JOURNAL OF TIME SERIES ANALYSIS
J. Time Ser. Anal. (2024)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1111/jtsa.12766

ORIGINAL ARTICLE

THE GRANGER–JOHANSEN REPRESENTATION THEOREM FOR INTEGRATED TIME SERIES ON BANACH SPACE

PHIL HOWLETT,[a] BRENDAN K. BEARE,[b] MASSIMO FRANCHI,[c] JOHN BOLAND[d] AND KONSTANTIN AVRACHENKOV[e]

[a]*Scheduling and Control Group (SCG), Centre for Industrial and Applied Mathematics (CIAM), UniSA STEM, University of South Australia, Mawson Lakes, SA, Australia*
[b]*School of Economics, University of Sydney, Camperdown, NSW, Australia*
[c]*Sapienza University of Rome, Rome, Italy*
[d]*Centre for Industrial and Applied Mathematics, UniSA, STEM, University of South Australia, Mawson Lakes, SA, Australia*
[e]*INRIA, Sophia Antipolis, France*

We prove an extended Granger–Johansen representation theorem (GJRT) for finite- or infinite-order integrated autoregressive time series on Banach space. We assume only that the resolvent of the autoregressive polynomial for the series is analytic on and inside the unit circle except for an isolated singularity at unity. If the singularity is a pole of finite order the time series is integrated of the same order. If the singularity is an essential singularity the time series is integrated of order infinity. When there is no deterministic forcing the value of the series at each time is the sum of an almost surely convergent stochastic trend, a deterministic term depending on the initial conditions and a finite sum of embedded white noise terms in the prior observations. This is the extended GJRT. In each case the original series is the sum of two separate autoregressive time series on complementary subspaces – a singular component which is integrated of the same order as the original series and a regular component which is not integrated. The extended GJRT applies to all integrated autoregressive processes irrespective of the spatial dimension, the number of stochastic trends and cointegrating relations in the system and the order of integration.

Received 21 August 2023; Accepted 20 July 2024

Keywords: AR time series; resolvent operators; GJRT; cointegrated series.

MOS subject classification: MSC[2020]:62M10; 91B84; 47A11; 47A55.

1. INTRODUCTION

We wish to establish an extended form of the Granger–Johansen representation theorem (GJRT) for vector autoregressive (AR) time series taking values in a Banach space. These series are often referred to as functional time series because, for instance, each real-valued integrable function on the unit interval of the real line can be represented by an infinite-dimensional vector whose components are the coefficients of the Fourier sine series. In this particular, case the sum of the magnitudes of the vector components will be finite. Thus we can study the space of integrable functions on the unit interval by considering the vector space of all infinite-dimensional vectors where the sum of the magnitudes of the components is finite. Consequently the study of vector AR processes on Banach space can be regarded as a generalized form of the study of AR time series for functions.

An important early contribution to the analysis of functional time series is Bosq (2000), where a theoretical treatment of linear processes in Banach and Hilbert spaces is developed. In particular the derivation of laws of large numbers and central limit theorems allows estimation and inference for infinite-dimensional stationary AR models. These models also enable direct representation of the dynamics of infinite-dimensional objects, such as a

*Correspondence to: Phil Howlett, Scheduling and Control Group (SCG), Centre for Industrial and Applied Mathematics (CIAM), UniSA STEM, University of South Australia, Mawson Lakes, SA, Australia.
E-mail: phil.howlett@unisa.edu.au

time series of continuous functions on a compact spatial domain, and they allow greater generality for modelling of conditional means and variances (Horváth and Kokoszka, 2012).

Economic applications of functional time series include studies on the term structure of interest rates (Kargin and Onatski, 2008), intraday volatility (Gabrys *et al.*, 2013; Hörmann *et al.*, 2013) and the human influence on climate (Chang *et al.*, 2020). Additional applications can be found in the recent monograph by Kokoszka and Reimherr (2017).

One might expect the analysis of functional time series to begin by testing for stationarity. In this regard Horváth *et al.* (2014); Kokoszka and Young (2016) and Kokoszka and Reimherr (2017) suggest extended forms of the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test of stationarity proposed originally for single variable series by Kwiatkowski *et al.* (1992) while Aue and van Delft (2020) devise a test of stationarity in the frequency domain. However, the literature on this topic is somewhat sparse. An alternative approach by Chang *et al.* (2016) proposed a test of non-stationarity based on the generalized eigenvalues for the covariance operator of the observed values and on the long-run covariance operator for the associated first differences. If the dimension of the non-stationary subspace is finite then the dimension is called the number of common stochastic trends. In a subsequent working paper, Hu and Park (2016) consider an infinite-dimensional AR(1) process with a compact operator and show that the common trends representation comprises a finite number of stochastic trends each integrated of order one and an infinite-dimensional cointegrating space. Hu and Park also propose an estimator for the functional autoregressive operator. More recently Nielsen *et al.* (2022) suggest a statistical procedure to determine the dimension of the non-stationary subspace of cointegrated functional time series taking values in a Hilbert space of square-integrable functions defined on a compact interval. As in Chang *et al.* (2016) they assume that the cointegrated series has a finite-dimensional non-stationary subspace while the stationary subspace is infinite dimensional. That is, there are infinitely many linearly independent cointegrating relations. The test is applied to several empirical examples: age-specific US employment rates, Australian temperature curves and Ontario electricity demand. For more details of functional time series we refer to extended reviews in Beare and Seo (2020), Franchi and Paruolo (2020) and Nielsen *et al.* (2022).

