

# COINTEGRATION IN FUNCTIONAL AUTOREGRESSIVE PROCESSES

MASSIMO FRANCHI  
*Sapienza University of Rome*

PAOLO PARUOLO  
*European Commission, Joint Research Centre*

This article defines the class of  $\mathcal{H}$ -valued autoregressive (AR) processes with a unit root of finite type, where  $\mathcal{H}$  is a possibly infinite-dimensional separable Hilbert space, and derives a generalization of the Granger–Johansen Representation Theorem valid for any integration order  $d = 1, 2, \dots$ . An existence theorem shows that the solution of an AR process with a unit root of finite type is necessarily integrated of some finite integer order  $d$ , displays a common trends representation with a finite number of common stochastic trends, and it possesses an infinite-dimensional cointegrating space when  $\dim \mathcal{H} = \infty$ . A characterization theorem clarifies the connections between the structure of the AR operators and (i) the order of integration, (ii) the structure of the attractor space and the cointegrating space, (iii) the expression of the cointegrating relations, and (iv) the triangular representation of the process. Except for the fact that the dimension of the cointegrating space is infinite when  $\dim \mathcal{H} = \infty$ , the representation of AR processes with a unit root of finite type coincides with the one of finite-dimensional VARs, which can be obtained setting  $\mathcal{H} = \mathbb{R}^p$  in the present results.

## 1. INTRODUCTION

The theory of time series that take values in infinite-dimensional separable Hilbert spaces, or infinite-dimensional  $\mathcal{H}$ -valued processes, is receiving increasing attention in econometrics. Infinite-dimensional  $\mathcal{H}$ -valued processes allow to represent directly the dynamics of infinite-dimensional objects, such as Lebesgue square-integrable functions on a compact domain. In this way, they allow greater modeling generality with respect to models for conditional means and variances, see, e.g., Horváth and Kokoszka (2012).

---

The article benefited from useful comments from the Editor, Peter C.B. Phillips, three anonymous referees and conference and seminar participants at 2018 NBER-NSF Time Series Conference, University of California San Diego, University of Bologna and ICEEE 2019, University of Lecce. The idea of the present article was 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. During the revision of the article in September 2018, the first author visited the Department of Economics, University of California San Diego, and the hospitality of Brendan K. Beare is gratefully acknowledged. The first author acknowledges partial financial support from MIUR PRIN grant 2010J3LZEN. The views expressed in this article do not necessarily reflect those of the authors' institutions. Address correspondence to Paolo Paruolo, European Commission, Joint Research Centre, Via E. Fermi 2749, I-21027 Ispra (VA), Italy; e-mail: paolo.paruolo@ec.europa.eu.

Economic applications of functional time series analysis include studies on the term structure of interest rates, see Kargin and Onatski (2008), and intraday volatility, see Hörmann, Horváth, and Reeder (2013) and Gabrys, Hörmann, and Kokoszka (2013); additional applications can be found in the recent monograph by Kokoszka and Reimherr (2017) and in the review article by Hörmann and Kokoszka (2012).

One notable special case is given by  $\mathcal{H}$ -valued processes  $x_t = \psi(f_t)$ , where  $f_t$  is a generic probability density function (pdf) and  $\psi$  is an invertible transformation, see Petersen and Müller (2016), Beare (2017), and Seo and Beare (2019).<sup>1</sup> Modeling dynamics of an entire density or parts of a density is of practical interest in modeling income distributions, see, e.g., Bourguignon, Ferreira, and Lustig (2005), Piketty (2014), and Chang, Kim, and Park (2016b).

An important early contribution to the theory 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 infinite-dimensional  $\mathcal{H}$ -valued stationary autoregressive (AR) models.

Let  $x_t$  be an infinite-dimensional  $\mathcal{H}$ -valued process and let  $\langle \cdot, \cdot \rangle$  be the inner product on  $\mathcal{H}$ ; as observed in Hu and Park (2016), the inner product  $\langle v, x_t \rangle$  for some  $v \in \mathcal{H}$  is the generalization of a linear combination in  $\mathbb{R}^P$ ; this is called the  $v$ -characteristic of  $x_t$ . The simplest form of cointegration for the  $\mathcal{H}$ -valued process  $x_t$  corresponds to a process integrated of order one ( $I(1)$ ), i.e., a random walk type process, together with some stationary  $v$ -characteristic of  $x_t$ .

Recently, Chang et al. (2016b) applied Functional Principal Components Analysis (FPCA) directly on the space of densities for individual earnings and intra-month distributions of stock returns. They found evidence of unit root persistence in a handful of  $v$ -characteristics of these cross-sectional distributions. The framework proposed by Chang et al. (2016b) 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, and the potential unit roots are identified and tested through FPCA.

The representation of infinite-dimensional  $\mathcal{H}$ -valued AR processes with unit roots has been recently considered in the literature. Hu and Park (2016) consider infinite-dimensional  $\mathcal{H}$ -valued AR(1) processes with a compact operator and prove that an extension of the Granger–Johansen Representation Theorem, see Theorem 4.2 in Johansen (1996), holds in the  $I(1)$  case. The corresponding common trends representation, or functional Beveridge–Nelson decomposition, displays a finite number of  $I(1)$  stochastic trends and an infinite-dimensional

<sup>1</sup> Density functions do not form a vector space with the standard addition and multiplication operations. This difficulty can be overcome by the transformation approach of Petersen and Müller (2016) and/or by redefining the basic operations of addition and multiplication, as in Seo and Beare (2019) and references thereof.

cointegrating space. They further propose an estimator for the functional autoregressive operator which builds on the results in Chang et al. (2016b).

Beare, Seo, and Seo (2017) consider infinite-dimensional  $\mathcal{H}$ -valued AR( $k$ ),  $k \geq 1$ , with compact operators if  $k > 1$  and no compactness assumption if  $k = 1$ , and show that the counterpart of the conditions of the Granger–Johansen Representation Theorem are sufficient to prove the existence an  $I(1)$  representation with cointegration. If  $k > 1$ , the number of  $I(1)$  stochastic trends is finite and the dimension of the cointegrating space is infinite, whereas if  $k = 1$  this is not necessarily the case.

To obtain the common trends representation of infinite-dimensional  $\mathcal{H}$ -valued AR( $k$ ),  $k \geq 1$ , with compact operators, Beare and Seo (2019) are the first to employ a theorem on the inversion of analytic operator functions in Gohberg, Goldberg, and Kaashoek (1990).<sup>2</sup> They also present results on the  $I(2)$  case that show that the number of  $I(2)$  stochastic trends is finite and the dimension of the cointegrating space is infinite.

Finally, Chang, Hu, and Park (2016a) consider infinite-dimensional  $\mathcal{H}$ -valued AR( $k$ ) processes possessing an error correction form with a 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, they show that the Granger–Johansen Representation Theorem continues to hold.

This article introduces the class of AR processes with a unit root of finite type, which contains infinite-dimensional  $\mathcal{H}$ -valued ARs with compact operators as a special case, and derives a generalization of the Granger–Johansen Representation Theorem valid for any integration order  $d = 1, 2, \dots$ . Necessary and sufficient conditions for AR processes with a unit root of finite type to generate cointegrated  $I(d)$  processes are provided, in a parallel way with respect to the  $I(d)$  representation results in the finite-dimensional VAR case derived in Franchi and Paruolo (2019). When  $d > 1$ , the notion of polynomial cointegration, see Granger and Lee (1989), corresponds to the existence of some linear combination of  $v$ -characteristics of  $x_t$  and other  $v$ -characteristics of  $\Delta^j x_t$ ,  $j = 1, \dots, d - 1$  that is integrated of order less than  $d$ .

More specifically, an existence theorem is provided, which shows that the solution of an AR process with a unit root of finite type is necessarily  $I(d)$  for some finite integer  $d$ , and it displays a common trends representation with a finite number of common stochastic trends of the type of (cumulated) bilateral random walks and a cointegrating space which is infinite-dimensional when  $\dim \mathcal{H} = \infty$ . This result is a direct consequence of the Analytic Fredholm Theorem, see Gohberg et al. (1990), that was first employed in Beare and Seo (2019) in the context of infinite-dimensional  $\mathcal{H}$ -valued ARs with compact operators.

<sup>2</sup> This theorem is usually called the Analytic Fredholm Theorem, and it used here to discuss the existence of a common trends representation in Section 3.

Despite these important implications, this existence result does not address a number of central issues, such as the connections between the structure of the AR operators and (i) the order of integration of the process, (ii) the structure of the attractor space and the cointegrating space, and (iii) the expression of the cointegrating relations.

The characterization of these links in the generic  $I(d)$  case constitutes the main contribution of the present article. More specifically, a necessary and sufficient condition for the order of integration  $d$  is given in terms of the decomposition of the space  $\mathcal{H}$  into the direct sum of  $d + 1$  orthogonal subspaces  $\tau_h$ ,  $h = 0, \dots, d$ , that are expressed recursively in terms of the AR operators. This condition is called the “POLE( $d$ ) condition”, because it is a necessary and sufficient condition for the inverse of an operator function  $A(z)$  to have a pole of order  $d$  at  $z = 1$ .

A crucial feature of the present POLE( $d$ ) condition is that the subspaces in the orthogonal direct sum decomposition  $\mathcal{H} = \tau_0 \oplus \tau_1 \oplus \dots \oplus \tau_d$ ,  $\tau_d \neq \{0\}$ , identify the directions in which the properties of the process differ. Specifically, for any nonzero  $v \in \tau_0$ , for which it is found that  $\dim \tau_0 = \infty$  when  $\dim \mathcal{H} = \infty$ , one can combine  $v$ -characteristics of  $x_t$  with other  $v$ -characteristics of  $\Delta^j x_t$ ,  $j = 1, \dots, d - 1$ , to obtain  $I(0)$  polynomial cointegrating relations.

For  $h = 1, \dots, d - 2$  and any nonzero  $v \in \tau_h$ , which is found to have dimension  $0 \leq \dim \tau_h < \infty$ , one can combine  $v$ -characteristics of  $x_t$  with other  $v$ -characteristics of  $\Delta^j x_t$ ,  $j = 1, \dots, d - h + 1$ , to obtain  $I(h)$  polynomial cointegrating relations. In  $\tau_{d-1}$ , with  $0 \leq \dim \tau_{d-1} < \infty$ , one has  $v$ -characteristics that are  $I(d - 1)$  and do not allow for polynomial cointegration. Finally, all nonzero  $v$ -characteristics  $v \in \tau_d$  do not present cointegration and  $0 < \dim \tau_d < \infty$ . These results parallel the ones in the triangular representation in the finite-dimensional case  $\mathcal{H} = \mathbb{R}^p$  discussed in Phillips (1991a) and Stock and Watson (1993); see also Franchi and Paruolo (2019).

These results show that conditions and properties of AR processes with a unit root of finite type extend those that apply in the usual finite-dimensional VAR case; in fact, setting  $\mathcal{H} = \mathbb{R}^p$  in the present results one finds the  $I(1)$  and  $I(2)$  results in Johansen (1996), and for the generic  $I(d)$  case, one finds the results in Franchi and Paruolo (2019). This shows that except for the fact that the dimension of the cointegrating space is infinite when  $\dim \mathcal{H} = \infty$ , the infinite-dimensionality of  $\mathcal{H}$  does not introduce additional elements in the representation analysis of AR processes with a unit root of finite type.

The present results are based on an orthogonal decomposition of the embedding Hilbert space and on the corresponding orthogonal projections. Orthogonal and nonorthogonal projections are well-known concepts in econometrics, where the choice between the two is usually discussed in terms of estimation efficiency; see Phillips (1991b) for how these arguments are modified for spectral GLS regressions methods in a cointegration context. In the context of the present representation theory, results can be obtained using either orthogonal, as in this article, or nonorthogonal projections, as done in Hu and Park (2016), Chang et al. (2016a),

Beare et al. (2017), and Beare and Seo (2019). The present choice of orthogonal projections is found to ease exposition and to allow for a characterization of the generic  $I(d)$  case.

The rest of the article is organized as follows: Section 2 presents basic definitions and concepts, Section 3 discusses the assumption of unit root of finite type and reports an initial existence result, Section 4 provides a characterization of  $I(1)$  and  $I(2)$  AR processes with a unit root of finite type, and Section 5 extends the analysis to the general  $I(d)$ ,  $d = 1, 2, \dots$ , case. Section 6 concludes.

Three Appendices collect background definitions, novel inversion results and proofs of the statements in the article. Specifically, Appendix A reviews notions on operators acting on a separable Hilbert space  $\mathcal{H}$  and on  $\mathcal{H}$ -valued random variables; Appendix B presents novel results on the inversion of a meromorphic operator function and Appendix C reports proofs of the results in the text.

**2.  $\mathcal{H}$ -VALUED LINEAR PROCESS, ORDER OF INTEGRATION AND COINTEGRATION**

This section introduces the notions of weakly stationary, white noise, linear, integrated, and cointegrated processes that take values in a possibly infinite-dimensional separable complex Hilbert space  $\mathcal{H}$ , where separable means that  $\mathcal{H}$  admits a countable orthonormal basis.<sup>3</sup> Basic definitions of operators acting on  $\mathcal{H}$  and of  $\mathcal{H}$ -valued random variables are reported in Appendix A. Time is indexed by  $t$ , which is discrete,  $t \in \mathbb{Z} = (\dots, -1, 0, 1, \dots)$ . The time-lag operator is denoted by  $L$  and  $\Delta = 1 - L$  is the difference operator; hence, for  $x_t \in \mathcal{H}$ , one has  $Lx_t = x_{t-1}$  and  $\Delta x_t = x_t - x_{t-1}$ .

**2.1. Definitions**

The definitions of weakly stationary and white noise process are taken from Bosq (2000, Definitions 2.4, 3.1, 7.1), whereas those of linear, integrated, and cointegrated process are adapted from Johansen (1996); they are similar to those employed in Chang et al. (2016b), Beare et al. (2017), Beare and Seo (2019). The definitions of expectation  $E(\cdot)$ , covariance operator, and cross-covariance function used in the following are reported in Appendix A.2.

**DEFINITION 2.1** (Weakly stationary process). *An  $\mathcal{H}$ -valued stochastic process  $\{\varepsilon_t, t \in \mathbb{Z}\}$  is said to be weakly stationary if (i)  $0 < E(\|\varepsilon_t\|^2) < \infty$ , (ii)  $E(\varepsilon_t)$  and the covariance operator of  $\varepsilon_t$  do not depend on  $t$ , and (iii) the cross-covariance function of  $\varepsilon_t$  and  $\varepsilon_s$ ,  $c_{\varepsilon_t, \varepsilon_s}(h, v)$ , is such that  $c_{\varepsilon_t, \varepsilon_s}(h, v) = c_{\varepsilon_{t+u}, \varepsilon_{s+u}}(h, v)$  for all  $h, v \in \mathcal{H}$  and all  $s, t, u \in \mathbb{Z}$ .*

The notion of  $\mathcal{H}$ -valued white noise is presented next.

<sup>3</sup> All results in the article are valid also for real separable Hilbert spaces; in the following examples are taken from the latter.

**DEFINITION 2.2** (White noise process). *An  $\mathcal{H}$ -valued weakly stationary stochastic process  $\{\varepsilon_t, t \in \mathbb{Z}\}$  is said to be white noise if (i)  $E(\varepsilon_t) = 0$  and (ii)  $c_{\varepsilon_t, \varepsilon_s}(h, v) = 0$  for all  $h, v \in \mathcal{H}$  and all  $s \neq t, s, t \in \mathbb{Z}$ , where  $c_{\varepsilon_t, \varepsilon_s}(h, v)$  is the cross-covariance function of  $\varepsilon_t$  and  $\varepsilon_s$ ; it is called strong white noise if (i) holds, and (ii) is replaced by the requirement that  $\varepsilon_t$  is an i.i.d. sequence of  $\mathcal{H}$ -valued random variables.*

In the following, it is assumed that any white noise is nondegenerate, i.e., that the probability that  $\varepsilon_t$  belongs to a strict subspace of  $\mathcal{H}$  for all  $t$  is equal to zero. Note that by definition any strong white noise is white noise, and any white noise process is weakly stationary. The same property holds for linear combinations of lags of a white noise process with suitable weights; this leads to the class of linear processes, introduced in Definition 2.3 below, where the following notation is employed:  $D(z_0, \rho)$  denotes the open disc  $\{z \in \mathbb{C} : |z - z_0| < \rho\}$  with center  $z_0 \in \mathbb{C}$  and radius  $0 < \rho \in \mathbb{R}$ ,  $\mathcal{H}_1, \mathcal{H}_2$  are separable Hilbert spaces and  $\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}$  indicates the set of bounded linear operators  $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$  with norm  $\|A\|_{\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}} = \sup_{\|v\|=1} \|Av\|$ ; an operator function  $B(z) = \sum_{n=0}^{\infty} B_n(z - z_0)^n$ , where  $B_n \in \mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}$ , is said to be absolutely convergent on  $D(z_0, \rho)$  if  $\sum_{n=0}^{\infty} \|B_n\|_{\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}} |z - z_0|^n < \infty$  for all  $z \in D(z_0, \rho)$ .<sup>4</sup> When  $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$ , the simplified notation  $\mathcal{L}_{\mathcal{H}}$  is used for  $\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}$ .

**DEFINITION 2.3** (Linear process). *Let  $\{\varepsilon_t, t \in \mathbb{Z}\}$  be an  $\mathcal{H}_1$ -valued (strong) white noise; an  $\mathcal{H}_2$ -valued stochastic process  $\{x_t, t \in \mathbb{Z}\}$  with  $\mu_t = E(x_t)$  is said to be a linear process if*

$$x_t = \mu_t + \sum_{n=0}^{\infty} B_n \varepsilon_{t-n}, \quad B_n \in \mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}, \quad B_0 \neq 0,$$

where  $B(z) = \sum_{n=0}^{\infty} B_n z^n, z \in \mathbb{C}$ , is absolutely convergent on the open disc  $D(0, \rho)$  for some  $\rho > 1$ .

Usually, the notion of strong white noise is used in the definition of linear processes; the only consequence of strong versus weak white noise on the representations is that the cumulation of  $\{\varepsilon_t, t \in \mathbb{Z}\}$  gives rise to a random walk with increments that are independent versus just uncorrelated.

Observe also that Definition 2.3 allows the white noise  $\{\varepsilon_t, t \in \mathbb{Z}\}$  and the linear process to live in different Hilbert spaces  $\mathcal{H}_1, \mathcal{H}_2$ . In particular, this is useful when considering  $v$ -characteristics  $w_t = \langle v, x_t \rangle$  of a linear process  $x_t = \sum_{n=0}^{\infty} B_n \varepsilon_{t-n}$  with  $x_t, \varepsilon_t \in \mathcal{H}_1$  say, for which one has  $w_t \in \mathcal{H}_2 = \mathbb{C}$  (or  $\mathcal{H}_2 = \mathbb{R}$ ) and  $w_t = \sum_{n=0}^{\infty} B_n^* \varepsilon_{t-n}$  where  $B_n^* \varepsilon_{t-n} = \langle v, B_n \varepsilon_{t-n} \rangle$  and  $B_n^* \in \mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}$ ; hence, in this case, the white noise process  $\varepsilon_t$  is an element of a possibly infinite-dimensional

<sup>4</sup> Note that  $\sum_{n=0}^{\infty} \|B_n\|_{\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}} |z - z_0|^n < \infty$  for all  $z \in D(z_0, \rho)$  implies that  $\sum_{n=0}^{\infty} B_n(z - z_0)^n$  converges in the operator norm to  $B(z) \in \mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}$  for all  $z \in D(z_0, \rho)$ , i.e.,  $\|B(z) - \sum_{n=0}^N B_n(z - z_0)^n\|_{\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}} \rightarrow 0$  as  $N \rightarrow \infty$ .

