Sapienza Università di Roma

Dipartimento di Scienze Statistiche
Sezione di Statistica Economica ed Econometria

Massimo Franchi **Paolo Paruolo**

Cointegration in functional autoregressive processes

DSS Empirical Economics and Econometrics Working Papers Series

DSS-E3 WP 2017/5

Dipartimento di Scienze Statistiche
Sezione di Statistica Economica ed Econometria
“Sapienza” Università di Roma
P.le A. Moro 5 – 00185 Roma - Italia

http://www.dss.uniroma1.it

Cointegration in functional autoregressive processes

MASSIMO FRANCHI AND PAOLO PARUOLO

DECEMBER 20, 2017

Abstract. This paper derives a generalization of the Granger-Johansen Representation Theorem valid for H-valued autoregressive (AR) processes, where H is an infinite dimensional separable Hilbert space, under the assumption that 1 is an eigenvalue of finite type of the AR operator function and that no other non-zero eigenvalue lies within or on the unit circle. A necessary and sufficient condition for integration of order $d = 1, 2, \ldots$ is given in terms of the decomposition of the space H into the direct sum of $d+1$ closed subspaces τ_h, $h = 0, \ldots, d$, each one associated with components of the process integrated of order h. These results mirror the ones recently obtained in the finite dimensional case, with the only difference that the number of cointegrating relations of order 0 is infinite.

1. Introduction

The theory and applications of time series that take values in infinite dimensional separable Hilbert spaces, or H-valued processes, are recently gaining increasingly attention in econometrics. They allow to represent directly the dynamics of infinite-dimensional objects, such as bounded continuous function on a compact.

An important early contribution to the literature of functional time series is Bosq (2000), where a theoretical treatment of linear processes in Banach and Hilbert spaces is developed. There, emphasis is given to the derivations of laws of large numbers and central limit theorems that allow to discuss estimation and inference for H-valued autoregressive (AR) models.

Empirical applications of functional time series analysis include studies on the term structure of interest rates, Kargin and Onatski (2008), and on intraday volatility, Hörmann et al. (2013) and Gabrys et al. (2013). The recent monographs by Horváth and Kokoszka (2012) and Kokoszka and Reimherr (2017) and the review in Hörmann and Kokoszka (2012) report additional examples.

Key words and phrases. Functional autoregressive process, Unit roots, Cointegration, Common Trends, Granger-Johansen Representation Theorem.

M. Franchi, Sapienza University of Rome, P.le A. Moro 5, 00185 Rome, Italy; e-mail: `massimo.franchi@uniroma1.it`.

P. Paruolo, European Commission, Joint Research Centre (JRC), Via E.Fermi 2749, I-3027 Ispra (VA), Italy; e-mail: `paolo.paruolo@ec.europa.eu`, **corresponding author**.

The ideas in the present paper were conceived while the first author was visiting the Department of Economics, Indiana University, in January 2017; the hospitality of Yoosoon Chang and Joon Park is gratefully acknowledged.

In a recent paper, Chang et al. (2016) present statistical tools for functional time series that are integrated of order one, $I(1)$, and find evidence of unit root persistence and cointegration in the time series of the several coordinates of the cross-sectional distributions of individual earnings and the intra-month distributions of stock returns.[1] The framework proposed by Chang et al. (2016) has (by construction) a finite number of $I(1)$ stochastic trends and an infinite dimensional cointegrating space. The theory is developed starting from the infinite moving average representation of the first differences of the process; the potential unit roots are identified and tested through a functional principal components analysis.

In a different vein, Hu and Park (2016) start from an AR process of order one with compact AR operator and provide an $I(1)$ condition that extends the Granger-Johansen Representation Theorem, see Theorem 4.2 in Johansen (1996), to $I(1)$ H-valued AR(1) processes. The corresponding common trends representation, or functional Beveridge-Nelson decomposition, displays a finite number of $I(1)$ stochastic trends and an infinite dimensional cointegrating space. They further propose an estimator for the functional autoregressive operator which builds on Chang et al. (2016).

Beare et al. (2017) consider an H-valued AR(s), $s > 1$, process with compact AR operators and provide a reformulation of the Johansen $I(1)$ condition that extends the Granger-Johansen Representation Theorem to this more general setup. Again here, the number of $I(1)$ stochastic trends is finite and the dimension of the cointegrating space is infinite.

Finally, Chang et al. (2016) consider an error correction form with compact error correction operator and show that in this case the number of $I(1)$ stochastic trends is infinite and the dimension of the cointegrating space is finite. Moreover, the Granger-Johansen Representation Theorem continues to holds in a form similar to the finite dimensional case.

The present paper provides an extension of the representation results in Beare et al. (2017) and Hu and Park (2016) for H-valued AR processes in the generic $I(d)$, $d = 1, 2, \ldots$ case, under the assumption that 1 is an eigenvalue of finite type of the AR operator function and that no other non-zero eigenvalue of the AR operator lies within or on the unit circle. It is found that the conditions and properties in the H-valued cointegrated AR processes coincide with those the finite dimensional AR processes, except for the fact that the number of cointegrating relations of order 0 is infinite for H-valued cointegrated AR processes.

The assumption that 1 is an eigenvalue of finite type of the AR operator function means that the inverse of the AR operator function has an isolated pole at 1, as in the finite dimensional case. The assumptions in Hu and Park (2016) and Beare et al. (2017) imply that 1 is an eigenvalue of finite type (but not viceversa), and hence the present analysis applies to those setups as special cases.

[1] Beare (2017) argues that the nonnegativity property of densities is not compatible with the postulated unit root nonstationarity. Nevertheless, while concurring that the framework in Chang et al. (2016) is useful for generic H-valued processes.

A necessary and sufficient condition for $I(d)$, $d = 1, 2, \ldots$, in terms of the decomposition of the space H into the direct sum of $d+1$ closed subspaces, that are defined recursively from the AR operators, $H = \tau_0 \oplus \tau_1 \oplus \cdots \oplus \tau_d$, where τ_0 is closed and infinite dimensional and τ_h, is closed and finite dimensional for $h = 1, \ldots, d$. The infinite dimensional subspace $\tau_0 \oplus \tau_1 \oplus \cdots \oplus \tau_{d-1}$ is the cointegrating space of x_t while the finite dimensional subspace τ_d is the attractor space of the $I(d)$ stochastic trends. Hence an H-valued cointegrated AR process x_t has a finite number of $I(d)$ stochastic trends and infinitely many cointegrating relations.

The properties of $\langle v, x_t \rangle$ vary with v in H: for $v \in \tau_0$, which is infinite dimensional, one can combine $\langle v, x_t \rangle$ with differences $\Delta^n x_t$ for $n = 1, \ldots, d-1$ to find $I(0)$ polynomial cointegrating relations, for $v \in \tau_1$, which is finite dimensional (and can as well be equal to 0), one can combine $\langle v, x_t \rangle$ with with differences $\Delta^n x_t$ for $n = 1, \ldots, d-2$ and find at most $I(1)$ polynomial cointegrating relations, and so on up to $v \in \tau_{d-1}$, which is finite dimensional (and possibly of 0 dimension), for which $\langle v, x_t \rangle \sim I(d-1)$ does not allow for polynomial cointegration. When $v \in \tau_d$, which is finite dimensional and different from 0, one has $\langle v, x_t \rangle \sim I(d)$, i.e. there is no cointegration.

For any v in the cointegrating space, the explicit expression of the coefficients of the polynomial cointegrating relations is provided in terms of operators that are defined recursively from the AR operators together with the sequence of τ_h.

The present results show that the infinite dimensionality of the space does not introduce additional elements in the analysis, apart from the fact that the number of $I(0)$ cointegrating relations is infinite. That is, the conditions and properties of H-valued cointegrated AR process coincide with those that apply in the finite dimensional case.

The present derivations parallel the development of the representation theory for finite dimensional autoregressive processes developed by Johansen (1996) for the I(1) and I(2) cases and extended in Franchi and Paruolo (2016) to I(d) processes. The present treatment makes extensive use of projection matrices in the ambient space, as well as of Moore-Penrose generalised inverses. These tools have direct counterparts both in the finite dimensional case in Franchi and Paruolo (2016) and in the present infinite dimensional space.

The rest of the paper is organized as follows: the remaining part of this introduction reports notation and preliminaries; Section 2 presents basic definitions and concepts and Section 3 reports an existence result. Section 4 provides a characterization of $I(1)$ and $I(2)$ cointegrated H-valued AR processes, Section 5 extends the result to the general $I(d)$, $d = 1, 2, \ldots$, case and Section 6 concludes. Appendix *A* reviews notions and results on separable Hilbert spaces and on operators acting on them, Appendix *B* presents basic facts on H-valued random variables and Appendix *C* contains proofs.

Notation and preliminaries. In the present paper H is an infinite dimensional separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$, where separable means that H admits a countable orthonormal basis. A random variable that takes values in H is said to be an H-valued random variable and a sequence of H-valued random variables is called an H-valued stochastic process.

Let $\mathcal{B}(H)$ be the (Banach) space of bounded linear operators from H to H and consider $A \in \mathcal{B}(H)$; the closed subspace $\{x \in H : Ax = 0\}$, written $\operatorname{Ker} A$, is called the kernel of A and the subspace $\{Ax : x \in H\}$, written $\operatorname{Im} A$, is called the image of A. The dimension of $\operatorname{Im} A$, written $\operatorname{rank} A$, is called the rank of A. The set $\{v \in H : \langle v, y\rangle = 0 \text{ for all } y \in S \subseteq H\}$ is called the orthogonal complement of S, written $S^\perp$, and P_S denotes the orthogonal projection on S, i.e. $P_S x = x$ for all $x \in S$ and $\operatorname{Ker} P_S = S^\perp$.

Let $z_0 \in \mathbb{C}$ and $0 < \rho \in \mathbb{R}$; the set $\{z \in \mathbb{C} : |z - z_0| < \rho\}$ is called the open disc centered in z_0 with radius ρ, written $D(z_0, \rho)$. If a power series $\sum_{n=0}^{\infty} A_n (z - z_0)^n$, $A_n \in \mathcal{B}(H)$, is absolutely convergent for all $z \in D(z_0, \rho)$, i.e. that $\sum_{n=0}^{\infty} \|A_n\| |z - z_0|^n < \infty$ for all $z \in D(z_0, \rho)$, then $\sum_{n=0}^{\infty} A_n (z - z_0)^n$ converges in the operator norm to $A(z) \in \mathcal{B}(H)$, i.e. that $\| \sum_{n=0}^{N} A_n (z - z_0)^n - A(z)\| \to 0$ as $N \to \infty$ for all $z \in D(z_0, \rho)$. In this case, the operator function $A(z) = \sum_{n=0}^{\infty} A_n (z - z_0)^n$ is said to be absolutely convergent on $D(z_0, \rho)$.

Appendix A reviews notions and results on separable Hilbert spaces and on operators acting on them and Appendix B presents the definitions of expectation, covariance and cross-covariance for H-valued random variables.

2. Basic definitions and concepts

This section introduces stochastic processes that take values in a separable Hilbert space H and presents the notions of weak stationarity, white noise, linear process, integration and cointegration. The definitions of weak stationarity and white noise are taken from Bosq (2000) while those of linear process, integration and cointegration are adapted from Johansen (1996) and are similar to those employed in Beare et al. (2017).

Definition 2.1 (Weak stationarity). *An H-valued stochastic process $\{\varepsilon_t, t \in \mathbb{Z}\}$ is said to be weakly stationary if i) $E(\|\varepsilon_t\|^2) < \infty$, ii) the expectation of ε_t does not depend on t and iii) the cross-covariance function of ε_t and ε_s is such that $c_{\varepsilon_t, \varepsilon_s}(h, x) = c_{\varepsilon_{t+u}, \varepsilon_{s+u}}(h, x)$ for all $h, x \in H$ and all $s, t, u \in \mathbb{Z}$.*

The basic notion of H-valued white noise is introduced next.

Definition 2.2 (White noise). *An H-valued stochastic process $\{\varepsilon_t, t \in \mathbb{Z}\}$ is said to be a white noise if i) $0 < E(\|\varepsilon_t\|^2) < \infty$, ii) the expectation of ε_t is equal to 0, iii) the covariance operator of ε_t does not depend on t, and iv) the cross-covariance function of ε_t and ε_s, $c_{\varepsilon_t, \varepsilon_s}(h, x)$, is equal to 0 for all $h, x \in H$ and all $s \neq t, s, t \in \mathbb{Z}$.*

It is evident from the definition that any white noise is weakly stationary. The same property applies to linear combinations of a white noise with suitable weights, which defines the class of linear processes.

Definition 2.3 (Linear process)**.** *An H-valued stochastic process $\{u_t, t \in \mathbb{Z}\}$ is said to be a linear process if*

$$u_t = \sum_{n=0}^{\infty} B_n \varepsilon_{t-n}, \qquad B_0 = I,$$

where $\{\varepsilon_t, t \in \mathbb{Z}\}$ is white noise, $B_n : H \to H$ is a bounded linear operator and $B(z) = \sum_{n=0}^{\infty} B_n z^n$ is absolutely convergent on $D(0, \rho)$ for some $\rho > 1$.

As discussed in Section 7.1 in Bosq (2000), existence and weak stationarity of $u_t = \sum_{n=0}^{\infty} B_n \varepsilon_{t-n}$ are guaranteed if $\sum_{n=0}^{\infty} \|B_n\|^2 < \infty$. Observe that the requirement that $B(z)$ is absolutely convergent $D(0, \rho)$ for some $\rho > 1$ is stronger. In fact, $\sum_{n=0}^{\infty} \|B_n\||z|^n < \infty$ for all $z \in D(0, \rho)$, $\rho > 1$, implies that $\sum_{n=0}^{\infty} \|B_n\|$ is finite, so that $\sum_{n=0}^{\infty} \|B_n\|^2$ is finite and u_t in Definition 2.3 is a well defined and weakly stationary process.

Moreover, the absolute convergence of $B(z)$ implies that $\sum_{n=0}^{\infty} B_n z^n = B(z) \in \mathcal{B}(H)$ for all $z \in D(0, \rho)$, $\rho > 1$, so that $B(1) = \sum_{n=0}^{\infty} B_n$ is a well defined bounded linear operator. This is needed for the notions of integration and cointegration in Definition 2.4 below.