Our primary concern here is to review the development of the GJRT for AR processes. According to Hansen (2005) the GJRT for a finite-dimensional cointegrated vector $AR(1)$ time series $\boldsymbol{x} \sim I(1)$ can be stated informally as follows.

> *A cointegrated vector autoregressive process can be decomposed into four components: a random walk, a strictly stationary process, a deterministic part, and a term that depends on the initial conditions.*

The original theorem proposed by Engle and Granger (1987) was a representation theorem for finite-dimensional cointegrated vector processes $\boldsymbol{x} \sim I(1)$ that connected the moving average, autoregressive, and error correction forms. The focus for Johansen was to find a more explicit representation for finite-dimensional $AR(p)$ processes $\boldsymbol{x} = \{\boldsymbol{x}(t)\}_{t\in\mathbb{Z}}$ satisfying

$$\boldsymbol{x}(t) = \Phi_1\boldsymbol{x}(t-1) + \cdots + \Phi_p\boldsymbol{x}(t-p) + \boldsymbol{\xi}(t), \tag{1.1}$$

for all $t \in \mathbb{Z}$ where $\Phi_1, \dots, \Phi_p \in \mathbb{R}^{n\times n}$ are matrix coefficients and $\boldsymbol{\xi}(t) \in \mathbb{R}^n$ is a strong white noise process. The specific focus on $AR(p)$ processes enabled Johansen to find necessary and sufficient conditions for $\boldsymbol{x} \sim I(1)$ (Johansen, 1991) and for $\boldsymbol{x} \sim I(2)$ (Johansen, 1992). In each case the conditions depend only on the properties of the matrix coefficients $\Phi_1, \dots, \Phi_p$.

Beare *et al.* (2017) used the model described in (1.1) for a process taking values in a Hilbert space H. In this case $\boldsymbol{x}(t), \boldsymbol{\xi}(t) \in H$ for all $t \in \mathbb{Z}$ and $\Phi_1, \dots, \Phi_p \in \mathcal{B}(H)$ are bounded linear operators. Beare et al. defined an autoregressive polynomial $\Phi : \mathbb{C} \to \mathcal{B}(H)$ by the formula

$$\Phi(z) = I - \sum_{i=1}^{p} \Phi_i z^i, \tag{1.2}$$

J. Time Ser. Anal. (2024)
Journal of Time Series Analysis published by John Wiley & Sons Ltd.
DOI: 10.1111/jtsa.12766

for each $z \in \mathbb{C}$. For $p = 1$ they show that if the spaces $\Phi(1)(H)$ and $\Phi'(1)\Phi(1)^{-1}(\{\mathbf{0}\})$ are closed complementary subspaces in H then the resolvent $\Phi(z)^{-1}$ has a simple pole at $z = 1$ and $\boldsymbol{x} \sim I(1)$. For $p > 1$ they reach the same conclusion using the additional assumption that $\Phi_1, \ldots, \Phi_p$ are compact. In a subsequent paper, Beare and Seo (2020) assume that $\Phi_1, \ldots, \Phi_p$ are compact and show that $\Phi(z)$ has a simple pole at $z = 1$ if and only if $\Phi(1)(H)$ and $\Phi'(1)\Phi(1)^{-1}(\{\mathbf{0}\})$ are closed complementary subspaces in H. They also show that if $\Phi(z)^{-1}$ is analytic for $z \in D_{1+\epsilon}(0) \setminus \{1\} = \{z \in \mathbb{C} | |z| < 1 + \epsilon\} \setminus \{1\}$ for some $\epsilon > 0$ and has a pole of order two at $z = 1$ with

$$\Phi(z)^{-1} = \Upsilon_{-2}/(z-1)^2 + \Upsilon_{-1}/(z-1) + \sum_{\ell \in \mathbb{N}_{-1}} \Upsilon_\ell (z-1)^\ell, \tag{1.3}$$

