formulated. In this step a characterization of minimality is used ascertaining that the restrictions can be imposed.

When imposing further restrictions to provide a unique realization it turns out convenient to use complex valued quantities to simplify the algebra. Consider $z_k = e^{i\omega_k}$ a unit modulus eigenvalue not equal to ± 1. Then the systems (A_k, B_k, C_k) with A_k in real Jordan normal form can be transformed to

$$A_{k,\mathbb{C}} = \begin{bmatrix} J_k & 0 \\ 0 & \overline{J}_k \end{bmatrix}, \qquad B_{k,\mathbb{C}} = \begin{bmatrix} B_{k,-} \\ \overline{B}_{k,-} \end{bmatrix}, \qquad C_{k,\mathbb{C}} = \begin{bmatrix} C_{k,-} & \overline{C}_{k,-} \end{bmatrix}, \tag{9}$$

where the matrices J_k have as their only eigenvalue $\overline{z}_k$ and are in the specific form given in (10) below. The notation $B_{k,-}$ and $C_{k,-}$ indicates the correspondence to complex eigenvalues with negative imaginary part. Clearly, in case that $z_k = \pm 1$ the real and complex representation coincide, i.e., $A_{k,\mathbb{C}} = J_k$, $B_{k,\mathbb{C}} = B_{k,-}$, and $C_{k,\mathbb{C}} = C_{k,-}$, where we here also use the notation with double subscript $(k,-)$ to indicate the correspondence to a representation with the A-matrix in this specific form. We use the following structure for the matrices $J_k \in \mathbb{C}^{d^k \times d^k}$, $k = 1, \ldots, l$:

$$J_k = \begin{bmatrix} \overline{z}_k I_{d_1^k} & \left[I_{d_1^k}, 0^{d_1^k \times \left(d_2^k - d_1^k\right)} \right] & 0 & \cdots & 0 \\ 0^{d_2^k \times d_1^k} & \overline{z}_k I_{d_2^k} & \left[I_{d_2^k}, 0^{d_2^k \times \left(d_3^k - d_2^k\right)} \right] & 0 & \vdots \\ 0 & 0 & \overline{z}_k I_{d_3^k} & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \left[I_{d_{h_k-1}^k}, 0^{d_{h_k-1}^k \times \left(d_{h_k}^k - d_{h_k-1}^k\right)} \right] \\ 0 & 0 & \cdots & 0 & \overline{z}_k I_{d_{h_k}^k} \end{bmatrix}, \tag{10}$$

where the integer h_k denotes the smallest integer such that $(J_k - I\overline{z}_k)^{h_k} = 0$. The integers $1 \le d_1^k \le d_2^k \le \ldots \le d_{h_k}^k$ denote the differences between the ranks of $(J_k - I\overline{z}_k)^{h_k - j}$ and $(J_k - I\overline{z}_k)^{h_k - j + 1}$ for $j = 1, \ldots, h_k$. By definition it holds that $d^k = \sum_{j=1}^{h_k} d_j^k$. For any given real matrix A_k with $z_k, \overline{z}_k$ of unit modulus as its only eigenvalues there is a similarity transformation to $A_{k,\mathbb{C}}$ with J_k in the given structure. This stems from the fact that J_k corresponds to the (unique) Jordan normal form with a specific ordering of the coordinates.[4]

The matrices $B_{k,-}$ and $C_{k,-}$ are partitioned according to the structure of J_k as given in (10), i.e., $B_{k,-} = [(B_{k,-}^1)', \ldots, (B_{k,-}^{h_k})']'$, $B_{k,-}^j \in \mathbb{C}^{d_j^k \times s}$, and $C_{k,-} = [C_{k,-}^1, \ldots, C_{k,-}^{h_k}]$, $C_{k,-}^j \in \mathbb{C}^{s \times d_j^k}$.

A real valued representation based on the specific complex representation for the kth subsystem, with k such that $0 < \omega_k < \pi$, is obtained by using the transformation matrix

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

$$T_{\mathbb{R},d} := \begin{bmatrix} 1 & & 1 & & \\ i & & -i & & \\ & 1 & & 1 & \\ & i & & -i & \\ & & \ddots & & \ddots \\ & & & \ddots & & \ddots \end{bmatrix} \in \mathbb{C}^{2d\times 2d}. \tag{11}$$

This results in

$$A_k = T_{\mathbb{R},d^k} A_{k,\mathbb{C}} T^{-1}_{\mathbb{R},d^k} \in \mathbb{R}^{2d^k\times 2d^k}, \qquad B_k = T_{\mathbb{R},d^k} B_{k,\mathbb{C}} \in \mathbb{R}^{2d^k\times s},$$
$$C_k = C_{k,\mathbb{C}} T^{-1}_{\mathbb{R},d^k} \in \mathbb{R}^{s\times 2d^k}, \tag{12}$$

with

$$A_k = \begin{bmatrix} I_{d_1^k}\otimes Q(\omega_k) & \left[I_{2d_1^k}, 0^{2d_1^k\times 2\left(d_2^k-d_1^k\right)}\right] & 0 & \cdots & 0 \\ 0^{2d_2^k\times 2d_1^k} & I_{d_2^k}\otimes Q(\omega_k) & \left[I_{2d_2^k}, 0^{2d_2^k\times 2\left(d_3^k-d_2^k\right)}\right] & 0 & \vdots \\ 0 & 0 & I_{d_3^k}\otimes Q(\omega_k) & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \left[I_{2d_{h_k-1}^k}, 0^{2d_{h_k-1}^k\times 2\left(d_{h_k}^k-d_{h_k-1}^k\right)}\right] \\ 0 & 0 & \cdots & 0 & I_{d_{h_k}^k}\otimes Q(\omega_k) \end{bmatrix}, \tag{13}$$

where $Q(\omega_k) := \begin{bmatrix} \cos(\omega_k) & -\sin(\omega_k) \\ \sin(\omega_k) & \cos(\omega_k) \end{bmatrix}$,

$$B_k = \begin{bmatrix} B_k^1 \\ \vdots \\ B_k^{h_k} \end{bmatrix}, \quad \text{with } B_k^j = \begin{bmatrix} 2\mathcal{R}\left(B^j_{k,-,1}\right) \\ -2\mathcal{I}\left(B^j_{k,-,1}\right) \\ \vdots \\ 2\mathcal{R}\left(B^j_{k,-,d_j^k}\right) \\ -2\mathcal{I}\left(B^j_{k,-,d_j^k}\right) \end{bmatrix} \quad \text{for } j = 1,\ldots,h_k, \tag{14}$$

with $B^j_{k,-,m}$ for $m = 1,\ldots,d_j^k$ denoting the mth row of $B^j_{k,-}$. Analogously,

$$C_k = \left[C_k^1,\ldots,C_k^{h_k}\right] \quad \text{with}$$
$$C_k^j = \left[\mathcal{R}\left(C^j_{k,-,1}\right), \mathcal{I}\left(C^j_{k,-,1}\right), \ldots, \mathcal{R}\left(C^j_{k,-,d_j^k}\right), \mathcal{I}\left(C^j_{k,-,d_j^k}\right)\right]$$
$$\text{for } j = 1,\ldots,h_k, \tag{15}$$

with $C^j_{k,-,m}$ for $m = 1,\ldots,d_j^k$ denoting the mth column of $C^j_{k,-}$.

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

For real valued unit roots of course no transformation combining pairs of complex conjugate matrices to result in real valued subsystem matrices is required. To have uniform notation we define for k such that $\omega_k \in \{0, \pi\}$:

$$A_k = A_{k,\mathbb{C}} = J_k \in \mathbb{R}^{d^k \times d^k}, \quad B_k = B_{k,\mathbb{C}} = B_{k,-} \in \mathbb{R}^{d^k \times s}, \quad \text{and}$$
$$C_k = C_{k,\mathbb{C}} = C_{k,-} \in \mathbb{R}^{s \times d^k}, \tag{16}$$

where $(J_k, B_{k,-}, C_{k,-})$ are obtained by the transformation corresponding to the reordering of the Jordan normal form of A_k to arrive at the format given in (10).

Using this notation and partitioning of the system matrices the first lemma characterizes minimality.

LEMMA 1. *Let (A, B, C) denote a state space system where the matrix $A = \operatorname{diag}(A_1, \ldots, A_l, A_\bullet)$ is block-diagonal with the matrices A_k, $k = 1, \ldots, l$ corresponding to unit modulus eigenvalues as discussed above and given in (12) respectively (16) and $|\lambda_{\max}(A_\bullet)| < 1$. The matrices B_k and C_k are also as discussed above and given in (12) respectively (16). Partition for $k = 1, \ldots, l$ $C_{k,-}^j = [C_{k,-}^{j,G}\ C_{k,-}^{j,E}]$ with $C_{k,-}^{j,E} \in \mathbb{C}^{s \times (d_j^k - d_{j-1}^k)}$ and $C_{k,-}^{j,G} \in \mathbb{C}^{s \times d_{j-1}^k}$ for $j = 1, \ldots, h_k$, with $d_0^k := 0$. Furthermore define $C_{k,-}^E = [C_{k,-}^{1,E}, \ldots, C_{k,-}^{h_k,E}] \in \mathbb{C}^{s \times d_{h_k}^k}$. Then it holds that*

(i) The system (A, B, C) is minimal if and only if for each $k = 1, \ldots, l$ the matrix $B_{k,-}^{h_k}$ has full row rank, $C_{k,-}^E$ has full column rank, and the stable subsystem $(A_\bullet, B_\bullet, C_\bullet)$ is minimal.

(ii) In case of minimality it holds that $d_{h_k}^k \leq s$ and that $C_{k,-}^{j,E}$ has full column rank for $k = 1, \ldots, l$ and $j = 1, \ldots, h_k$.

As has been mentioned, minimality and the A matrix in the described format are not sufficient to select a unique state space representation. The further necessary restrictions use the concept of positive upper triangular (p.u.t.) matrices. A matrix $M = [m_{i,j}]_{i=1,\ldots,c,\, j=1,\ldots,m} \in \mathbb{C}^{c \times m}$ is p.u.t. if there exist indices $1 \leq j_1 < j_2 < \cdots < j_c \leq m$, such that $m_{i,j} = 0$, $j < j_i$, $m_{i,j_i} \in \mathbb{R}$, $m_{i,j_i} > 0$, i.e., if M is of the form

$$\begin{bmatrix} 0 & \cdots & 0 & m_{1,j_1} & * & & \cdots & & * \\ 0 & & \cdots & & 0 & m_{2,j_2} & * & & \\ & & & & & & & & \\ 0 & & & \cdots & & & 0 & m_{c,j_c} & * \end{bmatrix}, \tag{17}$$

where the symbol $*$ indicates unrestricted entries. The additional restrictions needed to achieve identification are based on an extension of Ober (1996, Lem. 3.1, p. 129) from real valued to complex matrices (details are available upon request): It can be shown that for any matrix $X \in \mathbb{C}^{m \times m}$ with rank c, there exists a unique decomposition $X = CB$, with $C \in \mathbb{C}^{m \times c}$ and $B \in \mathbb{C}^{c \times m}$ with $C'C = I_c$ and where B is p.u.t.

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

THEOREM 2. *For every minimal state space system* (A, B, C) *with* $|\lambda_{\max}(A)| = 1$ *and* $|\lambda_{\max}(A - BC)| \leq 1$, *with the eigenvalues of the A matrix of unit modulus with non-negative imaginary part given by* $z_k = e^{i\omega_k}$, $k = 1, \ldots, l$, $0 \leq \omega_1 < \ldots < \omega_l \leq \pi$ *there exists a unique observationally equivalent minimal state space representation* $(\mathcal{A}, \mathcal{B}, \mathcal{C})$ *such that*

(i) $\mathcal{A} = \operatorname{diag}(\mathcal{A}_1, \ldots, \mathcal{A}_l, \mathcal{A}_\bullet)$ *where* $|\lambda_{\max}(\mathcal{A}_\bullet)| < 1$. *The matrices* $\mathcal{B} = [\mathcal{B}_1', \ldots, \mathcal{B}_l', \mathcal{B}_\bullet']'$ *and* $\mathcal{C} = [\mathcal{C}_1, \ldots, \mathcal{C}_l, \mathcal{C}_\bullet]$ *are partitioned accordingly.*

(ii) Using this partitioning it holds for $k = 1, \ldots, l$ *that*
- *for* $z_k = e^{i\omega_k}$ *with* $0 < \omega_k < \pi$: $(\mathcal{A}_k, \mathcal{B}_k, \mathcal{C}_k)$ *is as given in (12) with* $(\mathcal{A}_{k,\mathbb{C}}, \mathcal{B}_{k,\mathbb{C}}, \mathcal{C}_{k,\mathbb{C}})$ *as given in (9),*
- *for* $z_k = e^{i\omega_k}$ *with* $\omega_k \in \{0, \pi\}$: $(\mathcal{A}_k, \mathcal{B}_k, \mathcal{C}_k)$ *is as given in (16),*
with J_k *in the format defined in (10).*

(iii) Using the notation of Lemma 1 it holds that $(\mathcal{C}^E_{k,-})'\mathcal{C}^E_{k,-} = I_{d^k_{h_k}}$ *for* $k = 1, \ldots, l$ *and* $(\mathcal{C}^{j,G}_{k,-})'\mathcal{C}^{i,E}_{k,-} = 0$ *for* $1 \leq i \leq j$ *for* $j = 2, \ldots, h_k$ *and* $k = 1, \ldots, l$.