Finally note that the series obtained by termwise k times differentiation, $\sum_{n=k}^{\infty} n(n-1) \cdots (n-k+1) B_n z^{n-k}$, is absolutely convergent on $D(0, \rho)$, $\rho > 1$, and coincides with the k-th derivative of $B(z)$ for each $z \in D(0, \rho)$. Hence an absolutely convergent operator function is infinitely differentiable on $D(0, \rho)$. In the following $\Delta := 1 - L$ is the difference operator at frequency 0.

For simplicity, only the case of integration and cointegration at frequency 0 is considered in the following; similar definitions hold for any other root on the unit circle.

Definition 2.4 (Integrated and cointegrated processes at frequency 0)**.** *A linear process $u_t = B(L)\varepsilon_t$ is said to be integrated of order 0, written $u_t \sim I(0)$, if $B(1) = \sum_{n=0}^{\infty} B_n \neq 0$. An H-valued stochastic process $\{x_t, t = 0, 1, \dots\}$ is said to be integrated of order d (at frequency zero), written $x_t \sim I(d)$, if $\Delta^d x_t = B(L)\varepsilon_t \sim I(0)$. An $I(d)$ process x_t is said to be cointegrated (at frequency zero) if $B(1)$ is non-invertible; in this case any non-zero vector $v \in (\operatorname{Im} B(1))^{\perp}$ is such that $\langle v, x_t \rangle \sim I(d-j)$ for some $j > 0$ and any non-zero vector $v \in \operatorname{Im} B(1)$ is such that $\langle v, x_t \rangle \sim I(d)$. $(\operatorname{Im} B(1))^{\perp}$ is called the cointegrating space of x_t and $\operatorname{Im} B(1)$ is called the attractor space of the $I(d)$ trends.*

Observe that $I(0)$-ness implies weak stationarity but not viceversa and that a white noise is necessarily $I(0)$. Also note that the notions of integration and cointegration in Definition 2.4 are invariant to bounded linear invertible transformations of the process. That is, if $x_t \sim I(d)$ and A is a bounded linear invertible transformation then $Ax_t \sim I(d)$. Remark that the same invariance property holds replacing A with $A(z) = \sum_{n=0}^{\infty} A_n z^n$, $A_n \in \mathcal{B}(H)$, invertible and absolutely convergent for all $z \in D(0, \rho)$, $\rho > 1$.

The next definition introduces the class of processes that is studied in the present paper.

Definition 2.5 (Cointegrated H-valued AR at frequency 0)**.** *Consider an H-valued AR process of order s*

$$x_t = A_1^\circ x_{t-1} + \cdots + A_s^\circ x_{t-s} + \varepsilon_t,$$

where $A_n^\circ : H \to H$, $n = 1, \ldots, s$, is a bounded linear operator and $\{\varepsilon_t, t \in \mathbb{Z}\}$ is white noise as in Definition 2.2, and define the AR operator function $A(z) = I - A_1^\circ z - \cdots - A_s^\circ z^s$, $z \in \mathbb{C}$. An H-valued AR process $A(L)x_t = \varepsilon_t$ is said to be cointegrated (at frequency zero) if i) $A(1) \neq 0$, ii) $A(z)$ has an eigenvalue of finite type at $z = 1$, and iii) $A(z)$ is invertible in the punctured disc $D(0, \rho) \setminus \{1\}$ for some $\rho > 1$.

Remark 2.6. Beare et al. (2017) consider H-valued AR processes $\Phi(L)x_t = \varepsilon_t$, $\Phi(z) = I - \Phi_1 z - \cdots - \Phi_k z^s$, for which whenever $\Phi(z)$ is non-invertible one has either $z = 1$ or $|z| > 1$, and where Φ_n are compact if $s > 1$. Hu and Park (2016) consider H-valued AR processes $A(L)f_t = \varepsilon_t$, $A(z) = I - Az$, such that A is compact and 1 is an isolated eigenvalue of A.

The present setup contains the ones considered in Beare et al. (2017) and Hu and Park (2016) as special cases. In fact, because the sum of compact operators is compact, see Theorem 16.1 in Chapter II in Gohberg et al. (2003), and $I - B$ is Fredholm of index 0 if $B \in \mathcal{B}(H)$ is compact, see Theorem 4.2 in Chapter XV in Gohberg et al. (2003), $\Phi(1)$ in Beare et al. (2017) and $A(1)$ in Hu and Park (2016) are Fredholm of index 0 and 1 is isolated, i.e. 1 is an eigenvalue of finite type. For this reason both the processes considered by Beare et al. (2017) and Hu and Park (2016) are cointegrated H-valued AR process in the sense of Definition 2.5. Hu and Park (2016) also discuss estimators and asymptotic results for H-valued AR(1) models with a compact operator.

Remark 2.7. Chang et al. (2016), see their Assumption 2.1, study processes that satisfy $\Delta w_t = \Phi(L)\varepsilon_t = \sum_{n=0}^\infty \Phi_n \varepsilon_{t-n}$, where $\sum_{n=0}^\infty n\|\Phi_n\| < \infty$ and $\operatorname{Im}\Phi(1)$ is finite dimensional. As shown below, a cointegrated H-valued AR process necessarily meets these requirements. Hence their asymptotic analysis applies and their test can be employed in the present setup as well.

Remark 2.8. As shown in Chapter XI in Gohberg et al. (1990), the assumption that 1 is an eigenvalue of finite type of $A(z)$ implies that $A(1)$ is Fredholm of index 0, which means that $\dim\operatorname{Ker} A(1)$ and $\operatorname{codim}\operatorname{Im} A(1) = \dim(H/\operatorname{Im} A(1))$ are finite and equal. Because $\operatorname{Ker} A(1)$ is finite dimensional, it is complemented, see Theorem 5.7 in Chapter XI in Gohberg et al. (2003), and hence its orthogonal complement $(\operatorname{Ker} A(1))^\perp$ is closed. Moreover, because $H/\operatorname{Im} A(1)$ is finite dimensional, $\operatorname{Im} A(1)$ is closed, see Corollary 2.3 in Chapter XI in Gohberg et al. (1990). This shows that the subspaces $\operatorname{Im} A(1)$ and $(\operatorname{Ker} A(1))^\perp$ are closed and infinite dimensional while their orthogonal complements $(\operatorname{Im} A(1))^\perp$ and $\operatorname{Ker} A(1)$ are closed and finite dimensional. Hence $(\operatorname{Ker} A(1))^\perp = \operatorname{Im} A(1)^*$, where $A(1)^*$ is the adjoint of $A(1)$.

3. Existence results for cointegrated H-valued AR processes

This section presents an existence result in Theorem 3.1 for cointegrated H-valued AR processes $A(L)x_t = \varepsilon_t$. This result is a direct consequence of Theorem A.1 in Appendix A. Theorem 3.1 shows that $x_t \sim I(d)$ for some $d = 1, 2, \dots$ has necessarily finitely many $I(d)$ trends, finitely many $I(j)$, $1 \le j \le d-1$, cointegrating relations and infinitely many $I(0)$ cointegrating relations. Apart from the fact that the number of $I(0)$ cointegrating relations is infinite, this mirrors the finite dimensional case in Franchi and Paruolo (2017).

Theorem 3.1 (Existence results for cointegrated H-valued AR processes)**.** *Let $A(L)x_t = \varepsilon_t$ be a cointegrated H-valued AR process and assume that x_t is $I(d)$ for some $d = 1, 2, \dots$. Then there exist integers $0 = j_0 < j_1 < j_2 < \cdots < j_w = d$ and projections $P_0, P_1, \dots, P_w$ such that*

$$H = \operatorname{Im} P_0 \oplus \operatorname{Im} P_1 \oplus \cdots \oplus \operatorname{Im} P_w,$$

where $\operatorname{Im} P_0, \operatorname{Im} P_1, \dots, \operatorname{Im} P_w$ are closed, $\dim \operatorname{Im} P_0 = \infty$ and $0 < \dim \operatorname{Im} P_h < \infty$, $h = 1, \dots, w$. In this case,

$$x_t = C_0 s_{d,t} + C_1 s_{d-1,t} + \cdots + C_{d-1} s_{1,t} + C^\star(L)\varepsilon_t + v_0, \qquad \operatorname{Im} C_0 = \operatorname{Im} P_w,$$

where $s_{h,t} = \sum_{i=1}^t s_{h-1,i} \sim I(h)$, $s_{0,t} = \varepsilon_t$, $C^\star(L)\varepsilon_t$ is a linear process, and v_0 collects initial values. This shows that $\operatorname{Im} P_0 \oplus \operatorname{Im} P_1 \oplus \cdots \oplus \operatorname{Im} P_{w-1}$ is the cointegrating space of x_t and $\operatorname{Im} P_w$ is the attractor space of the $I(d)$ trends. Moreover, there exists an operator function $G(z) = \sum_{n=0}^\infty G_n(1-z)^n$, $G_0 = I$, invertible and absolutely convergent for all $z \in D(0, \rho)$, $\rho > 1$, such that

$$\langle v, x_t \rangle + \sum_{n=1}^{d-j_h-1} \langle v, G_n \Delta^n x_t \rangle$$

is $I(j_h)$ for all $v \in \operatorname{Im} P_h$.

Remark 3.2. In the $I(1)$ case, one has $w = 1$, $0 = j_0 < j_1 = 1$ in Theorem 3.1 and

$$x_t = C_0 \sum_{i=1}^t \varepsilon_i + C^\star(L)\varepsilon_t + v_0, \qquad \operatorname{Im} C_0 = \operatorname{Im} P_1, \qquad H = \operatorname{Im} P_0 \oplus \operatorname{Im} P_1,$$

where $C^\star(L)\varepsilon_t$ is a linear process, v_0 collects initial values, $\operatorname{Im} P_h$, $h = 0, 1$ is closed, $\dim \operatorname{Im} P_0 = \infty$ and $0 < \dim \operatorname{Im} P_1 < \infty$. Because $\operatorname{Im} C_0 = \operatorname{Im} P_1$ one has that $\operatorname{Im} P_0$ is the cointegrating space of x_t and $\operatorname{Im} P_1$ is the attractor space of the $I(1)$ trends. Moreover,

$$\langle v, x_t \rangle \sim I(0) \text{ for all } v \in \operatorname{Im} P_0, \qquad \langle v, x_t \rangle \sim I(1) \text{ for all } v \in \operatorname{Im} P_1.$$

Observe that the number of $I(1)$ trends in x_t is equal to $\dim \operatorname{Im} P_1$ and the number of $I(0)$ cointegrating relations is equal to $\dim \operatorname{Im} P_0$, so that x_t has finitely many $I(1)$ trends and infinitely many $I(0)$ cointegrating relations. This fact is also documented in Hu and Park (2016) and in Beare et al. (2017). Apart from the fact that the number of $I(0)$ cointegrating relations is infinite, this echoes the finite dimensional case, see Theorem 4.2 in Johansen (1996).

Remark 3.3. In the $I(2)$ case, one has either $w = 1$ or $w = 2$. If $w = 1$, $0 = j_0 < j_1 = 2$ and

$$x_t = C_0 \sum_{s=1}^{t} \sum_{i=1}^{s} \varepsilon_i + C_1 \sum_{i=1}^{t} \varepsilon_i + C^\star(L)\varepsilon_t + v_0, \qquad \operatorname{Im} C_0 = \operatorname{Im} P_1, \qquad H = \operatorname{Im} P_0 \oplus \operatorname{Im} P_1,$$

where $C^\star(L)\varepsilon_t$ is a linear process, v_0 collects initial values, $\operatorname{Im} P_h$, $h = 0, 1$ is closed, $\dim \operatorname{Im} P_0 = \infty$ and $0 < \dim \operatorname{Im} P_1 < \infty$. Because $\operatorname{Im} C_0 = \operatorname{Im} P_1$ one has that $\operatorname{Im} P_0$ is the cointegrating space of x_t and $\operatorname{Im} P_1$ is the attractor space of the $I(2)$ trends. Moreover,

$$\langle v, x_t\rangle + \langle v, G_1 \Delta x_t\rangle \sim I(0) \text{ for all } v \in \operatorname{Im} P_0, \qquad \langle v, x_t\rangle \sim I(2) \text{ for all } v \in \operatorname{Im} P_1.$$

Observe that the number of $I(2)$ trends in x_t is equal to $\dim \operatorname{Im} P_1$ and the number of $I(0)$ cointegrating relations is equal to $\dim \operatorname{Im} P_0$, so that x_t has finitely many $I(2)$ trends and infinitely many $I(0)$ cointegrating relations. Futher observe that v is either not cointegrating (when $v \in \operatorname{Im} P_1$) or it allows for polynomial cointegration (when $v \in \operatorname{Im} P_0$) of order 0.

In the I(2) case when $w = 2$, this does not hold; in fact, one has that $0 = j_0 < j_1 = 1 < j_2 = 2$ and

$$x_t = C_0 \sum_{s=1}^{t} \sum_{i=1}^{s} \varepsilon_i + C_1 \sum_{i=1}^{t} \varepsilon_i + C^\star(L)\varepsilon_t + v_0, \qquad \operatorname{Im} C_0 = \operatorname{Im} P_2, \qquad H = \operatorname{Im} P_0 \oplus \operatorname{Im} P_1 \oplus \operatorname{Im} P_2,$$

where $C^\star(L)\varepsilon_t$ is a linear process, v_0 collects initial values, $\operatorname{Im} P_h$, $h = 0, 1, 2$ is closed, $\dim \operatorname{Im} P_0 = \infty$ and $0 < \dim \operatorname{Im} P_h < \infty$, $h = 1, 2$. Because $\operatorname{Im} C_0 = \operatorname{Im} P_2$ one has that $\operatorname{Im} P_0 \oplus \operatorname{Im} P_1$ is the cointegrating space of x_t and $\operatorname{Im} P_2$ is the attractor space of the $I(2)$ trends. Moreover,

$$\langle v, x_t\rangle + \langle v, G_1 \Delta x_t\rangle \sim I(0) \text{ for all } v \in \operatorname{Im} P_0, \qquad \langle v, x_t\rangle \sim I(1) \text{ for all } v \in \operatorname{Im} P_1$$

and $\langle v, x_t\rangle \sim I(2)$ for all $v \in \operatorname{Im} P_2$, so that there are cointegrating vectors that allow for polynomial cointegration (when $v \in \operatorname{Im} P_0$) of order 0 and others that don't (when $v \in \operatorname{Im} P_1$). Observe that the number of $I(2)$ trends in x_t is equal to $\dim \operatorname{Im} P_2$, the number of $I(1)$ cointegrating relations is equal to $\dim \operatorname{Im} P_1$ and the number of $I(0)$ cointegrating relations is equal to $\dim \operatorname{Im} P_0$. Hence x_t has finitely many $I(2)$ trends, finitely many $I(1)$ cointegrating relations and infinitely many $I(0)$ cointegrating relations. Apart from the fact that the number of $I(0)$ cointegrating relations is infinite, this mimics the finite dimensional case, see Theorem 4.6 in Johansen (1996).