for all $z \in D_{0,\epsilon}(1) = \{z \in \mathbb{C} | 0 < |z-1| < \epsilon\}$ then $\boldsymbol{x} \sim I(2)$. They use the Moore–Penrose inverse to derive a formula for the coefficient Υ_{-2} and a more complicated formula for Υ_{-1}. Franchi and Paruolo (2020) extend these results to show that if $\Phi(z)^{-1}$ has a pole of order d at $z = 1$ with a corresponding unit root of finite type then $\boldsymbol{x} \sim I(d)$. The assumption that $\Phi_1, \ldots, \Phi_p$ are compact and the alternative assumption that the unit root for $\Phi(z)$ is of finite type are each sufficient to ensure that the subspace $\Upsilon_{-1}(H) \subset H$ is finite dimensional. Thus determination of the generalized eigenspace for $\Phi(1)$ and the associated Laurent series coefficients is reduced to a matrix problem. The procedure proposed by Franchi and Paruolo (2020) is essentially equivalent to earlier work by Avrachenkov *et al.* (2001), Avrachenkov *et al.* (2013), Franchi and Paruolo (2016) and Howlett (1982) on the inversion of matrix power series. See also Howlett *et al.* (2009). The works by Beare *et al.* (2017), Beare and Seo (2020) and Franchi and Paruolo (2020) are restricted to Hilbert spaces where $\Upsilon_{-1}(H)$ is finite dimensional.

Seo (2023) extends the results in Beare and Seo (2020) to AR processes $\boldsymbol{x} \sim I(1) \vee I(2)$ taking values in a Banach space X. The main results depend on two critical assumptions – for $\boldsymbol{x} \sim I(1) \vee I(2)$, that the subspaces $\Phi(1)(X)$ and $\Phi(1)^{-1}(\{\mathbf{0}\})$ can each be complemented in the space X and, for $\boldsymbol{x} \sim I(2)$, that certain additional key subspaces in X can also be complemented. Complementation of closed subspaces is not guaranteed in Banach space. Consequently, Seo (2023) leaves major theoretical questions unanswered. Seo considers only first and second-order poles of the resolvent operator at $z = 1$ and does not consider isolated essential singularities.

It is well known that a necessary and sufficient condition for an AR process $\boldsymbol{x}$ to be integrated of order $d < \infty$ is that $\Phi(z)^{-1}$ has a pole of order d at $z = 1$. In the early papers on cointegration a plethora of seemingly *ad hoc* conditions were used collectively to determine the order of integration. In some sense, the collective nature of these conditions obscured the fact that the underlying necessary and sufficient condition was simply the order of the pole.

The aim of this article is to obtain an extended form of the GJRT that is valid for all AR processes on Banach space. In this regard it is sufficient to establish the extended form for first-order vector autoregressive AR(1) processes[1] on Banach space. We will use recent research on the inversion of operator pencils (Albrecht *et al.*, 2014; Albrecht *et al.*, 2020) that has not previously been used for the analysis of integrated time series.

By confining our attention to the naturally complemented subspaces defined by the spectral projections for the autoregressive polynomial at $z = 1$, we remove the need for restrictive assumptions about complementation of certain subspaces or their finite dimensions as in Beare *et al.* (2017), Beare and Seo (2020), Franchi and Paruolo (2020), and Seo (2023). The theory presented in this article applies to all integrated autoregressive processes irrespective of the dimension of the space, the number of stochastic trends and cointegrating relations in the system, and the order of integration. Hence the present results are a substantial extension of those reported in Beare *et al.* (2017), Beare and Seo (2020), Franchi and Paruolo (2020), and Seo (2023). We provide several particular examples to illustrate the key structural issues including a composite graph showing the simulated evolution of an archetypal $I(\infty)$ process.

The rest of the article is organized as follows. Section 2 outlines the key components of the new representation and summarizes the significance of our contribution, Section 3 describes the mathematical tools used in the solution procedure and in Section 4 these tools are applied to solution of the time series model. Section 4 also includes

[1] Any AR(p) process on a space X can be modeled as an AR(1) process using augmented autoregressive coefficients on the product space X^p.

DOI: 10.1111/jtsa.12766
Journal of Time Series Analysis published by John Wiley & Sons Ltd.

formal statements of the main results. Section 5 shows that existing $I(1)$ and $I(2)$ representations are special cases of the extended GJRT. The general structures are illustrated in Section 6 via three examples and in Section 7 the new methods are applied to solve a non-autonomous problem for dynamic control of electrical power generation in a wind farm. In Section 8, we draw some brief conclusions. In the Appendix A our notation and terminology are listed and explained in Section A, the proofs of all formally stated results are given in Section B and the relevant properties of the weighted Volterra (integral) operator used in the wind farm problem are justified in Section C.

2. MAIN RESULTS

Our main contribution, presented formally as Proposition 4.2, is an extended GJRT on Banach space that applies to integrated AR(1) series of any spatial dimension, any number of stochastic trends and cointegrating relations, and any order of integration. Let $A_0, A_1 \in \mathcal{B}(X, Y)$ and suppose that $\boldsymbol{x} = \{\boldsymbol{x}(t)\}_{t\in\mathbb{Z}} \in X^{\mathbb{Z}}$ satisfies

$$A_0\boldsymbol{x}(t) + A_1\boldsymbol{x}(t-1) = \boldsymbol{\xi}(t), \tag{2.1}$$

for all $t \in \mathbb{N} - 1$ where $\boldsymbol{x}(-1) = \boldsymbol{c}$ and $\boldsymbol{\xi} = \{\boldsymbol{\xi}(t)\}_{t\in\mathbb{Z}} \in Y^{\mathbb{Z}}$ is a strong white noise. The autoregressive polynomial is $A(z) = A_0 + A_1 z = B_0 + B_1(z-1) \in \mathcal{B}(X, Y)$ for all $z \in \mathbb{C}$ where $B_0 = A_0 + A_1 \in \mathcal{B}(X, Y)$ and $B_1 = A_1 \in \mathcal{B}(X, Y)$ and the resolvent is $R(z) = A(z)^{-1} \in \mathcal{B}(Y, X)$ for all $z \in \mathbb{C} \setminus \sigma$ where σ is the spectral set for $A(z)$.