Hilbert space  $\mathcal{H}_1$  (as  $x_t$ ), whereas its  $v$ -characteristics  $w_t$  is an element of a one-dimensional Hilbert space  $\mathbb{C}$  (or  $\mathbb{R}$ ).<sup>5</sup>

As discussed in Section 7.1 in Bosq (2000), existence and weak stationarity of  $\sum_{n=0}^{\infty} B_n \varepsilon_{t-n}$  are guaranteed by the square summability condition  $\sum_{n=0}^{\infty} \|B_n\|_{\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}}^2 < \infty$ . Observe that the requirement that  $B(z)$  is absolutely convergent on  $D(0, \rho)$  for some  $\rho > 1$  is stronger. In fact,  $\sum_{n=0}^{\infty} \|B_n\|_{\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}} |z|^n < \infty$  for all  $z \in D(0, \rho)$ ,  $\rho > 1$ , implies  $\sum_{n=0}^{\infty} \|B_n\|_{\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}} < \infty$  and hence  $\sum_{n=0}^{\infty} \|B_n\|_{\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}}^2 < \infty$ . This shows that  $x_t - \mu_t$  in Definition 2.3 is well defined and weakly stationary.

Moreover, because  $B(z)$  is a bounded linear operator for all  $z \in D(0, \rho)$ ,  $\rho > 1$ ,  $B(1)$  is a bounded linear operator. Finally, note that  $B(z)$  is infinitely differentiable on  $D(0, \rho)$ ,  $\rho > 1$ , and the series obtained by termwise  $k$  times differentiation,  $\sum_{n=k}^{\infty} n(n-1)\dots(n-k+1)B_n z^{n-k}$ , is absolutely convergent and coincides with the  $k$ -th derivative of  $B(z)$  for each  $z \in D(0, \rho)$ . Hence,  $\sum_{n=k}^{\infty} n(n-1)\dots(n-k+1)\|B_n\|_{\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}} < \infty$ , which for  $k = 1$  reads  $\sum_{n=1}^{\infty} n\|B_n\|_{\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}} < \infty$ ; this condition is employed in Chang et al. (2016b) and in Beare and Seo (2019).

The notions of integration and cointegration are introduced next.

**DEFINITION 2.4 (Order of integration).** *A linear process  $x_t = \mu_t + B(L)\varepsilon_t$  is said to be integrated of order 0, written  $x_t \sim I(0)$ , if  $B(1) \neq 0$ . If  $\Delta^d x_t$  is  $I(0)$  for some finite integer  $d = 1, 2, \dots$ ,  $\{x_t, t \in \mathbb{Z}\}$  is said to be integrated of order  $d$ , indicated  $x_t \sim I(d)$ .*

This definition coincides with Definition 3.3 in Johansen (1996) of an  $I(d)$  process for the special case  $\mathcal{H} = \mathbb{R}^p$ .

Observe that a white noise process is  $I(0)$  and that an  $I(0)$  process is weakly stationary. To see that a weakly stationary process is not necessarily  $I(0)$ , take for instance  $u_t = \Delta \varepsilon_t = \varepsilon_t - \varepsilon_{t-1}$ ; this process is weakly stationary, with  $B(1) = 0$  and hence it does not satisfy the definition of an  $I(0)$  process, showing that the two concepts do not coincide. The distinction between weak stationarity and  $I(0)$ -ness is relevant for the definition of order of integration: in fact, the cumulation of an  $I(0)$  process is necessarily  $I(1)$ , whereas the cumulation of a stationary process is not necessarily so.

Following Hu and Park (2016), one can define the  $v$ -characteristic of  $x_t$  as the scalar process  $\langle v, x_t \rangle$ , for any  $v \in \mathcal{H}$ . Orthogonality with respect to the inner product leads to the following definition of orthogonal complement  $S^\perp$  of a subspace  $S \subseteq \mathcal{H}$ ,  $S^\perp = \{v \in \mathcal{H} : \langle v, s \rangle = 0 \text{ for all } s \in S\}$ . From Definition 2.4, one can see that a generic  $v$ -characteristic of  $x_t \sim I(d)$  is itself at most integrated of order  $d$ ; the case when a  $v$ -characteristic of  $x_t \sim I(d)$  is integrated of lower order  $b < d$  is associated with the notion of cointegration.

<sup>5</sup> The same construction is used in Beare and Seo (2019) Section 3.

**DEFINITION 2.5** (Cointegrated process). *An  $I(d)$  process  $x_t$  is said to be cointegrated (respectively polynomially cointegrated) if there exists a nonzero  $v \in \mathcal{H}$  such that  $\langle v, x_t \rangle$  (respectively  $\langle v, x_t \rangle + \sum_{n=1}^{d-b-1} \langle v, G_n \Delta^n x_t \rangle$  for some  $G_n$ ) is  $I(b)$  for some  $b < d$ ; in this case  $v \neq 0$  is called a cointegrating vector. The closed span of all cointegrating vectors is called the cointegrating space and its orthogonal complement is called the attractor space.*

As in the usual finite-dimensional case,  $x_t$  is cointegrated if there exists a  $v$ -characteristic of  $x_t$  that has lower order of integration than the original process. The cointegrating space contains all nonzero  $v \in \mathcal{H}$  that correspond to a  $v$ -characteristic of  $x_t$  with lower order of integration than the original process  $x_t$  and the null vector  $0 \in \mathcal{H}$ . On the other hand, the attractor space is the subspace where the dominant  $I(d)$  trends take values.

When both the attractor and the cointegrating space have finite dimension,  $\mathcal{H}$  is finite-dimensional, so that the standard results in the literature apply. The case in which the attractor space is infinite-dimensional corresponds to a process with an infinite number of  $I(d)$  stochastic trends; for  $d = 1$ , this case has been discussed in Chang et al. (2016a) and in Beare et al. (2017) for  $k = 1$ ; this case is not covered by the present results.

Most of the contributions in the literature have studied instead the case of an attractor space of finite dimension and a cointegrating space of infinite dimension. This is the setup studied for  $d = 1$  in Chang et al. (2016b), Hu and Park (2016) and Beare et al. (2017) for  $k > 1$  and in Beare and Seo (2019) for  $d = 1, 2$  and  $k \geq 1$ . This is the setting considered in the present article as well, and it is motivated also by the next example.

## 2.2. Yield Curve Example

As an example of an infinite-dimensional separable Hilbert space of economic interest, consider the yield curve  $x_{\circ,t}(s)$ , where  $s$  denotes maturity and  $t$  time. In this subsection  $t$  is omitted, unless needed for clarity.

Let  $\mathcal{H}$  be the set of Lebesgue measurable real functions  $x_{\circ}(s)$  such that  $\int_0^{s_{\max}} x_{\circ}^2(s) ds < \infty$ , where  $s_{\max}$  is the maximal maturity and  $\int g(s) ds$  indicates a Lebesgue integral with respect to the Lebesgue measure. One can rescale the maturity  $s$  into  $u = s/s_{\max}$  and define the rescaled yield curve  $x(u)$  by  $x(u) = x_{\circ}(u \cdot s_{\max})$ , with  $u \in (0, 1]$  and  $\int_0^1 x^2(u) du < \infty$ . The vector space operations on  $\mathcal{H}$  are defined in a natural way as  $(x + y)(u) = x(u) + y(u)$  and  $(\alpha x)(u) = \alpha x(u)$  where  $\alpha \in \mathbb{R}$ . Next, define the inner product

$$\langle x, y \rangle = \int_0^1 x(u)y(u) du. \quad (2.1)$$

This set of Lebesgue square-integrable functions with inner product (2.1), when identifying functions which are equal almost everywhere, is a separable real Hilbert space, see, e.g., Kokoszka and Reimherr (2017, p. 214).

The yield curve is often described in terms of the three features of level, slope, and curvature, see, e.g., Cochrane and Piazzesi (2005). These features of the



yield curve can be associated with the following  $v$ -characteristics of  $x$ . Define  $\pi_{j,1}, \dots, \pi_{j,j}$  as a partition of the unit interval  $(0, 1]$  into  $j$  segments  $\pi_{j,i}$  of length  $1/j$ ,  $\pi_{j,i} = (\frac{i-1}{j}, \frac{i}{j}]$ , and let  $1_{\{u \in \pi_{j,i}\}}$  be the indicator function that takes value one when  $u \in \pi_{j,i}$  and equals 0 otherwise.

Next define the following  $v$  functions

$$v_0(u) = 1_{\{u \in \pi_{1,1}\}}, \quad v_1(u) = \frac{1}{2} (1_{\{u \in \pi_{2,2}\}} - 1_{\{u \in \pi_{2,1}\}}),$$

$$v_2(u) = \frac{1}{4} (1_{\{u \in \pi_{4,4}\}} - 1_{\{u \in \pi_{4,3}\}}) - \frac{1}{4} (1_{\{u \in \pi_{4,2}\}} - 1_{\{u \in \pi_{4,1}\}}),$$

and observe that they belong to  $\mathcal{H}$ , because they are Lebesgue square-integrable functions. Finally, let  $x$  denote the rescaled yield curve and note that

$$\langle v_0, x \rangle = \int_0^1 v_0(u)x(u)du = \int_0^1 x(u)du,$$

$$\langle v_1, x \rangle = \int_0^1 v_1(u)x(u)du = \frac{1}{2} \left( \int_{\frac{1}{2}}^1 x(u)du - \int_0^{\frac{1}{2}} x(u)du \right),$$

$$\langle v_2, x \rangle = \int_0^1 v_2(u)x(u)du = \frac{1}{4} \left( \int_{\frac{3}{4}}^1 x(u)du - \int_{\frac{1}{2}}^{\frac{3}{4}} x(u)du \right) - \frac{1}{4} \left( \int_{\frac{1}{4}}^{\frac{1}{2}} x(u)du - \int_0^{\frac{1}{4}} x(u)du \right).$$

One can see that  $\langle v_0, x \rangle$  computes the average yield curve, and hence can be associated with the level of the yield curve. Similarly,  $\langle v_1, x \rangle$  computes the difference between the average yield on the longer maturities and the one on the shorter maturities; hence, it can be associated with the slope of the yield curve. Finally,  $\langle v_2, x \rangle$  computes the difference of the slopes on the longer maturities and the shorter maturities; hence, it can be associated with the curvature of the yield curve.

This shows that  $v_0, v_1, v_2$  define interesting  $v$ -characteristics for the yield curve  $x$ . If the yield curve  $x$  is modeled as a functional time series,  $x_t$ , then it is interesting to ask questions of the type: “what is the order of integration of the level (or slope, or curvature) of the yield curve?”. These questions translate into “what is the order of integration of the  $v_j$ -characteristics,  $j = 0, 1, 2$ , of the yield curve  $x_t$ ?”.

This illustrates how interesting hypotheses can be formulated in this context; other types of hypotheses can be formulated in a similar way. Moreover, it is of interest to determine how many and which characteristics are nonstationary, which corresponds to estimating the (dimension of the) attractor space in the  $I(1)$  case.

It appears natural in this context to assume (or test) that there are only a finite number of factors driving the dynamics of the yield curve. This seems to be a reasonable assumption also beyond the case of the yield curve; this setup is the one studied in the present article.

### 3. AR PROCESS WITH A UNIT ROOT OF FINITE TYPE

This section introduces the class of  $\mathcal{H}$ -valued ARs that is studied in the present article, called AR processes with a unit root of finite type, where  $\mathcal{H}$  is a possibly infinite-dimensional separable Hilbert space. It also presents an existence result about their common trends representation, which shows that the solution of an AR process with a unit root of finite type is necessarily  $I(d)$  for some finite integer  $d$ , displays a common trends representation with a finite number of common stochastic trends of the type of (cumulated) bilateral random walks and it possesses an infinite-dimensional cointegrating space when  $\dim \mathcal{H} = \infty$ .<sup>6</sup> The relations of AR processes with a unit root of finite type with the infinite-dimensional  $\mathcal{H}$ -valued ARs studied in literature are also discussed in this section, and an example of an AR process with a unit root of finite type with a noncompact AR operator is given.

#### 3.1. Main Assumption

Consider an  $\mathcal{H}$ -valued AR process

$$x_t = A_1^\circ x_{t-1} + \dots + A_k^\circ x_{t-k} + \varepsilon_t, \quad A_n^\circ \in \mathcal{L}\mathcal{H}, \quad t \in \mathbb{Z}, \tag{3.1}$$

where  $\mathcal{H}$  is a possibly infinite-dimensional separable Hilbert space,  $x_t, \varepsilon_t \in \mathcal{H}$ ,  $\{\varepsilon_t, t \in \mathbb{Z}\}$  is white noise, and the operator function

$$A(z) = I - \sum_{h=1}^k A_h^\circ z^h, \quad z \in \mathbb{C}, \quad A(1) \neq 0, \tag{3.2}$$

is noninvertible at  $z = 1$  and invertible in the punctured disc  $D(0, \rho) \setminus \{1\}$  for some  $\rho > 1$ ; here  $I$  indicates the identity operator in  $\mathcal{L}\mathcal{H}$ .

This requirement restricts attention to unit roots at frequency zero, corresponding to the point  $z = 1$  on the unit disc. Note that there is no loss of generality in assuming that  $A(1) \neq 0$ . In fact, if  $A(1) = 0$ , one can factorize  $(1 - z)^s$  from  $A(z)$ ,  $A(z) = (1 - z)^s \tilde{A}(z)$  for some  $\tilde{A}(1) \neq \{0\}$  and some  $s > 0$ , and rewrite the AR equations  $A(L)x_t = \varepsilon_t$  as  $\tilde{A}(L)y_t = \varepsilon_t$  for  $y_t = \Delta^s x_t$ .

The notion of eigenvalue of finite type, see Gohberg et al. (1990, Sect. XI.8), is central in the present setup and it is reported next. For any  $A \in \mathcal{L}\mathcal{H}$  the subspace  $\{v \in \mathcal{H} : Av = 0\}$ , written  $\text{Ker } A$ , is called the kernel of  $A$  and the subspace  $\{Av : v \in \mathcal{H}\}$ , written  $\text{Im } A$ , is called the image of  $A$ . The dimension of  $\text{Im } A$ , written  $\dim \text{Im } A$ , is called the rank of  $A$ .

**DEFINITION 3.1** (Eigenvalue of finite type).  *$A(z)$  in (3.2) is said to have an eigenvalue of finite type at  $z_0 \in \mathbb{C}$  if*

<sup>6</sup> This result is a direct consequence of the Analytic Fredholm Theorem, reported in Theorem A.1 in Appendix A.1, and first employed in Beare and Seo (2019) in the context of infinite-dimensional  $\mathcal{H}$ -valued ARs with compact operators.

- (i)  $A_0 = A(z_0)$  is Fredholm of index  $n - q$ , where  $n = \dim \text{Ker} A_0 < \infty$  and  $q = \dim(\text{Im} A_0)^\perp < \infty$ ,
- (ii)  $A_0 v = 0$  for some nonzero  $v \in \mathcal{H}$ ,
- (iii)  $A(z)$  is invertible for all  $z$  in some punctured disc  $D(z_0, \delta) \setminus \{z_0\}$ ,  $\delta > 0$ .

Direct consequences of this definition are listed in the following remark.

**Remark 3.2.** If  $A(z)$  has an eigenvalue of finite type at  $z = z_0$ ,  $A_0 = A(z_0)$  is necessarily Fredholm of index 0, see Gohberg et al. (1990, Sect. XI.8). Combining this with (i) and (ii) in Definition 3.1 one thus has that  $0 < \dim \text{Ker} A_0 = \dim(\text{Im} A_0)^\perp < \infty$ . Note that when  $\mathcal{H}$  is finite-dimensional any operator is Fredholm of index 0 and any eigenvalue is of finite type. Moreover, if  $A(z)$  has an eigenvalue of finite type at  $z = z_0$ ,  $\text{Im}(A_0)$  is necessarily closed, see Theorem 2.1 in Gohberg, Goldberg, and Kaashoek (2003, Sect. 15.2). When the image of a bounded operator is closed, the generalized maximal Tseng inverse of  $A_0$ , written  $A_0^+$ , satisfies the Moore–Penrose conditions, see Theorem 3 in Ben-Israel and Greville (2003, Chap. 9). In the following, whenever the generalized maximal Tseng inverse is used, it always coincides with the Moore–Penrose inverse because it is applied to bounded operators with a closed image, as it is the case for  $A_0$ , and it is simply referred to as “generalized inverse”.

The key assumption is introduced next.

**Assumption 3.3** (AR process with a unit root of finite type).  $A(z)$  has an eigenvalue of finite type at  $z = 1$ . In this case,  $A(L)x_t = \varepsilon_t$  in (3.1) is said to be an AR process with a unit root of finite type.

That is, an AR process with a unit root of finite type is such that  $A(z)$  is invertible for all  $z \in D(0, \rho) \setminus \{1\}$  for some  $\rho > 1$ ,  $A_0 = A(1) \neq 0$ ,  $0 < \dim \text{Ker} A_0 = \dim(\text{Im} A_0)^\perp < \infty$ , and  $\text{Im} A_0$  is closed. Note that when  $\mathcal{H}$  is finite-dimensional an AR process with a unit root of finite type coincides with a cointegrated VAR.

### 3.2. Existence of a Common Trends Representation

Under Assumption 3.3, one can apply the Analytic Fredholm Theorem of Section XI.8 of Gohberg et al. (1990), reported in Theorem A.1 in Appendix A.1, and first employed in Beare and Seo (2019) in the context of infinite-dimensional  $\mathcal{H}$ -valued ARs with compact operators. These results guarantee that there exists a finite integer  $d = 1, 2, \dots$  and operators  $C_n$ ,  $n = 0, 1, \dots$ , with finite rank for  $n = 0, \dots, d - 1$ , such that

$$A(z)^{-1} = \sum_{n=0}^{\infty} C_n (1 - z)^{n-d}, \quad z \in D(0, \rho) \setminus \{1\}, \quad \rho > 1, \tag{3.3}$$

so that the inverse of  $A(z)$  has a pole of finite order  $d$  at  $z = 1$ . Note that (3.3) can be written as

$$A(z)^{-1} = \sum_{n=0}^{d-1} C_n(1-z)^{n-d} + C^o(z), \quad z \in D(0, \rho) \setminus \{1\}, \quad \rho > 1,$$

where  $C^o(z) = \sum_{n=d}^{\infty} C_n(1-z)^{n-d}$  is absolutely convergent on  $D(0, \rho)$  for some  $\rho > 1$ .