The same structure applies in general: from Theorem 3.1 one has that the number of $I(d)$ trends is equal to $\dim \operatorname{Im} P_d$ and the number of $I(j_h)$ cointegrating relations is equal to $\dim \operatorname{Im} P_h$, $h = 0, \ldots, w-1$. Because $\dim \operatorname{Im} P_0 = \infty$ and $0 < \dim \operatorname{Im} P_h < \infty$, $h = 1, \ldots, w$, x_t has finitely many $I(d)$ trends, finitely many $I(j_h)$, $h = 1, \ldots, w-1$, cointegrating relations and infinitely many $I(0)$ cointegrating relations. As in the $I(1)$ and $I(2)$ cases above, the finite and the infinite dimensional cases are similar, apart from the fact that the number of $I(0)$ cointegrating relations is infinite, see Theorem 3.3 in Franchi and Paruolo (2017).

Finally observe that Theorem 3.1 provides information about the existence of the projections P_h, the dimensions of their images, the orders of integration j_h, and the G_n operators which are relevant to describe the properties of a generic $I(d)$ cointegrated H-valued AR process. However, it is completely silent about how those relevant quantities are related to the structure of the AR operators and how one can construct them in practice. Their form is described in Section 4 for the $I(1)$ and $I(2)$ cases and in Section 5 for the generic $I(d)$ case.

4. A characterization of $I(1)$ and $I(2)$ cointegrated H-valued AR processes

This section presents a characterization of $I(1)$ and $I(2)$ cointegrated H-valued AR processes $A(L)x_t = \varepsilon_t$. The I(1) case parallels the results in Beare et al. (2017), while the results for the I(2) case are novel for the infinite dimensional case. All the relevant quantities are expressed in terms of the coefficients of the expansion of the AR operator function $A(z) = I - \sum_{n=1}^{s} A_n^\circ z^n$ around 1,

$$A(z) = \sum_{h=0}^{\infty} A_h(1-z)^h, \qquad A_h = \begin{cases} I - \sum_{n=1}^{s} A_n^\circ & \text{for } h = 0 \\ (-1)^{h+1} \sum_{n=0}^{s-h} \binom{n+h}{h} A_{n+h}^\circ & \text{for } h = 1, 2, \dots \end{cases}. \tag{4.1}$$

In Theorem 4.1 below a necessary and sufficient condition for $x_t \sim I(1)$ is given in terms of the decomposition of the space into the sum of two closed subspaces, $H = \tau_0 \oplus \tau_1$, that are defined using A_0 and A_1, see (4.2) and (4.3) below. The infinite dimensional cointegrating space coincides with τ_0 and the finite dimensional attractor space of the $I(1)$ trends with τ_1. In Remark 4.2 the equivalence with the condition in Beare et al. (2017) is proved.

In Theorem 4.4 a necessary and sufficient condition for $x_t \sim I(2)$ is given in terms of the decomposition of the space into the sum of three closed subspaces, $H = \tau_0 \oplus \tau_1 \oplus \tau_2$, where τ_2 is defined in (4.4) below using A_0, A_1 and A_2. The infinite dimensional cointegrating space coincides with $\tau_0 \oplus \tau_1$ and τ_2 is the finite dimensional attractor space of the $I(2)$ trends. In τ_0, which is infinite dimensional, one finds the cointegrating vectors that allow for polynomial cointegration and in τ_1, which is finite dimensional (and can as well be equal to 0), those that don't allow for polynomial cointegration.

The following notation is employed: consider A_0 in (4.1) and define

$$\zeta_0 = \operatorname{Im} A_0, \qquad \zeta_0^\perp = (\operatorname{Im} A_0)^\perp, \qquad \tau_0 = (\operatorname{Ker} A_0)^\perp, \qquad \tau_0^\perp = \operatorname{Ker} A_0. \tag{4.2}$$

As shown in Remark 2.8, the assumption that 1 is an eigenvalue of finite type of $A(z)$ implies that the subspaces ζ_0 and τ_0 are closed and infinite dimensional, so that $\tau_0 = (\operatorname{Ker} A_0)^\perp = \operatorname{Im} A_0^*$, while their orthogonal complements $\zeta_0^\perp$ and $\tau_0^\perp$ are closed and finite dimensional. In the following, P_x indicates the orthogonal projection on x.

Theorem 4.1 (A characterization of $I(1)$ cointegrated H-valued AR processes). *Consider a cointegrated H-valued AR process $A(L)x_t = \varepsilon_t$, where $A(z)$ is written as in (4.1). Let ζ_0 and τ_0 be as in (4.2) and define*

$$\zeta_1 = \operatorname{Im} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}, \qquad \tau_1 = (\operatorname{Ker} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp})^\perp. \tag{4.3}$$

Then x_t is $I(1)$ if and only if

$$H = \tau_0 \oplus \tau_1,$$

called the I(1) condition, where τ_0 *and* τ_1 *are closed,* $\dim \tau_0 = \infty$ *and* $0 < \dim \tau_1 < \infty$*. In this case,*

$$x_t = C_0 \sum_{i=1}^{t} \varepsilon_i + C^\star(L)\varepsilon_t + v_0, \qquad \operatorname{Im} C_0 = \tau_1,$$

where $C^\star(L)\varepsilon_t$ *is a linear process and* v_0 *collects initial values. This shows that* τ_0 *is the cointegrating space of* x_t *and* τ_1 *is the attractor space of the* $I(1)$ *trends. Moreover,*

$$\langle v, x_t \rangle \sim I(0) \text{ for all } v \in \tau_0$$

and $\langle v, x_t \rangle \sim I(1)$ *for all* $v \in \tau_1$*. Finally, the* $I(1)$ *condition can be equivalently stated as* $H = \zeta_0 \oplus \zeta_1$*, where* ζ_0 *and* ζ_1 *are closed,* $\dim \zeta_0 = \infty$ *and* $0 < \dim \zeta_1 < \infty$*.*

Remark 4.2. As shown next, the $I(1)$ condition in Theorem 4.1 is equivalent to the $I(1)$ condition in Beare et al. (2017), $H = \zeta_0 \oplus A_1 \tau_0^\perp$, see their equation (4.15). First assume that $H = \zeta_0 \oplus A_1 \tau_0^\perp$; then $I = P_{\zeta_0} + P_{A_1 \tau_0^\perp}$, so that $I - P_{\zeta_0} = P_{\zeta_0^\perp} = P_{A_1 \tau_0^\perp}$. This implies that $A_1 \tau_0^\perp = P_{\zeta_0^\perp} A_1 \tau_0^\perp = \operatorname{Im} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} = \zeta_1$. This shows that $H = \zeta_0 \oplus \zeta_1$, i.e. the $I(1)$ condition in Theorem 4.1 holds. Conversely, assume that $H = \zeta_0 \oplus \zeta_1$. Because $\zeta_1 = \operatorname{Im} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}$, one has that $\zeta_0^\perp = \zeta_1 \subseteq \operatorname{Im} A_1 P_{\tau_0^\perp}$ and hence $\dim \zeta_0^\perp \leq \dim A_1 \tau_0^\perp$; because $\dim A_1 \tau_0^\perp \leq \dim \tau_0^\perp = \dim \zeta_0^\perp$, one thus has $\zeta_1 = \operatorname{Im} A_1 P_{\tau_0^\perp} = A_1 \tau_0^\perp$. This shows that $H = \zeta_0 \oplus A_1 \tau_0^\perp$, i.e. the $I(1)$ condition in Beare et al. (2017) holds.

Remark 4.3. In the finite dimensional case $H = \mathbb{R}^p$, Franchi and Paruolo (2016) show that the $I(1)$ condition in Theorem 4.2 in Johansen (1996) can be equivalently stated as $\mathbb{R}^p = \zeta_0 \oplus \zeta_1 = \tau_0 \oplus \tau_1$, where $\zeta_h = \operatorname{span}(\alpha_h)$, $\tau_h = \operatorname{span}(\beta_h)$, $h = 0, 1$, and the bases α_h, β_h are defined by the rank factorizations $A_0 = \alpha_0 \beta_0'$ and $P_{\alpha_0^\perp} A_1 P_{\beta_0^\perp} = \alpha_1 \beta_1'$, i.e. they are full column rank matrices that respectively span the column space and the row space of the corresponding matrix. Apart from the fact that $\dim \zeta_0 = \dim \tau_0 = \operatorname{rank} A_0$ is finite when $H = \mathbb{R}^p$, this mirrors what happens in the present infinite dimensional case.

The $I(2)$ case is presented next.

Theorem 4.4 (A characterization of $I(2)$ cointegrated H-valued AR processes)**.** *Consider a cointegrated* H*-valued AR process* $A(L)x_t = \varepsilon_t$*, where* $A(z)$ *is written as in* (4.1)*. Let* ζ_0 *and* τ_0 *be as in* (4.2)*, let* ζ_1 *and* τ_1 *as in* (4.3) *and let* $\mathscr{X}_2 = \zeta_0 \oplus \zeta_1$ *and* $\mathscr{T}_2 = \tau_0 \oplus \tau_1$*; further define*

$$\zeta_2 = \operatorname{Im} P_{\mathscr{X}_2^\perp} A_{2,1} P_{\mathscr{T}_2^\perp}, \qquad \tau_2 = (\operatorname{Ker} P_{\mathscr{X}_2^\perp} A_{2,1} P_{\mathscr{T}_2^\perp})^\perp, \qquad A_{2,1} = A_2 - A_1 Q_0 A_1, \qquad Q_0 = A_0^+, \tag{4.4}$$

where Q_0 *is the generalized inverse of* A_0*. Then* x_t *is* $I(2)$ *if and only if*

$$H = \tau_0 \oplus \tau_1 \oplus \tau_2,$$

called the I(2) condition, where τ_0*,* τ_1 *and* τ_2 *are closed,* $\dim \tau_0 = \infty$*,* $0 \leq \dim \tau_1 < \infty$*, and* $0 < \dim \tau_2 < \infty$*. In this case,*

$$x_t = C_0 \sum_{s=1}^{t} \sum_{i=1}^{s} \varepsilon_i + C_1 \sum_{i=1}^{t} \varepsilon_i + C^\star(L)\varepsilon_t + v_0, \qquad \operatorname{Im} C_0 = \tau_2,$$

where $C^\star(L)\varepsilon_t$ *is a linear process and* v_0 *collects initial values. This shows that* $\tau_0 \oplus \tau_1$ *is the cointegrating space of* x_t *and* τ_2 *is the attractor space of the* $I(2)$ *trends. Moreover,*

$$\langle v, x_t \rangle + \langle v, Q_0 A_1 \Delta x_t \rangle \sim I(0) \text{ for all } v \in \tau_0, \qquad \langle v, x_t \rangle \sim I(1) \text{ for all } v \in \tau_1$$

and $\langle v, x_t \rangle \sim I(2)$ *for all* $v \in \tau_2$*. Finally, the* $I(2)$ *condition can be equivalently stated as* $H = \zeta_0 \oplus \zeta_1 \oplus \zeta_2$*, where* ζ_0*,* ζ_1 *and* ζ_2 *are closed,* $\dim \zeta_0 = \infty$*,* $0 \leq \dim \zeta_1 < \infty$*, and* $0 < \dim \zeta_2 < \infty$*.*

Remark 4.5. As shown in Theorem 3 in Chapter 9 in Ben-Israel and Greville (2003), if $A \in \mathcal{B}(H)$ and $\operatorname{Im} A$ is closed, then its generalized inverse A^+ exists and it is the unique solution of the system $AA^+A = A, A^+AA^+ = A^+, (AA^+)^* = AA^+, (A^+A)^* = A^+A$. The assumption that 1 is an eigenvalue of finite type of $A(z)$ implies that $\zeta_0 = \operatorname{Im} A_0$ is closed, see Remark 2.8; hence Q_0 exists and it is unique.

Remark 4.6. In the finite dimensional case $H = \mathbb{R}^p$, Franchi and Paruolo (2016) show that the $I(2)$ condition in Theorem 4.6 in Johansen (1996) can be equivalently stated as $\mathbb{R}^p = \zeta_0 \oplus \zeta_1 \oplus \zeta_2 = \tau_0 \oplus \tau_1 \oplus \tau_2$, where $\zeta_h = \operatorname{span}(\alpha_h)$, $\tau_h = \operatorname{span}(\beta_h)$, $h = 0, 1, 2$, and the bases α_h, β_h are defined by the rank factorizations $A_0 = \alpha_0 \beta_0'$, $P_{\alpha_0^\perp} A_1 P_{\beta_0^\perp} = \alpha_1 \beta_1'$ and $P_{a_2^\perp}(A_2 - A_1 Q_0 A_1) P_{b_2^\perp} = \alpha_2 \beta_2'$, where $a_2 = (\alpha_0, \alpha_1)$, $b_2 = (\beta_0, \beta_1)$ and $Q_0 = \beta_0(\beta_0'\beta_0)^{-1}(\alpha_0'\alpha_0)^{-1}\alpha_0'$. Again here, apart from the fact that $\dim \zeta_0 = \dim \tau_0 = \operatorname{rank} A_0$ is finite when $H = \mathbb{R}^p$, this is exactly what happens in the infinite dimensional case.