Assumption 2.1. The resolvent $R(z) = A(z)^{-1}$ is analytic for $z \in [D_1(0)^a \setminus \{1\}] \cup D_{0,\delta}(1) = \{z \in \mathbb{C} | |z| \leq 1$ or $0 < |z-1| < \delta\} \setminus \{1\}$ for some $\delta > 0$ and has an isolated singularity at $z = 1$.

The extended representation: The series $\boldsymbol{x} = \{\boldsymbol{x}(t)\}_{t\in\mathbb{N}-1}$ is the sum of a stochastic trend

$$\sum_{k\in\mathbb{N}} (-1)^k T_{-k}\Delta^{-k}\boldsymbol{\xi}_+(t), \qquad T_{-k} = (-1)^{k-1}(T_{-1}B_0)^{k-1}T_{-1}, \tag{2.2}$$

where Δ is the difference operator, Δ^{-1} is the inverse difference or cumulation operator, $\boldsymbol{\xi}_+(t) = \boldsymbol{\xi}(t)$ for $t \in \mathbb{N} - 1$ and $\boldsymbol{\xi}(t) = \mathbf{0}$ otherwise, and T_{-1} is the residue of $R(z)$ at $z = 1$, a deterministic component

$$(-1)R_t B_1 \boldsymbol{c}, \qquad R_t = (-1)^t (A_0^{-1}A_1)^t A_0^{-1}, \tag{2.3}$$

and a finite weighted sum of strong white noise terms

$$\sum_{s=0}^{t} W_s\, \boldsymbol{\xi}(t-s), \qquad W_s = R_s + (I_X - T_{-1}B_0)^{-s-1}T_{-1}. \tag{2.4}$$

Additional observations: Our new results extend the GJRT in various ways.

1. If $R(z)$ has an isolated singularity at $z = 1$ then Lemma 3.1 shows that $R(z)$ is analytic for $z \in [D_1(0)^a \setminus \{1\}] \cup D_{0,\delta}(1)$ for some $\delta > 0$ if and only if there is some $\epsilon > 0$ such that $R(z)$ is analytic for $z \in D_{1+\epsilon}(0) \setminus \{1\}$.
2. The Laurent series $R(z) = \sum_{j\in\mathbb{Z}} T_j(z-1)^j$ for $z \in D_{0,\epsilon}(1)$ is found by solving the fundamental equations (3.1) and (3.2) for $\{T_j\}_{j\in\mathbb{Z}}$ subject to the magnitude constraints (3.3). The resolvent $R(z)$ is analytic for $z \in D_{0,\epsilon}(1)$ if and only if the fundamental equations have a basic solution $\{T_{-1}, T_0\}$. The resolvent is completely determined by the basic solution. A general solution procedure for the fundamental equations is available in separable Banach space but simpler methods can always be used in finite dimensional problems.
3. There is a closed formula $R(z) = [(z-1)I_X + T_{-1}B_0]^{-1}T_{-1} + [I_X + T_0B_1(z-1)]^{-1}T_0$ for the resolvent when $z \in D_{0,\epsilon}(1)$. The singular part is $R_{\sin}(z) = [(z-1)I_X + T_{-1}B_0]^{-1}T_{-1} = \sum_{k\in\mathbb{N}} T_{-k}/(z-1)^k$ when $z \neq 1$ where $T_{-k} = (-1)^{k-1}(T_{-1}B_0)^{k-1}T_{-1}$. The regular part is $R_{\text{reg}}(z) = [I_X + T_0B_1(z-1)]^{-1}T_0 = \sum_{\ell\in\mathbb{N}-1} T_\ell(z-1)^\ell$ when $z \in D_\epsilon(1)$ where $T_\ell = (-1)^\ell(T_0B_1)^\ell T_0$. These formulae are new to time series analysis.