(iv) Partitioning $\mathcal{B}^{h_k}_{k,-}$ *from* $\mathcal{B}_{k,-} = [(\mathcal{B}^1_{k,-})', \ldots, (\mathcal{B}^{h_k}_{k,-})']'$ *as* $\mathcal{B}^{h_k}_{k,-} = [(\mathcal{B}^{h_k,1}_{k,-})', \ldots, (\mathcal{B}^{h_k,h_k}_{k,-})']'$ *with* $\mathcal{B}^{h_k,j}_{k,-} \in \mathbb{C}^{(d^k_j - d^k_{j-1}) \times s}$ *it holds that* $\mathcal{B}^{h_k,j}_{k,-}$ *is positive upper triangular (p.u.t.) for* $d^k_j > d^k_{j-1}$ *for* $j = 1, \ldots, h_k$ *and* $k = 1, \ldots, l$.

(v) $(\mathcal{A}_\bullet, \mathcal{B}_\bullet, \mathcal{C}_\bullet)$ *is in echelon canonical form (cf. HD, Thm. 2.5.2, p. 61).*

Throughout the paper, for state space representations in canonical form calligraphic typesetting is used.

The canonical form presented in Theorem 2 highlights that the *eigenvalue structure* (for unit modulus eigenvalues) of the A-matrix as given in the matrices J_k carries more information than provided by the unit root structure as given in Definition 1. This motivates the definition of the state space unit root structure, which is useful for the characterization of model sets with common unit root properties.

DEFINITION 2. *A unit root process* $(y_t)_{t \in \mathbb{Z}}$ *with a canonical state space representation as given in Theorem 2 has* state space unit root structure

$$\Omega_S = \left\{ \left(\omega_1, \left(d^1_1, \ldots, d^1_{h_1} \right) \right), \ldots, \left(\omega_l, \left(d^l_1, \ldots, d^l_{h_l} \right) \right) \right\},$$

where $0 \leq d^k_1 \leq d^k_2 \leq \ldots \leq d^k_{h_k} \leq s$ *for* $k = 1, \ldots, l$.

5. COINTEGRATION AND POLYNOMIAL COINTEGRATION

The unique representation of Theorem 2 is developed to allow for convenient analysis of cointegration and polynomial cointegration, as will be discussed in this section. Before turning to a definition of cointegration and polynomial cointegration, we start by defining a semi-ordering of unit root structures. Denote for

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

a unit root structure Ω by $F(\Omega)$ the set of distinct unit root frequencies included, i.e., $F(\Omega) := \{\omega_1, \ldots, \omega_l\}$.

DEFINITION 3. *Let* $\tilde{\Omega} := ((\tilde{\omega}_1, \tilde{h}_1), \ldots, (\tilde{\omega}_{\tilde{l}}, \tilde{h}_{\tilde{l}}))$ *and* $\Omega := ((\omega_1, h_1), \ldots, (\omega_l, h_l))$ *be two unit root structures. Then* $\tilde{\Omega} \preceq \Omega$ *if*

(i) $F(\tilde{\Omega}) \subseteq F(\Omega)$.
(ii) For all $\omega \in F(\tilde{\Omega})$ *for* $\tilde{k}$ *and* k *such that* $\tilde{\omega}_{\tilde{k}} = \omega_k = \omega$ *it holds that* $\tilde{h}_{\tilde{k}} \leq h_k$.

Further $\tilde{\Omega} \prec \Omega$ *if* $\tilde{\Omega} \preceq \Omega$ *and* $\tilde{\Omega} \neq \Omega$.

For two unit root structures $\tilde{\Omega} \preceq \Omega$ *define the decrease* $\delta_k(\Omega, \tilde{\Omega})$ *of the integration order at frequency* ω_k*, for* $k = 1, \ldots, l$*, as*

$$\delta_k(\Omega, \tilde{\Omega}) := \begin{cases} h_k - \tilde{h}_{\tilde{k}}, & \exists \tilde{k} : \tilde{\omega}_{\tilde{k}} = \omega_k \in F(\tilde{\Omega}), \\ h_k, & \omega_k \notin F(\tilde{\Omega}). \end{cases}$$

Clearly, the empty unit root structure is the smallest element in this semi-ordering, i.e., $\Omega_0 := \{\} \prec \Omega$ for any nonempty unit root structure Ω.

Let $(y_t)_{t\in\mathbb{Z}}$ have unit root structure Ω and consider a transformed process $\beta(L)'(y_t)_{t\in\mathbb{Z}} = (\sum_{j=0}^{q} \beta_j' y_{t-j})_{t\in\mathbb{Z}}$. Here $\beta(z) = \sum_{j=0}^{q} \beta_j z^j$, $\beta_j \in \mathbb{R}^s$ for $j = 0, \ldots, q$ and $\beta_q \neq 0$.[5] By potentially accounting for a solution of the homogenous equation $D(L)y_{t,h} = 0$, where $D(L)$ corresponds to Ω, the process $\beta(L)'(y_t)_{t\in\mathbb{Z}}$ has a unit root structure equal to or smaller than Ω in the semi-ordering defined above.

Our definition of cointegration and polynomial cointegration is based on the introduced semi-ordering of unit root structures.

DEFINITION 4. *An* s*-dimensional unit root process* $(y_t)_{t\in\mathbb{Z}}$ *with unit root structure* Ω *is called* cointegrated of order $(\Omega, \tilde{\Omega})$*, where* $\tilde{\Omega} \prec \Omega$*, if there exists a vector* $\beta \in \mathbb{R}^s, \beta \neq 0$*, such that* $(\beta' y_t)_{t\in\mathbb{Z}}$ *has unit root structure* $\tilde{\Omega}$*. In this case the vector* β *is called* cointegrating vector (CIV) of order $(\Omega, \tilde{\Omega})$.

An s*-dimensional unit root process* $(y_t)_{t\in\mathbb{Z}}$ *with unit root structure* Ω *is called* polynomially cointegrated of order $(\Omega, \tilde{\Omega})$*, where* $\tilde{\Omega} \prec \Omega$*, if there exists a vector polynomial* $\beta(z) = \sum_{m=0}^{q} \beta_m z^m$*,* $\beta_m \in \mathbb{R}^s$*,* $m = 0, \ldots, q$*,* $\beta_q \neq 0$*, for some integer* $1 \leq q < \infty$ *such that*

(i) $\beta(L)'(y_t)_{t\in\mathbb{Z}}$ *has unit root structure* $\tilde{\Omega}$*,*
(ii) $\max_{k=1,\ldots,l} \|\beta(e^{i\omega_k})\| \delta_k(\Omega, \tilde{\Omega}) \neq 0$.

In this case the vector polynomial $\beta(z)$ *is called* polynomial cointegrating vector (PCIV) of order $(\Omega, \tilde{\Omega})$.

The restriction formulated in (ii) in the above definition rules out vector polynomials that reduce the unit root structure merely by simply differencing the process.

For a clearer distinction we sometimes refer to cointegration as *static* cointegration. Equivalently, we can interpret cointegration as polynomial cointegration with polynomial degree 0 (excluded in the above definition).

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

It is clear that a unit root process $(y_t)_{t\in\mathbb{Z}}$ can be statically and polynomially cointegrated of different orders. However, for any CIV or PCIV for a process $(y_t)_{t\in\mathbb{Z}}$, its cointegration respectively polynomial cointegration order is unique. This follows immediately from the uniqueness of the unit root structure of unit root processes (see the discussion at the end of Section 2).

In the considered state space representation the particular solution $(y_{t,p})_{t\in\mathbb{Z}}$ is a static linear function of the state and the error process. Therefore the cointegration and polynomial cointegration properties of $(y_{t,p})_{t\in\mathbb{Z}}$ are closely related to the respective properties of the state. The following theorem shows that the state process corresponding to the canonical representation developed in Theorem 2 has a particularly simple unit root structure and cointegration properties.

THEOREM 3. *Consider a minimal state space representation as given in Theorem 2 and let $(x_t)_{t\in\mathbb{Z}}$ be the state process corresponding to the particular solution $(y_{t,p})_{t\in\mathbb{Z}}$ as given in (8) with unit root structure $((\omega_1,h_1),\dots,(\omega_l,h_l))$.*

Consider the state process partitioned according to the partitioning of the system matrices given in Theorem 2, i.e., $x_t=[x_{t,1}',\dots,x_{t,l}',x_{t,\bullet}']'$. For $k=1,\dots,l$ further partition $x_{t,k}=[(x_{t,k}^1)',\dots,(x_{t,k}^{h_k})']'$, where using the indices d_j^k as before, $x_{t,k}^j\in\mathbb{R}^{2d_j^k}$ for $j=1,\dots,h_k$ for k such that $0<\omega_k<\pi$ and $x_{t,k}^j\in\mathbb{R}^{d_j^k}$ for $j=1,\dots,h_k$ for k such that $\omega_k\in\{0,\pi\}$.

Then the following properties hold:

(i) The processes $(x_{t,k}^j)_{t\in\mathbb{Z}}$ have unit root structure $((\omega_k,h_k-j+1))$ for $j=1,\dots,h_k$ and $k=1,\dots,l$.

(ii) The processes $(x_{t,k}^j)_{t\in\mathbb{Z}}$ are not cointegrated for $j=1,\dots,h_k$ and $k=1,\dots,l$.

For $\omega_k=0$ the processes $(x_{t,k}^j)_{t\in\mathbb{Z}}$ are referred to as stochastic trends of order h_k-j+1 and for $\omega_k\neq 0$ they are referred to as stochastic cycles of order h_k-j+1.

In combination with the rank restrictions on the system matrices stemming from minimality, these properties of the (sub-blocks of the) state $(x_t)_{t\in\mathbb{Z}}$ considerably facilitate cointegration and polynomial cointegration analysis for $(y_{t,p})_{t\in\mathbb{Z}}$. Consider the canonical representation $(\mathcal{A},\mathcal{B},\mathcal{C})$ developed in Theorem 2, then

$$y_{t,p}=\mathcal{C}x_t+\varepsilon_t=\sum_{k=1}^{l}\sum_{j=1}^{h_k}\mathcal{C}_k^j x_{t,k}^j+\mathcal{C}_\bullet x_{t,\bullet}+\varepsilon_t. \tag{18}$$

Since $\mathcal{C}_k^1\neq 0$ for $k=1,\dots,l$ according to the rank restrictions implied by minimality (see Lemma 1) the integration properties $(y_{t,p})_{t\in\mathbb{Z}}$ and $(x_t)_{t\in\mathbb{Z}}$ coincide.

Furthermore, the processes $(\mathcal{C}_k^j x_{t,k}^j)_{t\in\mathbb{Z}}$ have the same integration properties as unit root processes with unit root structures $((\omega_k,h_k-j+1))$ for $\mathcal{C}_k^j\neq 0$ for $j=1,\dots,h_k$ and $k=1,\dots,l$. Therefore, any vector $\beta\neq 0$ such that $\beta'[\mathcal{C}_k^1,\dots,\mathcal{C}_k^j]=0$ and $\beta'\mathcal{C}_k^{j+1}\neq 0$ leads to a reduction of the integration order of

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

$(\beta' y_{t,p})_{t\in\mathbb{Z}}$ at frequency ω_k from h_k to $h_k - j$. Thus, cointegration is characterized by orthogonality to sub-blocks of $\mathcal{C}_k$ for $k = 1, \ldots, l$. Clearly, a vector $\beta \neq 0$ can lead to a reduction of the integration orders at several or even all unit root frequencies.

Note in addition that knowledge concerning the unit root structure of coordinate processes of $(y_{t,p})_{t\in\mathbb{Z}}$ simply translates into zero restrictions on the corresponding entries in $\mathcal{C}$ and is thus simple to implement in a parameterization.

Also, polynomial cointegrating relationships can be recovered easily from the developed canonical representation. For any $h \in \mathbb{N}$, it holds that

$$y_{t,p} = \mathcal{C}\mathcal{A}^h x_{t-h} + \sum_{j=1}^{h} \mathcal{C}\mathcal{A}^{j-1}\mathcal{B}\varepsilon_{t-j} + \varepsilon_t. \tag{19}$$

For some $q \in \mathbb{N}$, which will be assigned a specific value below, define $Y^-_{t,q} := [y'_{t,p}, y'_{t-1,p}, \ldots, y'_{t-q,p}]'$ and $E^-_{t,q} := [\varepsilon'_t, \varepsilon'_{t-1}, \ldots, \varepsilon'_{t-q}]'$. Furthermore, use as short-hand notation $\mathcal{C}_u := [\mathcal{C}_1, \ldots, \mathcal{C}_l]$, $\mathcal{A}_u := \operatorname{diag}(\mathcal{A}_1, \ldots, \mathcal{A}_l)$, and $x_{t,u} := [x'_{t,1}, \ldots, x'_{t,l}]'$. Now consider equation (19) for the stacked vector $Y^-_{t,q}$ with the index h from (19) chosen in each block-row to give the state x_{t-q} in all equations:

$$Y^-_{t,q} = \underbrace{\begin{bmatrix} \mathcal{C}_u\mathcal{A}_u^q \\ \mathcal{C}_u\mathcal{A}_u^{q-1} \\ \vdots \\ \mathcal{C}_u \end{bmatrix}}_{\tilde{\mathcal{O}}_{q,u}} x_{t-q,u} + \underbrace{\begin{bmatrix} \mathcal{C}_\bullet\mathcal{A}_\bullet^q \\ \mathcal{C}_\bullet\mathcal{A}_\bullet^{q-1} \\ \vdots \\ \mathcal{C}_\bullet \end{bmatrix}}_{\tilde{\mathcal{O}}_{q,\bullet}} x_{t-q,\bullet} + \underbrace{\begin{bmatrix} I & \mathcal{C}\mathcal{B} & \ldots & \mathcal{C}\mathcal{A}^{q-2}\mathcal{B} & \mathcal{C}\mathcal{A}^{q-1}\mathcal{B} \\ 0 & I & \mathcal{C}\mathcal{B} & \ldots & \mathcal{C}\mathcal{A}^{q-2}\mathcal{B} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \mathcal{C}\mathcal{B} \\ 0 & \ldots & \ldots & 0 & I \end{bmatrix}}_{\mathcal{E}_q} E^-_{t,q}. \tag{20}$$

This equation defines $\tilde{\mathcal{O}}_{q,u}, \tilde{\mathcal{O}}_{q,\bullet}$ and $\mathcal{E}_q$. Consider a vector polynomial $\beta(z) = \sum_{m=0}^{q} \beta_m z^m, \beta_m \in \mathbb{R}^s, m = 0, \ldots, q$ and

$$\begin{aligned} \beta(L)' y_{t,p} &= \sum_{m=0}^{q} \beta'_m y_{t-m,p} = \begin{bmatrix} \beta'_0 \beta'_1 \ldots \beta'_q \end{bmatrix} Y^-_{t,q} \\ &= \vec{\beta}' \left(\tilde{\mathcal{O}}_{q,u} x_{t-q,u} + \tilde{\mathcal{O}}_{q,\bullet} x_{t-q,\bullet} + \mathcal{E}_q E^-_{t,q} \right), \end{aligned} \tag{21}$$

where $\vec{\beta} := [\beta'_0, \beta'_1, \ldots, \beta'_q]'$.