This implies that the solution of the AR equations is  $I(d)$  for some finite integer  $d$ . Moreover, because the operators that make up the principal part of  $A(z)^{-1}$  around  $z = 1$  have finite rank,  $x_t$  displays a common trends representation with a finite number of common stochastic trends of the type of (cumulated) bilateral random walks, as reported in Theorem 3.5 below.

To state Theorem 3.5, the cumulation operator  $\mathcal{S}$  is introduced, following Gregoir (1999).

**DEFINITION 3.4** (Cumulation operator  $\mathcal{S}$ ). *For an  $\mathcal{H}$ -valued generic process  $\{w_t, t \in \mathbb{Z}\}$  the cumulation operator  $\mathcal{S}$  is defined as*

$$\mathcal{S}w_t = 1_{(t \geq 1)} \cdot \sum_{i=1}^t w_i - 1_{(t \leq -1)} \cdot \sum_{i=t+1}^0 w_i. \tag{3.4}$$

When  $w_t = \varepsilon_t$  is white noise, the notation  $s_{h,t} = \mathcal{S}^h \varepsilon_t$ ,  $h = 1, 2, \dots$ , is employed, and  $s_{h,t}$  is called the  $(h - 1)$ -fold cumulated bilateral random walk.

Observe that by definition  $\mathcal{S}$  assigns value 0 to the cumulated process at time 0. In fact, applying the definition, see also Properties 2.1, 2.2 in Gregoir (1999), one has

$$\Delta \mathcal{S}w_t = w_t, \quad \mathcal{S} \Delta w_t = w_t - w_0, \quad t \in \mathbb{Z}. \tag{3.5}$$

Equation (3.5) shows that  $\mathcal{S}$  applied to  $\Delta w_t$  regenerates the level of the process  $w_t$ , up to a constant; this parallels the constant of integration in indefinite integrals. The cumulation operator  $\mathcal{S}$  is hence the inverse of the difference operator  $\Delta$  up a constant.

Note that when  $w_t = \varepsilon_t$  is white noise, (3.4) implies that  $s_{1,t} = \mathcal{S}\varepsilon_t$  is a bilateral  $\mathcal{H}$ -valued random walk, see Bosq (2000, Example 1.9 on page 20); because  $\Delta s_{1,t} = \Delta \mathcal{S}\varepsilon_t = \varepsilon_t$  is  $I(0)$ , this shows that  $s_{1,t}$  is  $I(1)$ . Similarly, for  $h = 2, 3, \dots$ ,  $s_{h,t} = \mathcal{S}s_{h-1,t}$  is the  $(h - 1)$ -fold cumulation of the bilateral random walk  $s_{1,t} \sim I(1)$  and hence it is  $I(h)$ .

The following result connects AR processes with a unit root of finite type with the existence of a common trend representation in terms of stochastic trends of the above type.

**THEOREM 3.5** (Existence of a common trends representation). *Let  $A(L)x_t = \varepsilon_t$  be an AR process with a unit root of finite type; then there exist a finite integer  $d = 1, 2, \dots$  and operators  $C_n$ ,  $n = 0, 1, \dots$ , with finite rank for  $n = 0, \dots, d - 1$ , such that  $x_t$  has common trends representation*

$$x_t = C_0 s_{d,t} + C_1 s_{d-1,t} + \dots + C_{d-1} s_{1,t} + y_t + \mu_t, \quad t \in \mathbb{Z}, \tag{3.6}$$

where  $s_{h,t} = S^h \varepsilon_t \sim I(h)$  is the  $(h - 1)$ -fold cumulation of the bilateral random walk  $s_{1,t} \sim I(1)$ ,  $y_t = C^\circ(L)\varepsilon_t$  is a linear process with  $C^\circ(z) = \sum_{n=d}^\infty C_n(1 - z)^{n-d}$ , and  $\mu_t = \sum_{n=0}^{d-1} v_n t^n$  is a polynomial in  $t$  whose coefficients  $v_0, \dots, v_{d-1} \in \mathcal{H}$  depend on the initial values of  $x_t, y_t, \varepsilon_t$  for  $t = -d, \dots, 0$ .

In the common trends representation (3.6), the operators  $C_0, C_1, \dots, C_{d-1}$  have finite rank; this implies that  $x_t$  depends only on a finite number of bilateral (cumulated) random walks. In fact, these common stochastic trends are selected from  $s_{h,t} \sim I(h)$ ,  $h = 1, \dots, d$ , by the finite rank operators  $C_0, C_1, \dots, C_{d-1}$  that load onto  $x_t$  only a finite number of characteristics from  $s_{h,t}$ ,  $h = 1, \dots, d$ .

Theorem 3.5 implies a number of properties for AR processes with a unit root of finite type, some of which are listed in the following corollary, namely, that  $d$  (the order of the pole of the inverse of  $A(z)$  at  $z = 1$ ) is finite, the process is cointegrated, the number of common trends is finite, and the dimension of the cointegrating space is infinite when  $\dim \mathcal{H} = \infty$ .

**COROLLARY 3.6** (Cointegration properties). *Let  $A(L)x_t = \varepsilon_t$  be an AR process with a unit root of finite type; then*

- (i)  $x_t \sim I(d)$  for some finite integer  $d = 1, 2, \dots$ ,
- (ii)  $x_t$  is cointegrated,
- (iii)  $\text{Im } C_0$  is the finite-dimensional attractor space,
- (iv)  $(\text{Im } C_0)^\perp$  is the cointegrating space, which is infinite-dimensional when  $\dim \mathcal{H} = \infty$ .

Despite these important implications of Theorem 3.5, these existence results do not address a number of important issues, such as the connection between the structure of  $A(z)$  and the order of integration  $d$  of the process. In fact, one cannot determine the order of integration  $d$  of the solution of the AR equations using Theorem 3.5. Moreover, Theorem 3.5 does not specify the connection between  $\text{Im } C_0$  and the AR operators, so that one does not know how the attractor space and the cointegrating space are related to the AR operators. Finally, as the relations among the finite rank operators  $C_0, C_1, \dots, C_{d-1}$  are not specified, Theorem 3.5 is mostly silent about the structure of the cointegrating relations.

These additional characterization results form the main contribution of the present article and are presented in Section 5 for the generic  $I(d)$  case. For ease of presentation, Section 4 starts from the cases of  $I(1)$  and  $I(2)$   $\mathcal{H}$ -valued AR processes.

### 3.3. Relations with the Literature

Before turning to the characterization results, the present subsection discusses the relationship between Assumption 3.3 and the assumptions employed in the literature. An example in the next section illustrates the differences.

The following proposition discusses the relation with Chang et al. (2016b), who study  $I(1)$  infinite-dimensional  $\mathcal{H}$ -valued processes  $x_t$  satisfying  $\Delta x_t = B(L)\varepsilon_t$ , where  $\sum_{n=1}^{\infty} n \|B_n\|_{\mathcal{L}\mathcal{H}} < \infty$  and  $\dim \text{Im } B(1) < \infty$ .

**PROPOSITION 3.7** ( $I(1)$  AR processes with a unit root of finite type). *Let  $A(L)x_t = \varepsilon_t$  be an AR process with a unit root of finite type with  $d = 1$ ; then  $\Delta x_t = B(L)\varepsilon_t$ , where  $B(z) = \sum_{n=0}^{\infty} B_n z^n$ ,  $z \in \mathbb{C}$ , is such that  $\sum_{n=1}^{\infty} n \|B_n\|_{\mathcal{L}\mathcal{H}} < \infty$  and  $\text{Im } B(1)$  is finite-dimensional. The converse does not necessarily hold.*

This shows that  $I(1)$  AR processes with a unit root of finite type necessarily satisfy Assumption 2.1 in Chang et al. (2016b); hence, their asymptotic analysis applies and their test can be employed in the present setup.

The next proposition discusses the relation with Hu and Park (2016), who consider (3.1) with  $\dim \mathcal{H} = \infty$ ,  $k = 1$  and compact  $A_1^\circ$ . Similarly, Beare et al. (2017) consider (3.1) with  $\dim \mathcal{H} = \infty$  and compact  $A_1^\circ, \dots, A_k^\circ$  if  $k > 1$  and Beare and Seo (2019) consider (3.1) with  $\dim \mathcal{H} = \infty$  and compact  $A_1^\circ, \dots, A_k^\circ$  for  $k \geq 1$ .

**PROPOSITION 3.8** (Compactness and AR processes with a unit root of finite type). *Assume that  $A_1^\circ, \dots, A_k^\circ$ ,  $k \geq 1$ , in (3.1) are compact; then (3.1) is an AR process with a unit root of finite type. The converse does not necessarily hold.*

This shows that the present results can be applied to the setups of Hu and Park (2016), Beare et al. (2017) for  $k > 1$  and Beare and Seo (2019). Beare et al. (2017) also consider  $x_t = A_1^\circ x_{t-1} + \varepsilon_t$  with no compactness assumption on  $A_1^\circ$ . This case is not covered by the present results; see also Proposition 4.7 below.

Finally, Chang et al. (2016a) consider an error correction form with compact error correction operator and show that in this case the number of  $I(1)$  common trends is infinite and the dimension of the cointegrating space is finite. This case is not covered by the present results.

### 3.4. Example of a Noncompact Operator

This section illustrates the relevance of Assumption 3.3 with an example, which is considered again in Sections 4.3 and 4.4 to illustrate the characterization results in the  $I(1)$  and in the  $I(2)$  cases.

Consider a real infinite-dimensional  $\mathcal{H}$ -valued AR(1) process  $x_t = A_1^\circ x_{t-1} + \varepsilon_t$ , where  $A_1^\circ$  is a band operator. Band operators are defined as follows: let  $\varphi_1, \varphi_2, \dots$  be an orthonormal basis of  $\mathcal{H}$  and let  $(a_{ij})$ , where  $a_{ij} = \langle A\varphi_j, \varphi_i \rangle$ , be the matrix representation of  $A \in \mathcal{L}\mathcal{H}$  corresponding to  $\varphi_1, \varphi_2, \dots$ , see, e.g., Gohberg et al. (2003, Sect. 2.4);  $A \in \mathcal{L}\mathcal{H}$  is called a band operator if all nonzero entries in its matrix representation  $(a_{ij})$  are in a finite number of diagonals parallel to the main diagonal, i.e., there exists an integer  $N$  such that  $a_{ij} = 0$  if  $|i - j| > N$ , see, e.g., Gohberg et al. (2003, Sect. 2.16).

Note that a band operator is compact if and only if  $\lim_{i,j \rightarrow \infty} a_{ij} = 0$ , see Theorem 16.4 in Gohberg et al. (2003, Sect. 2.16). Finally, let  $w_{i,t} = \langle \varphi_i, w_t \rangle$  be the  $i$ -th coordinate of the process  $w_t$ , where  $w_t = x_t, \varepsilon_t$ . Note also that  $A(z) = A_0 + A_1(1 - z)$  with  $A_0 = I - A_1^\circ$ ,  $A_1 = A_1^\circ$  and  $A_n = 0$  for  $n = 2, 3, \dots$ , in  $A(z) = \sum_{n=0}^\infty A_n(1 - z)^n$ ; see (4.1) below.

Consider the matrix representation  $(a_{ij})$  of  $A_1^\circ$  and assume that  $a_{ij} = 0$  for  $|i - j| > 0$  and  $a_{ii} = \alpha_i$ , where  $\alpha_i \in \mathbb{R}$ ,  $\alpha_1 = 1$  and  $0 < |\alpha_i| < 1$ ,  $i = 2, 3, \dots$ , without  $\alpha_i \rightarrow 0$  as  $i \rightarrow \infty$ . Note that  $A_1^\circ$  is a band operator. Observe that  $x_t = A_1^\circ x_{t-1} + \varepsilon_t$  reads

$$x_{1,t} = x_{1,t-1} + \varepsilon_{1,t}, \quad x_{i,t} = \alpha_i x_{i,t-1} + \varepsilon_{i,t}, \quad i = 2, 3, \dots$$

Note that  $A_1^\circ$  is not compact because  $\alpha_i$  does not tend to 0 as  $i \rightarrow \infty$ . Next note that  $A(z)$  is invertible for all  $z \in D(0, \rho) \setminus \{1\}$  for some  $\rho > 1$  and consider the matrix representation of  $A_1^\circ = A_1$  and  $A_0 = I - A_1^\circ$ , i.e.,

$$A_1^\circ = A_1 = \begin{pmatrix} 1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \ddots \end{pmatrix}, \quad A_0 = \begin{pmatrix} 0 & & & \\ & 1 - \alpha_2 & & \\ & & \ddots & \\ & & & \ddots \end{pmatrix}, \tag{3.7}$$

where unspecified entries are equal to 0, and compute

$$(\text{Im } A_0)^\perp = (\overline{\text{sp}}\{\varphi_2, \varphi_3, \dots\})^\perp = \text{sp}\{\varphi_1\}, \quad \text{Ker } A_0 = \text{sp}\{\varphi_1\},$$

where  $\text{sp}\{\cdot\}$  and  $\overline{\text{sp}}\{\cdot\}$  indicate the span of the set of vectors in curly brackets and its closure respectively. Because  $0 < \dim \text{Ker } A_0 = \dim(\text{Im } A_0)^\perp < \infty$ , this shows that  $A_0$  is Fredholm of index 0, so that Assumption 3.3 holds and  $x_t = A_1^\circ x_{t-1} + \varepsilon_t$  is an AR process with a unit root of finite type with noncompact operator.

#### 4. A CHARACTERIZATION OF $I(1)$ AND $I(2)$ AR PROCESSES WITH A UNIT ROOT OF FINITE TYPE

This section presents a characterization of  $I(1)$  and  $I(2)$  AR processes with a unit root of finite type. The  $I(1)$  case parallels the results in Hu and Park (2016), Beare et al. (2017) for  $k > 1$ , Beare and Seo (2019), and it is discussed in Theorem 4.1. The results for the  $I(2)$  case are novel, and they are given in Theorem 4.8.

To state the characterization results, it is useful to expand the operator function  $A(z) = I - \sum_{h=1}^k A_h^\circ z^h$  around 1, obtaining

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

where empty sums are defined to be 0; note that  $A_n = 0$  for  $n > k$ , see (3.1).

4.1. *I*(1) Case

Let  $A_n$  be as in (4.1) and define

$$S_0 = A_0, \quad \zeta_0 = \text{Im } S_0, \quad \tau_0 = (\text{Ker } S_0)^\perp, \quad (4.2)$$

$$S_1 = P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp}, \quad \zeta_1 = \text{Im } S_1, \quad \tau_1 = (\text{Ker } S_1)^\perp, \quad (4.3)$$

where  $P_\eta$  is the orthogonal projection on  $\eta$ , i.e.,  $P_\eta^2 = P_\eta$ ,  $\text{Im } P_\eta = \eta$  and  $\text{Ker } P_\eta = \eta^\perp$ .

Observe that

$$\zeta_1 \subseteq \zeta_0^\perp, \quad \tau_1 \subseteq \tau_0^\perp$$

by construction; that is,  $\zeta_1$  is orthogonal to  $\zeta_0$  and  $\tau_1$  is orthogonal to  $\tau_0$ . Moreover, because 1 is an eigenvalue of finite type, one has  $0 < \dim \tau_0^\perp = \dim \zeta_0^\perp < \infty$ , see Remark 3.2, so that the subspaces  $\zeta_1, \tau_1$  are finite-dimensional and the subspaces  $\zeta_0, \tau_0$  are infinite-dimensional when  $\dim \mathcal{H} = \infty$ . Note that  $S_0$  and  $S_1$  both have closed images ( $\zeta_0$  and  $\zeta_1$ ), thanks to Remark 3.2 and the fact that  $\zeta_1$  is finite-dimensional. This implies that the generalised inverses of  $S_0$  and  $S_1$  exist and satisfy the usual Moore–Penrose equations.

In the following, the orthogonal direct sum decomposition

$$\mathcal{H} = \tau_0 \oplus \tau_1, \quad \tau_1 \neq \{0\}, \quad (4.4)$$

is called the POLE(1) condition. Note that  $\{0\}$  in (4.4) indicates the subspace of  $\mathcal{H}$  of dimension 0, simply containing the 0 element; in other words condition (4.4) requires  $\tau_1$  to complement  $\tau_0$  and to have positive dimension. In the following,  $a \Rightarrow b$  indicates that  $a$  implies  $b$ .

**THEOREM 4.1** (A characterization of *I*(1) AR processes with a unit root of finite type). *Consider an AR process with a unit root of finite type  $A(L)x_t = \varepsilon_t$  and let  $\tau_0, \tau_1$  be as in (4.2), (4.3) respectively; then  $x_t$  is *I*(1) if and only if the POLE(1) condition in (4.4) holds. In this case, the common trends representation of  $x_t$  is found by setting  $d = 1$  in (3.6), i.e.,  $x_t = C_0 s_{1,t} + y_t + \mu_t$ . Moreover,  $\text{Im } C_0 = \tau_1$  is the finite-dimensional attractor space,  $\tau_0$  is the cointegrating space, which is infinite-dimensional when  $\dim \mathcal{H} = \infty$ , and for any nonzero  $v \in \mathcal{H}$  one has*

$$v \in \tau_0 \quad \Rightarrow \quad \langle v, x_t \rangle \sim I(0), \quad (4.5)$$

$$v \in \tau_1 \quad \Rightarrow \quad \langle v, x_t \rangle \sim I(1), \quad (4.6)$$

where  $\tau_1 = \tau_0^\perp \neq \{0\}$ .

Some remarks on Theorem 4.1 are in order.

**Remark 4.2.** An AR process with a unit root of finite type generates an *I*(1) process if and only if  $\tau_1 = \tau_0^\perp \neq \{0\}$ , i.e.,  $\dim \tau_1 > 0$ . In fact, the common trends representation  $x_t = C_0 s_{1,t} + y_t + \mu_t$  implies that

$$\langle v, x_t \rangle = \langle v, C_0 s_{1,t} \rangle + \langle v, y_t \rangle + \langle v, \mu_t \rangle.$$



Because  $\text{Im } C_0 = \tau_1 = \tau_0^\perp$ , for any nonzero  $v \in \tau_0$  one has  $\langle v, C_0 u \rangle = 0$  for all  $u \in \mathcal{H}$ , and hence also when  $u$  equals  $s_{1,t}$ , which shows that  $\langle v, x_t \rangle$  is stationary. Hence  $\tau_0$  is the cointegrating space, which is infinite-dimensional when  $\dim \mathcal{H} = \infty$ , and its orthogonal complement  $\tau_1 = \tau_0^\perp$  is the finite-dimensional attractor space. Theorem 4.1 further shows that  $\langle v, x_t \rangle$  is not only stationary, but  $I(0)$  for any nonzero  $v \in \tau_0$ .

**Remark 4.3.** The orthogonal decomposition  $\mathcal{H} = \tau_0 \oplus \tau_1$  can be employed to characterize the order of integration of any  $v$ -characteristic of the process. In fact, (4.4) implies  $P_{\tau_0} + P_{\tau_1} = I$ , where  $P_{\tau_h}$  is the orthogonal projection onto  $\tau_h$ ; hence, for any  $v \in \mathcal{H}$  one has

$$\langle v, x_t \rangle = \langle v_0, x_t \rangle + \langle v_1, x_t \rangle, \quad v_h = P_{\tau_h} v \in \tau_h, \quad h = 0, 1,$$

where  $\langle v_0, x_t \rangle \sim I(0)$  by (4.5) and  $\langle v_1, x_t \rangle \sim I(1)$  by (4.6). Here, the  $I(1)$  component dominates, and one has  $\langle v, x_t \rangle \sim I(1)$  if and only if  $v_1 \neq 0$ .