Summing up, Theorem 4.1 shows that $d = 1$ if and only if 2 circumstances hold: the first, $\tau_0 \subset H$ (or equivalently $\zeta_0 \subset H$), establishes that $d > 0$ and the second, $\tau_1 = \tau_0^\perp$ (or equivalently $\zeta_1 = \zeta_0^\perp$), establishes that $d = 1$. In τ_0, which is infinite dimensional, one finds the $I(0)$ cointegrating components and in τ_1, which is finite dimensional and different from 0, those that are not cointegrating. The combinations of these circumstances makes up the I(1) conditions.

Similarly, Theorem 4.4 shows that $d = 2$ if and only if 3 circumstances hold: the first, $\tau_0 \subset H$ (or equivalently $\zeta_0 \subset H$), establishes that $d > 0$, the second, $\tau_1 \subset \tau_0^\perp$ (or equivalently $\zeta_1 \subset \zeta_0^\perp$), establishes that $d > 1$ and the third, $\tau_2 = (\tau_0 \oplus \tau_1)^\perp$ (or equivalently $\zeta_2 = (\zeta_0 \oplus \zeta_1)^\perp$) establishes that $d = 2$. In τ_0, which is infinite dimensional, one finds the cointegrating components that allow for polynomial cointegration of order 0, in τ_1, which is finite dimensional (and can as well be equal to 0), those that are $I(1)$ and don't allow for polynomial cointegration, and in τ_2, which is finite dimensional (and different from 0), those that are not cointegrating. The combinations of these circumstances makes up the I(2) conditions.

As shown in the next section, this construction is true in the general I(d) case.

5. The general result

This section extends the results in Section 4 to the general $I(d)$, $d = 1, 2, \dots,$ case. Theorem 5.4 provides a necessary and sufficient condition for an H-valued AR processes $A(L)x_t = \varepsilon_t$ to be $I(d)$ and it is shows that under this condition the space is decomposed into the sum of $d+1$ closed subspaces, $H = \tau_0 \oplus \tau_1 \oplus \cdots \oplus \tau_d$, that are defined in terms of $A_0, A_1, \dots, A_d$ in (4.1), see Definition 5.1 below. The infinite dimensional cointegrating space coincides with $\tau_0 \oplus \tau_1 \oplus \cdots \oplus \tau_{d-1}$ and τ_d is the finite dimensional attractor space of the $I(d)$ trends. In τ_0, which is infinite dimensional, one finds the cointegrating vectors that allow for polynomial cointegration of order 0, in τ_h, $h = 1, \dots, d-2$, which is finite dimensional and can as well be equal to 0, those that allow for polynomial cointegration of order h and in τ_{d-1}, which is finite dimensional and can as well be equal to 0, those that are $I(d-1)$ and don't allow for polynomial cointegration. Finally, in τ_d, which is finite dimensional and different from 0, those that are $I(d)$ and don't allow for cointegration.

Definition 5.1 (ζ_h, τ_h subspaces and $Q_h, A_{h,n}$ operators)**.** Consider a cointegrated H-valued AR process $A(L)x_t = \varepsilon_t$, where $A(z)$ is written as in (4.1). Let

$$\zeta_0 = \operatorname{Im} A_0, \qquad \tau_0 = (\operatorname{Ker} A_0)^\perp, \qquad Q_0 = A_0^+,$$

where Q_0 is the generalized inverse of A_0, and for $h = 1, 2, \dots$ define

$$S_h = P_{\mathscr{Z}_h^\perp} A_{h,1} P_{\mathscr{T}_h^\perp}, \qquad \zeta_h = \operatorname{Im} S_h, \qquad \tau_h = (\operatorname{Ker} S_h)^\perp, \qquad Q_h = S_h^+,$$

where Q_h is the generalized inverse of S_h,

$$\mathscr{Z}_h = \zeta_0 \oplus \cdots \oplus \zeta_{h-1}, \qquad \mathscr{T}_h = \tau_0 \oplus \cdots \oplus \tau_{h-1},$$

and

$$A_{h,n} = \begin{cases} A_n & \text{for } h = 1 \\ A_{h-1,n+1} + A_{h-1,1} \sum_{j=0}^{h-2} Q_j A_{j+1,n} & \text{for } h = 2, 3, \dots \end{cases}, \qquad n = 1, 2, \dots .$$

Remark 5.2. As shown in Remark 2.8, the assumption that 1 is an eigenvalue of finite type of $A(z)$ implies that the subspaces $\zeta_0 = \operatorname{Im} A_0$ and $\tau_0 = (\operatorname{Ker} A_0)^\perp = \operatorname{Im} A_0^*$ are closed and infinite dimensional while their orthogonal complements $\zeta_0^\perp = (\operatorname{Im} A_0)^\perp$ and $\tau_0^\perp = \operatorname{Ker} A_0 = (\operatorname{Im} A_0^*)^\perp$ are closed and finite dimensional. This implies that the subspaces ζ_h, τ_h, $h = 1, 2, \dots,$ are closed and finite dimensional. Because ζ_h, $h = 0, 1, \dots,$ is closed, the corresponding generalized inverse Q_h, $h = 0, 1, \dots,$ exists and it is unique, see Remark 4.5.

Remark 5.3. By construction, for $h \neq s$, ζ_s is orthogonal to ζ_h and τ_s is orthogonal to τ_h; moreover, it is possible that ζ_h, τ_h, $h \neq 0$, are equal to 0.

Theorem 5.4 (A characterization of $I(d)$ cointegrated H-valued AR processes)**.** *Consider a cointegrated H-valued AR process $A(L)x_t = \varepsilon_t$, where $A(z)$ is written as in (4.1), and let ζ_h, τ_h, Q_h, and $A_{h,n}$ be as in Definition 5.1. Then x_t is $I(d)$ if and only if*

$$H = \tau_0 \oplus \tau_1 \oplus \cdots \oplus \tau_d,$$

where $\tau_0, \tau_1, \ldots, \tau_d$ *are closed,* $\dim \tau_0 = \infty$*,* $0 \leq \dim \tau_h < \infty$ *for any* $h \neq 0, d$*, and* $0 < \dim \tau_d < \infty$*. In this case,*

$$x_t = C_0 s_{d,t} + C_1 s_{d-1,t} + \cdots + C_{d-1} s_{1,t} + C^\star(L)\varepsilon_t + v_0, \qquad \operatorname{Im} C_0 = \tau_d,$$

where $s_{h,t} = \sum_{i=1}^t s_{h-1,i} \sim I(h)$*,* $s_{0,t} = \varepsilon_t$*,* $C^\star(L)\varepsilon_t$ *is a linear process, and* v_0 *collects initial values. This shows that* $\tau_0 \oplus \tau_1 \cdots \oplus \tau_{d-1}$ *is the cointegrating space of* x_t *and* τ_d *is the attractor space of the* $I(d)$ *trends. Moreover,*

$$\langle v, x_t \rangle + \sum_{n=1}^{d-h-1} \langle v, Q_h A_{h+1,n} \Delta^n x_t \rangle \sim I(h) \text{ for all } v \in \tau_h, \qquad h = 0, \ldots, d-2,$$

$$\langle v, x_t \rangle \sim I(d-1) \text{ for all } v \in \tau_{d-1},$$

and $\langle v, x_t \rangle \sim I(d)$ *for all* $v \in \tau_d$*. Finally, the* $I(d)$ *condition can be equivalently stated as* $H = \zeta_0 \oplus \zeta_1 \oplus \cdots \oplus \zeta_d$*, where* $\zeta_0, \zeta_1, \ldots, \zeta_d$ *are closed,* $\dim \zeta_0 = \infty$*,* $0 \leq \dim \zeta_h < \infty$ *for any* $h \neq 0, d$*, and* $0 < \dim \zeta_d < \infty$*.*

Remark 5.5. In the finite dimensional case $H = \mathbb{R}^p$, Franchi and Paruolo (2016) show that $d = 1, 2, \ldots$ if and only if $\mathbb{R}^p = \zeta_0 \oplus \cdots \oplus \zeta_d = \tau_0 \oplus \cdots \oplus \tau_d$, where $\zeta_h = \operatorname{span}(\alpha_h)$, $\tau_h = \operatorname{span}(\beta_h)$, $h = 0, 1, \ldots$, and the bases α_h, β_h are defined by the rank factorizations $P_{a_h^\perp} A_{h,1} P_{b_h^\perp} = \alpha_h \beta_h'$, where $a_h = (\alpha_0, \ldots, \alpha_{h-1})$, $b_h = (\beta_0, \ldots, \beta_{h-1})$, $Q_h = \beta_h(\beta_h'\beta_h)^{-1}(\alpha_h'\alpha_h)^{-1}\zeta_h'$ and $A_{h,1}$ is as in Definition 5.1. Again here, apart from the fact that $\dim \zeta_0 = \dim \tau_0 = \operatorname{rank} A_0$ is finite when $H = \mathbb{R}^p$, this mirrors what happens in the infinite dimensional case. Hence the infinite dimensionality of the space does not introduce additional elements in the $I(d)$ analysis.

Summing up, Theorem 5.4 shows that $d = 1, 2, \ldots$ if and only if $d + 1$ conditions hold: the first, $\tau_0 \subset H$ (or equivalently $\zeta_0 \subset H$), establishes that $d > 0$ and for $h = 1, \ldots, d-1$, $\tau_h \subset (\tau_0 \oplus \cdots \oplus \tau_{h-1})^\perp$ (or equivalently $\zeta_h \subset (\zeta_0 \oplus \cdots \oplus \zeta_{h-1})^\perp$), establishes that $d > h$ and the last one, $\tau_d = (\tau_0 \oplus \cdots \oplus \tau_{d-1})^\perp$ (or equivalently $\zeta_d = (\zeta_0 \oplus \cdots \oplus \zeta_{d-1})^\perp$), establishes the value of $d = 1, 2, \ldots$. The infinite dimensional cointegrating space coincides with $\tau_0 \oplus \tau_1 \oplus \cdots \oplus \tau_{d-1}$ and τ_d is the finite dimensional attractor space of the $I(d)$ trends. In τ_0, which is infinite dimensional, one finds the cointegrating vectors that allow for polynomial cointegration of order 0, in τ_h, $h = 1, \ldots, d-2$, which is finite dimensional and can as well be equal to 0, those that allow for polynomial cointegration of order h and in τ_{d-1}, which is finite dimensional and can as well be equal to 0, those that are $I(d-1)$ and don't allow for polynomial cointegration. The coefficients of the polynomial cointegrating relations are $Q_h A_{h+1,n}$, which are calculated recursively as in Definition 5.1.

6. Conclusion

The present paper characterizes the cointegration properties of an H-valued AR process $A(L)x_t = \varepsilon_t$ under the assumptions that *i)* $A(1) \neq 0$, *ii)* $A(z)$ has an eigenvalue of finite type at $z = 1$, and *iii)* $A(z)$ is invertible in the punctured disc $D(0, \rho) \setminus \{1\}$ for some $\rho > 1$.

A necessary and sufficient condition for $x_t \sim I(d)$, $d = 1, 2, \ldots,$ is given and it is shown that under this condition the space is decomposed into the sum of $d+1$ closed subspaces that are defined recursively from the AR operators, $H = \tau_0 \oplus \tau_1 \oplus \cdots \oplus \tau_d$, where τ_0 is closed and infinite dimensional and τ_h, $h = 1, \ldots, d$ is closed and finite dimensional. The infinite dimensional subspace $\tau_0 \oplus \tau_1 \oplus \cdots \oplus \tau_{d-1}$ is the cointegrating space of x_t while the finite dimensional subspace τ_d is the attractor space of the $I(d)$ stochastic trends. Hence an H-valued cointegrated AR process has a finite number of $I(d)$ trends and infinitely many cointegrating relations.

The properties of $\langle v, x_t \rangle$ vary with v in the cointegrating space: for $v \in \tau_0$, which is infinite dimensional, one can combine $\langle v, x_t \rangle$ with differences of the process and find $I(0)$ polynomial cointegrating relations, for $v \in \tau_1$, which is finite dimensional and can as well be equal to 0, one can combine $\langle v, x_t \rangle$ with differences and find at most $I(1)$ polynomial cointegrating relations, and so on up to $v \in \tau_{d-1}$, which is finite dimensional and can as well be equal to 0, for which $\langle v, x_t \rangle \sim I(d-1)$ does not allow for polynomial cointegration. For any v in the cointegrating space, the explicit expression $Q_h A_{h+1,n}$ of the coefficients of the polynomial cointegrating relations is provided in terms of operators that are defined by the same recursion of the τ_h.

The present results show that the infinite dimensionality of the space does not introduce additional elements in the analysis, under the assumption that 1 is an eigenvalue of finite type of the AR operator function and that no other non-zero eigenvalue of the AR operator lies within or on the unit circle. That is, apart from the fact that the number of $I(0)$ cointegrating relations is infinite, the conditions and properties of H-valued cointegrated AR process coincide with those that apply in the finite dimensional case.

Appendix A. Separable Hilbert spaces and operators acting on them

This section reviews definitions (typewritten in italics) and basic facts on separable Hilbert spaces and on operators that act on them. The material is based on Chapter I, II and XV in Gohberg et al. (2003) and Chapter XI in Gohberg et al. (1990). This section also introduces the basic system (A.2), which is central in the proofs in Appendix *C*, and a useful intermediate result in Lemma *A*.2.

Let H be an Hilbert space; H is called *separable* if there exist vectors $v_1, v_2, \dots$ which span a subspace dense in H, i.e. every vector in H is the limit of a sequence of vectors in $\mathrm{span}(v_1, v_2, \dots)$. It can be shown that (only) separable Hilbert spaces have countable orthonormal bases and that a closed subspace of a separable Hilbert space is separable. A separable Hilbert space H is said to be the *direct sum* of subspaces M and N, written $H = M \oplus N$, if every vector $v \in H$ has a unique representation of the form $v = x + y$, where $x \in M$ and $y \in N$. The dimension of N, written $\operatorname{codim} M$, is called the *codimension* of M and M is said to be *complemented* if N is closed.