By construction the latter two terms, $\vec{\beta}'(\tilde{\mathcal{O}}_{q,\bullet} x_{t-q,\bullet} + \mathcal{E}_q E^-_{t,q})$, are stationary. Thus, the unit root structure of $\beta(L)'(y_{t,p})_{t\in\mathbb{Z}}$ is identical to the unit root structure of $(\vec{\beta}'\tilde{\mathcal{O}}_{q,u} x_{t,u})_{t\in\mathbb{Z}}$, where both processes can be shown to possess a unit root structure (see Theorem 4 below).

Consequently, the key to understanding the polynomial cointegration properties of $(y_{t,p})_{t\in\mathbb{Z}}$ is to understand the integration properties of linear transformations

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

of $(\tilde{\mathcal{O}}_{q,u}x_{t,u})_{t\in\mathbb{Z}}$. This in turn is facilitated by the result of Theorem 3 showing that the individual blocks $(x^j_{t,k})_{t\in\mathbb{Z}}$ have unit root structures $((\omega_k, h_k - j + 1))$ and are not cointegrated. This implies that a *necessary* condition for polynomial cointegration of a certain order, i.e., a corresponding reduction in the integration orders, is orthogonality (of $\vec{\beta}$) to sub-blocks of $\tilde{\mathcal{O}}_{q,u}$. Thus, partition

$$\tilde{\mathcal{O}}_{q,u} = \left[\tilde{\mathcal{O}}^1_{q,1}, \tilde{\mathcal{O}}^2_{q,1}, \ldots, \tilde{\mathcal{O}}^{h_1}_{q,1}, \tilde{\mathcal{O}}^1_{q,2}, \ldots, \tilde{\mathcal{O}}^{h_l}_{q,l}\right], \tag{22}$$

with $\tilde{\mathcal{O}}^j_{q,k} \in \mathbb{R}^{(q+1)s\times d^k_j}$ for k such that $\omega_k \in \{0, \pi\}$ and $j = 1, \ldots, h_k$ and $\tilde{\mathcal{O}}^j_{q,k} \in \mathbb{R}^{(q+1)s\times 2d^k_j}$ for k such that $0 < \omega_k < \pi$ and $j = 1, \ldots, h_k$. Using the above partitioning, the following theorem now clarifies the reductions in integration orders achieved by dynamic linear transformations applied to $(y_{t,p})_{t\in\mathbb{Z}}$.

THEOREM 4. *Let $(y_{t,p})_{t\in\mathbb{Z}}$ be a unit root process with unit root structure $\Omega = ((\omega_1, h_1), \ldots, (\omega_l, h_l))$ defined as the particular solution (8) of the state space equations with the canonical state space representation $(\mathcal{A}, \mathcal{B}, \mathcal{C})$ as given in Theorem 2 and state process $(x_t)_{t\in\mathbb{Z}}$ as given in Theorem 3. Let again $\beta(L) = \sum_{m=0}^q \beta_m z^m$, with $\beta_m \in \mathbb{R}^s$ for $m = 0, \ldots, q$ and $\vec{\beta} = [\beta_0', \ldots, \beta_q']'$.*

(i) Then the process $\beta(L)'(y_{t,p})_{t\in\mathbb{Z}}$ has unit root structure $\tilde{\Omega} = ((\tilde{\omega}_1, \tilde{h}_1), \ldots, (\tilde{\omega}_{\tilde{l}}, \tilde{h}_{\tilde{l}})) \preceq \Omega$ if and only if:

- *For each k such that $\delta_k(\Omega, \tilde{\Omega}) = 0$, it holds that $\vec{\beta}'\tilde{\mathcal{O}}^1_{q,k} \neq 0$.*
- *For each k such that $0 < \delta_k(\Omega, \tilde{\Omega}) < h_k$, it holds that $\vec{\beta}'\tilde{\mathcal{O}}^j_{q,k} = 0$ for $j = 1, \ldots, \delta_k(\Omega, \tilde{\Omega})$ and $\vec{\beta}'\tilde{\mathcal{O}}^{\delta_k(\Omega,\tilde{\Omega})+1}_{q,k} \neq 0$.*
- *For each k such that $\delta_k(\Omega, \tilde{\Omega}) = h_k$, it holds that $\vec{\beta}'\tilde{\mathcal{O}}^j_{q,k} = 0$ for $j = 1, \ldots, h_k$.*

(ii) Denote again with H the degree of the polynomial $D(z)$ as given in Definition 1. Then there exists a vector polynomial $\beta(z) = \sum_{m=0}^q \beta_m z^m$, $\beta_m \in \mathbb{R}^s$, $m = 0, \ldots, q$ such that $\beta(L)'(y_{t,p})_{t\in\mathbb{Z}}$ has unit root structure $\tilde{\Omega}$ if and only if there exists a vector polynomial $\tilde{\beta}_{H-1}(z) = \sum_{m=0}^{H-1} \tilde{\beta}_m z^m$, $\tilde{\beta}_m \in \mathbb{R}^s$, $m = 0, \ldots, H-1$ such that $\tilde{\beta}_{H-1}(L)'(y_{t,p})_{t\in\mathbb{Z}}$ has unit root structure $\tilde{\Omega}$.

The conditions formulated in Theorem 4 do not incorporate the condition $\max_{k=1,\ldots,l} ||\beta(e^{i\omega_k})||\delta_k(\Omega, \tilde{\Omega}) \neq 0$ given in Definition 4. Thus, clearly the theorem only characterizes necessary conditions for polynomial cointegration. The second item shows that only polynomials up to degree $H-1$ have to be considered. A detailed characterization of polynomial cointegration, including the definition of *minimal degree* PCIV, is presented elsewhere due to space constraints. Let us here only briefly compare a few aspects of polynomial cointegration as defined in this paper with the definition of Gregoir (1999a). First, the definition of polynomial cointegration in Definition 4 extends the notion of multicointegration as defined by Yoo (1986) and Granger and Lee (1989a, 1989b). This is not the

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

case for polynomial cointegration as defined in Gregoir (1999a, Def. 3.1). Second, the definition of Gregoir includes only one unit root frequency. This implies that, in cases where the considered process is integrated at more than one frequency, the stochastic properties of the resulting dynamically transformed process are unclear, since possible reductions of the integration orders at the other frequencies are left unspecified. Third, the definition of Gregoir links the order of polynomial cointegration to the polynomial degree of the corresponding polynomial. This appears to be unsatisfactory since the relevant property of the PCIV is the achieved reduction in the unit root structure of the transformed process and not the polynomial degree of the PCIV. Fourth, the definition of Gregoir implies that a process that is polynomially cointegrated of degree m is also polynomially cointegrated of degrees $j = 1, \dots, m-1$. Thus, that definition does not include any notion of maximality of the reduction in the integration order.

6. SOME ILLUSTRATIVE EXAMPLES

In this section we illustrate the results of this paper for the four most widely used cases, namely I(1) AR and ARMA processes, an I(2) example (which we use to exemplify the construction of the canonical representation), and MFI(1) processes.

Remark 2. To facilitate comparison with related results in the literature, especially in the AR literature, we include non-zero (but zero mean) states $x_{1,u}$ in the discussion. It is straightforward to see that this adds a term $\mathcal{C}_u \mathcal{A}_u^{t-1} x_{1,u}, t \in \mathbb{Z}$ to the output process. Note that all results stated, in particular the integration and cointegration properties of both the components of the state process as well as the resulting output process, are not affected by this addition.

Furthermore, in the discussion of the examples we focus on the solution only for $t \in \mathbb{N}$, as is commonly done in large parts of the cointegration literature. This is typically accompanied in the AR cointegration literature by conditioning on initial observations $y_0, \dots, y_{1-H}$, which in our setting corresponds to conditioning on $x_{1,u}$. All results discussed below hold for the processes defined for $t \in \mathbb{Z}$.

6.1. I(1) AR Processes in the State Space Framework

This section is devoted to an illustration of the links between the autoregressive and state space frameworks using the well-known notation of Johansen (1995) for AR processes. Johansen (1995, (2.1), p. 11) uses AR processes defined as the solution of

$$y_t = \Pi_1 y_{t-1} + \Pi_2 y_{t-2} + \cdots + \Pi_p y_{t-p} + \varepsilon_t, \tag{23}$$

with $\Pi_j \in \mathbb{R}^{s \times s}$ for $j = 1, \dots, p$ for some starting values $y_0, \dots, y_{-p+1}$, abstaining here from deterministic components.

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

One state space representation (referred to as companion form in the AR literature, see, e.g., Johansen, 1995, p. 15) for AR processes is given by

$$y_t = \underbrace{\begin{bmatrix}\Pi_1 & \Pi_2 & \dots & \Pi_p\end{bmatrix}}_{C} x_t + \varepsilon_t,$$

$$x_{t+1} = \underbrace{\begin{bmatrix} \Pi_1 & \Pi_2 & \dots & \dots & \Pi_p \\ I_s & 0 & \cdots & \cdots & 0 \\ 0 & I_s & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_s & 0 \end{bmatrix}}_{A} x_t + \underbrace{\begin{bmatrix} I_s \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{B} \varepsilon_t.$$

with $x_t := [y'_{t-1}, \dots, y'_{t-p}]' \in \mathbb{R}^{ps}$ for a given initial state $x_1 := [y'_0, \dots, y'_{-p+1}]'$.

Under the assumption that the above AR equation (23) has an I(1) solution (see below for the corresponding constraints on the parameters), the above state space representation can be transformed to highlight the integration and cointegration properties. Define, parallelling Hansen (2005), with a similar line of thought also used in Phillips (1998), the following matrices (where empty entries indicate zeros):

$$\alpha^* := \begin{bmatrix} \alpha & \Gamma_1 & \dots & \Gamma_{p-1} \\ & I_s & & \\ & & \ddots & \\ & & & I_s \end{bmatrix}, \qquad \beta^* := \begin{bmatrix} \beta & I_s & & & \\ & -I_s & I_s & & \\ & & \ddots & \ddots & \\ & & & -I_s & I_s \\ & & & & -I_s \end{bmatrix},$$

$$\alpha^*_\perp := \begin{bmatrix} \alpha_\perp \\ -\Gamma'_1\alpha_\perp \\ \vdots \\ -\Gamma'_{p-1}\alpha_\perp \end{bmatrix}, \qquad \beta^*_\perp := \begin{bmatrix} \beta_\perp \\ \vdots \\ \beta_\perp \end{bmatrix}.$$

Here, as usual, $\alpha\beta' = \Pi = \sum_{j=1}^p \Pi_j - I_s$, with $\alpha, \beta \in \mathbb{R}^{s\times(s-c)}$ of full rank $s-c$ for some $0 \le c \le s$ and $\Gamma_i := -\sum_{j=i+1}^p \Pi_j$, $i = 1, \dots, p-1$. Using these matrices the above state equation can be written as $x_{t+1} = (I_{ps} + \alpha^*(\beta^*)')x_t + B\varepsilon_t$. Johansen (1995, Thm. 4.2, p. 49) shows that the above equation (23) has an I(1) solution under the assumption that $\det(\alpha'_\perp \Gamma \beta_\perp) \neq 0$, with $\Gamma := I_s - \sum_{j=1}^{p-1} \Gamma_i$ and $\alpha_\perp, \beta_\perp \in \mathbb{R}^{s\times c}$ full rank matrices such that $\alpha'\alpha_\perp = 0$ and $\beta'\beta_\perp = 0$. Under this constraint the matrix $T^{-1} := [\alpha^*_\perp, \beta^*]'$ is invertible; see Hansen (2005, Lem. A.4). Use this transformation matrix to obtain $\tilde{A} = T^{-1}AT$, $\tilde{B} = T^{-1}B$, $\tilde{C} = CT$, and $\tilde{x}_t = T^{-1}x_t$, which results in

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

$$\tilde{A} = \begin{bmatrix} I_c & 0 \\ 0 & I_{ps-c} + (\beta^*)'\alpha^* \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} \alpha'_\perp \\ \beta' \\ I_s \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} \beta_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1} & \tilde{C}_\bullet \end{bmatrix}. \tag{24}$$

Hansen (2005, Lem. A.2) shows that $|\lambda_{\max}(I_{ps-c} + (\beta^*)'\alpha^*)| < 1$. Note that the representation (24) need not be minimal. The index c coincides with d^1 as used in the theorems. Partition the state $\tilde{x}_t = [\tilde{x}_{t,u}, \tilde{x}_{t,\bullet}]$ according to the block structure of $\tilde{A}$. Then, considering the solution of (24) for $t \in \mathbb{N}$ leads to the well-known result

$$y_t = \beta_\perp \left(\alpha'_\perp \Gamma \beta_\perp\right)^{-1} \left(\alpha'_\perp \sum_{j=1}^{t-1} \varepsilon_{t-j} + \tilde{x}_{1,u} \right) + w_t, \tag{25}$$

with $w_t = \varepsilon_t + \tilde{C}_\bullet (I_{ps-c} + (\beta^*)'\alpha^*)^{t-1} \tilde{x}_{1,\bullet} + \sum_{j=1}^{t-1} \tilde{C}_\bullet (I_{ps-c} + (\beta^*)'\alpha^*)^{j-1} \tilde{B} \varepsilon_{t-j}$. By appropriate choice of $\tilde{x}_{1,\bullet}$ the process w_t can be made stationary. This representation (with the specific choice concerning $\tilde{x}_{1,\bullet}$), often referred to as Granger representation (see also Johansen, 1995, Thm. 4.2), shows that the cointegrating space is spanned by the column space of β.