Journal of Time Series Analysis published by John Wiley & Sons Ltd. DOI: 10.1111/jtsa.12766

4. The coefficients $(-1)^k T_{-k}$ for the loadings on the stochastic trend components, R_t for the deterministic components and W_s for the weights of the embedded noise terms are easily calculated using (2.2), (2.3) and (2.4).
5. The resolvent $R(z)$ may have a pole of order $d \in \mathbb{N}$ at $z = 1$ with $\boldsymbol{x} \sim I(d)$ or an isolated essential singularity at $z = 1$ with $\boldsymbol{x} \sim I(\infty)$. In the former case $T_{-1}B_0$ is nilpotent of degree d and the stochastic trend reduces to a finite sum. In the latter case $T_{-1}B_0$ is quasi-nilpotent with $(T_{-1}B_0)^k \neq 0$ for all $k \in \mathbb{N}$ but with $\|(T_{-1}B_0)^k\|^{1/k} \to 0$ as $k \to \infty$.
6. The operators $P = T_{-1}B_1 \in \mathcal{B}(X)$ and $Q = B_1T_{-1} \in \mathcal{B}(Y)$ are the spectral separation projections for $A(z)$ at $z = 1$. The direct sum decompositions $X = P(X) \oplus P^c(X)$ and $Y = Q(Y) \oplus Q^c(Y)$ separate the singular and regular parts of $R(z)$ with $R_{\sin}(z)A(z) = P$ and $A(z)R_{\sin}(z) = Q$ while $R_{\text{reg}}(z)A(z) = P^c$ and $A(z)R_{\text{reg}}(z) = Q^c$. It follows that $T_{-k} \in \mathcal{B}(Q(Y), P(X))$ for all $k \in \mathbb{N}$ and $T_\ell \in \mathcal{B}(Q^c(Y), P^c(X))$ for all $\ell \in \mathbb{N} - 1$.
7. The stochastic trend lies entirely in $P(X)$ and depends only on previous values of the noise component in $Q(Y)$. The weighted sum of embedded noise terms lies entirely in $P^c(X)$ and depends only on previous values of the noise component in $Q^c(Y)$. The term $R_tB_1\boldsymbol{c}$ can be written as the sum of $U_tB_1P\boldsymbol{c} \in P(X)$ and $W_tB_1P^c\boldsymbol{c} \in P^c(X)$.
8. The AR(1) time series $\boldsymbol{x} \in X^{\mathbb{Z}}$ can be expressed as the sum of two separate AR(1) time series $\boldsymbol{x}_{\sin} \in [P(X)]^{\mathbb{Z}}$ and $\boldsymbol{x}_{\text{reg}} \in [P^c(X)]^{\mathbb{Z}}$. The singular component $\boldsymbol{x}_{\sin}$ is defined by $(QA_0P)\boldsymbol{x}_{\sin}(t) + (QA_1P)\boldsymbol{x}_{\sin}(t-1) = Q\boldsymbol{\xi}(t)$ for $t \in \mathbb{Z}$ with $\boldsymbol{x}_{\sin}(-1) = P\boldsymbol{c}$ and the regular component $\boldsymbol{x}_{\text{reg}}$ is defined by $(Q^cA_0P^c)\boldsymbol{x}_{\text{reg}}(t) + (Q^cA_1P^c)\boldsymbol{x}_{\text{reg}}(t-1) = Q^c\boldsymbol{\xi}(t)$ for $t \in \mathbb{Z}$ with $\boldsymbol{x}_{\text{reg}}(-1) = P^c\boldsymbol{c}$. The resolvent $R_{\sin}(z) = PR(z)Q$ is analytic for all $z \neq 1$ and the extended resolvent $R_{\text{reg}}(z) = P^cR(z)Q^c$ is analytic for $z \in D_{1+\epsilon}(0)$. If $\boldsymbol{x} \sim I(d)$ for some $d \in \mathbb{N} \cup \{\infty\}$ then $\boldsymbol{x}_{\sin} \sim I(d)$. In all cases $\boldsymbol{x}_{\text{reg}} \sim I(0)$.
9. The key subspaces $P(X), P^c(X), Q(Y), Q^c(Y)$ need not be finite-dimensional and there are no restrictions on the operators $B_0, B_1 \in \mathcal{B}(X, Y)$. The subspaces $T_{-1}(Y) \subset P(X)$ and $[T_{-1}(Y)]^\perp \subset X^*$ are the attractor space and cointegrating space respectively.
10. If $\boldsymbol{x} \in \mathbb{R}^n$ then $R(z) = \operatorname{adj}[B_0 + B_1(z-1)]/\det[B_0 + B_1(z-1)] \in \mathbb{R}^{n\times n}$ and the form of the Laurent series $R(z) = \sum_{j\in\mathbb{N}-n} T_j(z-1)^j$ is known with $T_{-k} = 0$ for $k > n$. The matrix formula gives no direct indication that $R(z)$ is completely determined by $\{T_{-1}, T_0\}$.
11. Our decision to restrict discussion to *strictly stationary* noise processes is one of convenience. Analogous results remain true if the noise process is *weakly stationary*. In Banach space a process $\boldsymbol{\xi} \in X^{\mathbb{Z}}$ is said to be weakly stationary (Bosq, 2000, definition 2.4, p. 65) if $\mathbb{E}[\|\boldsymbol{\xi}(t)\|^2] < \infty$ for all $t \in \mathbb{Z}$, $\mathbb{E}[\boldsymbol{\xi}(t)] = \boldsymbol{\mu}$ does not depend on $t \in \mathbb{Z}$ and
$$\mathbb{E}\left[\langle \boldsymbol{\xi}(s+t) - \boldsymbol{\mu}, \boldsymbol{f}\rangle \overline{\langle \boldsymbol{\xi}(s) - \boldsymbol{\mu}, \boldsymbol{g}\rangle}\right] = c_{f,\,g}(t),$$
depends only on $t \in \mathbb{Z}$ for each fixed pair of linear functionals $\boldsymbol{f}, \boldsymbol{g} \in X^*$ and all $s \in \mathbb{Z}$.