Let H be a separable Hilbert space with inner product $\langle \cdot , \cdot \rangle$ and norm $\| \cdot \|$; a function A which maps H into H, $A : H \to H$, is called a *linear operator* if for all $x, y \in H$ and $c \in \mathbb{C}$, $A(x + y) = Ax + Ay$ and $A(cx) = cAx$, where Az and $A(z)$ both indicate the action of A on z. A linear operator $A : H \to H$ is called *bounded* if $\sup_{\|x\|=1} \|Ax\| < \infty$ and its norm $\|A\|$ is given by $\sup_{\|x\|=1} \|Ax\|$. The set of bounded linear operators which map H into H is denoted by $\mathcal{B}(H)$.

An operator $A \in \mathcal{B}(H)$ is said to be *invertible* if there exists an operator $B \in \mathcal{B}(H)$ such that $BAx = ABx = x$ for every $x \in H$; in this case B is called the *inverse* of A, written A^{-1}. $A \in \mathcal{B}(H)$ is said to be a *projection* if $A^2 = A$ and it can be shown that if A is a projection, $\operatorname{Ker} A$ is complemented, i.e. $H = \operatorname{Im} A \oplus \operatorname{Ker} A$ and $\operatorname{Im} A = \operatorname{Ker}(I - A)$ is closed.

An operator $A \in \mathcal{B}(H)$ is said to be *Fredholm of index* $n(A) - d(A)$ if the numbers $n(A) = \dim \operatorname{Ker} A$ and $d(A) = \operatorname{codim} \operatorname{Im} A$ are finite. It can be shown that if H is finite dimensional, any operator is Fredholm of index 0. Let $D(u, \rho) = \{z \in \mathbb{C} : |z - u| < \rho\}$ be the open disc centered in u with radius $\rho > 0$ and let $F : D(u, \rho) \to \mathcal{B}(H)$ be an absolutely convergent operator function; a point $z_0 \in D(u, \rho)$ is said to be an *eigenvalue of finite type* of $F(z)$ if $F(z_0)$ is Fredholm, $F(z_0)x = 0$ for some non-zero $x \in H$ and $F(z)$ is invertible for all z in some punctured disc $D(z_0, \rho) \setminus \{z_0\}$. It can be shown that if z_0 is an eigenvalue of finite type then $F(z_0)$ is Fredholm of index 0.

Let $T, W : D(u, \rho) \to \mathcal{B}(H)$ be absolutely convergent operator functions; $T(z)$ is said to be *locally equivalent* at z_0 to $W(z)$ if *i)* $T(z) = E(z)W(z)G(z)$ in some disc $D(z_0, \rho)$ and *ii)* $E(z)$ and $G(z)$ are invertible and absolutely convergent on $D(z_0, \rho)$.

Theorem A.1. *Let $T : D(u, \rho) \to \mathcal{B}(H)$ be an absolutely convergent operator function. Assume that there exists $z_0 \in D(u, \rho)$ such that $T(z_0) \neq 0$ is Fredholm of index 0 and non-invertible. Then $T(z)$ is locally equivalent at z_0 to*

$$W(z) = W_0 + W_1(z - z_0)^{m_1} + \cdots + W_s(z - z_0)^{m_s}, \qquad W_h W_j = \delta_{hj} W_h, \qquad \sum_{h=0}^{s} W_h = I,$$

where $m_1 \leq m_2 \leq \cdots \leq m_s$ are positive integers, δ_{hj} is the Kronecker delta, $W_0, W_1, \ldots, W_s$ are mutually disjoint projections that decompose the identity, W_0 is Fredholm of index 0 and $W_1, \ldots, W_s$ have rank one. That is, there exists $\rho > 0$ such that

$$T(z) = E(z)W(z)G(z) \text{ for all } z \in D(z_0, \rho),$$

where $E(z)$ and $G(z)$ are invertible and absolutely convergent on $D(z_0, \rho)$. Hence $T(z)^{-1} = G(z)^{-1}W(z)^{-1}E(z)^{-1}$ has a pole of order $d = m_s$ at z_0 and it admits representation

$$T(z)^{-1} = \sum_{n=0}^{\infty} U_n (z - z_0)^{n-d}, \qquad z \in D(z_0, \rho) \setminus \{z_0\}, \tag{A.1}$$

where $U_0, \ldots, U_{d-1}$ are operators of finite rank and U_d is Fredholm of index 0.

Proof. See Theorem 8.1, Corollary 8.4 and eq.(2) in section XI.9 in Gohberg et al. (1990). ∎

Consistently with the terminology employed in the finite dimensional case, see Gohberg et al. (1993), the operator function $W(z)$, the positive integers $m_1 \leq m_2 \leq \cdots \leq m_s$ and the operator functions $E(z), G(z)$ are respectively called the *local Smith factorization*, the *partial multiplicities* and *extended canonical system of root functions* of $T(z)$ at z_0.

Expanding $T(z)$ around z_0 as $T(z) = \sum_{n=0}^{\infty} T_n(z - z_0)^n$ and considering (A.1), one writes the identity $T(z)T(z)^{-1} = I = T(z)^{-1}T(z)$ as the following linear systems in the T_n, U_n operators

$$\begin{aligned}
T_0U_0 &= 0 = U_0T_0 \\
T_0U_1 + T_1U_0 &= 0 = U_0T_1 + U_1T_0 \\
&\vdots \\
T_0U_{d-1} + \cdots + T_{d-1}U_0 &= 0 = U_0T_{d-1} + \cdots + U_{d-1}T_0 \\
T_0U_d + T_1U_{d-1} + \cdots + T_dU_0 &= I = U_0T_d + U_1T_{d-1} + \cdots + U_dT_0 \\
T_0U_{d+1} + T_1U_d + \cdots + T_{d+1}U_0 &= 0 = U_0T_{d+1} + U_1T_d + \cdots + U_{d+1}T_0 \\
&\vdots
\end{aligned} \tag{A.2}$$

In the following, equations in the system (A.2) are indexed according to the highest value of the subscript of U_n; for instance the identity appears in equation d, which is the order of the pole.

The equations that derive from $T(z)T(z)^{-1} = I$ are called left versions and those that derive from $I = T(z)^{-1}T(z)$ are called right versions; for instance $T_0U_d + T_1U_{d-1} + \cdots + T_dU_0$ is the left version of equation d.

Lemma A.2. *Let $T(z)$ be as in Theorem A.1 and define $\zeta_0 = \operatorname{Im} T_0$ and $\tau_0 = (\operatorname{Ker} T_0)^{\perp}$. Then U_0 in (A.1) satisfies $U_0 = P_{\tau_0^{\perp}}U_0 = U_0P_{\zeta_0^{\perp}} = P_{\tau_0^{\perp}}U_0P_{\zeta_0^{\perp}} \neq 0$.*

Proof. The left version of equation 0, $T_0U_0 = 0$, implies that $\operatorname{Im} U_0 \subseteq \operatorname{Ker} T_0$. Let $\tau_0 = (\operatorname{Ker} T_0)^{\perp}$, so that $\tau_0^{\perp} = \operatorname{Ker} T_0$, and define the associated orthogonal projections P_{τ_0} and $P_{\tau_0^{\perp}}$. Clearly P_{τ_0} and $P_{\tau_0^{\perp}}$ are disjoint and decompose the identity. Hence $U_0 = P_{\tau_0}U_0 + P_{\tau_0^{\perp}}U_0 = P_{\tau_0^{\perp}}U_0$. Similarly, from the right version of equation 0, $U_0T_0 = 0$, one has $\operatorname{Im} T_0 \subseteq \operatorname{Ker} U_0$. Defining $\zeta_0 = \operatorname{Im} T_0$, so that $\zeta_0^{\perp} = (\operatorname{Im} T_0)^{\perp}$, and the associated orthogonal projections P_{ζ_0} and $P_{\zeta_0^{\perp}}$, one has $U_0 = U_0P_{\zeta_0} + U_0P_{\zeta_0^{\perp}} = U_0P_{\zeta_0^{\perp}}$ and hence the statement. ■

Appendix B. Random variables in separable Hilbert spaces

The following definitions are taken from Bosq (2000). Let H be a separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$, norm $\| \cdot \|$ and Borel σ-algebra $\sigma(H)$ and let $(\Omega, \mathcal{A}, P)$ be a probability space. A function $Z : \Omega \to H$ is called an H-valued random variable on $(\Omega, \mathcal{A}, P)$ if it is measurable, i.e. for every subset $S \in \sigma(H)$, $\{\omega : Z(\omega) \in S\} \in \mathcal{A}$. For a $\mathbb{C}$-valued random variable U on $(\Omega, \mathcal{A}, P)$, define $E(U) = \int_{\Omega} U(\omega)dP(\omega)$; the expectation of an H-valued random variable Z, written μ_Z, is defined as the unique element of H such that

$$E(\langle h, Z \rangle) = \langle h, \mu_Z \rangle \text{ for all } h \in H.$$

It can be shown that the existence of μ_Z is guaranteed by the condition $E(\|Z\|) < \infty$. The covariance function of an H-valued random variable Z is defined as

$$c_Z(h, x) = E(\langle h, Z - \mu_Z \rangle \langle x, Z - \mu_Z \rangle), \qquad h, x \in H.$$

It is immediate to see that $c_Z(h, x) = E(\langle h, Z \rangle \langle x, Z \rangle) - \langle h, \mu_Z \rangle \langle x, \mu_Z \rangle = E(\langle h, \langle x, Z \rangle Z \rangle) - \langle h, \mu_Z \rangle \langle x, \mu_Z \rangle$. If $E(\|\langle x, Z \rangle Z\|) < \infty$, the expectation of the H-valued random variable $\langle x, Z \rangle Z$ exists and it is the unique element of H such that $E(\langle h, \langle x, Z \rangle Z \rangle) = \langle h, \mu_{\langle x, Z \rangle Z} \rangle$ for all $h \in H$. One thus has

$$c_Z(h, x) = \langle h, \mu_{\langle x, Z \rangle Z} \rangle - \langle h, \mu_Z \rangle \langle x, \mu_Z \rangle, \qquad h, x \in H.$$

Because $\|\langle x, Z \rangle Z\| = |\langle x, Z \rangle| \|Z\| \leq \|x\| \|Z\|^2$, the existence of the covariance function of Z is guaranteed by the condition $E(\|Z\|^2) < \infty$. Define the operator $C_Z : H \to H$ that maps x into $\mu_{\langle x, Z \rangle Z}$ and rewrite the covariance function as

$$c_Z(h, x) = \langle h, C_Z(x) \rangle - \langle h, \mu_Z \rangle \langle x, \mu_Z \rangle, \qquad h, x \in H.$$

As C_z is completely determined by the covariance function, it is called the covariance operator of Z. Similarly, the cross-covariance function of two H-valued random variables Z and U is defined as

$$c_{Z,U}(h,x) = E(\langle h, Z-\mu_Z\rangle\langle x, U-\mu_U\rangle), \qquad h,x \in H.$$

This completely determines the cross-covariance operators of Z and U, $C_{Z,U}$ and $C_{U,Z}$, respectively defined as the mappings $x \mapsto \mu_{\langle x,Z\rangle U}$ and $x \mapsto \mu_{\langle x,U\rangle Z}$.

Appendix C. Proofs

Proof of Theorem 3.1. The result is a direct consequence of Theorem A.1 in Appendix A. By definition, a cointegrated H-valued AR process $A(L)x_t = \varepsilon_t$ is such that $A(1) \neq 0$, $A(z)$ has an eigenvalue of finite type at $z=1$ and $A(z)$ is invertible in some punctured disc $D(0,\rho)\setminus\{1\}$, $\rho > 1$. Hence $A(1)$ is Fredholm of index 0. Because $A : \mathbb{C} \to \mathcal{B}(H)$, one can apply Theorem A.1. This states that there exists $\rho > 0$ such that

$$A(z) = E(z)W(z)G(z) \text{ for all } z \in D(1,\rho), \tag{C.1}$$

where $E(z)$ and $G(z)$ are invertible and absolutely convergent on $D(1,\rho)$ and

$$W(z) = W_0 + W_1(1-z)^{m_1} + \cdots + W_s(1-z)^{m_s}, \qquad W_hW_j = \delta_{hj}W_h, \qquad \sum_{h=0}^{s} W_h = I,$$

where the positive integers $m_1 \leq m_2 \leq \cdots \leq m_s$ are the partial multiplicities of $A(z)$ at 1, $W_0, W_1, \ldots, W_s$ are mutually disjoint projections that decompose the identity, W_0 is Fredholm of index 0 and $W_1, \ldots, W_s$ have rank one.

Given that $G_0 = G(1)$ is invertible, one can normalize it to be equal to I, in fact $A(z) = E_0(z)W_0(z)G_0(z)$, where $E_0(z) = E(z)G_0$, $W_0(z) = G_0^{-1}W(z)G_0$, and $G_0(z) = G_0^{-1}G(z)$ share the same properties of $E(z)$, $W(z)$ and $G(z)$ in (C.1). In the following one can set $G(1) = I$. Moreover, $A(z)$ is invertible in the punctured disc $D(0,\rho)\setminus\{1\}$, $\rho > 1$, and because $A(z)^{-1} = G(z)^{-1}W(z)^{-1}E(z)^{-1}$, this implies that $E(z)$ and $G(z)$ are invertible and absolutely convergent for all $z \in D(0,\rho)$, $\rho > 1$.