One difference to the Granger representation given in Johansen (1995) is that the stochastic trends are given here by $\alpha'_\perp \sum_{j=1}^{t-1} \varepsilon_{t-j}$ compared to $\alpha'_\perp \sum_{j=0}^{t-1} \varepsilon_{t-j}$ in Johansen (1995). This minor difference stems from the fact that the stochastic trends are in our setup defined as $x_{t,1}$. It does not matter whether ε_t is included or not in the stochastic trend, but our definition implies that the transfer function $k_\bullet(z)$ is normalized to $k_\bullet(0) = I_s$.

Whether a state space representation of a cointegrated I(1) AR process is suitable for structural identification depends upon the nature of the restrictions. In situations where economic theory implies restrictions on the AR coefficients, e.g., zero restrictions, these restrictions need not necessarily be easily incorporated in the corresponding state space representation. This stems from the fact that while the outlined procedure provides a one-to-one mapping between the coefficients in the AR and the state space representation, this mapping is nonlinear. This implies that constraints easily imposed in the AR framework may be difficult to impose in the state space framework. Conversely, when an economic model is solved in the state space framework, the canonical form for state space representations derived in this paper may allow for a simpler and more direct imposition of restrictions than corresponding AR representations.

6.2. I(1) Unit Root Processes

I(1) unit root processes have unit root structure $((0, 1))$ and thus in the canonical representation $\mathcal{A} = \operatorname{diag}(I_d, \mathcal{A}_\bullet)$ for some $d > 0$. The matrices $\mathcal{B}$ and $\mathcal{C}$ are

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

partitioned accordingly, i.e., $\mathcal{B} = [\mathcal{B}_1', \mathcal{B}_\bullet']'$ and $\mathcal{C} = [\mathcal{C}_1, \mathcal{C}_\bullet]$ with $\mathcal{B}_1 \in \mathbb{R}^{d\times s}$ and $\mathcal{C}_1 \in \mathbb{R}^{s\times d}$, with s again denoting the dimension of the process considered. The restrictions formulated in Theorem 2 require that $\mathcal{C}_1'\mathcal{C}_1 = I_d$ and that $\mathcal{B}_1$ is p.u.t. and has full row rank. Furthermore, the stationary subsystem $(\mathcal{A}_\bullet, \mathcal{B}_\bullet, \mathcal{C}_\bullet)$ is in echelon canonical form. With the state partitioned accordingly in $x_t = [x_{t,1}', x_{t,\bullet}']'$, it is assumed that $x_{t,\bullet} = \sum_{j=1}^{\infty} \mathcal{A}_\bullet^{j-1}\mathcal{B}_\bullet\varepsilon_{t-j}$, $t \in \mathbb{Z}$. Consider as in the previous subsection $x_{1,1}$ a zero mean random variable, then similarly to (8) the solution of the state space equations is given by

$$\tilde{y}_{t,p} = \begin{cases} \mathcal{C}_1 x_{1,1} + \mathcal{C}_1\mathcal{B}_1 \sum_{j=1}^{t-1} \varepsilon_{t-j} + w_t, & t \geq 1, \\ \mathcal{C}_1 x_{1,1} - \mathcal{C}_1\mathcal{B}_1 \sum_{j=t}^{0} \varepsilon_{t-j} + w_t, & t < 1, \end{cases} \tag{26}$$

where here $w_t = k_\bullet(L)\varepsilon_t = \varepsilon_t + \sum_{j=1}^{\infty} \mathcal{C}_\bullet\mathcal{A}_\bullet^{j-1}\mathcal{B}_\bullet\varepsilon_{t-j}$. The solution for $t \geq 1$ extends the Granger representation for I(1) processes as given for AR processes in Johansen (1995, Thm. 4.2). Denote with $\mathcal{C}_{1,\perp} \in \mathbb{R}^{s\times(s-c)}$ a full rank matrix such that $\mathcal{C}_{1,\perp}'\mathcal{C}_{1,\perp} = I_{s-c}$ and $\mathcal{C}_1'\mathcal{C}_{1,\perp} = 0$. Then $(\mathcal{C}_{1,\perp}'\tilde{y}_{t,p})_{t\in\mathbb{Z}} = (\mathcal{C}_{1,\perp}'w_t)_{t\in\mathbb{Z}}$ and is hence stationary, i.e., the column space of $\mathcal{C}_{1,\perp}$ equals the cointegrating space.

A parameterization of $\mathcal{B}_1$ and $\mathcal{C}_1$ has to take into account the restrictions formulated in the canonical form. For given order n, d and locations of the restricted entries in the matrix $\mathcal{B}_1$ in p.u.t. form, the set of all matrices $C \in \mathbb{R}^{s\times d}$ such that $C'C = I_d$ as well as the set of all B matrices with the given p.u.t. structure have to be parameterized (with entries that are either positive or unrestricted). This setup allows one to incorporate restrictions on the cointegrating space, e.g., restricting certain coordinates of $(\tilde{y}_{t,p})_{t\in\mathbb{Z}}$ to be stationary or certain directions to lie in the cointegrating space, straightforwardly in the parameterization. For the stationary subsystem a parameterization that is suitable for usage with the echelon canonical form can be chosen.

6.3. An I(2) Example

In this section we consider a trivariate state space system corresponding to an I(2) process to illustrate how the constraints formulated in Theorems 1 and 2 lead to a canonical representation and to also highlight the advantages of the developed canonical form for cointegration analysis.

Thus, consider

$$A = \operatorname{diag}(J_1, A_\bullet) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0.5 \end{bmatrix}, \qquad B = \begin{bmatrix} B_1 \\ B_\bullet \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

$$C = [C_1, C_\bullet] = \begin{bmatrix} 1 & 0 & 1 & 0.2 \\ 0 & 1 & 1 & -0.4 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

Denote, using the notation established in the theorems, the first three columns of C by $C_1 = [C_1^{1,E}\ C_1^{2,G}\ C_1^{2,E}]$, in the example given by

$$C_1^{1,E} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad C_1^{2,G} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad C_1^{2,E} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

The assumption that the stable part of the system is in echelon canonical form leads to $\mathcal{A}_\bullet = 0.5$, $\mathcal{B}_\bullet = (0.2,\ 0,\ 0)$, and $\mathcal{C}_\bullet = (1,\ -2,\ 5)'$.

Next note that the matrix J_1 is already in the format given in (10). Consequently, in order to preserve this format by pre- and post-multiplication with nonsingular transformation matrices T_1, these matrices need to have the following structure (cf. Gantmacher, 1966, (14), p. 231):

$$T_1^{-1} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ 0 & t_{11} & 0 \\ 0 & t_{32} & t_{33} \end{bmatrix}, \quad \text{with } t_{11} \neq 0, t_{33} \neq 0. \tag{27}$$

Applying such a transformation matrix T_1^{-1} to the matrix C_1 gives

$$\tilde{C}_1 = C_1 T_1^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ 0 & t_{11} & 0 \\ 0 & t_{32} & t_{33} \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12}+t_{32} & t_{13}+t_{33} \\ 0 & t_{11}+t_{32} & t_{33} \\ 0 & t_{32} & t_{33} \end{bmatrix}. \tag{28}$$

Denote the columns of $\tilde{C}_1 = [\tilde{C}_1^{1,E}\ \tilde{C}_1^{2,G}\ \tilde{C}_1^{2,E}]$. The restrictions formulated for the canonical form in item (iii) of Theorem 2 are in the present example given by $(\tilde{C}_1^{1,E})'\tilde{C}_1^{1,E} = 1$, $(\tilde{C}_1^{2,E})'\tilde{C}_1^{2,E} = 1$, $(\tilde{C}_1^{2,E})'\tilde{C}_1^{1,E} = 0$, and $(\tilde{C}_1^{2,G})'[\tilde{C}_1^{1,E}\ \tilde{C}_1^{2,E}] = 0$. These restrictions imply $t_{11} = \pm 1$, $t_{13} = -t_{33}$, $t_{12} = -t_{32}$, and $t_{11} = -2t_{32}$. Thus, in order for these constraints to be fulfilled in $\tilde{C}_1$, a transformation matrix T_1 has to be of the following structure:

$$T_1 = \begin{bmatrix} t_{11} & 0 & t_{11} \\ 0 & t_{11} & 0 \\ 0 & 1/(2t_{33}) & 1/t_{33} \end{bmatrix}, \qquad T_1^{-1} = \begin{bmatrix} t_{11} & t_{11}/2 & -t_{33} \\ 0 & t_{11} & 0 \\ 0 & -t_{11}/2 & t_{33} \end{bmatrix}$$

with $t_{11} = \pm 1, t_{33} = \pm 1/\sqrt{2}$.

Now the restrictions formulated in item (iv) of Theorem 2, i.e., the p.u.t. structure of the second and third rows of $\tilde{B}_1 = T_1 B_1$, determine a unique transformation matrix (with $t_{11} = 1$, $t_{33} = 1/\sqrt{2}$)

$$T_1^* = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1/\sqrt{2} & \sqrt{2} \end{bmatrix}.$$

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

Therefore, the canonical representation is obtained by transforming the system (A, B, C) with the unique transformation matrix $T^* = \mathrm{diag}(T_1^*, T_\bullet^*)$ and is given by

$$\mathcal{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0.5 \end{bmatrix}, \qquad \mathcal{B} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 3/\sqrt{2} & \sqrt{2} & 1/\sqrt{2} \\ 0.2 & 0 & 0 \end{bmatrix}, \qquad \mathcal{C} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1/2 & 1/\sqrt{2} & -2 \\ 0 & -1/2 & 1/\sqrt{2} & 5 \end{bmatrix}.$$

The canonical representation is well suited for studying the cointegration properties of the corresponding unit root solution process $(\tilde{y}_{t,p})_{t\in\mathbb{Z}}$. As in the other examples we consider only the solution for $t \in \mathbb{N}$ for the value $x_1 = [x_{1,1}', x_{1,\bullet}]'$ with $x_{1,\bullet} = \sum_{j=1}^{\infty} \mathcal{A}_\bullet^{j-1} \mathcal{B}_\bullet \varepsilon_{1-j} = 0.2 \sum_{j=1}^{\infty} 0.5^{j-1} \varepsilon_{1-j,1}$, with $\varepsilon_{t,1}$ denoting the first coordinate of ε_t. Denoting the first three rows of $\mathcal{B}$ by $\mathcal{B}_1^j$, $j = 1, 2, 3$, we obtain

$$\begin{aligned} \tilde{y}_{t,p} &= \mathcal{C}_1 J_1^{t-1} x_{1,1} + \mathcal{C}_1^{1,E} \mathcal{B}_1^2 \sum_{i=1}^{t-1} \sum_{j=1}^{i-1} \varepsilon_{i-j} + \underbrace{\left(\mathcal{C}_1^{1,E} \mathcal{B}_1^1 + \mathcal{C}_1^{2,G} \mathcal{B}_1^2 + \mathcal{C}_1^{2,E} \mathcal{B}_1^3 \right)}_{\mathcal{C}_1 \mathcal{B}_1} \\ &\quad \times \sum_{j=1}^{t-1} \varepsilon_{t-j} + w_t, \qquad (29) \\ &= \begin{pmatrix} 1 & (t-1) & 0 \\ 0 & 1/2 & 1/\sqrt{2} \\ 0 & -1/2 & 1/\sqrt{2} \end{pmatrix} x_{1,1} + \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \sum_{i=1}^{t-1} \sum_{j=1}^{i-1} \varepsilon_{i-j} + \begin{pmatrix} 2 & 1 & 0 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \sum_{j=1}^{t-1} \varepsilon_{t-j} + w_t, \end{aligned}$$

with $w_t = k_\bullet(L)\varepsilon_t = \varepsilon_t + \sum_{j=1}^{\infty} \mathcal{C}_\bullet \mathcal{A}_\bullet^{j-1} \mathcal{B}_\bullet \varepsilon_{t-j} = \begin{pmatrix} 0.2 \\ -0.4 \\ 1 \end{pmatrix} \sum_{j=1}^{\infty} 0.5^{j-1} \varepsilon_{t-j,1}$ stationary. The representation (29) immediately displays cointegration and is analogous to the representation developed for the autoregressive case in (Johansen, 1997, (10), p. 437). The difference is again that the stochastic trends in the canonical state space representation only include errors up to ε_{t-1}.