**Remark 4.4.** Franchi and Paruolo (2016) study the finite-dimensional case  $\mathcal{H} = \mathbb{R}^p$  (or  $\mathcal{H} = \mathbb{C}^p$ ); they show that the  $I(1)$  condition in Theorem 4.2 in Johansen (1996) can be equivalently stated as  $\mathbb{R}^p = \zeta_0 \oplus \zeta_1 = \tau_0 \oplus \tau_1$ ,  $\zeta_1 \neq \{0\}$ , and  $\tau_1 \neq \{0\}$ , where  $\zeta_h = \text{sp}\{\alpha_h\}$ ,  $\tau_h = \text{sp}\{\beta_h\}$ ,  $h = 0, 1$ , and the bases  $\alpha_h, \beta_h$  are defined by the rank factorizations  $A_0 = \alpha_0 \beta_0'$  and  $P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} = \alpha_1 \beta_1'$ ; here  $\text{sp}\{a\}$  indicates  $\text{sp}\{a_1, \dots, a_k\}$  when the argument  $a$  of  $\text{sp}\{a\}$  is a matrix with  $k$  columns  $a_i$ ,  $a = (a_1, \dots, a_k)$ . In this case, observe that  $\alpha_h, \beta_h$  are full-column-rank matrices that respectively span the column space  $\zeta_h$  and the row space  $\tau_h$  of the corresponding matrix. Except for the fact that  $\dim \zeta_0 = \dim \tau_0$  is finite when  $\mathcal{H} = \mathbb{R}^p$ , this mirrors what happens when  $\dim \mathcal{H} = \infty$ .

**Remark 4.5.** The POLE(1) condition in (4.4) is equivalent to  $\tau_1 = \tau_0^\perp \neq \{0\}$ . Moreover, Theorem B.4 in Appendix B shows that it can be equivalently stated as (i)  $\mathcal{H} = \zeta_0 \oplus \zeta_1$ ,  $\zeta_1 \neq \{0\}$ , (ii)  $\zeta_1 = \zeta_0^\perp \neq \{0\}$ , (iii)  $\text{Im } C_0 = \tau_1$ , (iv)  $\text{Ker } C_0 = \zeta_0$ .

The POLE(1) condition is next compared to equivalent conditions in the literature. Beare et al. (2017, Definition 4.3) say that  $A(z)$  satisfies the ‘Johansen  $I(1)$  condition’ if  $\text{Im } A_0$  and  $A_1 \text{Ker } A_0$  are closed and

$$\mathcal{H} = \text{Im } A_0 \oplus A_1 \text{Ker } A_0, \tag{4.7}$$

where  $A_0, A_1$  are as in (4.1). The condition (4.7) is a nonorthogonal direct sum decomposition. The next proposition shows the equivalence of the orthogonal direct sum (POLE(1)) condition in (4.4) and the nonorthogonal direct sum (‘Johansen  $I(1)$ ’) condition in (4.7).

**PROPOSITION 4.6** (Equivalence of (4.4) and (4.7)). *Let  $A(L)x_t = \varepsilon_t$  be an AR process with a unit root of finite type; then the Johansen  $I(1)$  condition in (4.7) is equivalent to the POLE(1) condition in (4.4).*

Beare et al. (2017, Theorem 4.1) has further connections with Theorem 4.1 of the present article in case  $k = 1$ , when (4.7) reduces to

$$\mathcal{H} = \text{Im } A_0 \oplus \text{Ker } A_0 \tag{4.8}$$

Their Theorem 4.1 places no compactness assumption on  $A_1^\circ$  when  $k = 1$  and they state that the ‘‘Johansen  $I(1)$  condition’’ (4.8) is sufficient to imply the common trends representation (3.6) with  $d = 1$  and  $\text{Im } C_0 = \text{Ker } A_0$ . The following proposition clarifies the connection between AR processes with a unit root of finite type and their result for  $k = 1$ , showing that if  $\text{Ker } A_0$  is finite-dimensional, then the ‘‘Johansen  $I(1)$  condition’’ (4.8) is a necessary and sufficient condition for the common trends representation (3.6).

**PROPOSITION 4.7** (AR(1) case with finite-dimensional  $\text{Ker } A_0$ ). *Consider the infinite-dimensional  $\mathcal{H}$ -valued AR(I) process  $x_t = A_1^\circ x_{t-1} + \varepsilon_t$  and let  $\mathcal{H} = \text{Im } A_0 \oplus \text{Ker } A_0$ . If  $\text{Ker } A_0$  is finite-dimensional then  $x_t = A_1^\circ x_{t-1} + \varepsilon_t$  is an AR process with a unit root of finite type, and the ‘‘Johansen  $I(1)$  condition’’ (4.8) (or equivalently the POLE(1) condition in (4.4)) is necessary and sufficient for  $x_t \sim I(1)$ .*

One can observe that the case with  $k = 1$  and infinite-dimensional  $\text{Ker } A_0$ , which corresponds to an infinite-dimensional attractor space, is treated in Beare et al. (2017, Thm. 4.1) but it is not covered by the present results.

**4.2.  $I(2)$  Case**

The  $I(2)$  case is considered next. Let  $A_n, \zeta_0, \tau_0$  and  $\zeta_1, \tau_1$  be as in (4.1), (4.2), and (4.3) respectively and define

$$S_2 = P_{\mathcal{Z}_2^\perp} A_{2,1} P_{\mathcal{T}_2^\perp}, \quad \zeta_2 = \text{Im } S_2, \quad \tau_2 = (\text{Ker } S_2)^\perp, \tag{4.9}$$

where  $\mathcal{Z}_2 = \zeta_0 \oplus \zeta_1, \mathcal{T}_2 = \tau_0 \oplus \tau_1$ , and  $A_{2,1} = A_2 - A_1 A_0^+ A_1$ . Recall that the generalized inverse  $A_0^+$  exists and it is unique, see Remark 3.2.

Observe that

$$\zeta_2 \subseteq (\zeta_0 \oplus \zeta_1)^\perp, \quad \tau_2 \subseteq (\tau_0 \oplus \tau_1)^\perp$$

by construction; that is, for  $0 \leq j < h, \zeta_h$  is orthogonal to  $\zeta_j$ , and  $\tau_h$  is orthogonal to  $\tau_j$ . Moreover, because  $0 < \dim \zeta_0^\perp = \dim \tau_0^\perp < \infty$ , the subspaces  $\zeta_1, \zeta_2, \tau_1, \tau_2$  are finite-dimensional and the subspaces  $\zeta_0, \tau_0$  are infinite-dimensional when  $\dim \mathcal{H} = \infty$ . Again note that  $S_0, S_1$ , and  $S_2$  have closed images ( $\zeta_0, \zeta_1$ , and  $\zeta_2$ ), thanks to Remark 3.2 and the fact that  $\zeta_1, \zeta_2$  are finite-dimensional. This implies that the generalised inverses of  $S_0, S_1$ , and  $S_2$  exist and satisfy the usual Moore–Penrose equations.

In the following, the orthogonal direct sum decomposition

$$\mathcal{H} = \tau_0 \oplus \tau_1 \oplus \tau_2, \quad \tau_2 \neq \{0\}, \tag{4.10}$$

is called the POLE(2) condition.

**THEOREM 4.8** (A characterization of  $I(2)$  AR processes with a unit root of finite type). *Consider an AR process with a unit root of finite type  $A(L)x_t = \varepsilon_t$ ; let  $\tau_0, \tau_1, \tau_2$  be as in (4.2), (4.3), (4.9) respectively and let  $A_0^+$  be the generalized inverse of  $A_0$ ; then  $x_t$  is  $I(2)$  if and only if the POLE(2) condition in (4.10) holds. In this case, the common trends representation of  $x_t$  is found by setting  $d = 2$  in (3.6), i.e.,  $x_t = C_0s_{2,t} + C_1s_{1,t} + y_t + \mu_t$ . Moreover,  $\text{Im } C_0 = \tau_2$  is the finite-dimensional attractor space,  $\tau_0 \oplus \tau_1$  is the cointegrating space, which is infinite-dimensional when  $\dim \mathcal{H} = \infty$ , and for any nonzero  $v \in \mathcal{H}$  one has*

$$v \in \tau_0 \quad \Rightarrow \quad \langle v, x_t \rangle + \langle v, A_0^+ A_1 \Delta x_t \rangle \sim I(0), \tag{4.11}$$

$$v \in \tau_1 \quad \Rightarrow \quad \langle v, x_t \rangle \sim I(1), \tag{4.12}$$

$$v \in \tau_2 \quad \Rightarrow \quad \langle v, x_t \rangle \sim I(2), \tag{4.13}$$

where  $\tau_1 \subset \tau_0^\perp$  and  $\tau_2 = (\tau_0 \oplus \tau_1)^\perp \neq \{0\}$ .

Some remarks on Theorem 4.8 are in order.

**Remark 4.9.** An AR process with a unit root of finite type generates an  $I(2)$  process if and only if  $\tau_2 = (\tau_0 \oplus \tau_1)^\perp \neq \{0\}$ . The common trends representation of  $x_t$  shows that the  $I(2)$  stochastic trends  $s_{2,t}$  are loaded into the process by  $C_0$ ; because  $\text{Im } C_0 = \tau_2 = (\tau_0 \oplus \tau_1)^\perp$ , for any nonzero  $v \in \tau_0 \oplus \tau_1$  one has  $\langle v, C_0u \rangle = 0$  for all  $u \in \mathcal{H}$ , and hence also for  $u = s_{2,t}$ ; this implies that  $\langle v, x_t \rangle$  is at most  $I(1)$ , i.e.,  $\tau_0 \oplus \tau_1$  is the cointegrating space and its orthogonal complement  $\tau_2 = (\tau_0 \oplus \tau_1)^\perp$  is the finite-dimensional attractor space. Note that when  $\dim \mathcal{H} = \infty$  one has  $\dim \tau_0 = \infty$  and hence the cointegrating space is infinite-dimensional.

**Remark 4.10.** Theorem 4.8 further shows that the cointegrating space is partitioned into  $\tau_0 \oplus \tau_1$ ; in  $\tau_0$ , which is infinite-dimensional when  $\dim \mathcal{H} = \infty$ , one finds the  $v$ -characteristics that allow for  $I(0)$  polynomial cointegration and in  $\tau_1$ , with  $0 \leq \dim \tau_1 < \infty$ , those that do not allow for polynomial cointegration. Specifically, any nonzero  $v_0 \in \tau_0$ , if one combines levels and first differences as in  $\langle v_0, x_t \rangle + \langle v_0, A_0^+ A_1 \Delta x_t \rangle$ , one finds an  $I(0)$  process. The contribution of  $\langle v_0, A_0^+ A_1 \Delta x_t \rangle$  may be irrelevant; there are situations, in fact, depending on the specific form of  $v_0, A_0, A_1$  where  $\langle v_0, A_0^+ A_1 u \rangle$  is equal to 0 for any  $u$ ; in this case one would have  $\langle v_0, x_t \rangle \sim I(0)$ . On the other hand, polynomial cointegration cannot happen in the  $\tau_1$  subspace, in which every nonzero  $v_1 \in \tau_1$  is such that  $\langle v_1, x_t \rangle \sim I(1)$ . Apart from the fact that the dimension of  $\tau_0$  is infinite when  $\dim \mathcal{H} = \infty$ , this parallels the finite-dimensional case, see Theorem 4.6 in Johansen (1996).

**Remark 4.11.** The orthogonal direct sum decomposition  $\mathcal{H} = \tau_0 \oplus \tau_1 \oplus \tau_2$  can be employed to characterize the order of integration of any  $v$ -characteristic of the process. In fact, (4.10) implies  $P_{\tau_0} + P_{\tau_1} + P_{\tau_2} = I$ , where  $P_{\tau_h}$  is the orthogonal projection onto  $\tau_h$ ; hence for any  $v \in \mathcal{H}$  one has  $\langle v, x_t \rangle = \langle v_0, x_t \rangle + \langle v_1, x_t \rangle +$

$\langle v_2, x_t \rangle$ , where  $v_h = P_{\tau_h} v \in \tau_h$ , and  $\langle v_h, x_t \rangle \sim I(h)$  by (4.12) and (4.13) for  $h = 1, 2$ . Taking these results together with (4.11), one also concludes that  $\langle v_0, x_t \rangle$  is either  $I(0)$  or  $I(1)$ . Here, the  $I(2)$  trend dominates, and hence one has  $\langle v, x_t \rangle \sim I(2)$  if and only if  $v_2 \neq 0$ .

**Remark 4.12.** In the finite-dimensional case  $\mathcal{H} = \mathbb{R}^p$ , Franchi and Paruolo (2016) show that the  $I(2)$  condition in Theorem 4.6 in Johansen (1996) can be equivalently stated as  $\mathbb{R}^p = \zeta_0 \oplus \zeta_1 \oplus \zeta_2 = \tau_0 \oplus \tau_1 \oplus \tau_2$ ,  $\zeta_2 \neq \{0\}$ , and  $\tau_2 \neq \{0\}$ , where  $\zeta_h = \text{sp}\{\alpha_h\}$ ,  $\tau_h = \text{sp}\{\beta_h\}$ ,  $h = 0, 1, 2$ , and the bases  $\alpha_h, \beta_h$  are defined by the rank factorizations  $A_0 = \alpha_0 \beta_0'$ ,  $P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} = \alpha_1 \beta_1'$ , and  $P_{\mathcal{Z}_2^\perp} A_{2,1} P_{\mathcal{T}_2^\perp} = \alpha_2 \beta_2'$  where  $A_{2,1} = A_2 - A_1 A_0^+ A_1$ ,  $A_0^+ = (\alpha_0 \beta_0')^+ = \bar{\beta}_0 \bar{\alpha}_0'$ , and  $\bar{\eta} = \eta(\eta' \eta)^{-1}$  for a generic full-column-rank matrix  $\eta$ . This shows that the infinite-dimensional case parallels the finite-dimensional one, apart from the fact that  $\dim \zeta_0 = \dim \tau_0$  is finite when  $\mathcal{H} = \mathbb{R}^p$  and infinite when  $\dim \mathcal{H} = \infty$ .

**Remark 4.13.** The POLE(2) condition in (4.10) is equivalent to  $\tau_2 = (\tau_0 \oplus \tau_1)^\perp \neq \{0\}$ . Moreover, Theorem B.4 in Appendix B shows that it can be equivalently stated as (i)  $\mathcal{H} = \zeta_0 \oplus \zeta_1 \oplus \zeta_2$ ,  $\zeta_2 \neq \{0\}$ , (ii)  $\zeta_2 = (\zeta_0 \oplus \zeta_1)^\perp \neq \{0\}$ , (iii)  $\text{Im } C_0 = \tau_2$ , (iv)  $\text{Ker } C_0 = \zeta_0 \oplus \zeta_1$ .

**4.3. Illustration:  $I(1)$  Example**

Consider the setup in Section 3.4. Here, the analysis should deliver that  $x_t$  is  $I(1)$ , the attractor space coincides with  $\text{sp}\{\varphi_1\}$  and the cointegrating space with  $\overline{\text{sp}}\{\varphi_2, \varphi_3, \dots\}$ . Since  $\langle v, x_t \rangle$  is  $I(0)$  for any nonzero  $v \in \overline{\text{sp}}\{\varphi_2, \varphi_3, \dots\}$  and  $\langle v, x_t \rangle$  is  $I(1)$  for any nonzero  $v \in \text{sp}\{\varphi_1\}$ , the analysis should further convey that  $\tau_0 = \overline{\text{sp}}\{\varphi_2, \varphi_3, \dots\}$  and  $\tau_1 = \text{sp}\{\varphi_1\}$ .

From (3.7), one has

$$\begin{aligned} \zeta_0 &= \text{Im } A_0 = \overline{\text{sp}}\{\varphi_2, \varphi_3, \dots\}, \\ \tau_0 &= (\text{Ker } A_0)^\perp = (\text{sp}\{\varphi_1\})^\perp = \overline{\text{sp}}\{\varphi_2, \varphi_3, \dots\}, \\ \zeta_1 &= \text{Im } P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} = \text{sp}\{\varphi_1\}, \\ \tau_1 &= (\text{Ker } P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp})^\perp = (\overline{\text{sp}}\{\varphi_2, \varphi_3, \dots\})^\perp = \text{sp}\{\varphi_1\}. \end{aligned}$$

This shows that  $\mathcal{H} = \tau_0 \oplus \tau_1$ ,  $\tau_1 \neq \{0\}$ , so that the POLE(1) condition in (4.4) holds and Theorem 4.1 applies: the common trends representation of  $x_t$  is found by setting  $d = 1$  in (3.6),  $\text{Im } C_0 = \tau_1 = \text{sp}\{\varphi_1\}$  is the finite-dimensional attractor space and  $\tau_0 = \overline{\text{sp}}\{\varphi_2, \varphi_3, \dots\}$  is the infinite-dimensional cointegrating space.

**4.4. Illustration:  $I(2)$  Example**

Consider gain the setup in Section 3.4, with the following modifications. In the matrix representation  $(a_{ij})$  of  $A_1^\circ$ , assume that  $a_{ij} = 0$  for  $|i - j| > 1$ ,  $a_{12} = 1$  and  $a_{ii} = \alpha_i$ , where  $\alpha_i \in \mathbb{R}$ ,  $\alpha_1 = \alpha_2 = \alpha_3 = 1$  and  $0 < |\alpha_i| < 1$ ,  $i = 4, 5, \dots$  without

$\alpha_i \rightarrow 0$  as  $i \rightarrow \infty$ . Here,  $A_1^\circ$  is not compact but  $x_t = A_1^\circ x_{t-1} + \varepsilon_t$  is an AR process with a unit root of finite type. Observe that  $x_t = A_1^\circ x_{t-1} + \varepsilon_t$  reads

$$\begin{aligned} x_{1,t} &= x_{1,t-1} + x_{2,t-1} + \varepsilon_{1,t}, & x_{2,t} &= x_{2,t-1} + \varepsilon_{2,t}, & x_{3,t} &= x_{3,t-1} + \varepsilon_{3,t}, \\ x_{i,t} &= \alpha_i x_{i,t-1} + \varepsilon_{i,t}, & & & & i = 4, 5, \dots \end{aligned}$$

Hence, the analysis should deliver that  $x_t$  is  $I(2)$ , the attractor space coincides with  $\text{sp}\{\varphi_1\}$  and the cointegrating space with  $\overline{\text{sp}}\{\varphi_2, \varphi_3, \dots\}$ . Next, note that  $\langle v, x_t \rangle$  is  $I(0)$  for any nonzero  $v \in \overline{\text{sp}}\{\varphi_4, \varphi_5, \dots\}$  and  $\langle v, x_t \rangle$  is  $I(1)$  for any nonzero  $v \in \text{sp}\{\varphi_2, \varphi_3\}$ . Moreover, because  $\Delta x_{1,t} = x_{2,t-1} + \varepsilon_{1,t} = x_{2,t} - \varepsilon_{2,t} + \varepsilon_{1,t}$ , one has that  $x_{2,t} - \Delta x_{1,t}$  is  $I(0)$ , i.e.,  $\langle \varphi_2, x_t \rangle - \langle \varphi_1, \Delta x_t \rangle$  is  $I(0)$ , so that  $\langle \varphi_2, x_t \rangle$  allows for polynomial cointegration while  $\langle \varphi_3, x_t \rangle$  does not. Hence, the analysis should further convey that  $\tau_0 = \overline{\text{sp}}\{\varphi_2, \varphi_4, \varphi_5, \dots\}$ ,  $\tau_1 = \text{sp}\{\varphi_3\}$ ,  $\tau_2 = \text{sp}\{\varphi_1\}$ ,  $\langle \varphi_2, A_0^+ A_1 \Delta x_{i,t} \rangle = -\Delta x_{i,t}$ , and  $\langle \varphi_i, A_0^+ A_1 \Delta x_{i,t} \rangle = 0$  for  $i = 4, 5, \dots$ , as shown below.