3. REPRESENTATION OF THE RESOLVENT AND THE $\mathcal{M}$-TRANSFORM

We consider the $AR(1)$ process defined in (2.1). The autoregressive polynomial is a bounded linear operator pencil $A(z) = A_0 + A_1z = B_0 + B_1(z-1) \in \mathcal{B}(X, Y)$ for all $z \in \mathbb{C}$. Albrecht *et al.* (2014) showed that the resolvent $R(z) = A(z)^{-1} \in \mathcal{B}(Y, X)$ exists and can be expressed as a Laurent series $R(z) = \sum_{j\in\mathbb{Z}} T_j(z-1)^j$ for all $z \in D_{0,\,\epsilon}(1)$ if and only if the coefficients $\{T_j\}_{j\in\mathbb{Z}} \in \mathcal{B}(Y, X)^{\mathbb{Z}}$ satisfy a system of left and right fundamental equations
$$T_{j-1}B_1 + T_jB_0 = \begin{cases} I_X & \text{if } j = 0 \\ 0_X & \text{if } j \in \mathbb{Z},\ j \neq 0, \end{cases} \tag{3.1}$$
and
$$B_1T_{j-1} + B_0T_j = \begin{cases} I_Y & \text{if } j = 0 \\ 0_Y & \text{if } j \in \mathbb{Z},\ j \neq 0, \end{cases} \tag{3.2}$$

DOI: 10.1111/jtsa.12766 *Journal of Time Series Analysis* published by John Wiley & Sons Ltd.

where the coefficients $\{T_j\}_{j\in\mathbb{Z}}$ also satisfy the magnitude constraints

$$\lim_{k\to\infty}\|T_{-k}\|^{1/k}=0 \quad \text{and} \quad \lim_{\ell\to\infty}\|T_\ell\|^{1/\ell}\le 1/\epsilon. \tag{3.3}$$

If (3.1), (3.2) and (3.3) are all satisfied then

$$T_{-k}=(-1)^{k-1}(T_{-1}B_0)^{k-1}T_{-1}\in\mathcal{B}(Y,X), \tag{3.4}$$

for all $k\in\mathbb{N}$ and

$$T_\ell=(-1)^{\ell}(T_0B_1)^{\ell}T_0\in\mathcal{B}(Y,X), \tag{3.5}$$

for all $\ell\in\mathbb{N}-1$. Thus the series is completely determined by the basic solution $\{T_{-1},T_0\}$. Now we can write $R(z)=R_{\sin}(z)+R_{\text{reg}}(z)$ where the singular part of the Laurent series

$$R_{\sin}(z)=[I_X(z-1)+T_{-1}B_0]^{-1}T_{-1}=\sum_{k\in\mathbb{N}}T_{-k}(z-1)^{-k}, \tag{3.6}$$

converges for all $z\neq 1$ and the regular part of the Laurent series

$$R_{\text{reg}}(z)=[I_X+T_0B_1(z-1)]^{-1}T_0=\sum_{\ell\in\mathbb{N}-1}T_\ell(z-1)^{\ell}, \tag{3.7}$$

converges for all $z\in D_\epsilon(1)$. Formal statements of these results can be found in Albrecht *et al.* (2014) and Albrecht *et al.* (2020). The next result, Lemma 3.1, shows that the necessary and sufficient conditions for the above representation are satisfied if and only if $R(z)$ is analytic for $z\in[D_1(0)^a\setminus\{1\}]\cup D_{0,\delta}(1)$ for some $\delta>0$ and has an isolated singularity at $z=1$. Lemma 3.1 is proved in Section B of the Appendix.

Lemma 3.1. Assumption 2.1 holds if and only if $R(z)$ is analytic on a set $D_{1+\epsilon}(0)\setminus\{1\}$ for some $\epsilon>0$ and has an isolated singularity at $z=1$.