Concerning the state process corresponding to this representation it holds (cf. Theorem 3 or straightforward computation in this example) that in $x_t = [x_{t,1}^1, x_{t,1}^2, x_{t,1}^3, x_{t,\bullet}]'$, the first coordinate $(x_{t,1}^1)_{t\in\mathbb{Z}}$ is I(2), $(x_{t,1}^2)_{t\in\mathbb{Z}}$ and $(x_{t,1}^3)_{t\in\mathbb{Z}}$ are both I(1) and not cointegrated and $(x_{t,\bullet})_{t\in\mathbb{Z}}$ is stationary.

Every vector $\beta \in \mathbb{R}^3$, $\beta \neq 0$ such that $\beta' \mathcal{C}_1^{1,E} = 0$ reduces the integration order in $(\beta' \tilde{y}_{t,p})_{t\in\mathbb{Z}}$ from 2 to 1. Note that, since in the example $\mathcal{C}_1\mathcal{B}_1$ has full rank, no cointegrating vectors that reduce the integration order from 2 to 0 exist.

Consider next polynomial cointegration. According to Theorem 4 polynomials $\beta(L)$ that result in a stationary process $\beta(L)'(\tilde{y}_{t,p})_{t\in\mathbb{Z}}$ can be found with $q = 1$ using the equation

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

$$[\beta_0', \beta_1'] \begin{bmatrix} \mathcal{C}_1 J_1 \\ \mathcal{C}_1 \end{bmatrix} = 0.$$

A basis for the corresponding three-dimensional solution space is given by the following three vectors:

$$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}.$$

The first two vectors correspond to the polynomials $\beta(z) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \Delta_0(z)$ and $\beta(z) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \Delta_0(z)$; i.e., they are just multiples of the differencing filter $\Delta_0(z)$. Thus they do not fulfill condition (ii) of Definition 4 and are consequently not PCIVs. The third solution vector is a PCIV in the sense of Definition 4 and is given by

$$\beta(z) = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} z = \begin{bmatrix} \Delta_0(z) \\ -1 \\ 1 \end{bmatrix}.$$

This PCIV describes a cointegrating relationship between the first difference of the first coordinate and the second and third coordinates of $(\tilde{y}_{t,p})_{t\in\mathbb{Z}}$, all of which are I(1) processes. Thus, it is a multicointegrating relationship in the sense of Granger and Lee (1989a). By reparameterizing a polynomial $\beta(z) = \beta_0 + \beta_1 z$ as $\tilde{\beta}_0 + \tilde{\beta}_1 \Delta_0(z)$, one can directly focus on the multicointegration formulation as given by Granger and Lee (1989a).

The multicointegrating relationship can of course also be computed directly from either the solution given in (29) or by looking at the solution in terms of the state. For brevity we consider $y_{t,p}$; i.e., we ignore the first term in (29). As discussed, reduction from I(2) to I(1) requires that $\tilde{\beta}_0' \mathcal{C}_1^{1,E} = 0$; i.e., the first coordinates of $y_{t,p}$ needs to be wiped out. The I(1) components of the second and third coordinates of $y_{t,p}$ are given by $1/2x_{t,1}^2 + 1/\sqrt{2}x_{t,1}^3$ and $-1/2x_{t,1}^2 + 1/\sqrt{2}x_{t,1}^3$. Only the first coordinate of $\Delta_0(L)y_{t,p}$ is an I(1) process and its I(1) component is given by $x_{t-1,1}^2$. Combining these observations with the properties of the state in the canonical representation given in Theorem 3, which in the example imply no cointegration between $(x_{t,1}^2)_{t\in\mathbb{Z}}$ and $(x_{t,1}^3)_{t\in\mathbb{Z}}$, leads to the same result concerning multicointegration as computed above.

We close the discussion of the I(2) example by showing, as has been mentioned throughout the text, that the state space framework allows for simple

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

parameterizations of minimal state space systems. Thus, we now want to parameterize all state space systems of order $n = 4$ for three-dimensional unit root processes with state space unit root structure given according to the matrix $\mathcal{A}$. Hence, we now consider the matrices $\mathcal{B}$ and $\mathcal{C}$ as unknown. Since our focus is cointegration analysis we only discuss a parameterization of $\mathcal{B}_1$ and $\mathcal{C}_1$ and do not consider the stable part in detail. The constraints formulated in item (iii) of Theorem 2 are in the present situation given by $(\mathcal{C}_1^{1,E})'\mathcal{C}_1^{1,E} = 1$, $(\mathcal{C}_1^{2,E})'\mathcal{C}_1^{2,E} = 1$, $(\mathcal{C}_1^{2,E})'\mathcal{C}_1^{1,E} = 0$, and $(\mathcal{C}_1^{2,G})'[\mathcal{C}_1^{1,E}\ \mathcal{C}_1^{2,E}] = 0$. Given that all three vectors are three-dimensional, a parameterization of $\mathcal{C}_1^{1,E}$ requires two parameters using, e.g., polar coordinates when taking the constraint into account. Restricting a coordinate of $(y_t)_{t\in\mathbb{Z}}$ to be I(1) rather than I(2) corresponds to a zero restriction and thus reduces the parameter set to be one-dimensional. Since $\mathcal{C}_1^{2,E}$ needs to be normalized and orthogonal to $\mathcal{C}_1^{1,E}$ the corresponding parameter space is one-dimensional. Restricting a coordinate of $(y_t)_{t\in\mathbb{Z}}$ to be stationary requires the corresponding row of $\mathcal{C}_1$ to be equal to 0, i.e., introduces further orthogonality constraints. The remaining vector is determined up to a scalar factor from the first two vectors via orthogonality, thus $\mathcal{C}_1^{2,G} = \gamma\,\overline{\mathcal{C}}_1^{2,G}$ with $(\overline{\mathcal{C}}_1^{2,G})'\overline{\mathcal{C}}_1^{2,G} = 1$ determined from $\mathcal{C}_1^{1,E}$ and $\mathcal{C}_1^{2,E}$ (up to the orientation, which can be fixed by choosing the first nonzero entry to be positive). A necessary and sufficient condition for the presence of multicointegration is that $\gamma \neq 0$; compare also the discussion in the example above. This leads to a simple null hypothesis for the occurrence of multicointegration. Furthermore, the second and third rows of $\mathcal{B}_1$ need to be p.u.t. with one parameterization for such matrices discussed, e.g., in Bauer and Wagner (2006, Lem. 10).

6.4. Multiple Frequency I(1) Processes

As a final example we consider multiple frequency I(1) processes, i.e., unit root processes with unit root structure $((\omega_1, 1), \ldots, (\omega_l, 1))$. The primary use of MFI(1) processes is for seasonal unit root and cointegration analysis, where, e.g., for quarterly observations unit roots may be present at the frequencies 0, $\pi/2$ and π (see Ahn and Reinsel, 1994; Johansen and Schaumburg, 1999).

According to Theorem 2 a minimal state space representation of the following form exists for $(y_{t,p})_{t\in\mathbb{Z}}$:

$$y_{t,p} = \begin{bmatrix}\mathcal{C}_1 \cdots \mathcal{C}_l\ \mathcal{C}_\bullet\end{bmatrix} x_t + \varepsilon_t, \tag{30}$$

$$\begin{bmatrix} x_{t+1,1} \\ \vdots \\ x_{t+1,l} \\ x_{t+1,\bullet} \end{bmatrix} = \begin{bmatrix} \mathcal{A}_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \mathcal{A}_l & 0 \\ 0 & \ldots & 0 & \mathcal{A}_\bullet \end{bmatrix} \begin{bmatrix} x_{t,1} \\ \vdots \\ x_{t,l} \\ x_{t,\bullet} \end{bmatrix} + \begin{bmatrix} \mathcal{B}_1 \\ \vdots \\ \mathcal{B}_1 \\ \mathcal{B}_\bullet \end{bmatrix} \varepsilon_t. \tag{31}$$

For k such that $0 < \omega_k < \pi$ the subsystems $(\mathcal{A}_k, \mathcal{B}_k, \mathcal{C}_k)$ are related to pairs of complex conjugate matrices as in (12) and for k such that $\omega_k \in \{0, \pi\}$ there

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

is no distinction required between real and complex system matrices, as summarized for notational clarity in (16). The corresponding matrices $J_k = \overline{z_k} I_{d_k}$ for $k = 1, \ldots, l$ and the additional constraints imposed by the canonical representation are $(\mathcal{C}_{k,-})'\mathcal{C}_{k,-} = I_{d^k}$ and $\mathcal{B}_{k,-}$ is p.u.t. and has full row rank for $k = 1, \ldots, l$. The stable subsystem $(\mathcal{A}_\bullet, \mathcal{B}_\bullet, \mathcal{C}_\bullet)$ is in echelon canonical form.

In the MFI(1) case it holds for frequencies $0 < \omega_k < \pi$ that $\mathcal{A}_k = I_{d^k} \otimes Q(\omega_k)$. Consider here the case that $\omega_1 = 0$ and $\omega_l = \pi$; then the solution (again considered only for $t \in \mathbb{N}$ and $x_1 = [x'_{1,u}, x'_{1,\bullet}]'$) of the above state space system is given by

$$\begin{aligned}\tilde{y}_{t,p} = {} & \mathcal{C}_1\mathcal{B}_1 \sum_{j=1}^{t-1} \varepsilon_{t-j} + \sum_{j=1}^{t-1} \mathcal{C}_2 \left(I_{d^2} \otimes Q(\omega_2)^{j-1}\right)\mathcal{B}_2\varepsilon_{t-j} \\ & + \cdots + \sum_{j=1}^{t-1} \mathcal{C}_{l-1} \left(I_{d^{l-1}} \otimes Q(\omega_{l-1})^{j-1}\right)\mathcal{B}_{l-1}\varepsilon_{t-j} \\ & + \mathcal{C}_l\mathcal{B}_l \sum_{j=1}^{t-1} (-1)^{j-1}\varepsilon_{t-j} + \tilde{S}_t + w_t, \end{aligned} \tag{32}$$

with $\tilde{S}_t = \mathcal{C}_1 x_{1,1} + \mathcal{C}_2(I_{d^2} \otimes Q(\omega_2)^{t-1})x_{1,2} + \cdots + \mathcal{C}_l(-1)^{t-1}x_{1,l}$ summarizing the effect of the state components $x_{1,u} = [x'_{1,1}, \ldots, x'_{1,l}]'$ and $w_t = \varepsilon_t + \sum_{j=1}^{\infty} \mathcal{C}_\bullet \mathcal{A}_\bullet^{j-1} \mathcal{B}_\bullet \varepsilon_{t-j}$, with again $x_{1,\bullet} = \sum_{j=1}^{\infty} \mathcal{A}_\bullet^{j-1}\mathcal{B}_\bullet\varepsilon_{1-j}$.

Thus, the canonical state space representation immediately leads to a Granger-type representation decomposing the process $(\tilde{y}_{t,p})_{t\in\mathbb{Z}}$ into the stochastic trends $(\mathcal{C}_1\mathcal{B}_1 \sum_{j=1}^{t-1} \varepsilon_{t-j})$ and cycles $(\mathcal{C}_k \sum_{j=1}^{t-1} (I_{d^k} \otimes Q(\omega_k)^{j-1})\mathcal{B}_k\varepsilon_{t-j}$, for $k = 2, \ldots, l-1$, $\mathcal{C}_l\mathcal{B}_l \sum_{j=1}^{t-1}(-1)^{j-1}\varepsilon_{t-j})$ at the different unit root frequencies, the effects of the state $x_{1,u}$ $(\tilde{S}_t)$ and a stationary process (w_t). The relationship between the state space system matrices and the Granger representation (32) is much simpler than, e.g., the relation between AR coefficients and the corresponding Granger representation as derived for the AR case in Johansen and Schaumburg (1999). This implies that it is easier to define a parameterization in the state space framework than in the AR or ARMA framework; see also Bauer and Wagner (2006).

Theorem 3 shows that in the MFI(1) case the stochastic trends respectively cycles $(x_{t,k})_{t\in\mathbb{Z}}$ are not cointegrated. Thus, cointegration occurs for vectors $\beta \neq 0$ such that $\beta'\mathcal{C}_k = 0$ for at least one $k = 1, \ldots, l$.

In the MFI(1) case additionally so-called *dynamic cointegrating* relationships, i.e., polynomial cointegrating relationships of polynomial degree 1 may occur, for unit root frequencies $0 < \omega_k < \pi$. Let $\beta(L) = \beta_0 + \beta_1 L$, with $\beta_0, \beta_1 \in \mathbb{R}^s$ and $\beta_0, \beta_1 \neq 0$. Considering one unit root frequency we obtain

$$\left(\beta_0' + \beta_1' L\right)\mathcal{C}_k x_{t,k} = \beta_0'\mathcal{C}_k\varepsilon_{t-1} + \left[\beta_0'\mathcal{C}_k(I_{d^k} \otimes Q(\omega_k)) + \beta_1'\mathcal{C}_k\right] x_{t-1,k},$$

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

using $x_{t,k} = (I_{d^k} \otimes Q(\omega_k))x_{t-1,k} + \mathcal{B}_k \varepsilon_{t-1}$. Therefore, the polynomial $\beta(L)$ eliminates the stochastic cycle $(x_{t,k})_{t\in\mathbb{Z}}$ in $\beta(L)'(\tilde{y}_{t,p})_{t\in\mathbb{Z}}$ and is a PCIV, since it automatically fulfills condition (ii) of Definition 4, if and only if

$$[\beta_0' \; \beta_1'] \begin{bmatrix} \mathcal{C}_k(I_{d^k} \otimes Q(\omega_k)) \\ \mathcal{C}_k \end{bmatrix} = 0.$$

The above matrix of dimension $2s \times 2d^k$, with $d^k \leq s$ (cf. Theorem 1), has full column rank $2d^k$ due to minimality. This leaves space for a $2(s-d^k)$-dimensional subspace of dynamic cointegrating relationships.