Consider the matrix representation of  $A_1^\circ = A_1$  and  $A_0 = I - A_1^\circ$ , i.e.,

$$A_1^\circ = A_1 = \begin{pmatrix} 1 & 1 & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \alpha_4 & \\ & & & & \ddots \end{pmatrix}, \quad A_0 = \begin{pmatrix} 0 & -1 & & & \\ & 0 & & & \\ & & 0 & & \\ & & & 1 - \alpha_4 & \\ & & & & \ddots \end{pmatrix},$$

where empty entries are equal to 0. Compute

$$\begin{aligned} \zeta_0 &= \text{Im } A_0 = \overline{\text{sp}}\{\varphi_1, \varphi_4, \varphi_5, \dots\}, \\ \tau_0 &= (\text{Ker } A_0)^\perp = (\text{sp}\{\varphi_1, \varphi_3\})^\perp = \overline{\text{sp}}\{\varphi_2, \varphi_4, \varphi_5, \dots\}, \end{aligned}$$

so that  $\zeta_0^\perp = \text{sp}\{\varphi_2, \varphi_3\}$  and  $\tau_0^\perp = \text{sp}\{\varphi_1, \varphi_3\}$ ; because  $0 < \dim \text{Ker } A_0 = \dim(\text{Im } A_0)^\perp < \infty$ , this shows that  $A_0$  is Fredholm of index 0 and because  $A(z) = I - A_1^\circ z$  is invertible for all  $z \in D(0, \rho) \setminus \{1\}$  for some  $\rho > 1$ , and  $x_t = A_1^\circ x_{t-1} + \varepsilon_t$  is an AR process with a unit root of finite type with noncompact operator.

Next compute

$$\begin{aligned} \zeta_1 &= \text{Im } P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp} = \text{sp}\{\varphi_3\}, \\ \tau_1 &= (\text{Ker } P_{\zeta_0^\perp} A_1 P_{\tau_0^\perp})^\perp = (\overline{\text{sp}}\{\varphi_1, \varphi_2, \varphi_4, \varphi_5, \dots\})^\perp = \text{sp}\{\varphi_3\}. \end{aligned}$$

This shows that  $\tau_1$  is strictly contained in  $\tau_0^\perp$ , so that the POLE(1) condition in (4.4) does not hold and the process is  $I(d)$  for some finite  $d = 2, 3, \dots$

Now consider  $P_{\mathcal{Z}_2^\perp} A_{2,1} P_{\mathcal{T}_2^\perp}$  in (4.9); since  $\mathcal{Z}_2 = \zeta_0 \oplus \zeta_1 = \overline{\text{sp}}\{\varphi_1, \varphi_3, \varphi_4, \dots\}$  and  $\mathcal{T}_2 = \tau_0 \oplus \tau_1 = \overline{\text{sp}}\{\varphi_2, \varphi_3, \dots\}$ , one has  $\mathcal{Z}_2^\perp = \text{sp}\{\varphi_2\}$  and  $\mathcal{T}_2^\perp = \text{sp}\{\varphi_1\}$ . Note that  $A_2 = 0$  and hence  $A_{2,1} = -A_1 A_0^+ A_1$ ; thus,  $P_{\mathcal{Z}_2^\perp} A_{2,1} P_{\mathcal{T}_2^\perp} = -P_{\text{sp}\{\varphi_2\}} A_1 A_0^+ A_1 P_{\text{sp}\{\varphi_1\}}$  and because  $P_{\text{sp}\{\varphi_2\}} A_1 = P_{\text{sp}\{\varphi_2\}}$  and  $A_1 P_{\text{sp}\{\varphi_1\}} = P_{\text{sp}\{\varphi_1\}}$ , one has  $P_{\mathcal{Z}_2^\perp} A_{2,1} P_{\mathcal{T}_2^\perp} = -P_{\text{sp}\{\varphi_2\}} A_0^+ P_{\text{sp}\{\varphi_1\}}$ . Next, the matrix representation of  $A_0^+$  is investigated; from Lemma B.1 one has  $\text{Ker } A_0^+ = (\text{Im } A_0)^\perp$ ,

$A_0^+ A_0 = P_{(\text{Ker } A_0)^\perp}$  and because  $(\text{Im } A_0)^\perp = \text{sp}\{\varphi_2, \varphi_3\}$  and  $(\text{Ker } A_0)^\perp = \overline{\text{sp}}\{\varphi_2, \varphi_4, \varphi_5, \dots\}$  one has

$$A_0^+ = \begin{pmatrix} 0 & & & & & \\ -1 & 0 & & & & \\ & 0 & & & & \\ & & \frac{1}{1-\alpha_4} & & & \\ & & & \ddots & & \end{pmatrix}.$$

This implies that the only nonzero element in the matrix representation of  $P_{\mathcal{X}_2^\perp} A_{2,1} P_{\mathcal{X}_2^\perp} = -P_{\text{sp}\{\varphi_2\}} A_0^+ P_{\text{sp}\{\varphi_1\}}$  is a one in row 2 and column 1, so that

$$\begin{aligned} \zeta_2 &= \text{Im } P_{\mathcal{X}_2^\perp} A_{2,1} P_{\mathcal{X}_2^\perp} = \text{sp}\{\varphi_2\}, \\ \tau_2 &= (\text{Ker } P_{\mathcal{X}_2^\perp} A_{2,1} P_{\mathcal{X}_2^\perp})^\perp = (\overline{\text{sp}}\{\varphi_2, \varphi_3, \dots\})^\perp = \text{sp}\{\varphi_1\}. \end{aligned}$$

Hence,  $\mathcal{H} = \tau_0 \oplus \tau_1 \oplus \tau_2$ ,  $\tau_2 \neq \{0\}$ , i.e., the POLE(2) condition in (4.10) holds and Theorem 4.8 applies: the common trends representation of  $x_t$  is found by setting  $d = 2$  in (3.6),  $\text{Im } C_0 = \tau_2 = \text{sp}\{\varphi_1\}$  is the finite-dimensional attractor space and  $\tau_0 \oplus \tau_1 = \overline{\text{sp}}\{\varphi_2, \varphi_3, \dots\}$  is the infinite-dimensional cointegrating space. Moreover, for any nonzero  $v \in \mathcal{H}$  one has

$$\begin{aligned} v \in \tau_0 = \overline{\text{sp}}\{\varphi_2, \varphi_4, \varphi_5, \dots\} &\Rightarrow \langle v, x_t \rangle + \langle v, A_0^+ A_1 \Delta x_t \rangle \sim I(0), \\ v \in \tau_1 = \text{sp}\{\varphi_3\} &\Rightarrow \langle v, x_t \rangle \sim I(1), \\ v \in \tau_2 = \text{sp}\{\varphi_1\} &\Rightarrow \langle v, x_t \rangle \sim I(2). \end{aligned}$$

Note that  $\tau_1 \subset \tau_0^\perp$  and  $\tau_2 = (\tau_0 \oplus \tau_1)^\perp \neq \{0\}$ . Moreover,  $\langle \varphi_2, A_0^+ A_1 \Delta x_t \rangle = -\Delta x_{1,t} - \Delta x_{2,t}$  and  $\langle \varphi_i, A_0^+ A_1 \Delta x_t \rangle = \frac{\alpha_i}{1-\alpha_i} \Delta x_{i,t}$ , for  $i = 4, 5, \dots$ ; hence  $\langle \varphi_2, x_t \rangle + \langle \varphi_2, A_0^+ A_1 \Delta x_t \rangle = x_{2,t} - \Delta x_{1,t} - \Delta x_{2,t}$  and, for  $i = 4, 5, \dots$ ,  $\langle \varphi_i, x_t \rangle + \langle \varphi_i, A_0^+ A_1 \Delta x_t \rangle = x_{i,t} + \frac{\alpha_i}{1-\alpha_i} \Delta x_{i,t}$ . This shows that  $\langle v, x_t \rangle + \langle v, A_0^+ A_1 \Delta x_t \rangle$  contains stationary terms ( $\Delta x_{2,t}$  and  $\frac{\alpha_i}{1-\alpha_i} \Delta x_{i,t}$ ) that are not necessary for cointegration; these can be eliminated by considering  $\langle v, A_0^+ A_1 P_{\tau_2} \Delta x_t \rangle$  instead of  $\langle v, A_0^+ A_1 \Delta x_t \rangle$ , as in the finite-dimensional  $I(2)$  case, see Theorem 4.6 in Johansen (1996).

### 5. A CHARACTERIZATION OF $I(d)$ AR PROCESSES WITH A UNIT ROOT OF FINITE TYPE

This section extends the results in Section 4 to the general  $I(d)$ ,  $d = 1, 2, \dots, < \infty$ , case. Theorem 5.3 provides a necessary and sufficient condition for AR processes with a unit root of finite type to be  $I(d)$  and it is shown that under this condition the space  $\mathcal{H}$  is decomposed into the direct sum of  $d + 1$  orthogonal subspaces  $\tau_h$ ,  $\mathcal{H} = \tau_0 \oplus \tau_1 \oplus \dots \oplus \tau_d$ ,  $\tau_d \neq \{0\}$ , that are defined in terms of the AR operators  $A_0, A_1, \dots, A_d$  in (4.1), see Definition 5.1 below.

The finite-dimensional attractor space coincides with  $\tau_d$  and  $\tau_0 \oplus \tau_1 \oplus \dots \oplus \tau_{d-1}$  is the cointegrating space, which is infinite-dimensional when  $\dim \mathcal{H} = \infty$ . In  $\tau_0$ , with  $\dim \tau_0 = \infty$  when  $\dim \mathcal{H} = \infty$ , one finds the  $v$ -characteristics that allow for polynomial cointegration of order 0 and in  $\tau_h$ ,  $h = 1, \dots, d - 2$ , with  $0 \leq \dim \tau_h < \infty$ , those that allow for  $I(h)$  polynomial cointegration. In  $\tau_{d-1}$ , with  $0 \leq \dim \tau_{d-1} < \infty$ , those that are  $I(d - 1)$  and do not allow for polynomial cointegration. Finally, any nonzero  $v \in \tau_d$ , with  $0 < \dim \tau_d < \infty$ , is such that  $\langle v, x_t \rangle \sim I(d)$ . The results in Section 4 are found as special cases for  $d = 1$  and  $d = 2$ .

Before stating the results, some definitions are introduced.

**DEFINITION 5.1** ( $S_h, \zeta_h, \tau_h$ , and  $A_{h,n}$ ). Consider an AR process with a unit root of finite type  $A(L)x_t = \varepsilon_t$ , where  $A(z) = \sum_{n=0}^{\infty} A_n(1 - z)^n$  is as in (4.1). Let

$$S_0 = A_0, \quad \zeta_0 = \text{Im } S_0, \quad \tau_0 = (\text{Ker } S_0)^\perp$$

and for  $h = 1, 2, \dots$  define

$$S_h = P_{\mathcal{L}_h^\perp} A_{h,1} P_{\mathcal{T}_h^\perp}, \quad \zeta_h = \text{Im } S_h, \quad \tau_h = (\text{Ker } S_h)^\perp, \tag{5.1}$$

where

$$\mathcal{L}_h = \zeta_0 \oplus \dots \oplus \zeta_{h-1}, \quad \mathcal{T}_h = \tau_0 \oplus \dots \oplus \tau_{h-1} \tag{5.2}$$

and

$$A_{h,n} = \begin{cases} A_n & \text{for } h = 1 \\ A_{h-1,n+1} - A_{h-1,1} \sum_{j=0}^{h-2} S_j^+ A_{j+1,n} & \text{for } h = 2, 3, \dots \end{cases}, \quad n = 1, 2, \dots \tag{5.3}$$

A few remarks on Definition 5.1 are in order.

**Remark 5.2.** First note that for  $h = 1, 2$  (5.1), (5.2), and (5.3) deliver (4.3) and (4.9) respectively. Next, observe that for  $h = 1, 2, \dots$  one has

$$\zeta_h \subseteq (\zeta_0 \oplus \dots \oplus \zeta_{h-1})^\perp, \quad \tau_h \subseteq (\tau_0 \oplus \dots \oplus \tau_{h-1})^\perp \tag{5.4}$$

by construction; that is, for  $0 \leq j < h$ ,  $\zeta_h$  is orthogonal to  $\zeta_j$  and  $\tau_h$  is orthogonal to  $\tau_j$ . Moreover, because  $0 < \dim \zeta_0^\perp = \dim \tau_0^\perp < \infty$ , for  $h = 1, 2, \dots$  the subspaces  $\zeta_h$  and  $\tau_h$  are finite-dimensional and possibly of dimension equal to 0 and the subspaces  $\zeta_0, \tau_0$  are infinite-dimensional when  $\dim \mathcal{H} = \infty$ .

Note also that, as  $h$  increases, the finite-dimensional subspaces  $\mathcal{L}_h^\perp = (\zeta_0 \oplus \dots \oplus \zeta_{h-1})^\perp$  and  $\mathcal{T}_h^\perp = (\tau_0 \oplus \dots \oplus \tau_{h-1})^\perp$  have nonincreasing dimension and, because  $0 < \dim \zeta_0^\perp = \dim \tau_0^\perp < \infty$ , they will eventually have dimension 0. This shows that only a finite number of  $\zeta_h, \tau_h$  are nonzero. Let  $s$  be the value of  $h$  such

that  $\mathcal{X}_s^\perp \neq \{0\}$ ,  $\mathcal{T}_s^\perp \neq \{0\}$  and  $\mathcal{X}_{s+1}^\perp = \mathcal{T}_{s+1}^\perp = \{0\}$ . As shown in Theorem B.4 in Appendix B, the integer  $s$  is precisely the order of the pole of  $A(z)^{-1}$  at  $z = 1$ .

Finally, observe that the generalized inverse of  $S_h, S_h^+$ , exists and it is unique for  $h = 0, 1, \dots$ , because  $\text{Im } S_h, h = 0, 1, \dots$ , is closed; in fact,  $S_0$  is Fredholm of index 0, see Remark 3.2, and  $\dim \text{Im } S_h < \infty$  for  $h = 1, 2, \dots$ .

In the following, the orthogonal direct sum decomposition

$$\mathcal{H} = \tau_0 \oplus \tau_1 \oplus \dots \oplus \tau_d, \quad \tau_d \neq \{0\}, \tag{5.5}$$

is called the POLE( $d$ ) condition.

**THEOREM 5.3** (A characterization of  $I(d)$  AR processes with a unit root of finite type). *Consider an AR process with a unit root of finite type  $A(L)x_t = \varepsilon_t$  and let  $S_h, \tau_h$  and  $A_{h,n}$  be as in Definition 5.1; then  $x_t$  is  $I(d)$  if and only if the POLE( $d$ ) condition in (5.5) holds. In this case, the common trends representation of  $x_t$  is found in (3.6), i.e.,  $x_t = \sum_{n=0}^{d-1} C_n s_{d-n,t} + y_t + \mu_t$ . Moreover,  $\text{Im } C_0 = \tau_d$  is the finite-dimensional attractor space,  $\tau_0 \oplus \tau_1 \oplus \dots \oplus \tau_{d-1}$  is the cointegrating space, which is infinite-dimensional when  $\dim \mathcal{H} = \infty$ , and for any nonzero  $v \in \mathcal{H}$  and for  $h = 0, 1, \dots, d$ , one has*

$$v \in \tau_h \quad \Rightarrow \quad \langle v, x_t \rangle + \sum_{n=1}^{d-h-1} \langle v, S_h^+ A_{h+1,n} \Delta^n x_t \rangle \sim I(h), \tag{5.6}$$

where empty sums are defined to be 0,  $\tau_h \subset (\tau_0 \oplus \dots \oplus \tau_{h-1})^\perp$  for  $h = 1, \dots, d-1$  and  $\tau_d = (\tau_0 \oplus \dots \oplus \tau_{d-1})^\perp \neq \{0\}$ .

**Remark 5.4.** Theorem 5.3 provides a full description of the properties of an  $I(d)$  AR process with a unit root of finite type for a generic  $d = 1, 2, \dots < \infty$ . For  $d = 1$  and  $d = 2$  one finds the  $I(1)$  and  $I(2)$  cases discussed in Theorems 4.1 and 4.8. Observe that all the relevant quantities in Theorem 5.3 are expressed in terms of the AR operators via Definition 5.1.

**Remark 5.5.** The cointegrating relations in (5.6) provide information that parallels the triangular representation for finite-dimensional process discussed in Phillips (1991a) and Stock and Watson (1993); see also Franchi and Paruolo (2019, Corollary 4.6).

**Remark 5.6.** An AR process with a unit root of finite type generates an  $I(d)$  process if and only if  $\tau_d = (\tau_0 \oplus \dots \oplus \tau_{d-1})^\perp \neq \{0\}$ . The common trends representation of  $x_t$  shows that the  $I(d)$  stochastic trends  $s_{d,t}$  are loaded into the process by  $C_0$ ; because  $\text{Im } C_0 = \tau_d = (\tau_0 \oplus \dots \oplus \tau_{d-1})^\perp$ , for any nonzero  $v \in \tau_0 \oplus \dots \oplus \tau_{d-1}$  one has  $\langle v, C_0 u \rangle = 0$  for all  $u \in \mathcal{H}$ , and hence also for  $u = s_{d,t}$ ; this implies that  $\langle v, x_t \rangle$  is at most  $I(d-1)$ , i.e.,  $\tau_0 \oplus \dots \oplus \tau_{d-1}$  is the cointegrating space and its orthogonal complement  $\tau_d = (\tau_0 \oplus \dots \oplus \tau_{d-1})^\perp$  is the finite-dimensional attractor space. Note that when  $\dim \mathcal{H} = \infty$  one has  $\dim \tau_0 = \infty$  and hence the cointegrating space is infinite-dimensional.