Our preferred method of solution is the Maclaurin transform, which we shall refer to as an $\mathcal{M}$-transform. The $\mathcal{M}$-transform of the stochastic time series $\boldsymbol{u}\in X^{\mathbb{Z}}$ with $\mathbb{E}[\|\boldsymbol{u}(\omega,t)\|]=\int_\Omega\|\boldsymbol{u}(\omega,t)\|\mathrm{d}\mu(\omega)\le c/r^t$ for some $c,r>0$ and all $t\in\mathbb{N}-1$ is an almost surely convergent random power series denoted by $\mathcal{M}[\boldsymbol{u}](z)=\boldsymbol{U}(z):\Omega\to X$ and defined by

$$\boldsymbol{U}(\omega,z)=\sum_{t\in\mathbb{N}-1}\boldsymbol{u}(\omega,t)z^t, \tag{3.8}$$

for each $z\in D_r(0)\subset\mathbb{C}$ and all $\omega\in\Omega$. We justify the almost sure convergence as follows. Choose $z\in D_r(0)$. Now $\mathbb{E}[\|\boldsymbol{U}(\omega,z)\|]\le\sum_{t\in\mathbb{N}-1}\mathbb{E}[\|\boldsymbol{u}(\omega,t)\|]|z|^t\le\sum_{t\in\mathbb{N}-1}c(|z|/r)^t=c/(1-|z|/r)<\infty$. Therefore $\|\boldsymbol{U}(\omega,z)\|<\infty$ for almost all $\omega\in\Omega$. See Bosq (2000, lemma 7.1, pp. 182–183). Clearly $\boldsymbol{U}(z)=\mathcal{M}[\boldsymbol{u}](z)=\mathcal{M}[\boldsymbol{u}_+](z)=\boldsymbol{U}_+(z)$ because the transform uses no information for $t<0$. If the time series is only observed for $t\in\mathbb{N}-1$ the $\mathcal{M}$-transform captures all of the observed information. Because $\boldsymbol{u}_+(t-s)=0$ for $s>t$ it follows from (A1) and (A2) in Section A of the Appendix that

$$\Delta^{-k}\boldsymbol{u}_+(t)=\sum_{s\in\mathbb{N}-1}\binom{k+s-1}{s}\boldsymbol{u}_+(t-s)=\sum_{s=0}^{t}\binom{k+s-1}{s}\boldsymbol{u}(t-s), \tag{3.9}$$

for $k\in\mathbb{N}$ and

$$\Delta^{\ell}\boldsymbol{u}_+(t)=\sum_{s=0}^{\ell}\binom{\ell}{s}(-1)^s\boldsymbol{u}_+(t-s)=\sum_{s=0}^{\min\{\ell,t\}}\binom{\ell}{s}(-1)^s\boldsymbol{u}(t-s), \tag{3.10}$$

Journal of Time Series Analysis published by John Wiley & Sons Ltd.
DOI: 10.1111/jtsa.12766

for $\ell \in \mathbb{N}-1$ and all $t \in \mathbb{N}-1$. Therefore

$$\begin{aligned}\mathcal{M}\left[\{\Delta^{-k}\boldsymbol{u}_+(t)\}_{t\in\mathbb{N}-1}\right](z) &= \sum_{t=0}^{\infty}\left[\sum_{s=0}^{t}\binom{k+s-1}{s}\boldsymbol{u}(t-s)\right]z^t \\ &= \sum_{s=0}^{\infty}\binom{k+s-1}{s}z^s\left[\sum_{\tau=0}^{\infty}\boldsymbol{u}(\tau)z^\tau\right] = (1-z)^{-k}\boldsymbol{U}(z),\end{aligned} \tag{3.11}$$

for $z \in D_r(0)$ and each $k \in \mathbb{N}$ and

$$\begin{aligned}\mathcal{M}\left[\{\Delta^{\ell}\boldsymbol{u}_+(t)\}_{t\in\mathbb{N}-1}\right](z) &= \sum_{t=0}^{\infty}\left[\sum_{s=0}^{\min\{\ell,t\}}\binom{\ell}{s}(-1)^s\boldsymbol{u}(t-s)\right]z^t \\ &= \sum_{s=0}^{\ell}\binom{\ell}{s}(-1)^s z^s\left[\sum_{\tau=0}^{\infty}\boldsymbol{u}(\tau)z^\tau\right] = (1-z)^{\ell}\boldsymbol{U}(z),\end{aligned} \tag{3.12}$$

for $z \in D_r(0)$ and each $\ell \in \mathbb{N}-1$. The $\mathcal{M}$-transform satisfies a standard convolution identity. Let $\{V_t\}_{t\in\mathbb{N}-1} \in \mathcal{B}(Y,X)^{\mathbb{N}-1}$ with $\|V_t\| \le c/r^t$ for some $c, r > 0$. The corresponding $\mathcal{M}$-transform $S(z) = \sum_{t\in\mathbb{N}-1} V_t z^t \in \mathcal{B}(Y,X)$ is well defined for all $z \in D_r(0)$. Suppose $\boldsymbol{\xi} \in Y^{\mathbb{Z}}$ is a strong white noise with $\mathcal{M}$-transform $\boldsymbol{\Xi}(z) = \sum_{j\in\mathbb{N}-1}\boldsymbol{\xi}(t)x^t$ for $z \in D_1(0)$. If we define the convolution $\boldsymbol{w} = V \star \boldsymbol{\xi} \Leftrightarrow \boldsymbol{w}(t) = V_t \star \boldsymbol{\xi}(t)$ by setting

$$\boldsymbol{w}(t) = \sum_{s=0}^{t} V_s\,\boldsymbol{\xi}(t-s), \tag{3.13}$$

for all $t \in \mathbb{N}-1$ then $\boldsymbol{W}(z) = S(z)\boldsymbol{\Xi}(z) \Leftrightarrow \mathcal{M}[V \star \boldsymbol{\xi}](z) = \mathcal{M}[V](z)\mathcal{M}[\boldsymbol{\xi}](z)$ for all $z \in D_r(0) \cap D_1(0)$.