7. SUMMARY AND CONCLUSIONS

This paper has demonstrated that the state space framework is convenient for representing unit root processes, where the definition of unit root processes is confined in this paper to ARMA processes to allow for finite order state space representations. A canonical form for unit root processes with unit roots with integer integration orders at any frequency that clearly reveals the integration, cointegration, and polynomial cointegration properties of the process has been developed. This representation is convenient for cointegration analysis because of the particularly simple unit root and cointegration properties of the associated state process. These properties depend to a large extent on the eigenvalue structure of the sub-blocks of the $\mathcal{A}$-matrix in the canonical representation. The corresponding information is summarized in the newly defined state space unit root structure. The paper has also provided a new definition of polynomial cointegration that overcomes some discussed limitations of existing definitions and coincides in the I(2) case with the definitions of multicointegration of Yoo (1986) and Granger and Lee (1989a, 1989b). As has been mentioned, a detailed discussion of polynomial cointegration based on the developed canonical representation as well as a comparison to existing definitions is provided elsewhere.

To illustrate that the developed state space representation is, compared to the ARMA framework, advantageous for understanding the dynamic properties of unit root processes, several examples that cover the cases of major interest in empirical analysis have been discussed. The discussion highlights that the state space representation is closely related to the so-called Granger representation. Already in the I(2) and MFI(1) cases the relation of the AR parameters to the Granger representation is rather complicated, compare Johansen (1997) and Johansen and Schaumburg (1999), whereas it is an immediate result of the state space representation for any unit root structure. Furthermore, the I(2) and MFI(1) examples highlight another advantage of the state space representation, namely that (static and polynomial) cointegrating relationships can simply be recovered via orthogonality relationships. Related to the previous observation note furthermore that the state space representation allows one to easily restrict the unit root structure of coordinates of $(y_t)_{t\in\mathbb{Z}}$ via zero restrictions on certain entries of the $\mathcal{C}$-matrix of the

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

canonical representation. Imposing such restrictions is a much more challenging task in the ARMA framework.

NOTES

1. To avoid dealing with such complexities Gregoir (1999b) considers a simple estimator based on principal components. As discussed in Gregoir (1999b, Sec. 1), such estimators are not efficient. A simple parameterization, e.g., based on the canonical form developed in the present paper, will allow one to study properties of pseudo maximum likelihood estimators and their efficiency properties for processes with unit roots at arbitrary frequencies with integer integration orders.

2. The results obtained in Bauer and Wagner (2006) for unit root processes where the integration orders of all unit roots are equal to one show that the resultant parameterization indeed allows for the implementation of numerically simple pseudo maximum likelihood estimation procedures.

3. The integer $m \leq p + H$ accounts for lags potentially necessary in the multivariate case due to the usage of the Smith form for $\tilde{a}(z)$; see HD for details.

4. Note that a similar idea has been used in Sims, Stock, and Watson (1990) for I(d) processes, where in that paper the definition of I(d) processes is not restricted to ARMA processes (compare Remark 1).

5. Note for completeness that q as used here is in no connection with the moving average order used above.

REFERENCES

Ahn, S.K. & G.C. Reinsel (1994) Estimation of partially nonstationary vector autoregressive models with seasonal behavior. *Journal of Econometrics* 62, 317–350.

Aoki, M. (1987) *State Space Modeling of Time Series*. Springer.

Aoki, M. & A. Havenner (1989) A method for approximate representation of vector valued time series and its relation to two alternatives. *Journal of Econometrics* 42, 181–199.

Aoki, M. & A. Havenner (1991), State space modeling of multiple time series. *Econometric Reviews* 10(1), 1–59.

Bauer, D. & M. Wagner (2006) Asymptotic Theory of Pseudo Maximum Likelihood Estimators for Multiple Frequency I(1) Processes. Mimeo.

Bauer, D. & M. Wagner (2007) Autoregressive Approximations of Multiple Frequency I(1) Processes. Mimeo.

Gantmacher, F. (1966) *Matrizentheorie*. Springer.

Granger, C.W.J. & T. Lee (1989a) Investigation of production, sales and inventory relationship using multicointegration and non-symmetric error correction models. *Journal of Applied Econometrics* 4, 145–159.

Granger, C.W.J. & T. Lee (1989b) Multicointegration. In G. Rhodes & T. Fomby (eds.), *Advances in Econometrics: Cointegration, Spurios Regression and Unit Roots*, pp. 71–84. JAI Press.

Gregoir, S. (1999a) Multivariate time series with various hidden unit roots, part I. *Econometric Theory* 15, 435–468.

Gregoir, S. (1999b) Multivariate time series with various hidden unit roots, part II. *Econometric Theory* 15, 469–518.

Gregoir, S. & G. Laroque (1994) Polynomial cointegration: Estimation and test. *Journal of Econometrics* 63, 183–214.

Hannan, E.J. & M. Deistler (1988) *The Statistical Theory of Linear Systems*. Wiley.

Hansen, P.R. (2005) Granger's representation theorem: A closed form expression for I(1) processes. *The Econometrics Journal* 8, 23–38.

Harvey, A. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

Hylleberg, S., R.F. Engle, C.W.J. Granger, & B. Yoo (1990) Seasonal integration and cointegration. *Journal of Econometrics* 44, 215–238.

Johansen, S. (1995) *Likelihood-Based Inference in Cointegrated Vector Auto-Regressive Models*. Oxford University Press.

Johansen, S. (1997) Likelihood analysis of the I(2) model. *Scandinavian Journal of Statistics* 24, 433–462.

Johansen, S. & E. Schaumburg (1999) Likelihood analysis of seasonal cointegration. *Journal of Econometrics* 88, 301–339.

Kailath, T. (1980) *Linear Systems*. Prentice-Hall.

Lee, H. (1992) Maximum likelihood inference on cointegration and seasonal cointegration. *Journal of Econometrics* 54, 1–47.

Ober, R. (1996) Balanced canonical forms. In S. Bittanti and G. Picci (eds.), *Identification, Adaptation, Learning: The Science of Learning Models from Data*, pp. 120–183. Springer.

Phillips, P.C.B. (1998) Impulse response and forecast error variance asymptotics in nonstationary VARs. *Journal of Econometrics* 83, 21–56.

Phillips, P.C.B. & B.C. Hansen (1990) Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57, 99–125.

Sims, C.A., J.H. Stock, & M.W. Watson (1990) Inference in linear time series models with some unit roots. *Econometrica* 58, 113–144.

Stock, J.H. & M.W. Watson (1993) A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61, 783–820.

Yoo, B. (1986) Multi-cointegrated time series and generalized error correction models. Technical report, University of California San Diego.

APPENDIX: Proofs

Proof of Theorem 1. The unit root process $(y_t)_{t\in\mathbb{Z}}$ is a solution of the ARMA system $(\tilde{a}(z), b(z))$. The ARMA system corresponds to the transfer function $\tilde{c}(z) = \tilde{a}(z)^{-1}b(z)$, for which a minimal state space representation (A, B, C) with the A matrix in block-diagonal form $A = \mathrm{diag}(A_u, A_\bullet)$, where all eigenvalues of A_u are of unit modulus and $|\lambda_{\max}(A_\bullet)| < 1$, exists. Let $(y_{t,p})_{t\in\mathbb{Z}}$ denote the particular solution as given in (8).

In order to show the first part of the theorem, it has to be shown that $(y_{t,p})_{t\in\mathbb{Z}}$ is a unit root process and that it solves $\tilde{a}(L)y_{t,p} = b(L)\varepsilon_t, t \in \mathbb{Z}$. Since $\tilde{c}(z) = \tilde{a}(z)^{-1}b(z) = \sum_{j=0}^{\infty}\tilde{C}_j z^j$, we obtain from a comparison of power series coefficients that $\sum_{i=0}^{\min(l,H+p)}\tilde{A}_i\tilde{C}_{l-i} = B_l$ where $B_l = 0, l > q$, with $\tilde{A}_j, B_j$ as in (7). Let $\tilde{H} = \max(q, H + p)$ and denote by n the order of the state space system. Now using $l = \tilde{H} + 1, \tilde{H} + 2, \tilde{H} + n$ one obtains $\sum_{i=0}^{\tilde{H}}\tilde{A}_i\tilde{C}_{l-i} = 0$ where in all cases $l - i > 0$. This implies $\sum_{i=0}^{\tilde{H}}\tilde{A}_i CA^{l-i-1}B = 0$ and thus, since minimality of (A, B, C) implies that $[B, AB, A^2B, \ldots, A^{n-1}B]$ has full row rank, one obtains $\sum_{i=0}^{\tilde{H}}\tilde{A}_i CA^{\tilde{H}-i} = 0$.

For the particular solution (8) of the minimal state space system (A, B, C) it holds that $y_{t-j,p} = CA^{\tilde{H}-j}x_{t-\tilde{H},p} + \sum_{i=0}^{\tilde{H}-j}\tilde{C}_i\varepsilon_{t-j-i}$. From this it is straightforwardly verified that

$$\sum_{j=0}^{\tilde{H}}\tilde{A}_j y_{t-j,p} = \sum_{l=0}^{\tilde{H}} B_l\varepsilon_{t-l}, \qquad t \in \mathbb{Z}.$$

Thus, $(y_{t,p})_{t\in\mathbb{Z}}$ is a solution of the ARMA system $(\tilde{a}(z), b(z))$ and it follows from the proof of the second part of the theorem below that it is a unit root process with the same

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

unit root structure as $(y_t)_{t\in\mathbb{Z}}$. Thus, $D(L)y_{t,p} = v_t, t \in \mathbb{Z}$. This in turn implies for $y_{t,h} = y_t - y_{t,p}$ that $D(L)y_{t,h} = D(L)(y_t - y_{t,p}) = 0, t \in \mathbb{Z}$. This finishes the proof of the first part of the theorem.

For the second part of the theorem now let (A, B, C) denote a minimal state space system with $|\lambda_{\max}(A)| = 1$ and $|\lambda_{\max}(A - BC)| \leq 1$ with $A = \mathrm{diag}(A_u, A_\bullet)$, with all eigenvalues of A_u having unit modulus and $|\lambda_{\max}(A_u)| < 1$ and B and C partitioned accordingly. In showing the claim it is convenient to transform the system (A, B, C) into the specific (more structured) block-diagonal format discussed in the main text at the beginning of Section 4. Thus, consider a transformed system as described by equations (9) to (16), where for notational brevity we denote the appropriately transformed system also as (A, B, C). The described transformation implies also a corresponding transformation of the state process associated with the considered solution, as has been discussed in Section 3. We are confident that this notational shortcut does not lead to any confusion and henceforth only consider the transformed system and state process. Also note again that considering such transformed systems only is not a restriction of generality. By construction the process $(y_{t,p})_{t\in\mathbb{Z}}$ given in (8) is a solution of the state space system (A, B, C) in the specific format considered and it thus only remains to show that it has unit root structure $((\omega_1, h_1), \ldots, (\omega_l, h_l))$. Thus, it remains to show that $D(L)(y_{t,p})_{t\in\mathbb{Z}}$, with $D(L) = \prod_{k=1}^{l} \Delta_{\omega_k}^{h_k}(L)$, is a stationary process with non-zero spectrum at the unit root frequencies $\omega_1, \ldots, \omega_l$. Due to the block-diagonal structure of A it holds that $y_{t,p} = \sum_{k=1}^{l} C_k x_{t,k} + C_\bullet x_{t,\bullet} + \varepsilon_t$, $t \in \mathbb{Z}$ (with $x_{t,k}$, $k = 1, \ldots, l$ denoting the components of $x_{t,u}$ corresponding to the partitioning of A).

We start by establishing stationarity, using

$$D(L)y_{t,p} = \sum_{k=1}^{l} C_k D(L)x_{t,k} + C_\bullet D(L)x_{t,\bullet} + D(L)\varepsilon_t. \tag{A.1}$$

The second and third terms above, $C_\bullet D(L)x_{t,\bullet}$ and $D(L)\varepsilon_t$, are stationary processes with zero spectrum at the unit root frequencies. This implies that only the first l summands need to be investigated further. For any $k = 1, \ldots, l$ we can write $D(z) = \left(\prod_{j=1, j\neq k}^{l} \Delta_{\omega_j}^{h_j}(z)\right) \Delta_{\omega_k}^{h_k}(z) = D_{\neg k}(z)\Delta_{\omega_k}^{h_k}(z)$ with $D_{\neg k}(z_k) \neq 0$.

We now consider a fixed index k that corresponds to a unit root frequency ω_k such that $0 < \omega_k < \pi$ (noting that the arguments for unit roots with frequency in $\{0, \pi\}$ are analogous and simpler). Since we now consider a fixed block we simplify notation by omitting the index k; i.e., we use for the moment for brevity $A = A_k$, $B = B_k$, $C = C_k$, $x_t = x_{t,k}$, $h = h_k$, $d_j = d_j^k$, $d = d^k$, and $\omega = \omega_k$.