**Remark 5.7.** Theorem 5.3 further shows that the cointegrating space is partitioned into  $\tau_0 \oplus \dots \oplus \tau_{d-1}$ ; in  $\tau_0$ , which is infinite-dimensional when  $\dim \mathcal{H} = \infty$ , one finds the cointegrating vectors that allow for  $I(0)$  polynomial cointegration, i.e., for any nonzero  $v \in \tau_0$ , one has  $\langle v, x_t \rangle + \sum_{n=1}^{d-1} \langle v, A_0^+ A_n \Delta^n x_t \rangle \sim I(0)$ . For any nonzero  $v \in \tau_h$ ,  $h = 1, \dots, d - 2$ , for which  $0 \leq \dim \tau_h < \infty$ , one finds those that allow for  $I(h)$  polynomial cointegration, i.e.,  $\langle v, x_t \rangle + \sum_{n=1}^{d-h-1} \langle v, S_h^+ A_{h+1,n} \Delta^n x_t \rangle \sim I(h)$ . In  $\tau_{d-1}$ , with  $0 \leq \dim \tau_{d-1} < \infty$ , every nonzero  $v$ -characteristic of  $x_t$  is  $I(d - 1)$  and does not allow for polynomial cointegration, and in  $\tau_d$ , with  $0 < \dim \tau_d < \infty$ , every nonzero  $v$ -characteristic of  $x_t$  is  $I(d)$ . Apart from the fact that the dimension of  $\tau_0$  is infinite when  $\dim \mathcal{H} = \infty$ , this parallels the finite-dimensional case, see Theorem 4.3 in Franchi and Paruolo (2019).

**Remark 5.8.** The orthogonal direct sum decomposition  $\mathcal{H} = \tau_0 \oplus \tau_1 \oplus \dots \oplus \tau_d$  can be employed to characterize the order of integration of any  $v$ -characteristic of the process. In fact, (5.5) implies  $P_{\tau_0} + P_{\tau_1} + \dots + P_{\tau_d} = I$ , where  $P_{\tau_h}$  is the orthogonal projection onto  $\tau_h$ ; hence for any  $v \in \mathcal{H}$  one has  $\langle v, x_t \rangle = \langle v_0, x_t \rangle + \langle v_1, x_t \rangle + \dots + \langle v_d, x_t \rangle$ , where  $v_h = P_{\tau_h} v \in \tau_h$ , where (5.6) implies that  $\langle v_h, x_t \rangle$  is integrated of order  $I(d - 1)$  or less for  $h = 0, 1, \dots, d - 1$ ; these components are dominated by  $\langle v_d, x_t \rangle \sim I(d)$ . Hence  $\langle v, x_t \rangle \sim I(d)$  if and only if  $v_d \neq 0$ .

**Remark 5.9.** In the finite-dimensional case  $\mathcal{H} = \mathbb{R}^p$ , Franchi and Paruolo (2016) show that  $d = 1, 2, \dots$  if and only if  $\mathbb{R}^p = \zeta_0 \oplus \dots \oplus \zeta_d = \tau_0 \oplus \dots \oplus \tau_d$ , where  $\zeta_h = \text{sp}\{\alpha_h\}$ ,  $\tau_h = \text{sp}\{\beta_h\}$ ,  $h = 0, 1, \dots$ , and the bases  $\alpha_h, \beta_h$  are defined by the rank factorizations  $P_{\mathcal{F}_h^\perp} A_{h,1} P_{\mathcal{F}_h^\perp} = \alpha_h \beta_h'$ , where  $A_{h,1}$  is as in Definition 5.1 with  $S_h^+ = (\alpha_h \beta_h')^+ = \tilde{\beta}_h \tilde{\alpha}_h'$ . Apart from the fact that  $\dim \zeta_0 = \dim \tau_0$  is finite when  $\mathcal{H} = \mathbb{R}^p$  and infinite when  $\dim \mathcal{H} = \infty$ , the general case in this article coincides with the finite-dimensional case discussed in that article.

**Remark 5.10.** The POLE( $d$ ) condition in (5.5) is equivalent to  $\tau_d = (\tau_0 \oplus \dots \oplus \tau_{d-1})^\perp \neq \{0\}$ . Moreover, Theorem B.4 in Appendix B shows that it can be equivalently stated as (i)  $\mathcal{H} = \zeta_0 \oplus \zeta_1 \oplus \dots \oplus \zeta_d, \zeta_d \neq \{0\}$ , (ii)  $\zeta_d = (\zeta_0 \oplus \dots \oplus \zeta_{d-1})^\perp \neq \{0\}$ , (iii)  $\text{Im } C_0 = \tau_d, (iv) \text{ Ker } C_0 = \tau_0 \oplus \dots \oplus \tau_{d-1}$ .

In order to complete the discussion of the relation of the present results with the existing literature, the equivalence of the POLE( $d$ ) condition in (5.5) to the condition in Hu and Park (2016) reported in equation (5.7) below is discussed.

Hu and Park (2016) consider an infinite-dimensional  $\mathcal{H}$ -valued AR(1) process  $x_t = A_1^\circ x_{t-1} + \varepsilon_t$  with  $A_1^\circ$  compact, formulate an  $I(d)$  condition and then study the  $I(1)$  case. In order to state their  $I(d)$  condition, they employ the nonorthogonal direct sum decomposition  $\mathcal{H} = \mathcal{H}_P \oplus \mathcal{H}_T$ , where  $\mathcal{H}_P$  is the finite-dimensional image of the Riesz projection associated with the isolated eigenvalue  $z = 1$  and  $\mathcal{H}_T$  is the infinite-dimensional image of the Riesz projections associated with the remaining stable eigenvalues. Using the nonorthogonal projections associated with the nonorthogonal direct sum decomposition  $\mathcal{H} = \mathcal{H}_P \oplus \mathcal{H}_T$ , they decompose the process into  $x_t = x_t^P + x_t^T$ , where  $x_t^X = A_X x_t^X + \varepsilon_t^X \in \mathcal{H}_X$  and  $A_X = A_1^\circ |_{\mathcal{H}_X}$  is the restriction of  $A_1^\circ$  to  $\mathcal{H}_X, X = T, P$ . Their  $I(d)$  condition is stated as

$A_P - I$  is a nilpotent matrix of order  $d$ , (5.7)

i.e.,  $(A_P - I)^{d-1} \neq 0$  and  $(A_P - I)^d = 0$ , which simplifies to  $A_P = I$  in the  $I(1)$  case studied in that article.

**PROPOSITION 5.11** (Equivalence of (5.5) and (5.7)). *Let  $x_t = A_1^\circ x_{t-1} + \varepsilon_t$  be an AR process with a unit root of finite type; then the  $I(d)$  condition in (5.7) is equivalent to the POLE( $d$ ) condition in (5.5).*

## 6. CONCLUSION

The present article characterizes the cointegration properties of infinite-dimensional  $\mathcal{H}$ -valued AR processes  $A(L)x_t = \varepsilon_t$  such that  $A(z)$  has an eigenvalue of finite type at  $z = 1$  and it is invertible in the punctured disc  $D(0, \rho) \setminus \{1\}$  for some  $\rho > 1$ . It is shown that these processes, called AR processes with a unit root of finite type, are necessarily integrated of finite integer order  $d$ ,  $I(d)$ , and necessarily have a finite number of  $I(d)$  trends and an infinite-dimensional cointegrating space when  $\dim \mathcal{H} = \infty$ . This is in line with the setup employed in many contributions in the literature and seems to be a relevant framework for applications.

A necessary and sufficient condition on the AR operators that establishes the value of  $d$  is given in terms of the orthogonal direct sum decomposition  $\mathcal{H} = \tau_0 \oplus \tau_1 \oplus \dots \oplus \tau_d$ ,  $\tau_d \neq \{0\}$ , where  $\tau_0$  is infinite-dimensional when  $\dim \mathcal{H} = \infty$  and  $0 \leq \dim \tau_h < \infty$ ,  $h = 1, \dots, d$ , with strict inequality for  $h = d$ .

A full description of how the properties of the  $v$ -characteristic  $\langle v, x_t \rangle$  vary with  $v \in \mathcal{H}$  is given: in  $\tau_0$ , one can combine  $\langle v, x_t \rangle$  with differences of the process and find  $I(0)$  polynomial cointegrating relations; in  $\tau_1$ , one can combine  $\langle v, x_t \rangle$  with differences and find at most  $I(1)$  polynomial cointegrating relations, and so on up to  $\tau_{d-2}$ , in which one can combine  $\langle v, x_t \rangle$  with differences and find at most  $I(d-2)$  polynomial cointegrating relations. Finally, any nonzero  $v \in \tau_{d-1}$  is such that  $\langle v, x_t \rangle$  is  $I(d-1)$  and does not allow for polynomial cointegration and any nonzero  $v \in \tau_d$  is such that  $\langle v, x_t \rangle$  is  $I(d)$ . This shows that the subspace  $\tau_0 \oplus \tau_1 \oplus \dots \oplus \tau_{d-1}$ , which is infinite-dimensional when  $\dim \mathcal{H} = \infty$ , is the cointegrating space, whereas the finite-dimensional subspace  $\tau_d$  is the attractor space.

For any nonzero  $v$  in the cointegrating space, the expression of the polynomial cointegrating relations is provided in terms of operators that are defined recursively in terms of the AR operators together with the  $\tau_h$  subspaces.

The present results show that, under the assumption that 1 is an eigenvalue of finite type of the AR operator function, the infinite-dimensionality of the space  $\mathcal{H}$  does not introduce additional elements in the representation theory. That is, apart from the fact that the dimension of the cointegrating space is infinite when  $\dim \mathcal{H} = \infty$ , conditions and properties of AR processes with a unit root of finite type coincide with those that apply in the usual finite-dimensional VAR case, which is obtained setting  $\mathcal{H} = \mathbb{R}^p$  in the present results.

## 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, & W. Seo (2017) Cointegrated linear processes in Hilbert space. *Journal of Time Series Analysis* 38, 1010–1027.
- Beare, B. & W. Seo (2019) Representation of I(1) and I(2) autoregressive Hilbertian processes. *Econometric Theory*, <https://doi.org/10.1017/S0266466619000276>.
- Ben-Israel, A. & T. Greville (2003) *Generalized Inverses: Theory and Applications*, 2nd ed. Springer.
- Bosq, D. (2000) *Linear Processes in Function Spaces*. Springer-Verlag.
- Bourguignon, F., F.H. Ferreira, & N. Lustig (2005) *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*. World Bank and Oxford University Press
- Chang, Y., B. Hu, & J. Park (2016a). On the Error Correction Model for Functional Time Series with Unit Roots. Mimeo, Indiana University.
- Chang, Y., C. Kim, & J. Park (2016b) Nonstationarity in time series of state densities. *Journal of Econometrics* 192, 152–167.
- Cochrane, J.H. & M. Piazzesi (2005) Bond risk premia. *American Economic Review* 95(1), 138–160.
- Franchi, M. & 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. & P. Paruolo (2019) A general inversion theorem for cointegration. *Econometric Reviews* 38, 1176–1201.
- Gabrys, R., S. Hörmann, & P. Kokoszka (2013) Monitoring the intraday volatility pattern. *Journal of Time Series Econometrics* 5, 87–116.
- Gohberg, I., S. Goldberg, & M. Kaashoek (1990) *Classes of Linear Operators, vol. 1. Operator Theory*. Birkhäuser Verlag.
- Gohberg, I., S. Goldberg, & M. Kaashoek (2003) *Basic Classes of Linear Operators*. Birkhäuser.
- Granger, C.W.J. & T.H. Lee (1989) Investigation of production, sales and inventory relationships using multicointegration and non-symmetric error correction models. *Journal of Applied Econometrics* 4, S145–S159.
- Gregoir, S. (1999) Multivariate time series with various hidden unit roots, Part I. *Econometric Theory* 15, 435–468.
- Hörmann, S., L. Horváth, & R. Reeder (2013) A functional version of the ARCH model. *Econometric Theory* 29, 267–288.
- Hörmann, S. & P. Kokoszka (2012) Functional time series. *Handbook of Statistics* 30, 157–186.
- Horn, R. & C. Johnson (2013) *Matrix Analysis*. Cambridge University Press.
- Horváth, L. & P. Kokoszka (2012) *Inference for Functional Data with Application*. Springer.
- Howlett, P., K. Avrachenkov, C. Pearce, & V. Ejov (2009) Inversion of analytically perturbed linear operators that are singular at the origin. *Journal of Mathematical Analysis and Applications* 353, 68–84.
- Hu, B. & 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. & A. Onatski (2008) Curve forecasting by functional autoregression. *Journal of Multivariate Analysis* 99, 2508–2526.
- Kokoszka, P. & M. Reimherr (2017) *Introduction to Functional Data Analysis*. Chapman and Hall.
- Petersen, A. & H.-G. Müller (2016) Functional data analysis for density functions by transformation to a Hilbert space. *Annals of Statistics* 44(1), 183–218.
- Phillips, P. (1991a) Optimal inference in cointegrated systems. *Econometrica* 59, 283–306.
- Phillips, P. (1991b). Spectral regression for cointegrated time series. In W. Barnett (ed.), *Nonparametric and Semiparametric Methods in Economics and Statistics*, pp. 413–435. Cambridge University Press.

- Piketty, T. (2014) *Capital in the Twenty-First Century*. The Belknap Press of Harvard University Press.
- Seo, W. & B. Beare (2019) Cointegrated linear processes in Bayes Hilbert space. *Statistics and Probability Letters* 147, 90–95.
- Stock, J. & M. Watson (1993) A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61, 783–820.

## APPENDIX A: Notation and Background Results

In the present article,  $\mathcal{H}$  is a possibly infinite-dimensional separable Hilbert space, an  $\mathcal{H}$ -valued random variable is a random variable that takes values in  $\mathcal{H}$  and an  $\mathcal{H}$ -valued stochastic process is a sequence of  $\mathcal{H}$ -valued random variables. Section A.1 reviews notions and results on separable Hilbert spaces and on operators acting on them and Section A.2 presents the definitions of expectation and covariance operator for  $\mathcal{H}$ -valued random variables.

### A.1. Separable Hilbert Spaces and Operators Acting on Them

The material in this section is based on Chapters I, II in Gohberg et al. (2003) and Chapter XI in Gohberg et al. (1990). Consider separable Hilbert spaces  $\mathcal{H}, \mathcal{H}_1, \mathcal{H}_2$  with inner products  $\langle \cdot, \cdot \rangle$  and norms  $\|x\| = \langle x, x \rangle^{\frac{1}{2}}$ ;<sup>7</sup> a transformation  $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ , is called a linear operator if for all  $v, w \in \mathcal{H}_1$  and  $c \in \mathbb{C}$ ,  $A[v + w] = Av + Aw$  and  $A[cv] = cAv$ , where  $Au$  and  $A[u]$  both indicate the action of  $A$  on  $u \in \mathcal{H}$ . A linear operator  $A$  is called bounded if its norm  $\|A\|_{\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}} = \sup_{\|v\|=1} \|Av\|$  is finite and the set of bounded linear operators with norm  $\|\cdot\|_{\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}}$  is denoted by  $\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}$ . For any  $A \in \mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}$  the subspace  $\{v \in \mathcal{H}_1 : Av = 0\}$ , written  $\text{Ker } A$ , is called the kernel of  $A$  and the subspace  $\{Av : v \in \mathcal{H}_1\}$ , written  $\text{Im } A$ , is called the image of  $A$ . The dimension of  $\text{Im } A$ , written  $\dim \text{Im } A$ , is called the rank of  $A$ , written  $\text{rank } A$ . When  $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$ ,  $\mathcal{L}_{\mathcal{H}}$  is written in place of  $\mathcal{L}_{\mathcal{H}_1, \mathcal{H}_2}$ .

$\mathcal{H}$  is said to be the direct sum of subspaces  $S$  and  $U$ , written  $\mathcal{H} = S \oplus U$ , if  $S \cap U = 0$  and if every vector  $v \in \mathcal{H}$  can be written as  $v = s + u$ , where  $s \in S$  and  $u \in U$ . For  $U = S^\perp$ , one has the orthogonal direct sum  $\mathcal{H} = S \oplus S^\perp$ . The orthogonal projection on  $\eta$ , written  $P_\eta$ , is such that  $P_\eta \in \mathcal{L}_{\mathcal{H}}$ ,  $P_\eta^2 = P_\eta$ ,  $\text{Im } P_\eta = \eta$  and  $\text{Ker } P_\eta = \eta^\perp$ ; moreover,  $I = P_\eta + P_{\eta^\perp}$ .

An operator  $A \in \mathcal{L}_{\mathcal{H}}$  is said to be invertible if there exists an operator  $B \in \mathcal{L}_{\mathcal{H}}$  such that  $BAv = ABv = v$  for every  $v \in \mathcal{H}$ ; in this case  $B$  is called the inverse of  $A$ , written  $A^{-1}$ . An operator  $A \in \mathcal{L}_{\mathcal{H}}$  such that  $n = \dim \text{Ker } A < \infty$  and  $q = \dim(\text{Im } A)^\perp < \infty$  is said to be Fredholm of index  $n - q$ . Observe that if  $\mathcal{H}$  is finite-dimensional, any  $A \in \mathcal{L}_{\mathcal{H}}$  is Fredholm of index 0.

Corollary 8.4 in Section XI.8 in Gohberg et al. (1990) states that the inverse of an operator function that is Fredholm of index 0 and non-invertible at some isolated point has a pole at that point. Moreover, the operators that make up the principal part of its Laurent representation around that point have finite rank. If  $z_0$  is an eigenvalue of finite type of  $W(z)$ , see Definition 3.1, then  $z_0$  is an isolated singularity of  $W(z)^{-1}$ ,  $W(z_0)$  is Fredholm of index 0 and non-invertible at  $z_0$ , so that Theorem A.1 below applies.

<sup>7</sup> The same notation is used for the inner products and norms on  $\mathcal{H}, \mathcal{H}_1, \mathcal{H}_2$  but this should cause no confusion.

**THEOREM A.1.** *Let  $z_0$  be an eigenvalue of finite type of an operator function  $W(z)$ ; then there exist a finite integer  $d = 1, 2, \dots$  and operators  $U_n, n = 0, 1, \dots$ , with finite rank for  $n = 0, \dots, d - 1$ , such that*

$$W(z)^{-1} = \sum_{n=0}^{\infty} U_n(z - z_0)^{n-d}, \quad z \in D(z_0, \delta) \setminus \{z_0\},$$

where  $U_d$  is Fredholm of index 0.