Let w be the number of distinct partial multiplicities and organize them as in

$$\underbrace{m_1 = \cdots = m_{q_1}}_{=j_1} < \underbrace{m_{q_1+1} = \cdots = m_{q_1+q_2}}_{=j_2} < \cdots < \underbrace{m_{\sum_{i=1}^{w-1} q_i+1} = \cdots = m_s}_{=j_w},$$

where q_h is the number of partial multiplicities that are equal to the given value j_h. This leads to define the projections P_h as the sum of the projections W_n that load the same partial multiplicity into $W(z)$, i.e.

$$P_1 = W_1 + \cdots + W_{q_1}, \quad P_2 = W_{q_1+1} + \cdots + W_{q_1+q_2}, \quad \ldots \quad, \quad P_w = W_{\sum_{i=1}^{w-1} q_i+1} + \cdots + W_s,$$

so that $W(z) = \sum_{h=0}^{w} P_h(1-z)^{j_h}$, where $j_0 = 0$, $P_0 = W_0$ and $P_h = \sum_{n=1}^{j_h} W_{n+\sum_{i=1}^{h-1} q_i}$, $h = 1, \ldots, w$. Observe that $P_hP_j = \delta_{hj}P_h$ and $\sum_{h=0}^{w} P_h = I$, where P_0 has infinite rank $q_0 = \dim \operatorname{Im} P_0 = \infty$ and $P_1, \ldots, P_w$ have finite rank q_h. This leads to the direct sum decomposition

$$H = \operatorname{Im} P_0 \oplus \operatorname{Im} P_1 \oplus \cdots \oplus \operatorname{Im} P_w, \tag{C.2}$$

where $\operatorname{Im} P_0$ is closed because P_0 is a projection. Substituting (C.1) in $A(L)x_t = \varepsilon_t$ and rearraging one has

$$W(L)y_t = u_t, \qquad y_t = G(L)x_t, \qquad u_t = E(L)^{-1}\varepsilon_t,$$

and using $W(z) = \sum_{h=0}^{w} P_h(1-z)^{j_h}$ and $\sum_{h=0}^{w} P_h = I$ one finds

$$P_0 y_t + \Delta^{j_1} P_1 y_t + \dots + \Delta^{j_w} P_w y_t = P_0 u_t + P_1 u_t + \dots + P_w u_t,$$

where $P_h y_t$ and $P_h u_t$ belong to $\operatorname{Im} P_h$. As a consequence of the direct sum decomposition in (C.2), one has $\Delta^{j_h} P_h y_t = P_h u_t$ and hence $P_h y_t = P_h u_{j_h,t} + P_h v_{h,0}$, where $u_{j_h,t} \sim I(j_h)$ is the j_h-th cumulation of $u_t \sim I(0)$ and $v_{h,0}$ collects initial values. As $y_t = \sum_{h=0}^{w} P_h y_t = \sum_{h=0}^{w} P_h u_{j_h,t} + \sum_{h=0}^{w} P_h v_{h,0}$, one has

$$y_t = P_0 u_{j_0,t} + P_1 u_{j_1,t} + \dots + P_w u_{j_w,t} + v_0, \qquad u_{h,t} = \sum_{i=1}^{t} u_{h-1,i} \sim I(h), \qquad u_{0,t} = u_t \sim I(0),$$

where $v_0 = \sum_{h=0}^{w} P_h v_{h,0}$ collects initial values. This shows that $y_t = G(L)x_t \sim I(j_w)$, $q_h = \dim \operatorname{Im} P_h$ is the number of $I(j_h)$ processes in y_t and $\langle v, y_t \rangle \sim I(j_h)$ if and only if $v \in \operatorname{Im} P_h$. Because $G(1) = I$, $x_t = G(L)^{-1} y_t = y_t + F_1 \Delta y_t + \dots$, where $G(z)^{-1} = F(z) = \sum_{n=0}^{\infty} F_n(1-z)^n$, one has

$$x_t = (P_0 u_{j_0,t} + P_1 u_{j_1,t} + \dots + P_w u_{j_w,t} + v_0) + F_1 \Delta \left(P_0 u_{j_0,t} + P_1 u_{j_1,t} + \dots + P_w u_{j_w,t} + v_0\right) + \dots,$$

which shows that x_t is $I(j_w)$, $\langle v, x_t \rangle \sim I(j_w)$ if and only if $v \in \operatorname{Im} P_w$ and hence $\operatorname{Im} P_0 \oplus \operatorname{Im} P_1 \oplus \dots \oplus \operatorname{Im} P_{w-1}$ is the cointegrating space of x_t. Finally, let $n_h = j_w - j_h$ and expand $G(z)$ around 1 as $G(z) = \sum_{n=0}^{n_h-1} G_n(1-z)^n + (1-z)^{n_h} G^{\star}_{n_h}(z)$, where $G_0 = I$ and $G^{\star}_{n_h}(z)$ is absolutely convergent on $D(0,\rho)$, $\rho > 1$. Then

$$x_t + G_1 \Delta x_t + \dots + G_{n_h-1} \Delta^{n_h-1} x_t = P_0 u_{j_0,t} + P_1 u_{j_1,t} + \dots + P_w u_{j_w,t} + x_{j_h,t} + v_0,$$

where $x_{j_h,t} = -G^{\star}_{n_h}(L)\Delta^{n_h} x_t = -G_{n_h}\Delta^{n_h} x_t + \dots = -G_{n_h}\Delta^{n_h} P_w u_{j_w,t} + \dots = -G_{n_h} P_w u_{j_h,t} + \dots$ is at most integrated of order $j_w - n_h = j_h$, so that

$$P_h\left(x_t + G_1 \Delta x_t + \dots + G_{n_h-1}\Delta^{n_h-1} x_t\right) = P_h\left(u_{j_h,t} + x_{j_h,t}\right) + P_h v_0.$$

Because $u_{j_h,t} + x_{j_h,t} = (I - G_{n_h} P_w) u_{j_h,t} + \dots$, one has that $\langle v, (u_{j_h,t} + x_{j_h,t}) x \rangle \neq 0$ for any $v \in \operatorname{Im} P_h$ and any $x \in \operatorname{Im} P_0 \oplus \dots \oplus \operatorname{Im} P_{w-1}$. This shows that $\langle v, x_t \rangle + \sum_{n=1}^{n_h-1} \langle v, G_n \Delta^n x_t \rangle$ is $I(j_h)$ for all $v \in \operatorname{Im} P_h$ and completes the proof. ■

Proof of Theorem 4.1. Replacing T with A and U with C in system (A.2), for $d = 1$ one has

$$\begin{gathered} A_0 C_0 = 0 = C_0 A_0 \\ A_0 C_1 + A_1 C_0 = I = C_0 A_1 + C_1 A_0 \\ A_0 C_2 + A_1 C_1 + A_2 C_0 = 0 = C_0 A_2 + C_1 A_1 + C_2 A_0 \\ \vdots \end{gathered}$$

Because $P_{\zeta_0^\perp} A_0 = 0$ and $C_0 = P_{\tau_0^\perp} C_0$, see Lemma A.2, the left version of equation 1 implies that $(P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}) C_0 = P_{\zeta_0^\perp}$. This shows that if $x \in \operatorname{Im} P_{\zeta_0^\perp}$ then $x \in \operatorname{Im} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}$, i.e. $\operatorname{Im} P_{\zeta_0^\perp} \subseteq \operatorname{Im} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}$; because the reverse inclusion is clearly satisfied, one has that $\operatorname{Im} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} = \operatorname{Im} P_{\zeta_0^\perp}$, i.e. $\zeta_1 = \zeta_0^\perp$. Similarly, because $A_0 P_{\tau_0^\perp} = 0$ and $C_0 = C_0 P_{\zeta_0^\perp}$, see Lemma A.2, the right version of equation 1 implies that $C_0(P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}) = P_{\tau_0^\perp}$. Hence if $x \in \operatorname{Ker} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}$ then $x \in \operatorname{Ker} P_{\tau_0^\perp}$, i.e. $\operatorname{Ker} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} \subseteq \operatorname{Ker} P_{\tau_0^\perp}$; because the reverse inclusion is clearly satisfied, one has that $\operatorname{Ker} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} = \operatorname{Ker} P_{\tau_0^\perp}$, so that $(\operatorname{Ker} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp})^\perp = (\operatorname{Ker} P_{\tau_0^\perp})^\perp$, i.e. $\tau_1 = \tau_0^\perp$. This shows that if $d = 1$ then $\zeta_1 = \zeta_0^\perp$, $\tau_1 = \tau_0^\perp$, $C_0 = P_{\tau_1} C_0 P_{\zeta_1}$, $\operatorname{Im} C_0 = \tau_1$ and $\operatorname{Ker} C_0 = \zeta_0$. Finally note that if $d > 1$, because the identity is in equation d, the same analysis leads to $(P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}) C_0 = 0$ and $C_0(P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}) = 0$. In this case there exist a nonzero $x \in \operatorname{Im} C_0$ such that $x \in \operatorname{Ker} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}$ and a nonzero $y \in \operatorname{Im} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}$ such that $y \in \operatorname{Ker} C_0$. i.e. $\tau_1 \subset \tau_0^\perp$ and $\zeta_1 \subset \zeta_0^\perp$. This shows that $\tau_1 = \tau_0^\perp$ (or any of $\zeta_1 = \zeta_0^\perp$, $\operatorname{Im} C_0 = \tau_1$ and $\operatorname{Ker} C_0 = \zeta_0$) is a necessary and sufficient condition for $d = 1$.

Because $\Delta x_t = C(L)\varepsilon_t \sim I(0)$ and $\operatorname{Im} C_0 = \tau_1$, τ_1 is the attractor space of the $I(1)$ trends, i.e. $\langle v, x_t \rangle \sim I(1)$ for all $v \in \tau_1$, and $\tau_1^\perp = \tau_0$ is the cointegrating space of x_t. In order to prove that $\langle v, x_t \rangle \sim I(0)$ for all $v \in \tau_0$, write $P_{\tau_0} \Delta x_t = P_{\tau_0} C_0 \varepsilon_t + P_{\tau_0} C_1 \Delta \varepsilon_t + C^\star(L)\Delta^2 \varepsilon_t$, where here and in the following $C^\star(z), C^{\star\star}(z), C^{\star\star\star}(z)$ represent remaining terms. Because $P_{\tau_0} C_0 = 0$, $P_{\tau_0} x_t = P_{\tau_0} C_1 \varepsilon_t + C^\star(L)\Delta\varepsilon_t$ is a linear process. Next consider the generalized inverse of A_0, $Q_0 = A_0^+$, and note that $Q_0 A_0 = P_{\tau_0}$, see Lemma C.1 below. From the left version of equation 1 one has $Q_0 A_0 C_1 + Q_0 A_1 C_0 = Q_0$ and thus $P_{\tau_0} C_1 = Q_0(I - A_1 C_0)$; this implies that $\langle v, P_{\tau_0} C_1 x \rangle \neq 0$ for any $v \in \tau_0$ and any $x \in \zeta_0$, i.e. $\langle v, x_t \rangle \sim I(0)$ for all $v \in \tau_0$. ■

Proof of Theorem 4.4. Replacing T with A and U with C in system (A.2), for $d = 2$ one has

$$
\begin{gathered}
A_0 C_0 = 0 = C_0 A_0 \\
A_0 C_1 + A_1 C_0 = 0 = C_0 A_1 + C_1 A_0 \\
A_0 C_2 + A_1 C_1 + A_2 C_0 = I = C_0 A_2 + C_1 A_1 + C_2 A_0 \\
A_0 C_3 + A_1 C_2 + A_2 C_1 + A_3 C_1 = 0 = C_0 A_3 + C_1 A_2 + C_2 A_1 + C_3 A_0 \\
\vdots
\end{gathered}
$$

From the proof of Theorem 4.1 one has that if $d > 1$ the left versions of equations 0 and 1 lead to $A_0 C_0 = 0$ and $(P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}) C_0 = 0$ and the right versions to $C_0 A_0 = 0$ and $C_0(P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}) = 0$. Hence $\operatorname{Im} C_0 \subseteq (\operatorname{Ker} A_0 \oplus \operatorname{Ker} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}) = (\tau_0 \oplus \tau_1)^\perp = \mathscr{T}_2^\perp$ and $\operatorname{Ker} C_0 \supseteq (\operatorname{Im} A_0 \oplus \operatorname{Im} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}) = (\zeta_0 \oplus \zeta_1) = \mathscr{Z}_2$. Hence $C_0 = P_{\mathscr{T}_2^\perp} C_0 = C_0 P_{\mathscr{Z}_2^\perp} = P_{\mathscr{T}_2^\perp} C_0 P_{\mathscr{Z}_2^\perp}$. Because $P_{\zeta_0^\perp} A_0 = 0$, the left version of equation 2 implies that $P_{\zeta_0^\perp} A_1 C_1 + P_{\zeta_0^\perp} A_2 C_0 = P_{\zeta_0^\perp}$. Inserting $I = P_{\tau_0} + P_{\tau_0^\perp}$ in the first term one has $P_{\zeta_0^\perp} A_1 C_1 = P_{\zeta_0^\perp} A_1 P_{\tau_0} C_1 + P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} C_1$ and hence $P_{\mathscr{Z}_2^\perp} A_1 C_1 = P_{\mathscr{Z}_2^\perp} A_1 P_{\tau_0} C_1$, because $\operatorname{Im} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} = \zeta_1$. Because $Q_0 A_0 = P_{\tau_0}$, see Lemma C.1 below, the left version of equation 1 implies $P_{\tau_0} C_1 = -Q_0 A_1 C_0$; hence one has $P_{\zeta_0^\perp} A_1 C_1 = -P_{\zeta_0^\perp} A_1 Q_0 A_1 C_0 + P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} C_1$ and thus $P_{\zeta_0^\perp} A_1 C_1 + P_{\zeta_0^\perp} A_2 C_0 = P_{\zeta_0^\perp}$ leads to

$$P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} C_1 + P_{\zeta_0^\perp} A_{2,1} C_0 = P_{\zeta_0^\perp}, \qquad A_{2,1} = A_2 - A_1 Q_0 A_1. \tag{C.3}$$