Note first that $\Delta_\omega(A) = I_{\dim A} - 2\cos(\omega)A + A^2$, for A as given in (13), is (using the reduced notation) given by

$$\begin{bmatrix}
0 & \left[I_{d_1} \otimes V(\omega), 0^{2d_1\times 2(d_2-d_1)}\right] & \left[I_{2d_1}, 0^{2d_1\times 2(d_3-d_1)}\right] & \cdots & 0 \\
0 & 0 & \left[I_{d_2} \otimes V(\omega), 0^{2d_2\times 2(d_3-d_2)}\right] & \ddots & \vdots \\
0 & 0 & 0 & \ddots & \left[I_{2d_{h-2}}, 0^{2d_{h-2}\times 2(d_h-d_{h-2})}\right] \\
\vdots & \vdots & \ddots & \ddots & \left[I_{d_{h-1}} \otimes V(\omega), 0^{2d_{h-1}\times 2(d_h-d_{h-1})}\right] \\
0 & 0 & \cdots & 0 & 0
\end{bmatrix}, \tag{A.2}$$

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

with $V(\omega) = 2(Q(\omega) - \cos(\omega) I_2) = 2 \begin{bmatrix} 0 & -\sin(\omega) \\ \sin(\omega) & 0 \end{bmatrix}$ a nonsingular matrix (since $0 < \omega < \pi$). From the above representation (A.2) it follows that $\Delta_\omega(A)^h = 0$ and $\Delta_\omega(A)^j \neq 0$ for $j = 1, \ldots, h-1$. In case $h > 1$ it holds that

$$(\Delta_\omega(A))^{h-1} = \begin{bmatrix} I_{2d_1} \\ 0 \end{bmatrix} \left[0^{2d_1 \times 2(d-d_h)} \left[I_{d_1} \otimes V(\omega)^{h-1}, 0^{2d_1 \times 2(d_h - d_1)} \right] \right]. \tag{A.3}$$

Now consider

$$\begin{aligned} \Delta_\omega(L) x_t &= x_t - 2\cos(\omega) x_{t-1} + x_{t-2} = (A - 2\cos(\omega) I_{\dim A}) x_{t-1} + x_{t-2} + B\varepsilon_{t-1} \\ &= \left(I_{\dim A} - 2\cos(\omega) A + A^2 \right) x_{t-2} + v_t^{(1)} = \Delta_\omega(A) x_{t-2} + v_t^{(1)}, \end{aligned}$$

with $v_t^{(1)} = B\varepsilon_{t-1} + (A - 2\cos(\omega) I_{\dim A}) B\varepsilon_{t-2}$ an MA(1) process. Continuing with backward substitution leads to $\Delta_\omega^h(L) x_t = (\Delta_\omega(A))^h x_{t-2h} + v_t^{(h)}$, with $v_t^{(h)}$ an MA(2h-1) process. Since $(\Delta_\omega(A))^h = 0$ it follows that $\Delta_\omega^h(L)(x_t)_{t \in \mathbb{Z}}$ is a stationary process. The above argument shows, when considering again all indices $k = 1, \ldots, l$, that $D(L)(y_{t,p})_{t \in \mathbb{Z}}$ is stationary.

It remains to show that the spectrum of $D(L)(y_{t,p})_{t \in \mathbb{Z}}$ is non-zero at the unit root frequencies. The above representation (A.1) together with $D_{\neg k}(z_k) \neq 0$ and $D_{\neg k}(z_j) = 0$ for $j \neq k$ implies that for considering the spectral properties of $D(L)(y_{t,p})_{t \in \mathbb{Z}}$ at frequency ω_k it suffices to consider the term $C_k \Delta_{\omega_k}^{h_k}(L)(x_{t,k})_{t \in \mathbb{Z}}$.

Since we now again consider only one specific frequency $0 < \omega_k < \pi$ (the case $\omega_k \in \{0, \pi\}$ is analogous but simpler) we again use the simplified notation already used above by omitting the index k. The proof is by contradiction: Hence assume that $C \Delta_\omega(L)^h x_t$ is stationary with the spectrum equal to zero at frequency ω. Consequently the same holds for $(C^1)' C \Delta_\omega(L)^h x_t$, with C^1 denoting the first $2d_1$ columns of C. This implies that also $(C^1)' C \Delta_\omega(L)^{h-1} x_t$ is stationary. For $h > 1$ the above results, in particular (A.3), imply that

$$\begin{aligned} \Delta_\omega^{h-1}(L) x_t = &\begin{bmatrix} I_{2d_1} \\ 0 \end{bmatrix} \left[0^{2d_1 \times 2(d-d_h)} \left[I_{d_1} \otimes V(\omega)^{h-1}, 0^{2d_1 \times 2(d_h - d_1)} \right] \right] \\ &\times x_{t-2(h-1)} + v_t^{(h-1)}, \qquad t \in \mathbb{Z}. \end{aligned} \tag{A.4}$$

Thus

$$(C^1)' C^1 \left[0^{2d_1 \times 2(d-d_h)} \left[I_{d_1} \otimes V(\omega)^{h-1}, 0^{2d_1 \times 2(d_h - d_1)} \right] \right] x_{t-2(h-1)} + (C^1)' C v_t^{(h-1)},$$

$t \in \mathbb{Z}$ is stationary. Due to minimality (cf. Lemma 1(i)) the matrix C^1 is nonzero and consequently (A.4) implies that there exists a vector $\beta \neq 0$ such that $(\beta' x_t)_{t \in \mathbb{Z}}$ is stationary. For $h = 1$ the result is immediate.

Since $\beta' x_1 = 0$, due to the normalization $x_1 = 0$, stationarity implies that

$$0 = \beta' x_t = \beta' \sum_{j=1}^{t-1} A^{j-1} B \varepsilon_{t-j}, t \in \mathbb{N}.$$

Since $(\varepsilon_t)_{t \in \mathbb{Z}}$ is white noise with $\mathbb{E}(\varepsilon_t \varepsilon_t') > 0$ (cf. Definition 1) it follows that $\beta' A^j B = 0$, $j \geq 0$. This is a contradiction to minimality; compare HD (Thm. 2.3.3, p. 48). Therefore the

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

spectrum of $C\Delta_\omega(L)^h x_t$ must be nonzero at frequency ω. Combining the above arguments for all $\omega_1, \ldots, \omega_l$ concludes the proof. ■

Proof of Lemma 1. The proof of the lemma is based on the following result, known as the Popov-Belovich-Hautus (PBH) test, which can be found including a proof in Kailath (1980, Thm. 2.4-8., p. 135).

A state space system (A, B, C), with $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times s}, C \in \mathbb{R}^{s\times n}$ is not minimal if and only if there exists either an eigenvector x, $x \neq 0$, of A such that $Cx = 0$ or an eigenvector y, $y \neq 0$, of A' such that $y'B = 0$.

The PBH test will be used to characterize minimality for state space systems with the A-matrix in the block-diagonal format considered. Due to the block-diagonal structure of the A-matrix, eigenvectors of this matrix corresponding to unit modulus eigenvalues are of the form $\tilde{v}_k = [0, \ldots, 0, v_k', 0, \ldots, 0]'$ with $A_k v_k = \lambda_k v_k$. The same holds for eigenvalues of A'.

Henceforth we therefore consider only one typical block corresponding to a unit root frequency $0 < \omega_k < \pi$ and omit the index k for brevity in the rest of the proof; i.e., we use $A = A_k$, $J = J_k$, $B_- = B_{k,-}$, $C_- = C_{k,-}$, $v = v_k$, $h = h_k$, and $d_j = d_j^k$. For the blocks corresponding to unit root frequencies $\omega \in \{0, \pi\}$ the arguments are analogous but simpler since no transformation to complex valued quantities is required.

Due to the structure of A, eigenvectors of A and A' are of a particular form. Suppose that $v \neq 0$ is an eigenvector of A, i.e.,

$$Av = \lambda v \Leftrightarrow T_{\mathbb{R}}^{-1} A T_{\mathbb{R}} T_{\mathbb{R}}^{-1} v = \lambda T_{\mathbb{R}}^{-1} v \Leftrightarrow \begin{pmatrix} J & 0 \\ 0 & \overline{J} \end{pmatrix} v_{\mathbb{C}} = \lambda v_{\mathbb{C}},$$

where $v_{\mathbb{C}} = T_{\mathbb{R}}^{-1} v$ with $T_{\mathbb{R}}$ as given in (11) and J as given in (10). Since the only eigenvalue of J is $\overline{z}$ it follows that $v_{\mathbb{C}} = [w'\ 0']'$ with $Jw = \overline{z}w$ (with $\lambda = \overline{z}$) or $v_{\mathbb{C}} = [0'\ \overline{w}']'$ (with $\lambda = z$).

The structure of the matrix J (compare (10)) places 0-restrictions on w. Partition the vector $w = [(w_1^E)', (w_2^G)', (w_2^E)', \ldots, (w_h^G)', (w_h^E)']'$ with $w_j^E \in \mathbb{C}^{d_j - d_{j-1}}$ and $w_j^G \in \mathbb{C}^{d_{j-1}}$ for $j = 1 \ldots, h$, with $d_0 = 0$. For w to be an eigenvector of J it is necessary that $w_j^G = 0$ for $j = 1, \ldots, h$.

After transformation to complex quantities the constraint that $Cv = 0$ formulated in the PBH test reads as $[C_- \ \overline{C}_-]v_{\mathbb{C}} = 0$. Therefore, if C_-^E has full column rank $C_- w = C_-^E w^E = 0$, where $w^E = [(w_1^E), \ldots, (w_h^E)']'$ implies $w = 0$.

Entirely similar arguments apply to eigenvectors of A' that have the same structure when transformed by $T_{\mathbb{R}}$, as is seen by multiplication from the right with a vector $v_{\mathbb{C}}$ using the same notation for simplicity. For w to be an eigenvector of J' it has to hold that $w = [0, \ldots, 0, w_h']'$. This implies, using similar arguments as above, that B_-^h of full row rank implies $w_h' B_-^h = 0$, which leads to $w_h = 0$ and thus $w = 0$. Jointly the formulated constraints imply minimality for the subsystem.

Combining the subsystems (A_k, B_k, C_k) for $k = 1, \ldots, l$ shows that the restrictions formulated in the theorem indeed characterize minimality, since also the stable subsystem $(A_\bullet, B_\bullet, C_\bullet)$ is assumed to be minimal. The remaining two statements of the theorem follow immediately. ■

Proof of Theorem 2. Based on the specific minimal state space representation (A, B, C) the proof proceeds in showing that only one observationally equivalent state space representation $(\mathcal{A}, \mathcal{B}, \mathcal{C})$ exists, which fulfills all *additional* restrictions formulated in the

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

theorem as well as the restrictions formulated already in Lemma 1. This amounts to showing that if $(\mathcal{A},\mathcal{B},\mathcal{C})$ fulfills all restrictions and also $(T\mathcal{A}T^{-1},T\mathcal{B},\mathcal{C}T^{-1})$ fulfills all restrictions this implies that $T=I_n$.

As before, due to the block-diagonal structure of the A-matrix it suffices to consider one typical subsystem, since this block-diagonal structure is only preserved by block-diagonal transformation matrices $T=\mathrm{diag}(T_1,\ldots,T_l,T_\bullet)$; cf. Gantmacher (1966, (11), p. 231). For notational brevity we again omit the index k in the rest of the proof. With respect to the stable subsystem $(\mathcal{A}_\bullet,\mathcal{B}_\bullet,\mathcal{C}_\bullet)$ uniqueness is achieved by assuming it to be in echelon canonical form.

Note here as a remark that observational equivalence is usually discussed using only real valued nonsingular transformation matrices. Here we allow for complex valued transformation matrices and show that the real valued representation given in the formulation in the theorem is unique even when allowing for complex nonsingular transformation matrices. This is achieved by providing a unique complex valued representation, which is then transformed as given in (12) to obtain a unique real valued representation.

For a typical subsystem corresponding to one unit root we first describe the set of transformation matrices that leave the matrix A with the posited structure unchanged. This set can equivalently be described by characterizing the set of matrices that fulfill $T^{-1}JT=J$, with J as given in (10) and $T\in\mathbb{C}^{d\times d}$, using again simplified notation. The structure of J implies that this latter equation is equivalent to $TN=NT$, with $N=J-\bar{z}I_d$. Matrix equations of this form are discussed in Gantmacher (1966, Sect. 8.1). Denote with $\tilde{I}_{d_j}=[I_{d_j},0^{d_j\times(d_{j+1}-d_j)}]$ the non-zero blocks in N and partition $T=[T_{i,j}]_{i,j=1,\ldots,h}$ with $T_{i,j}\in\mathbb{C}^{d_i\times d_j}$ accordingly. Using this partitioning the equation $TN=NT$ implies several constraints (compare also Gantmacher, 1966, (14), p. 231). First it follows that $T_{h,j}\tilde{I}_{d_j}=0$ for $j=1,\ldots,h-1$, which implies that $T_{h,j}=0$ for $j=1,\ldots,h-1$. Second, the relation

$$T_{i,j}\tilde{I}_{d_j}=\tilde{I}_{d_i}T_{i+1,j+1} \tag{A.5}$$

for $i,j=1,\ldots,h-1$ implies

$$T_{i+1,j+1}=\begin{bmatrix} T_{i,j} & 0 \\ T^{(1)}_{i+1,j+1} & T^{(2)}_{i+1,j+1}\end{bmatrix}, \tag{A.6}$$

with $T^{(1)}_{i+1,j+1}\in\mathbb{C}^{(d_{i+1}-d_i)\times d_j}$ and $T^{(2)}_{i+1,j+1}\in\mathbb{C}^{(d_{i+1}-d_i)\times(d_{j+1}-d_j)}$ for $i,j=1,\ldots,h-1$. The above relation (A.6) together with $T_{h,j}=0$ for $j=1,\ldots,h-1$ implies that $T_{i,j}=0$ for $1\le j<i$. Thus, the matrix T is block upper triangular, i.e.,

$$T=\begin{bmatrix} T_{1,1} & T_{1,2} & \cdots & T_{1,h} \\ 0 & T_{2,2} & \cdots & T_{2,h} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & T_{h,h}\end{bmatrix},$$

with the blocks along the main diagonal related by (A.6).