*Proof.* See Section XI.8 in Gohberg et al. (1990). ■

### A.2. Random Variables in Separable Hilbert Spaces

The definitions in this section are taken from Chapter 1 in Bosq (2000). Let  $\mathcal{H}$  be a separable Hilbert space with inner product  $\langle \cdot, \cdot \rangle$ , norm  $\|x\| = \langle x, x \rangle^{1/2}$ , and Borel  $\sigma$ -algebra  $\sigma(\mathcal{H})$  and let  $(\Omega, \mathcal{A}, P)$  be a probability space. A function  $Z : \Omega \rightarrow \mathcal{H}$  is called an  $\mathcal{H}$ -valued random variable on  $(\Omega, \mathcal{A}, P)$  if it is measurable, i.e., for every subset  $S \in \sigma(\mathcal{H})$ ,  $\{\omega : Z(\omega) \in S\} \in \mathcal{A}$ . For a  $\mathbb{C}$ -valued random variable  $X$  on  $(\Omega, \mathcal{A}, P)$ , define  $E(X) = \int_{\Omega} X(\omega) dP(\omega)$ ; the expectation of an  $\mathcal{H}$ -valued random variable  $Z$ , written  $E(Z)$ , is defined as the unique element of  $\mathcal{H}$  such that

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

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

$$c_Z(v, w) = E(\langle v, Z - E(Z) \rangle \langle w, Z - E(Z) \rangle), \quad v, w \in \mathcal{H}.$$

It is immediate to see that  $c_Z(v, w) = E(\langle v, W \rangle) - \langle v, E(Z) \rangle \langle w, E(Z) \rangle$ , where  $W = \langle w, Z \rangle Z$ . If  $E(\|W\|) < \infty$ , one has

$$c_Z(v, w) = \langle v, E(W) \rangle - \langle v, E(Z) \rangle \langle w, E(Z) \rangle, \quad v, w \in \mathcal{H}, \quad W = \langle w, Z \rangle Z.$$

Because  $\|W\| = |\langle w, Z \rangle| \|Z\| \leq \|w\| \|Z\|^2$ , the existence of the covariance function of  $Z$  is guaranteed by the condition  $E(\|Z\|^2) < \infty$ . Define the operator  $C_Z : \mathcal{H} \rightarrow \mathcal{H}$  that maps  $w$  into  $E(W)$  and rewrite the covariance function as  $c_Z(v, w) = \langle v, C_Z w \rangle - \langle v, E(Z) \rangle \langle w, E(Z) \rangle, v, w \in \mathcal{H}$ .  $C_Z$  is fully determined by the covariance function and it is called the covariance operator of  $Z$ . Similarly, the cross-covariance function of two  $\mathcal{H}$ -valued random variables  $Z$  and  $U$  is defined as

$$c_{Z,U}(v, w) = E(\langle v, Z - E(Z) \rangle \langle w, U - E(U) \rangle), \quad v, w \in \mathcal{H}.$$

This also completely determines the cross-covariance operators of  $Z$  and  $U$ ,  $C_{Z,U}$  and  $C_{U,Z}$ , respectively, defined as the mappings  $w \mapsto E(\langle w, Z \rangle U)$  and  $w \mapsto E(\langle w, U \rangle Z)$ .

## APPENDIX B: Inversion of an Operator Function around a Singular Point

This Appendix presents novel results on the inversion of a meromorphic operator function; these are used in Appendix C for the proofs of the propositions in the text.

The inversion results are derived from system (B.1) below, also employed in Howlett, Avrachenkov, Pearce, and Ejov (2009). When the inverse  $A(z)^{-1}$  has a pole of order  $d$ , from the identity  $A(z)A(z)^{-1} = I = A(z)^{-1}A(z)$  one finds the following systems in the  $A_n, C_n$  operators defined in (4.1) and (3.3),

$$\begin{aligned}
 A_0C_0 &= 0 = C_0A_0 \\
 A_0C_1 + A_1C_0 &= 0 = C_0A_1 + C_1A_0 \\
 &\vdots \\
 A_0C_{d-1} + \dots + A_{d-1}C_0 &= 0 = C_0A_{d-1} + \dots + C_{d-1}A_0 \\
 A_0C_d + A_1C_{d-1} + \dots + A_dC_0 &= I = C_0A_d + C_1A_{d-1} + \dots + C_dA_0 \\
 A_0C_{d+1} + A_1C_d + \dots + A_{d+1}C_0 &= 0 = C_0A_{d+1} + C_1A_d + \dots + C_{d+1}A_0 \\
 &\vdots
 \end{aligned}
 \tag{B.1}$$

In the following, equations in system (B.1) are numbered according to the highest value of the subscript of  $C_n$  which is present in the equation. Note that the identity appears in equation  $d$ , which is the order of the pole. The equations derived from  $A(z)A(z)^{-1} = I$  are called left versions (and correspond to the left side of (B.1)) and those that derive from  $I = A(z)^{-1}A(z)$  are called right versions (and correspond to the right side of (B.1)). For instance  $A_0C_d + A_1C_{d-1} + \dots + A_dC_0 = I$  is called the left version of equation  $d$ . Finally, let  $P_\eta$  indicate the orthogonal projection operator on  $\eta$  and  $A^+$  and  $A^*$  respectively denote the generalized inverse and the adjoint of  $A$ .

LEMMA B.1. Consider Definition 5.1; then  $\text{Ker } S_h^+ = (\text{Im } S_h)^\perp, S_h^+S_h = P_{\tau_h}$  and  $S_h^+P_{\mathcal{X}_h^\perp} = S_h^+, h = 0, 1, \dots, d$ .

*Proof.* From Theorem 3 in Ben-Israel and Greville (2003, Chap. 9), one has  $S_h^+S_h = P_{\text{Im } S_h}$  and  $\text{Ker } S_h^+ = \text{Ker } S_h^*$  and from Theorem 11.4 in Gohberg et al. (2003, Chap. II) one has  $\text{Im } S_h^* = (\text{Ker } S_h)^\perp$  and  $\text{Ker } S_h^* = (\text{Im } S_h)^\perp$ , so that  $\text{Ker } S_h^+ = (\text{Im } S_h)^\perp$ . By Definition 5.1,  $(\text{Ker } S_h)^\perp = \tau_h$  and hence  $S_h^+S_h = P_{\tau_h}$ . Moreover, by Definition 5.1,  $(\text{Im } S_h)^\perp = \zeta_h^\perp \supseteq \zeta_0 \oplus \dots \oplus \zeta_{h-1} = \mathcal{X}_h$  and hence  $\mathcal{X}_h \subseteq \text{Ker } S_h^+$ , which implies  $S_h^+ = S_h^+P_{\mathcal{X}_h^\perp}$ . ■

LEMMA B.2 (Subspace decompositions of system (B.1)). Consider Definition 5.1 and further define  $P_{\mathcal{X}_0^\perp} = P_{\mathcal{T}_0} = I$ ; then the left version of equation  $n + h \leq d$  in system (B.1) implies

$$S_hC_n + P_{\mathcal{X}_h^\perp} \sum_{k=1}^n A_{h+1,k}C_{n-k} = \delta_{n+h,d}P_{\mathcal{X}_h^\perp}, \quad h = 0, 1, \dots, d - n,
 \tag{B.2}$$

where  $\delta_{h,j}$  is the Kronecker delta, which is equal to 1 if  $h = j$  and equal to 0 otherwise. Similarly, the right version of equation  $n + h \leq d$  in system (B.1) implies

$$C_nS_h + \sum_{k=1}^n C_{n-k}A_{h+1,k}P_{\mathcal{T}_h^\perp} = \delta_{n+h,d}P_{\mathcal{T}_h^\perp}, \quad h = 0, 1, \dots, d - n.
 \tag{B.3}$$

**Proof.** The proof of (B.2) is by induction and consists in showing that the left version of equation  $n \leq d$  in system (B.1) implies

$$S_h C_{n-h} + P_{\mathcal{X}_h^\perp} \sum_{k=1}^{n-h} A_{h+1,k} C_{n-h-k} = \delta_{n,d} P_{\mathcal{X}_h^\perp}, \quad h = 0, 1, \dots, n; \tag{B.4}$$

replacing  $n$  with  $n + h$  one finds (B.2). In order to show that (B.4) holds for  $h = 0$ , observe that the left version of equation  $n$  in system (B.1) reads  $A_0 C_n + \sum_{k=1}^n A_k C_{n-k} = \delta_{n,d} I$ . By definition,  $P_{\mathcal{X}_0^\perp} = I$ ,  $S_0 = A_0$ , and  $A_{1,k} = A_k$  and this shows that (B.4) holds for  $h = 0$ . Next assume that (B.4) holds for  $h = 0, \dots, \ell - 1$  for some  $1 \leq \ell \leq d$ ; one wishes to show that it also holds for  $h = \ell$ . First note that  $S_h^+ S_h = P_{\tau_h}$  and  $S_h^+ P_{\mathcal{X}_h^\perp} = S_h^+$ , see Lemma B.1; thus the induction assumption implies

$$P_{\tau_h} C_{n-h} + S_h^+ \sum_{k=1}^{n-h} A_{h+1,k} C_{n-h-k} = \delta_{n,d} S_h^+, \quad h = 0, 1, \dots, \ell - 1,$$

and replacing  $n$  with  $n - \ell + h$  and  $h$  with  $i$ , one has

$$P_{\tau_i} C_{n-\ell} = -S_i^+ \sum_{k=1}^{n-\ell} A_{i+1,k} C_{n-\ell-k} + \delta_{n-\ell+i,d} S_i^+, \quad i = 0, 1, \dots, \ell - 1.$$

Observe that for  $i = 0, 1, \dots, \ell - 1$  one has  $n - \ell + i \leq n - 1 < d$ ; hence  $\delta_{n-\ell+i,d} = 0$  and one finds

$$P_{\tau_i} C_{n-\ell} = -S_i^+ \sum_{k=1}^{n-\ell} A_{i+1,k} C_{n-\ell-k}, \quad i = 0, 1, \dots, \ell - 1. \tag{B.5}$$

Next write (B.4) for  $h = \ell - 1$ ,

$$S_{\ell-1} C_{n-\ell+1} + P_{\mathcal{X}_{\ell-1}^\perp} \sum_{k=1}^{n-\ell+1} A_{\ell,k} C_{n-\ell+1-k} = \delta_{n,d} P_{\mathcal{X}_{\ell-1}^\perp},$$

where  $\text{Im } S_{\ell-1} = \zeta_{\ell-1}$ , see Definition 5.1; applying  $P_{\mathcal{X}_\ell^\perp}$ , where  $\mathcal{X}_\ell = \zeta_0 \oplus \dots \oplus \zeta_{\ell-1}$ , one has  $P_{\mathcal{X}_\ell^\perp} S_{\ell-1} = 0$  and rearranging one finds

$$P_{\mathcal{X}_\ell^\perp} A_{\ell,1} C_{n-\ell} + P_{\mathcal{X}_\ell^\perp} \sum_{k=1}^{n-\ell} A_{\ell,k+1} C_{n-\ell-k} = \delta_{n,d} P_{\mathcal{X}_\ell^\perp}. \tag{B.6}$$

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

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

By Definition 5.1,  $P_{\mathcal{X}_\ell^\perp} A_{\ell,1} P_{\mathcal{T}_\ell^\perp} = S_\ell$ , so that  $U = S_\ell C_{n-\ell} + U_2$ . Substituting  $P_{\mathcal{T}_\ell} = P_{\tau_0} + \dots + P_{\tau_{\ell-1}}$  in  $U_2$ , one has  $U_2 = P_{\mathcal{X}_\ell^\perp} A_{\ell,1} \sum_{i=0}^{\ell-1} P_{\tau_i} C_{n-\ell}$  and by the induction

assumption, see (B.5), one finds

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

Substituting the expression of  $U_2$  into  $U = S_\ell C_{n-\ell} + U_2$  and using  $A_{\ell+1,k} = A_{\ell,k+1} - A_{\ell,1} \sum_{i=0}^{\ell-1} S_i^+ A_{i+1,k}$ , see Definition 5.1, one hence can rewrite (B.6) as

$$S_\ell C_{n-\ell} + P_{\mathcal{X}_\ell^\perp} \sum_{k=1}^{n-\ell} A_{\ell+1,k} C_{n-\ell-k} = \delta_{n,d} P_{\mathcal{X}_\ell^\perp}.$$

This shows that (B.4) holds for  $h = \ell$  and completes the proof of (B.2). A similar induction on the right version of system (B.1) leads to (B.3). ■

LEMMA B.3. Consider Definition 5.1; then  $\text{Im } C_0 \subseteq \mathcal{F}_d^\perp$  and  $\mathcal{Z}_d \subseteq \text{Ker } C_0$ .

*Proof.* For  $n = 0$ , (B.2) and (B.3) read

$$S_h C_0 = \delta_{h,d} P_{\mathcal{X}_h^\perp}, \quad C_0 S_h = \delta_{h,d} P_{\mathcal{F}_h^\perp} \quad h = 0, 1, \dots, d, \tag{B.7}$$

where  $S_h = P_{\mathcal{X}_h^\perp} A_{h,1} P_{\mathcal{F}_h^\perp}$ , see Definition 5.1. (B.7) implies  $S_h C_0 = C_0 S_h = 0$  for  $h = 0, 1, \dots, d - 1$ . From  $S_h C_0 = 0$ ,  $h = 0, 1, \dots, d - 1$ , one has  $\text{Im } C_0 \subseteq \text{Ker } S_h$  for  $h = 0, 1, \dots, d - 1$ , i.e.,  $\text{Im } C_0 \subseteq (\text{Ker } S_0 \cap \text{Ker } S_1 \cap \dots \cap \text{Ker } S_{d-1})$ . By Definition 5.1,  $\text{Ker } S_h = \tau_h^\perp$  and hence  $\text{Im } C_0 \subseteq (\tau_0^\perp \cap \tau_1^\perp \cap \dots \cap \tau_{d-1}^\perp) = (\tau_0 \oplus \tau_1 \oplus \dots \oplus \tau_{d-1})^\perp = \mathcal{F}_d^\perp$ . This proves the first statement.

From  $C_0 S_h = 0$ ,  $h = 0, 1, \dots, d - 1$ , one has  $\text{Im } S_h \subseteq \text{Ker } C_0$  for  $h = 0, 1, \dots, d - 1$ , i.e.,  $(\text{Im } S_0 \oplus \text{Im } S_1 \oplus \dots \oplus \text{Im } S_{d-1}) \subseteq \text{Ker } C_0$ . By Definition 5.1,  $\text{Im } S_h = \zeta_h$  and hence  $\zeta_0 \oplus \zeta_1 \oplus \dots \oplus \zeta_{d-1} = \mathcal{Z}_d \subseteq \text{Ker } C_0$ . ■

THEOREM B.4 (Order of the pole). Consider Definition 5.1. The following statements are equivalent:

- (i)  $A(z)^{-1}$  has a pole of order  $d$  at  $z = 1$ ,
- (ii) the identity is in equation  $d$  of system (B.1),
- (iii)  $\zeta_d = (\zeta_0 \oplus \dots \oplus \zeta_{d-1})^\perp \neq \{0\}$ ,
- (iv)  $\text{Ker } C_0 = \zeta_d^\perp$ ,
- (v)  $\tau_d = (\tau_0 \oplus \dots \oplus \tau_{d-1})^\perp \neq \{0\}$ ,
- (vi)  $\text{Im } C_0 = \tau_d$ .

*Proof.*

(i)  $\Leftrightarrow$  (ii) By definition.

(ii)  $\Rightarrow$  (iii)  $\Rightarrow$  (iv). Under (ii), one has  $h = d$  in the left equation in (B.7), i.e.,  $S_d C_0 = P_{\mathcal{X}_d^\perp}$ ,  $\mathcal{Z}_d^\perp \neq \{0\}$ ; by Definition 5.1,  $\text{Im } S_d \subseteq \mathcal{Z}_d^\perp$  and because  $\text{Im } S_d \subset \mathcal{Z}_d^\perp$  contradicts  $S_d C_0 = P_{\mathcal{X}_d^\perp}$ , one has  $\text{Im } S_d = \mathcal{Z}_d^\perp$ . By Definition 5.1,  $\text{Im } S_d = \zeta_d$  and  $\mathcal{Z}_d^\perp = (\zeta_0 \oplus \dots \oplus \zeta_{d-1})^\perp$ , and hence (iii). Moreover, by Lemma B.3,  $\mathcal{Z}_d \subseteq \text{Ker } C_0$  and because  $\mathcal{Z}_d \subset \text{Ker } C_0$  contradicts  $S_d C_0 = P_{\mathcal{X}_d^\perp}$ , one has  $\mathcal{Z}_d = \text{Ker } C_0$ . Using  $\mathcal{Z}_d = \zeta_d^\perp$ , see (iii), one finds (iv).



(iv)  $\Rightarrow$  (ii). Let  $\text{Ker } C_0 = \zeta_d^\perp$  and proceed by contradiction, assuming that the identity is not in equation  $d$ , so that the right equation in (B.7) reads  $C_0 S_d = 0$ , which implies  $\text{Im } S_d \subseteq \text{Ker } C_0$ , where  $\text{Im } S_d = \zeta_d$  and  $\text{Ker } C_0 = \zeta_d^\perp$ . Hence  $\zeta_d \subseteq \zeta_d^\perp$ , so that  $\zeta_d = \{0\}$  and thus  $\zeta_d^\perp = \mathcal{H}$ . This contradicts  $C_0 \neq 0$ , i.e., that the pole has order  $d$ , and proves that (ii) holds.

(ii)  $\Rightarrow$  (v)  $\Rightarrow$  (vi). Under (ii), one has  $h = d$  in the right equation in (B.7), i.e.,  $C_0 S_d = P_{\mathcal{F}_d^\perp}$ ,  $\mathcal{F}_d^\perp \neq \{0\}$ ; by Definition 5.1,  $\mathcal{F}_d \subseteq \text{Ker } S_d$  and because  $\mathcal{F}_d \subset \text{Ker } S_d$  contradicts  $C_0 S_d = P_{\mathcal{F}_d^\perp}$ , one has  $\mathcal{F}_d = \text{Ker } S_d$ . By Definition 5.1,  $\text{Ker } S_d = \tau_d^\perp$  and  $\mathcal{F}_d = \tau_0 \oplus \dots \oplus \tau_{d-1}$ , and hence (v). Moreover, by Lemma B.3,  $\text{Im } C_0 \subseteq \mathcal{F}_d^\perp$  and because  $\text{Im } C_0 \subset \mathcal{F}_d^\perp$  contradicts  $C_0 S_d = P_{\mathcal{F}_d^\perp}$ , one has  $\text{Im } C_0 = \mathcal{F}_d^\perp$ . Using  $\mathcal{F}_d^\perp = \tau_d$ , see (v), one finds (vi).

(vi)  $\Rightarrow$  (ii). Let  $\text{Im } C_0 = \tau_d$  and proceed by contradiction, assuming that the identity is not in equation  $d$ , so that the left equation in (B.7) reads  $S_d C_0 = 0$ , which implies  $\text{Im } C_0 \subseteq \text{Ker } S_d$ , where  $\text{Im } C_0 = \tau_d$  and  $\text{Ker } S_d = \tau_d^\perp$ . Hence  $\tau_d \subseteq \tau_d^\perp$ , so that  $\tau_d = \{0\}$ . This contradicts  $C_0 \neq 0$ , i.e., that the pole has order  $d$ , and proves that (ii) holds. ■

**THEOREM B.5** (Pole cancellations in  $A(z)^{-1}$ ). Consider Definition 5.1 and for  $h = 0, 1, \dots, d$  define

$$\gamma_h(z) = P_{\tau_h} + S_h^+ \sum_{n=1}^{d-h-1} A_{h+1,n} (1-z)^n; \tag{B.8}$$

then  $\gamma_h(z)A(z)^{-1}$  has a pole of order  $h$  at  $z = 1$ , i.e.,

$$\gamma_h(z)A(z)^{-1} = (1-z)^{-h} \tilde{\gamma}_h(z), \quad \tilde{\gamma}_h(1) \neq 0, \tag{B.9}$$

where  $\tilde{\gamma}_h(z)$  is absolutely convergent on  $D(0, \rho)$  for some  $\rho > 1$ . Moreover,  $\text{Im } \tilde{\gamma}_h(1) = \tau_h$ .