Because $\mathscr{X}_2 = \zeta_0 \oplus \zeta_1$ and $\operatorname{Im} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} = \zeta_1$, one has $P_{\mathscr{X}_2^\perp} P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} = 0$. Hence substituting $P_{\mathscr{X}_2^\perp} P_{\zeta_0^\perp} = P_{\mathscr{X}_2^\perp}$ and $C_0 = P_{\mathscr{T}_2^\perp} C_0$ in (C.3), one finds $(P_{\mathscr{X}_2^\perp} A_{2,1} P_{\mathscr{T}_2^\perp}) C_0 = P_{\mathscr{X}_2^\perp}$. This shows that $\operatorname{Im} P_{\mathscr{X}_2^\perp} \subseteq \operatorname{Im} P_{\mathscr{X}_2^\perp} A_{2,1} P_{\mathscr{T}_2^\perp}$; because the reverse inclusion is clearly satisfied, one has that $\operatorname{Im} P_{\mathscr{X}_2^\perp} A_{2,1} P_{\mathscr{T}_2^\perp} = \operatorname{Im} P_{\mathscr{X}_2^\perp}$, i.e. $\zeta_2 = \mathscr{X}_2^\perp$. Similarly, the right version of equation 2 implies that $C_0 (P_{\mathscr{X}_2^\perp} A_{2,1} P_{\mathscr{T}_2^\perp}) = P_{\mathscr{T}_2^\perp}$, which shows that $\operatorname{Ker} P_{\mathscr{X}_2^\perp} A_{2,1} P_{\mathscr{T}_2^\perp} \subseteq \operatorname{Ker} P_{\mathscr{T}_2^\perp}$; because the reverse inclusion is clearly satisfied, one has that $\operatorname{Ker} P_{\mathscr{X}_2^\perp} A_{2,1} P_{\mathscr{T}_2^\perp} = \operatorname{Ker} P_{\mathscr{T}_2^\perp}$, so that $(\operatorname{Ker} P_{\mathscr{X}_2^\perp} A_{2,1} P_{\mathscr{T}_2^\perp})^\perp = (\operatorname{Ker} P_{\mathscr{T}_2^\perp})^\perp$, i.e. $\tau_2 = \mathscr{T}_2^\perp$. This shows that if $d = 2$ then $\zeta_2 = \mathscr{X}_2^\perp$, $\tau_2 = \mathscr{T}_2^\perp$, $C_0 = P_{\tau_2} C_0 P_{\zeta_2}$, $\operatorname{Im} C_0 = \tau_2$ and $\operatorname{Ker} C_0 = \mathscr{X}_2$. Finally note that if $d > 2$, the same analysis leads to $(P_{\mathscr{X}_2^\perp} A_{2,1} P_{\mathscr{T}_2^\perp}) C_0 = 0$ and $C_0 (P_{\mathscr{X}_2^\perp} A_{2,1} P_{\mathscr{T}_2^\perp}) = 0$, i.e. to $\tau_2 \subset \mathscr{T}_2^\perp$ and $\zeta_2 \subset \mathscr{X}_2^\perp$. This shows that $\tau_2 = \mathscr{T}_2^\perp$ (or any of $\zeta_2 = \mathscr{X}_2^\perp$, $\operatorname{Im} C_0 = \tau_2$ and $\operatorname{Ker} C_0 = \mathscr{X}_2$) is a necessary and sufficient condition for $d = 2$.

Because $\Delta^2 x_t = C(L)\varepsilon_t \sim I(0)$ and $\operatorname{Im} C_0 = \tau_2$, τ_2 is the attractor space of the $I(2)$ trends, i.e. $\langle v, x_t \rangle \sim I(2)$ for all $v \in \tau_2$, and $\tau_2^\perp = \tau_0 \oplus \tau_1$ is the cointegrating space of x_t. In order to prove that $\langle v, x_t \rangle \sim I(1)$ for all $v \in \tau_1$, write $P_{\tau_1} \Delta^2 x_t = P_{\tau_1} C_0 \varepsilon_t + P_{\tau_1} C_1 \Delta \varepsilon_t + C^\star(L) \Delta^2 \varepsilon_t$. Because $P_{\tau_1} C_0 = 0$, $P_{\tau_1} \Delta x_t = P_{\tau_1} C_1 \varepsilon_t + C^\star(L) \Delta \varepsilon_t$ is a linear process. Next consider the generalized inverse of $P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}$, $Q_1 = (P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp})^+$, and note that $Q_1 (P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}) = P_{\tau_1}$ and $Q_1 P_{\zeta_0^\perp} = Q_1$, see Lemma *C.1* below. From (C.3) one thus has $P_{\tau_1} C_1 = Q_1 (I - A_{2,1} C_0)$; this implies that $\langle v, P_{\tau_1} C_1 x \rangle \neq 0$ for any $v \in \tau_1$ and any $x \in \zeta_1$, i.e. $\langle v, x_t \rangle \sim I(1)$ for all $v \in \tau_1$.

Finally, from $P_{\tau_0} \Delta^2 x_t = P_{\tau_0} C_0 \varepsilon_t + P_{\tau_0} C_1 \Delta \varepsilon_t + P_{\tau_0} C_2 \Delta^2 \varepsilon_t + C^\star(L) \Delta^3 \varepsilon_t$, $P_{\tau_0} C_0 = 0$ and $P_{\tau_0} C_1 = -Q_0 A_1 C_0$, one has $P_{\tau_0} \Delta^2 x_t = -Q_0 A_1 C_0 \Delta \varepsilon_t + P_{\tau_0} C_2 \Delta^2 \varepsilon_t + C^\star(L) \Delta^3 \varepsilon_t$. On the other hand, $Q_0 A_1 \Delta x_t = Q_0 A_1 C_0 \Delta \varepsilon_t + Q_0 A_1 C_1 \Delta^2 \varepsilon_t + C^{\star\star}(L) \Delta^3 \varepsilon_t$, and hence

$$P_{\tau_0} x_t + Q_0 A_1 \Delta x_t = (P_{\tau_0} C_2 + Q_0 A_1 C_1) \varepsilon_t + C^{\star\star\star}(L) \Delta \varepsilon_t$$

is a linear process. In order to see that $\langle v, x_t \rangle + \langle v, Q_0 A_1 \Delta x_t \rangle \sim I(0)$ for all $v \in \tau_0$, observe that the left version of equation 2 implies $Q_0 A_0 C_2 + Q_0 A_1 C_1 + Q_0 A_2 C_0 = Q_0$ and hence $P_{\tau_0} C_2 + Q_0 A_1 C_1 = Q_0 (I - A_2 C_0)$. Because $\langle v, (P_{\tau_0} C_2 + Q_0 A_1 C_1) x \rangle \neq 0$ for any $v \in \tau_0$ and any $x \in \zeta_0 \oplus \zeta_1$, one has $\langle v, x_t \rangle + \langle v, Q_0 A_1 \Delta x_t \rangle \sim I(0)$ for all $v \in \tau_0$. ∎

The proof of Theorem 5.4 is based on Lemma *C.2* below, which makes use of the following result.

Lemma C.1. *Let $\mathscr{X}_h$, $\mathscr{T}_h$, S_h, Q_h, and $A_{h,n}$ be as in Definition 5.1 and futher define $P_{\mathscr{X}_0^\perp} = P_{\mathscr{T}_0^\perp} = I$, $A_{0,1} = A_0$ and $S_0 = P_{\mathscr{X}_0^\perp} A_{0,1} P_{\mathscr{T}_0^\perp}$. Then $Q_h S_h = P_{\tau_h}$ and $Q_h P_{\mathscr{X}_h^\perp} = Q_h$ for $h = 0, 1, \ldots$.*

Proof. What follows holds for any $h = 0, 1, \ldots$. Recall that ζ_h and τ_h are closed and the corresponding generalized inverse Q_h exists and it is unique, see Remark 5.2. From Theorem 3 in Chapter 9 in Ben-Israel and Greville (2003), one has that if $A \in \mathcal{B}(H)$ and $\operatorname{Im} A$ is closed, then

$A^+A = P_{\mathrm{Im}\, A^*}$ and $\mathrm{Ker}\, A^+ = \mathrm{Ker}\, A^*$, so that $Q_h S_h = P_{\mathrm{Im}\, S_h^*}$ and $\mathrm{Ker}\, Q_h = \mathrm{Ker}\, S_h^*$. Because τ_h is closed, one has that $\tau_h = (\mathrm{Ker}\, S_h)^\perp = \mathrm{Im}\, S_h^*$ and hence $Q_h S_h = P_{\tau_h}$. Because ζ_h is closed, one has that $\zeta_h = \mathrm{Im}\, S_h = (\mathrm{Ker}\, S_h^*)^\perp$. Hence $\mathrm{Ker}\, Q_h = \mathrm{Ker}\, S_h^* = \zeta_h^\perp \supseteq \mathscr{Z}_h = \zeta_0 \oplus \cdots \oplus \zeta_{h-1}$, so that $(\mathrm{Ker}\, Q_h)^\perp \subseteq \mathscr{Z}_h^\perp$. This shows that $Q_h P_{\mathscr{Z}_h^\perp} = Q_h$ and completes the proof. ∎

Lemma C.2 (Subspace decomposition of system (A.2)). *Let ζ_h, τ_h, Q_h, and $A_{h,n}$ be as in Definition 5.1 and replace T with A and U with C in system (A.2). Then equation $n + h \leq d$ in system (A.2) can be written as*

$$P_{\tau_h} C_n + Q_h \sum_{k=1}^{n} A_{h+1,k} C_{n-k} = \delta_{n+h,d} Q_h, \qquad h = 0, 1, \ldots, d-n, \tag{C.4}$$

where δ_{hj} is the Kronecker delta, $P_{\mathscr{Z}_0^\perp} = P_{\mathscr{T}_0^\perp} = I$ and $A_{0,1} = A_0$. Moreover, $A(z)^{-1}$ has a pole of order d at $z = 1$ if and only if either of the following equivalent statements holds: i) $H = \zeta_0 \oplus \zeta_1 \oplus \cdots \oplus \zeta_d$, where $\zeta_0, \zeta_1, \ldots, \zeta_d$ are closed, $\dim \zeta_0 = \infty$, $0 \leq \dim \zeta_h < \infty$ for any $h \neq 0, d$, $0 < \dim \zeta_d < \infty$; ii) $H = \tau_0 \oplus \tau_1 \oplus \cdots \oplus \tau_d$, where $\tau_0, \tau_1, \ldots, \tau_d$ are closed, $\dim \tau_0 = \infty$, $0 \leq \dim \tau_h < \infty$ for any $h \neq 0, d$, $0 < \dim \tau_d < \infty$. Finally,

$$\langle v, \gamma_h(z) A(z)^{-1} \rangle, \qquad \gamma_h(z) = P_{\tau_h} + Q_h \sum_{n=1}^{d-h-1} A_{h+1,n} (1-z)^n, \tag{C.5}$$

has a pole of order h for all $v \in \tau_h$, $h = 0, 1, \ldots, d$.

Proof. The statement is divided into three parts: the first consists of (C.4), the second in i) and ii) being necessary and sufficient conditions for a pole of order d at $z = 1$ in $A(z)^{-1}$ and the third of (C.5). The proof is thus split into three parts. Replace T with A and U with C in system (A.2). The proof of the first part is by induction and consists in showing that the left version of equation $n \leq d$ in system (A.2) can be written as

$$\left(P_{\mathscr{Z}_h^\perp} A_{h,1} P_{\mathscr{T}_h^\perp}\right) C_{n-h} + P_{\mathscr{Z}_h^\perp} \sum_{k=1}^{n-h} A_{h+1,k} C_{n-h-k} = \delta_{n,d} P_{\mathscr{Z}_h^\perp}, \qquad h = 0, 1, \ldots, n, \tag{C.6}$$

where $P_{\mathscr{Z}_0^\perp} = P_{\mathscr{T}_0^\perp} = I$ and $A_{0,1} = A_0$. Replacing n with $n + h$ in (C.6), one finds

$$\left(P_{\mathscr{Z}_h^\perp} A_{h,1} P_{\mathscr{T}_h^\perp}\right) C_n + P_{\mathscr{Z}_h^\perp} \sum_{k=1}^{n} A_{h+1,k} C_{n-k} = \delta_{n+h,d} P_{\mathscr{Z}_h^\perp}, \qquad h = 0, 1, \ldots, d-n, \tag{C.7}$$

and hence (C.4), which follows from $Q_h(P_{\mathscr{Z}_h^\perp} A_{h,1} P_{\mathscr{T}_h^\perp}) = P_{\tau_h}$ and $Q_h P_{\mathscr{Z}_h^\perp} = Q_h$, see Lemma *C.1*.

In order to show that (C.6) holds for $h = 0$, observe that the left version of equation n in system (A.2) reads $A_0 C_n + \sum_{k=1}^n A_k C_{n-k} = \delta_{n,d} I$. By definition, $P_{\mathscr{Z}_0^\perp} = P_{\mathscr{T}_0^\perp} = I$, $A_{0,1} = A_0$ and $A_{1,k} = A_k$ and this shows that (C.6) holds for $h = 0$. Next assume that (C.6) holds for $h = 0, \ldots, \ell - 1$ for some $1 < \ell \leq d$; one wishes to show that it also holds for $h = \ell$. First note that $Q_h(P_{\mathscr{Z}_h^\perp} A_{h,1} P_{\mathscr{T}_h^\perp}) = P_{\tau_h}$ and $Q_h P_{\mathscr{Z}_h^\perp} = Q_h$, see Lemma *C.1*, and thus the induction assumption implies

$$P_{\tau_h} C_{n-h} + Q_h \sum_{k=1}^{n-h} A_{h+1,k} C_{n-h-k} = \delta_{n,d} Q_h, \qquad h = 0, 1, \ldots, \ell - 1. \tag{C.8}$$

Next write (C.6) for $h = \ell - 1$,

$$\left(P_{\mathscr{Z}_{\ell-1}^{\perp}} A_{\ell-1,1} P_{\mathscr{T}_{\ell-1}^{\perp}}\right) C_{n-\ell+1} + P_{\mathscr{Z}_{\ell-1}^{\perp}} \sum_{k=1}^{n-\ell+1} A_{\ell,k} C_{n-\ell+1-k} = \delta_{n,d} P_{\mathscr{Z}_{\ell-1}^{\perp}},$$

where $\operatorname{Im} P_{\mathscr{Z}_{\ell-1}^{\perp}} A_{\ell-1,1} P_{\mathscr{T}_{\ell-1}^{\perp}} = \zeta_{\ell-1}$; pre-multiplying by $P_{\mathscr{Z}_{\ell}^{\perp}}$, where $\mathscr{Z}_\ell = \zeta_0 \oplus \cdots \oplus \zeta_{\ell-1}$, and rearranging one finds

$$P_{\mathscr{Z}_{\ell}^{\perp}} A_{\ell,1} C_{n-\ell} + P_{\mathscr{Z}_{\ell}^{\perp}} \sum_{k=1}^{n-\ell} A_{\ell,k+1} C_{n-\ell-k} = \delta_{n,d} P_{\mathscr{Z}_{\ell}^{\perp}} \tag{C.9}$$