The unique form of the A-matrix based on the matrix J does not suffice to select a unique state space representation and hence the additional restrictions formulated in the theorem

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

are placed on the matrices B and C via placing restrictions on B_- and C_-. Consider consequently $T^{-1}B_-$ and C_-T given in partitioned format as

$$\mathcal{C}_- = C_-T = \left[C_-^1\ C_-^2 \cdots C_-^h\right] \begin{bmatrix} T_{1,1} & T_{1,2} & \cdots & T_{1,h} \\ 0 & T_{2,2} & \cdots & T_{2,h} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & T_{h,h} \end{bmatrix}, \tag{A.7}$$

$$\mathcal{B}_- = T^{-1}B_- = \begin{bmatrix} T_{1,1}^{-1} & * & \cdots & * \\ 0 & T_{2,2}^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & T_{h,h}^{-1} \end{bmatrix} \begin{bmatrix} B_-^1 \\ B_-^2 \\ \vdots \\ B_-^h \end{bmatrix}, \tag{A.8}$$

with $\mathcal{C}_-$ and $\mathcal{B}_-$ fulfilling all constraints formulated in the theorem and $*$ denoting unspecified entries. In the rest of the proof we now show that $T = I_d$ if also C_- and B_- defined from (A.7) and (A.8) fulfill all constraints. The proof is recursive: The first block-column of (A.7) equals $\mathcal{C}_-^1 = \mathcal{C}_-^{1,E} = C_-^{1,E}T_{1,1} = C_-^1T_{1,1}$ and the last block-row of (A.8) equals $\mathcal{B}_-^h = T_{h,h}^{-1}B_-^h$, with the partitioning as introduced in Lemma 1. The structure of the matrices $T_{j,j}$ described in (A.6) implies that

$$T_{h,h}^{-1} = \begin{bmatrix} T_{1,1}^{-1} & 0 \\ ** & * \end{bmatrix}.$$

The restrictions formulated in the theorem, $(\mathcal{C}_-^{1,E})'\mathcal{C}_-^{1,E} = I_{d_1}$ and $B_-^{h,1}$ p.u.t., now imply a unique decomposition $\mathcal{C}_-^{1,E}\mathcal{B}_-^{h,1} = \left(C_-^{1,E}T_{1,1}\right)\left(T_{1,1}^{-1}B_-^{h,1}\right)$ and consequently $T_{1,1} = I_{d_1}$. This implies that $\mathcal{C}_-^{1,E} = C_-^{1,E}$.

The argument proceeds with the next two blocks

$$T_{1,2} = \left[T_{1,2}^{(1)}\ T_{1,2}^{(2)}\right], \qquad T_{2,2} = \begin{bmatrix} I_{d_1} & 0 \\ T_{2,2}^{(1)} & T_{2,2}^{(2)} \end{bmatrix}.$$

In partitioned format the second block-column of (A.7) equals

$$\left[\mathcal{C}_-^{2,G},\ \mathcal{C}_-^{2,E}\right] = \left[\mathcal{C}_-^{1,E}T_{1,2}^{(1)} + C_-^{2,G} + C_-^{2,E}T_{2,2}^{(1)},\ \mathcal{C}_-^{1,E}T_{1,2}^{(2)} + C_-^{2,E}T_{2,2}^{(2)}\right], \tag{A.9}$$

using the already established relationship $\mathcal{C}_-^{1,E} = C_-^{1,E}$. The following constraints are imposed in the formulation of the theorem: $(\mathcal{C}_-^{2,G})'[\mathcal{C}_-^{1,E}\ \mathcal{C}_-^{2,E}] = 0$ and $[\mathcal{C}_-^{1,E}\ \mathcal{C}_-^{2,E}]'$ $[\mathcal{C}_-^{1,E}\ \mathcal{C}_-^{2,E}] = I_{d_1+d_2}$. By assumption these have to hold similarly for the blocks of the system matrix C_-; i.e., $(C_-^{2,G})'[\mathcal{C}_-^{1,E}\ C_-^{2,E}] = 0$ and $[\mathcal{C}_-^{1,E}\ C_-^{2,E}]'[\mathcal{C}_-^{1,E}\ C_-^{2,E}] = I_{d_1+d_2}$. These constraints imply $T_{1,2}^{(1)} = 0$ (by premultiplying the first block column of (A.9) with $(\mathcal{C}_-^{1,E})'$), $T_{1,2}^{(2)} = 0$ (by premultiplying the second block column of (A.9) with $(\mathcal{C}_-^{1,E})'$) and

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

$T_{2,2}^{(1)} = 0$ (by premultiplying the first block column of (A.9) with $(C_-^{2,E})'$). The latter in turn implies that

$$T_{h,h}^{-1} = \begin{bmatrix} I_{d_1} & 0 & 0 \\ 0 & \left(T_{2,2}^{(2)}\right)^{-1} & 0 \\ * & * & * \end{bmatrix},$$

which implies via (A.8) that $\mathcal{B}_-^{h,2} = (T_{2,2}^{(2)})^{-1} B_-^{h,2}$. Combining this with $\mathcal{C}_-^{2,E} = C_-^{2,E} T_{2,2}^{(2)}$, which follows from $T_{1,2}^{(2)} = 0$, implies again that there is a unique decomposition of $\mathcal{C}_-^{2,E}\mathcal{B}_-^{h,2} = C_-^{2,E} B_-^{h,2}$ that fulfills all constraints, i.e., with $\mathcal{B}_-^{h,2}$ p.u.t. and $(\mathcal{C}_-^{2,E})'\mathcal{C}_-^{2,E} = I_{d_2-d_1}$. This shows that $T_{2,2} = I_{d_2}$ and $T_{1,2} = 0$ and thus $[\mathcal{C}_-^{2,G}, \mathcal{C}_-^{2,E}] = [C_-^{2,G}, C_-^{2,E}]$.

The proof continues by induction. Suppose therefore that the constraints on the subblocks of T are fulfilled up to some index $2 < j < h$, i.e. $T_{i,i} = I_{d_i}$ for $i = 1, \ldots, j$ and $T_{i,g} = 0$, for $1 \le i, g \le j$ and $i \ne g$. Furthermore, we have

$$T_{1,j+1} = \begin{bmatrix} T_{1,j+1}^{(1)} & T_{1,j+1}^{(2)} \end{bmatrix}, \tag{A.10}$$

$$T_{i,j+1} = \begin{bmatrix} 0 & 0 \\ T_{i,j+1}^{(1)} & T_{i,j+1}^{(2)} \end{bmatrix}, \qquad 2 \le i \le j, \tag{A.11}$$

$$T_{j+1,j+1} = \begin{bmatrix} I_{d_j} & 0 \\ T_{j+1,j+1}^{(1)} & T_{j+1,j+1}^{(2)} \end{bmatrix}. \tag{A.12}$$

The two zero blocks in the upper block-row of $T_{i,j+1}$ follow from (A.5) and the fact that by the induction assumption $T_{i-1,j} = 0$ for the considered values of i. Computations analogous to those performed with (A.9) for the second block column for the jth block column of (A.7) show, by using the fulfillment of orthogonality conditions via pre-multiplication with $(\mathcal{C}_-^{i,E})'$, that $T_{i,j+1}^{(1)} = 0, i = 1, \ldots, j+1, T_{i,j+1}^{(2)} = 0, i = 1, \ldots, j$. Then the unique decomposition of $\mathcal{C}_-^{j+1,E}\mathcal{B}_-^{h,j+1} = C_-^{j+1,E} B_-^{h,j+1}$ concludes the induction step.

The result follows by combining all blocks: Uniqueness of the complex valued matrices $J_k, \mathcal{B}_{k,-}, \mathcal{C}_{k,-}$ implies unique real valued matrices $\mathcal{A}_k, \mathcal{B}_k, \mathcal{C}_k$ and thus by combination with the unique stable subsystem in echelon canonical form a unique real valued state space representation $(\mathcal{A}, \mathcal{B}, \mathcal{C})$. ■

Proof of Theorem 3. The proof relies on properties of the state process already established in the proof of Theorem 1.

(i) In the last part of the proof of Theorem 1 it has been shown that the first blocks $(x_{t,k}^1)_{t\in\mathbb{Z}}$ have unit root structures $((\omega_k, h_k))$; compare in particular (A.4). In fact the arguments there show that it is the first blocks $(x_{t,k}^1)_{t\in\mathbb{Z}}$ that have unit root structure $((\omega_k, h_k))$; compare in particular (A.4). Note that the matrix $\mathcal{A}_k$ has a specific recursive structure in the sense that omitting the first j block rows and the first j block columns leads to a matrix of the same structure with h_k reduced to $h_k - j$. This implies that entirely analogous arguments as used in the proof of Theorem 1 show that the processes $(x_{t,k}^j)_{t\in\mathbb{Z}}$ have unit root structures $((\omega_k, h_k - j + 1))$.

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press

(ii) We now consider again only the state for one fixed index k, which we again assume to correspond to a unit root frequency $0 < \omega_k < \pi$. The arguments are again analogous but simpler for unit roots with frequency $\omega_k \in \{0, \pi\}$. For notational brevity we again omit the index k. Partition the state x_t $(= x_{t,k})$ as

$$x_t = \begin{bmatrix} x_t^1 \\ x_t^2 \\ \vdots \\ x_t^h \end{bmatrix} = \begin{bmatrix} x_t^{1,E} \\ x_t^{2,G} \\ x_t^{2,E} \\ \vdots \\ x_t^{h,G} \\ x_t^{h,E} \end{bmatrix},$$

with $x_t^j \in \mathbb{R}^{2d_j}$, $x_t^{j,G} \in \mathbb{R}^{2d_{j-1}}$, and $x_t^{j,E} \in \mathbb{R}^{2(d_j - d_{j-1})}$ for $j = 1, \ldots, h$ and $d_0 = 0$. For the last block $(x_t^h)_{t\in\mathbb{Z}}$ with unit root structure $((\omega, 1))$ we already know from Theorem 1 that it is not cointegrated. Consider next the process $(x_t^{h-1})_{t\in\mathbb{Z}}$, which is given by $x_{t+1}^{h-1} = [I_{d_{h-1}} \otimes Q(\omega)]x_t^{h-1} + x_t^{h,G} + B^{h-1}\varepsilon_t$, $t \in \mathbb{Z}$. Prevalence of cointegration for $(x_{t+1}^{h-1})_{t\in\mathbb{Z}}$ is equivalent to the existence of $\beta \neq 0$ such that $(\beta' x_{t+1}^{h-1})_{t\in\mathbb{Z}}$ has unit root structure $((\omega, 1))$ or is stationary. The latter case is excluded as has been shown in the proof of Theorem 1. Using the block structure of the state and the A-matrix, similar derivations and expressions as have been used in the proof of Theorem 1 can be applied to the state equation describing only the sub-vector $[(x_t^{h-1})', (x_t^h)']'$. It holds that

$$\Delta_\omega(L)x_t^{h-1} = [I_{d_{h-1}} \otimes V(\omega)]x_{t-2}^{h,G} + \tilde{v}_t, \qquad t \in \mathbb{Z}, \tag{A.13}$$

where $(\tilde{v}_t)_{t\in\mathbb{Z}}$ is a stationary process. Since $[I_{d_{h-1}} \otimes V(\omega)]$ has full rank and $(x_t^{h,G})_{t\in\mathbb{Z}}$ is not cointegrated, this shows that there exists no vector $\beta \neq 0$ that reduces the unit root structure of $(x_t^{h-1})_{t\in\mathbb{Z}}$ from $((\omega, 2))$ to $((\omega, 1))$.

Similar arguments as just outlined can be applied in a recursive fashion to show that also none of the sub-vectors $(x_t^j)_{t\in\mathbb{Z}}$ for $j = h-2, \ldots, 1$ is cointegrated. ■

Proof of Theorem 4.

(i) It follows immediately from the representation (21) that $\beta(L)'y_{t,p} = \bar{\beta}'\mathcal{O}_{q,u}x_{t-q,u} + w_t$, with $w_t := \bar{\beta}'(\bar{\mathcal{O}}_{q,\bullet}x_{t-q,\bullet} + \mathcal{E}_q E_{t,q}^-)$ stationary as discussed in the main text. Now the properties of the state $(x_t)_{t\in\mathbb{Z}}$ derived in Theorem 3, most notably the absence of cointegration within the sub-blocks, show immediately that the orthogonality constraints posited in the formulation of the theorem are equivalent for the process $\beta(L)'(y_{t,p})_{t\in\mathbb{Z}}$ to have unit root structure $\tilde{\Omega}$.

(ii) If $q < H-1$ then setting $\beta_m = 0$ for $j = q+1, \ldots, H-1$ shows the claim and for $q > H-1$ the claim is shown by taking the componentwise remainder with respect to division by the polynomial $D(z)$ defined in Definition 1. ■

https://doi.org/10.1017/S026646661200014X Published online by Cambridge University Press