*Proof.* Applying  $S_h^+$  to (B.2) and using  $S_h^+ S_h = P_{\tau_h}$  and  $S_h^+ P_{\mathcal{F}_h^\perp} = S_h^+$ , see Lemma B.1, one finds

$$P_{\tau_h} C_n + S_h^+ \sum_{k=1}^n A_{h+1,k} C_{n-k} = \delta_{n+h,d} S_h^+, \quad h = 0, 1, \dots, d-n. \tag{B.10}$$

Write  $A(z)^{-1} = \sum_{n=0}^\infty C_n (1-z)^{n-d}$  as

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

and apply  $P_{\tau_h}$  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  $h = 0, \dots, d - 1$ . Setting  $n = 0$  in (B.10) one has  $P_{\tau_h} C_0 = 0$  and hence

$$P_{\tau_h} A(z)^{-1} = \sum_{n=1}^{d-h-1} P_{\tau_h} C_n (1-z)^{n-d} + (1-z)^{-h} P_{\tau_h} R_0(z). \tag{B.11}$$

From (B.10), for  $n \leq d - h$  one has  $P_{\tau_h} C_n = -S_h^+ \sum_{k=1}^n A_{h+1,k} C_{n-k} + \delta_{n+h,d} S_h^+$  and because  $\delta_{n+h,d} = 0$  for  $n = 1, \dots, d - h - 1$ , one has

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

Rearranging one thus finds

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

Next write

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

so that

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

Substituting in (B.11) and rearranging one thus finds  $\gamma_h(z) A(z)^{-1} = (1-z)^{-h} \tilde{\gamma}_h(z)$ , where

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

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

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

from (B.10) for  $n = d - h$  one finds  $P_{\tau_h} C_{d-h} + S_h^+ \sum_{k=1}^{d-h} A_{h+1,k} C_{d-h-k} = S_h^+$ , so that

$$\tilde{\gamma}_h(1) = S_h^+ (I - A_{h+1,d-h} C_0).$$

Hence  $\text{Im } \tilde{\gamma}_h(1) \subseteq \text{Im } S_h^+$ . From Lemma 3 in Ben-Israel and Greville (2003, Chap. 9) one has that  $\text{Im } S_h^+ = \mathcal{H} \cap (\text{Ker } S_h)^\perp = \mathcal{H} \cap \tau_h = \tau_h$ , so that  $\text{Im } \tilde{\gamma}_h(1) \subseteq \tau_h$ . Using  $S_h^+ S_h = P_{\tau_h}$  and  $C_0 S_h = 0$  one finds  $\tilde{\gamma}_h(1) S_h = P_{\tau_h}$ , so that  $\text{Im } \tilde{\gamma}_h(1) = \tau_h$ . This shows that  $\tilde{\gamma}_h(1) \neq 0$  and hence  $\gamma_h(z) A(z)^{-1}$  has a pole of order  $h = 0, \dots, d - 1$  at  $z = 1$ .

Next consider  $h = d$ . One has  $\gamma_d(z) = P_{\tau_d}$  and  $\gamma_d(z)A(z)^{-1} = P_{\tau_d}C_0(1 - z)^{-d} + P_{\tau_d} \sum_{n=1}^{\infty} C_n(1 - z)^{n-d}$ . Setting  $n = 0$  and  $h = d$  in (B.10) one finds that

$$\tilde{\gamma}_d(1) = P_{\tau_d}C_0 = S_d^+,$$

and hence  $\text{Im } \tilde{\gamma}_d(1) = \text{Im } S_d^+ = \tau_d$ , where the last equality follows as above from  $\text{Im } S_d^+ = \mathcal{H} \cap (\text{Ker } S_d)^\perp = \mathcal{H} \cap \tau_d = \tau_d$ . Hence  $\tilde{\gamma}_d(1)$  is different from 0, which shows that  $\gamma_d(z)A(z)^{-1} = P_{\tau_d}A(z)^{-1}$  has a pole of order  $d$  at  $z = 1$ . This completes the proof. ■

### APPENDIX C: Proofs

This Appendix contains proofs of the results in the text. The proof of Theorem 3.5 makes use of the following fact, which is proven in Franchi and Paruolo (2019): for  $t \in \mathbb{Z}$ , one has

$$S^s \Delta^h u_t = S^{s-h} u_t - \sum_{n=s-h}^{s-1} \varsigma_{n,t} \Delta^{h-s+n} u_0, \quad 0 < h \leq s, \tag{C.1}$$

where  $\varsigma_{n,t}$  is a polynomial of order  $n$  in  $t$ .

**Proof of Theorem 3.5.** The result is a direct consequence of Theorem A.1 in Appendix A.1. By definition, an AR process with a unit root of finite type  $A(L)x_t = \varepsilon_t$  is such that  $A(1) \neq 0$ ,  $A(z)$  has an eigenvalue of finite type at  $z = 1$  and  $A(z)$  is invertible in the punctured disc  $D(0, \rho) \setminus \{1\}$  for some  $\rho > 1$ . Letting  $z_0 = 1$  and  $C_n = U_n(-1)^{n-d}$ , Theorem A.1 states that there exist a finite integer  $d = 1, 2, \dots$  and operators  $C_n$ ,  $n = 0, 1, \dots$ , with finite rank for  $n = 0, \dots, d - 1$ , such that

$$A(z)^{-1} = \sum_{n=0}^{\infty} C_n(1 - z)^{n-d}, \quad z \in D(1, \delta) \setminus \{1\}. \tag{C.2}$$

Write  $A(z)^{-1} = \sum_{n=0}^{d-1} C_n(1 - z)^{n-d} + C^\circ(z)$ , where  $C^\circ(z) = \sum_{n=d}^{\infty} C_n(1 - z)^{n-d}$  is absolutely convergent on  $D(0, \rho)$  for some  $\rho > 1$ ; applying  $A(L)^{-1}$  on both sides of  $A(L)x_t = \varepsilon_t$  one finds the common trends representation  $x_t = C_0 s_{d,t} + C_1 s_{d-1,t} + \dots + C_{d-1} s_{1,t} + y_t + \mu_t$ , where  $s_{h,t} = S^h \varepsilon_t \sim I(h)$  is the  $h$ -fold integrated bilateral random walk,  $y_t = C^\circ(L)\varepsilon_t$  is a linear process and  $\mu_t = \sum_{n=0}^{d-1} v_n t^n$ ,  $v_n \in \mathcal{H}$ , is a polynomial in  $t$  whose coefficients  $v_0, \dots, v_{d-1} \in \mathcal{H}$  depend on the initial values of  $x_t, y_t, \varepsilon_t$  for  $t = -d, \dots, 0$ , see (C.1). ■

**Proof of Corollary 3.6.** The order  $d$  of the pole of the inverse of  $A(z)$  is finite by Theorem A.1 in Appendix A.1; this implies  $x_t \sim I(d)$  via (3.6) because the  $I(d)$  trend  $s_{d,t}$  is non-degenerate, i.e., the probability that it belongs to a strict subspace of  $\mathcal{H}$  for all  $t$  is equal to zero, and hence  $C_0 s_{d,t} \neq 0$  (with probability 1). One has  $\dim \text{Im } C_0 < \infty$  by Theorem A.1 in Appendix A.1. Observe that

$$\langle v, x_t \rangle = \langle v, C_0 s_{d,t} \rangle + \sum_{n=1}^{d-1} \langle v, C_n s_{d-n,t} \rangle + \langle v, y_t \rangle + \langle v, \mu_t \rangle,$$

where  $\langle v, C_n s_{d-n,t} \rangle$  is at most  $I(d-n)$ , for  $n = 0, \dots, d-1$ . One sees that  $\langle v, C_0 s_{d,t} \rangle = 0$  for all  $t$  if and only if  $v \in (\text{Im } C_0)^\perp$ , because  $s_{d,t}$  is non-degenerate. This shows that  $(\text{Im } C_0)^\perp$  is the cointegrating space and  $\text{Im } C_0$  is the attractor space. One has  $\dim(\text{Im } C_0)^\perp = \infty$  when  $\dim \mathcal{H} = \infty$  because it complements  $\text{Im } C_0$ , which is finite-dimensional. ■

**Proof of Proposition 3.7.** First note that for  $d = 1$ , Theorem 3.5 and Corollary 3.6 imply  $\Delta x_t = B(L)\varepsilon_t$ , where  $B(z) = \sum_{n=0}^\infty B_n z^n$  is absolutely convergent on  $D(0, \rho)$ ,  $\rho > 1$ ,  $B(1) = C_0 \neq 0$  and  $\text{Im } B(1) = \text{Im } C_0$  is finite-dimensional. Because  $B(z)$  is infinitely differentiable on  $D(0, \rho)$ ,  $\rho > 1$ , the series obtained by termwise differentiation coincides with the first derivative of  $B(z)$  for each  $z \in D(0, \rho)$ , and hence one has  $\sum_{n=1}^\infty n \|B_n\|_{\mathcal{L}\mathcal{H}} < \infty$ . ■

**Proof of Proposition 3.8.** Because the sum of compact operators is compact, see Theorem 16.1 in Gohberg et al. (2003, Chap. II), and if  $K$  is compact then  $I - K$  is Fredholm of index 0, see Theorem 4.2 in Gohberg et al. (2003, Chap. XV), then  $A_0 = I - \sum_{n=1}^k A_n^\circ$  is Fredholm of index 0. Because  $A_0$  is non-invertible, by the Fredholm alternative there exists a nonzero  $v \in \mathcal{H}$  such that  $A_0 v = 0$ , see Theorem 4.1 in Gohberg et al. (2003, Chap. XIII). Finally, since  $z = 1$  is assumed to be the only isolated singularity of  $A(z)^{-1}$  within  $D(0, \rho)$ ,  $\rho > 1$ , this shows that  $z = 1$  is a eigenvalue of finite type of  $A(z)$ . ■

**Proof of Theorem 4.1.** Set  $d = 1$  in Theorem 5.3. ■

**Proof of Proposition 4.6.** The notation  $\zeta_0 = \text{Im } A_0$  and  $\tau_0 = (\text{Ker } A_0)^\perp$ , see (4.2), is employed. In the present notation, (4.7) reads  $\mathcal{H} = \zeta_0 \oplus A_1 \tau_0^\perp$ , where by assumption of AR process with a unit root of finite type, see Remark 3.2, one has  $0 < \dim \tau_0^\perp = \dim \zeta_0^\perp < \infty$ . Observe that (4.4) is equivalent to the condition:

$$\text{the identity is in equation 1 of system (B.1),} \tag{C.3}$$

see (ii) in Theorem B.4 in Appendix B.

*Proof that (C.3) implies (4.7).* Let (C.3) hold, i.e.,  $A_0 C_1 + A_1 C_0 = I$ . This implies that for any  $v \in \mathcal{H}$  one has  $v = u + s$ , where  $u = A_0 C_1 v \in \zeta_0$  because  $\text{Im } A_0 = \zeta_0$  and  $s = A_1 C_0 v \in A_1 \tau_0^\perp$  because  $\text{Im } C_0 = \tau_0^\perp$ , see (v) and (vi) in Theorem B.4. It remains to show that  $\text{Im } A_0 C_1 \cap \text{Im } A_1 C_0 = \{0\}$ .

Assume that  $v \in \text{Im } A_0 C_1 \cap \text{Im } A_1 C_0$ , i.e.,  $v = A_0 C_1 w = A_1 C_0 q$ , for some  $w, q \in \mathcal{H}$ . Subtracting the two representations, one finds  $A_0 C_1 w - A_1 C_0 q = 0$ . Applying  $P_{\zeta_0^\perp}$  to both sides of this equation one finds  $0 = P_{\zeta_0^\perp} A_1 C_0 q = P_{\zeta_0^\perp} q$ , which implies that  $q \in \zeta_0$ ; here use is made of  $\text{Im } A_0 = \zeta_0$  and  $P_{\zeta_0^\perp} A_1 C_0 = P_{\zeta_0^\perp}$ , which follows from  $A_0 C_1 + A_1 C_0 = I$ . Substituting  $q \in \zeta_0$  in the second representation of  $v$ ,  $v = A_1 C_0 q$ , one finds that  $v = 0$  because  $\text{Ker } C_0 = \zeta_0$ , see (iii) and (iv) in Theorem B.4. Hence only  $v = 0$  is in  $\text{Im } A_0 C_1 \cap \text{Im } A_1 C_0$ , and this proves that (4.7) holds.

*Proof that (4.7) implies (C.3).* By contradiction, assume (C.3) does not hold, i.e.,  $A_0 C_1 + A_1 C_0 = 0$ . This implies that there exists a nonzero  $v \in \mathcal{H}$  such that  $v = u + s = 0$ , where  $u = A_0 C_1 v \in \zeta_0$  and  $s = A_1 C_0 v \in A_1 \tau_0^\perp$ , this is a contradiction to  $\mathcal{H} = \zeta_0 \oplus A_1 \tau_0^\perp$  and hence (4.4) holds. ■

**Proof of Proposition 4.7.** When  $k = 1$ , (4.7) reads  $\mathcal{H} = \text{Im } A_0 \oplus \text{Ker } A_0$ . By assumption,  $\text{Ker } A_0$  has finite dimension; hence  $\mathcal{H} = \text{Im } A_0 \oplus \text{Ker } A_0$  implies that  $(\text{Im } A_0)^\perp$  has finite dimension equal to  $\dim \text{Ker } A_0$ . This shows that  $A_0$  is Fredholm of index 0. Because  $\dim \text{Ker } A_0 > 0$ , there exists a nonzero  $v \in \mathcal{H}$  such that  $A_0 v = 0$  and since  $z = 1$  is assumed

to be the only isolated singularity of  $A(z)^{-1}$  within  $D(0, \rho)$ ,  $\rho > 1$ , this shows that  $z = 1$  is a eigenvalue of finite type of  $A(z)$ . ■

**Proof of Theorem 4.8.** Set  $d = 2$  in Theorem 5.3. ■

**Proof of Theorem 5.3.** The proof makes use of Theorem B.4 in Appendix B, which establishes the order of the pole of  $A(z)^{-1}$  at  $z = 1$ , and Theorem B.5 in Appendix B, which describes the pole cancellations that give rise to cointegration. The common trends representation of  $x_t \sim I(d)$  is found in (3.6) and, see Corollary 3.6,  $\text{Im } C_0$  and  $(\text{Im } C_0)^\perp$  are respectively the attractor space and the cointegrating space. By Theorem B.4 one has that  $x_t \sim I(d)$  if and only if  $\text{Im } C_0 = \tau_d = (\tau_0 \oplus \dots \oplus \tau_{d-1})^\perp \neq \{0\}$ ; thus  $\tau_d$  is the finite-dimensional attractor space and  $\tau_0 \oplus \dots \oplus \tau_{d-1}$  is the cointegrating space, which is infinite-dimensional when  $\mathcal{H}$  is infinite-dimensional.

Recall that from (B.9) in Theorem B.5 one has  $\gamma_h(z)A(z)^{-1} = (1 - z)^{-h}\tilde{\gamma}_h(z)$ , where  $\tilde{\gamma}_h(z) = \tilde{\gamma}_h(1) + (1 - z)\tilde{\gamma}_h^\circ(z)$ , say, is analytic and  $\text{Im } \tilde{\gamma}_h(1) = \tau_h$ . Hence  $\gamma_h(L)x_t = \gamma_h(L)A(L)^{-1}\varepsilon_t = \tilde{\gamma}_h(L)s_{h,t} + \eta_t$  where  $\eta_t$  depends on the initial values which can be chosen such that  $\langle v, \eta_t \rangle = 0$  for all  $v \in \tau_h$ . Note that  $\tilde{\gamma}_h(1)s_{h,t} \neq 0$  because  $s_{h,t} \sim I(h)$  is non-degenerate and  $\tilde{\gamma}_h(1) \neq 0$ ; hence  $\gamma_h(L)x_t \sim I(h)$ .

For  $v \in \tau_h$  one finds  $\langle v, \gamma_h(L)x_t \rangle = \langle v, \tilde{\gamma}_h(L)s_{h,t} \rangle$  where  $\langle v, \tilde{\gamma}_h(L)s_{h,t} \rangle = \langle v, \tilde{\gamma}_h(1)s_{h,t} \rangle + \langle \tilde{\gamma}_h^\circ(z)s_{h-1,t} \rangle$ , and  $\langle v, \tilde{\gamma}_h(1)s_{h,t} \rangle \sim I(h)$  because  $\text{Im } \tilde{\gamma}_h(1) = \tau_h$  and  $s_{h,t}$  is non-degenerate. This implies  $\langle v, \gamma_h(L)x_t \rangle \sim I(h)$ . Note that from (B.8) one has

$$\gamma_h(L)x_t = P_{\tau_h}x_t + \sum_{n=1}^{d-h-1} S_h^+ A_{h+1,n} \Delta^n x_t$$

and hence

$$\langle v, \gamma_h(L)x_t \rangle = \langle v, P_{\tau_h}x_t \rangle + \sum_{n=1}^{d-h-1} \langle S_h^+ A_{h+1,n} \Delta^n x_t \rangle = \langle v, x_t \rangle + \sum_{n=1}^{d-h-1} \langle S_h^+ A_{h+1,n} \Delta^n x_t \rangle.$$

In fact  $\langle v, P_{\tau_h}x_t \rangle = \langle P_{\tau_h}v, x_t \rangle = \langle v, x_t \rangle$ , where the first equality follows from the fact that  $P_{\tau_h}$  is an orthogonal projection and hence self-adjoint, and the second one from  $v \in \tau_h$ . This completes the proof. ■

**Proof of Proposition 5.11.** Recall that  $x_t = A_1^\circ x_{t-1} + \varepsilon_t$  is an AR process with a unit root of finite type if  $A(z) = I - A_1^\circ z$  has an eigenvalue of finite type at  $z = 1$ ; in this case, see Gohberg et al. (1990, p. 27–28),  $\mathcal{H} = \mathcal{H}_P \oplus \mathcal{H}_T$ , where  $\mathcal{H}_P$  is the finite-dimensional image of the Riesz projection associated with the isolated eigenvalue  $z = 1$ . Relative to this decomposition,  $I - A_1^\circ z$  admits the following operator matrix representation

$$I - A_1^\circ z = \begin{pmatrix} I - A_P z & \\ & I - A_T z \end{pmatrix},$$

where  $A_X = A_1^\circ|_{\mathcal{H}_X}$  is the restriction of  $A_1^\circ$  to  $\mathcal{H}_X$ ,  $X = T, P$ . Note that  $A_P$  acts on a finite-dimensional space and it has precisely one point in its spectrum, namely 1. So  $\mathcal{H}_P$  has a basis of eigenvectors and generalized eigenvectors such that the matrix of  $A_P$  relative to this basis has a Jordan normal form with 1 on the main diagonal.

Because  $A_P - I$  is a nilpotent matrix of order  $d$  if and only if the largest Jordan block of  $A_P$  has dimension  $d$ , see e.g., Horn and Johnson (2013, p. 181), and the size of the largest Jordan block is equal to  $d$  if and only if the POLE( $d$ ) condition holds, see Franchi and Paruolo (2019), one has the statement. ■