Next consider $\mathscr{T}_\ell = \tau_0 \oplus \cdots \oplus \tau_{\ell-1}$ and use projections, inserting $I = P_{\mathscr{T}_\ell} + P_{\mathscr{T}_\ell^{\perp}}$ between $A_{\ell,1}$ and $C_{n-\ell}$ in $U = P_{\mathscr{Z}_{\ell}^{\perp}} A_{\ell,1} C_{n-\ell}$; one finds

$$U = \left(P_{\mathscr{Z}_{\ell}^{\perp}} A_{\ell,1} P_{\mathscr{T}_{\ell}^{\perp}}\right) C_{n-\ell} + P_{\mathscr{Z}_{\ell}^{\perp}} A_{\ell,1} P_{\mathscr{T}_\ell} C_{n-\ell} = U_1 + U_2.$$

Substituting $P_{\mathscr{T}_\ell} = P_{\tau_0} + \cdots + P_{\tau_{\ell-1}}$, one has $U_2 = P_{\mathscr{Z}_{\ell}^{\perp}} A_{\ell,1} \sum_{i=0}^{\ell-1} P_{\tau_i} C_{n-\ell}$ and by the induction assumption, replacing n with $n - \ell + h$ and h with i in (C.8) and rearraging, one finds

$$P_{\tau_i} C_{n-\ell} = -Q_i \sum_{k=1}^{n-\ell} A_{i+1,k} C_{n-\ell-k} + \delta_{n-\ell+i,d} Q_i, \qquad i = 0, 1, \ldots, \ell - 1.$$

Because $n - \ell + i \leq n - 1 < d$, $\delta_{n-\ell+i,d} = 0$ for $i = 0, 1, \ldots, \ell - 1$ and hence substituting in U_2, one finds

$$U_2 = -P_{\mathscr{Z}_{\ell}^{\perp}} \sum_{k=1}^{n-\ell} \left(A_{\ell,1} \sum_{i=0}^{\ell-1} Q_i A_{i+1,k}\right) C_{n-\ell-k}.$$

Substituting $U = U_1 + U_2$ one hence rewrites (C.9) as

$$\left(P_{\mathscr{Z}_{\ell}^{\perp}} A_{\ell,1} P_{\mathscr{T}_{\ell}^{\perp}}\right) C_{n-\ell} + P_{\mathscr{Z}_{\ell}^{\perp}} \sum_{k=1}^{n-\ell} A_{\ell+1,k} C_{n-\ell-k} = \delta_{n,d} P_{\mathscr{Z}_{\ell}^{\perp}}$$

where $A_{\ell+1,k} = A_{\ell,k+1} - A_{\ell,1} \sum_{i=0}^{\ell-1} Q_i A_{i+1,k}$ by definition. This shows that (C.6) holds for $h = \ell$ and completes the proof of the first part of the statement.

The proof of the second part consists in showing that *i*) and *ii*) are necessary and sufficient conditions for a pole of order d at $z = 1$ in $A(z)^{-1}$. First consider *i*). Assume that $A(z)^{-1}$ has a pole of order d at $z = 1$. Setting $n = 0$ and $h = d$ in (C.7) one has $\left(P_{\mathscr{Z}_{d}^{\perp}} A_{d,1} P_{\mathscr{T}_{d}^{\perp}}\right) C_0 = P_{\mathscr{Z}_{d}^{\perp}}$, so that $\zeta_d = \operatorname{Im} P_{\mathscr{Z}_{d}^{\perp}} A_{d,1} P_{\mathscr{T}_{d}^{\perp}} \supseteq \operatorname{Im} P_{\mathscr{Z}_{d}^{\perp}} = \mathscr{Z}_d^{\perp} = (\zeta_0 \oplus \cdots \oplus \zeta_{d-1})^{\perp}$; because the reverse inclusion is clearly satisfied, one has that $\operatorname{Im} P_{\mathscr{Z}_{d}^{\perp}} A_{d,1} P_{\mathscr{T}_{d}^{\perp}} = \operatorname{Im} P_{\mathscr{Z}_{d}^{\perp}}$, i.e. $\zeta_d = (\zeta_0 \oplus \cdots \oplus \zeta_{d-1})^{\perp}$, and hence $H = \zeta_0 \oplus \zeta_1 \oplus \cdots \oplus \zeta_d$. Expand $A(z)$ around 1 as $A(z) = \sum_{n=0}^{\infty} A_n (1 - z)^n$, where $A_0 = A(1) \neq 0$ implies $\dim \zeta_0 > 0$. Because 1 is an eigenvalue of finite type, A_0 if Fredholm of index 0, which means that $\dim \operatorname{Ker} A_0$ and $\operatorname{codim} \operatorname{Im} A_0$ are finite and equal. Hence $\operatorname{codim} \operatorname{Im} A_0 = \dim(\zeta_1 \oplus \cdots \oplus \zeta_d)$ is finite dimensional (and closed) and hence it is complemented. That is, its orthogonal complement ζ_0 is closed (and infinite dimensional). This shows that $\zeta_0, \zeta_1, \ldots, \zeta_d$ are closed, $\dim \zeta_0 = \infty$ and $0 \leq \dim \zeta_h < \infty$ for any $h \neq 0$. Because $\dim \zeta_d = 0$ implies $H = \zeta_0 \oplus \zeta_1 \oplus \cdots \oplus \zeta_{d-1}$, one has $\dim \zeta_d > 0$. Finally observe that $H = \zeta_0 \oplus \zeta_1 \oplus \cdots \oplus \zeta_d$, where $\zeta_0, \zeta_1, \ldots, \zeta_d$ are closed, $\dim \zeta_0 = \infty$, $0 \leq \dim \zeta_h < \infty$ for any $h \neq 0, d$, $0 < \dim \zeta_d < \infty$, implies that the identity is in equation d in system (A.2), i.e. that $A(z)^{-1}$ has a pole of order d at $z = 1$. This completes the proof of i). Next consider ii). Observe that (C.7) is based on the left version of system (A.2). By performing a similar induction on the right version of system (A.2), one reaches the right counterpart of (C.7), which implies $C_0 \left(P_{\mathscr{Z}_h^\perp} A_{h,1} P_{\mathscr{T}_h^\perp} \right) = \delta_{h,d} P_{\mathscr{T}_h^\perp}$ for $h = 0, 1, \ldots, d$, so that $\tau_d^\perp = \operatorname{Ker} P_{\mathscr{T}_d^\perp} A_{d,1} P_{\mathscr{T}_d^\perp} \subseteq \operatorname{Ker} P_{\mathscr{T}_d^\perp} = \mathscr{T}_d = \tau_0 \oplus \cdots \oplus \tau_{d-1}$; because the reverse inclusion is clearly satisfied, one has that $\operatorname{Ker} P_{\mathscr{T}_d^\perp} A_{d,1} P_{\mathscr{T}_d^\perp} = \operatorname{Ker} P_{\mathscr{T}_d^\perp}$, i.e. $\tau_d^\perp = \tau_0 \oplus \cdots \oplus \tau_{d-1}$, and hence $H = \tau_0 \oplus \tau_1 \oplus \cdots \oplus \tau_d$. The statement follows by the same reasoning in the proof of i), replacing $\operatorname{codim} \operatorname{Im} A_0$ with $\dim \operatorname{Ker} A_0$ and ζ with τ. This completes the proof of the second part of the statement.

The proof of the third part proceeds as follows. Write $A(z)^{-1} = \sum_{n=0}^{\infty} C_n (1-z)^{n-d}$ as

$$A(z)^{-1} = C_0 (1-z)^{-d} + \sum_{n=1}^{d-h-1} C_n (1-z)^{n-d} + (1-z)^{-h} R_0(z), \qquad R_0(1) = C_{d-h},$$

and pre-multiply by P_{τ_h} to find

$$P_{\tau_h} A(z)^{-1} = P_{\tau_h} C_0 (1-z)^{-d} + \sum_{n=1}^{d-h-1} P_{\tau_h} C_n (1-z)^{n-d} + (1-z)^{-h} P_{\tau_h} R_0(z).$$

First consider $h = 0, \ldots, d-1$. Setting $n = 0$ in (C.4) one has $P_{\tau_h} C_0 = 0$ and hence

$$P_{\tau_h} A(z)^{-1} = \sum_{n=1}^{d-h-1} P_{\tau_h} C_n (1-z)^{n-d} + (1-z)^{-h} P_{\tau_h} R_0(z). \tag{C.10}$$

From (C.4), for $n \leq d-h$ one has $P_{\tau_h} C_n = -Q_h \sum_{k=1}^{n} A_{h+1,k} C_{n-k} + \delta_{n+h,d} Q_h$ and hence

$$\sum_{n=1}^{d-h-1} P_{\tau_h} C_n (1-z)^{n-d} = - \sum_{n=1}^{d-h-1} \left(Q_h \sum_{k=1}^{n} A_{h+1,k} C_{n-k} \right) (1-z)^{n-d},$$

because $\delta_{n+h,d} = 0$ for $n = 1, \ldots, d-h-1$. Rearraging one thus finds

$$\sum_{n=1}^{d-h-1} P_{\tau_h} C_n (1-z)^{n-d} = -Q_h \sum_{k=1}^{d-h-1} A_{h+1,k} \left(\sum_{n=k}^{d-h-1} C_{n-k} (1-z)^{n-d} \right).$$

Next write

$$(1-z)^k A(z)^{-1} = \left(\sum_{n=k}^{d-h-1} C_{n-k} (1-z)^{n-d} \right) + (1-z)^{-h} R_k(z), \qquad R_k(1) = C_{d-h-k},$$

so that

$$\sum_{n=1}^{d-h-1} P_{\tau_h} C_n (1-z)^{n-d} = - \left(Q_h \sum_{k=1}^{d-h-1} A_{h+1,k} (1-z)^k \right) A(z)^{-1} + (1-z)^{-h} Q_h \sum_{k=1}^{d-h-1} A_{h+1,k} R_k(z).$$

Substituting in (C.10) and rearraging one thus finds $\gamma_h(z) A(z)^{-1} = (1-z)^{-h} \widetilde{\gamma}_h(z)$, where

$$\gamma_h(z) = P_{\tau_h} + Q_h \sum_{k=1}^{d-h-1} A_{h+1,k} (1-z)^k, \qquad \widetilde{\gamma}_h(z) = P_{\tau_h} R_0(z) + Q_h \sum_{k=1}^{d-h-1} A_{h+1,k} R_k(z).$$

Note that, because $R_k(1) = C_{d-h-k}$, one has

$$\widetilde{\gamma}_h(1) = P_{\tau_h} C_{d-h} + Q_h \sum_{k=1}^{d-h-1} A_{h+1,k} C_{d-h-k};$$

from (C.4) for $n = d - h$ one finds $P_{\tau_h} C_{d-h} + Q_h \sum_{k=1}^{d-h} A_{h+1,k} C_{d-h-k} = Q_h$, so that $\widetilde{\gamma}_h(1) = Q_h(I - A_{h+1,d-h} C_0)$. Because $\langle v, \widetilde{\gamma}_h(1)x\rangle \neq 0$ for any $v \in \tau_h$ and any $x \in \zeta_h$, one has that $\langle v, \gamma_h(z)A(z)^{-1}\rangle$ has a pole of order h for all $v \in \tau_h$, $h = 0, \ldots, d-1$. Finally consider $h = d$. Setting $n = 0$ in (C.4) one has $P_{\tau_d} C_0 = Q_d$ and this shows that $\langle v, A(z)^{-1}\rangle$ has a pole of order d for all $v \in \tau_d$. This completes the proof. ■

Proof of Theorem 5.4. Direct consequence of Lemma *C*.2. ■

References

Beare, B. (2017). The Chang-Kim-Park model of cointegrated density-valued time series cannot accommodate a stochastic trend. *Econ Journal Watch 14*, 133–137.

Beare, B., J. Seo, and W. Seo (2017). Cointegrated Linear Processes in Hilbert Space. *Journal of Time Series Analysis 38*, 1010–1027.

Ben-Israel, A. and T. Greville (2003). *Generalized Inverses: Theory and Applications, 2nd ed.* Springer.

Bosq, D. (2000). *Linear Processes in Function Spaces.* Springer-Verlag.

Chang, Y., B. Hu, and J. Park (2016). On the Error Correction Model for Functional Time Series with Unit Roots. Mimeo, Indiana University.

Chang, Y., C. Kim, and J. Park (2016). Nonstationarity in time series of state densities. *Journal of Econometrics 192*, 152 – 167.

Franchi, M. and P. Paruolo (2016). Inverting a matrix function around a singularity via local rank factorization. *SIAM Journal of Matrix Analysis and Applications 37*, 774–797.

Franchi, M. and P. Paruolo (2017). A general inversion theorem for cointegration. DSS-E3 Working Paper 2017/3, Sapienza University of Rome.

Gabrys, R., S. Hörmann, and P. Kokoszka (2013). Monitoring the Intraday Volatility Pattern. *Journal of Time Series Econometrics 5*, 87–116.

Gohberg, I., S. Goldberg, and M. Kaashoek (1990). *Classes of linear operators, vol. 1.* Operator theory. Birkhäuser Verlag.

Gohberg, I., S. Goldberg, and M. Kaashoek (2003). *Basic Classes of Linear Operators.* Birkhäuser, Basel-Boston-Berlin.

Gohberg, I., M. Kaashoek, and F. van Schagen (1993). On the local theory of regular analytic matrix functions. *Linear Algebra and its Applications 182*, 9–25.

Hörmann, S., L. Horváth, and R. Reeder (2013). A functional version of the ARCH model. *Econometric Theory 29*, 267–288.

Hörmann, S. and P. Kokoszka (2012). Functional time series. *Handbook of Statistics 30*, 157–186.

Horváth, L. and P. Kokoszka (2012). *Inference for Functional Data with Application.* Springer.

Hu, B. and J. Park (2016). Econometric Analysis of Functional Dynamics in the Presence of Persistence. Mimeo, Indiana University.

Johansen, S. (1996). *Likelihood-based Inference in Cointegrated Vector Auto-Regressive Models.* Oxford University Press.

Kargin, V. and A. Onatski (2008). Curve forecasting by functional autoregression. *Journal of Multivariate Analysis 99*, 2508 – 2526.

Kokoszka, P. and M. Reimherr (2017). *Introduction to Functional Data Analysis.* Chapman and Hall.