4. RESOLUTION OF THE AR(1) TIME SERIES MODEL

We consider the AR(1) process defined in (2.1) by $A_0\boldsymbol{x}(t) + A_1\boldsymbol{x}(t-1) = \boldsymbol{\xi}(t)$ for each $t \in \mathbb{N}-1$ with $\boldsymbol{x}(-1) = \boldsymbol{c}$. It follows from this infinite collection of equations that

$$\begin{aligned} A_0\boldsymbol{x}(0) + A_1\boldsymbol{x}(-1) &= \boldsymbol{\xi}(0) \\ A_0\boldsymbol{x}(1)z + A_1\boldsymbol{x}(0)z &= \boldsymbol{\xi}(1)z \\ A_0\boldsymbol{x}(2)z^2 + A_1\boldsymbol{x}(1)z^2 &= \boldsymbol{\xi}(2)z^2 \\ \vdots &= \vdots, \end{aligned}$$

for each $z \in \mathbb{C}$. By summing the equations we can collect all observable information into a single algebraic equation

$$A(z)\boldsymbol{X}(z) = \boldsymbol{\Xi}(z) - B_1\,\boldsymbol{c}, \tag{4.1}$$

where $A(z) = A_0 + A_1 z$ is the autoregressive polynomial, $\boldsymbol{X}(z) = \mathcal{M}[\boldsymbol{x}](z)$ and $\boldsymbol{\Xi}(z) = \mathcal{M}[\boldsymbol{\xi}](z)$ are the respective $\mathcal{M}$-transforms of $\boldsymbol{x}$ and $\boldsymbol{\xi}$ and where we prefer to write $A_1 = B_1$ on the right-hand side. According to Lemma 3.1 we may assume that $R(z) = A(z)^{-1} \in \mathcal{B}(Y,X)$ is analytic on a set $D_{1+\epsilon}(0) \setminus \{1\}$ for some $\epsilon > 0$ and has an isolated singularity at $z = 1$. An intuitive resolution of (4.1) is therefore given by

$$\boldsymbol{X}(z) = R(z)[\boldsymbol{\Xi}(z) - B_1\boldsymbol{c}] = R_{\text{sin}}(z)[\boldsymbol{\Xi}(z) - B_1\boldsymbol{c}] + R_{\text{reg}}(z)[\boldsymbol{\Xi}(z) - B_1\boldsymbol{c}], \tag{4.2}$$

for all $z \in D_{1+\epsilon}(0) \setminus \{1\}$. Our main task is to justify this intuitive resolution and obtain a suitable representation for $\boldsymbol{x}(t) = \mathcal{M}^{-1}[\boldsymbol{X}](t)$ for $t \in \mathbb{N}-1$.

DOI: 10.1111/jtsa.12766
Journal of Time Series Analysis published by John Wiley & Sons Ltd.

4.1. A transform formula for the singular part of the resolvent

The function $R_{\sin}(z)$ defined in (3.6) is analytic for all $z \neq 1$. Therefore we can write

$$R_{\sin}(z) = \sum_{s\in\mathbb{N}-1} U_t z^t = \mathcal{M}[\{U_t\}_{t\in\mathbb{N}-1}](z), \tag{4.3}$$

as an $\mathcal{M}$-transform for all $z \in D_1(0)$ where the coefficients $\{U_t\}_{t\in\mathbb{N}-1}$ are given by the formula

$$U_t = (1/t!)\left[d^{\,t}R_{\sin}(z)/dz^t\right]_{z=0} = (-1)(I_X - T_{-1}B_0)^{-t-1}T_{-1}. \tag{4.4}$$

Consequently $U_t \in \mathcal{B}(Q(Y), P(X))$ for all $t \in \mathbb{N} - 1$.

4.2. A transform formula for the extended regular part of the resolvent

The resolvent $R(z) = (A_0 + A_1 z)^{-1}$ is analytic for $z \in D_1(0)$. Therefore A_0^{-1} is well defined and

$$R(z) = \sum_{t\in\mathbb{N}-1} R_t z^t = \mathcal{M}[\{R_t\}_{t\in\mathbb{N}-1}](z), \tag{4.5}$$

for all $z \in D_1(0)$ where

$$R_t = (1/t!)\left[d^{\,t}(A_0 + A_1 z)^{-1}/dz^t\right]_{z=0} = (-1)^t(A_0^{-1}A_1)^t A_0^{-1}, \tag{4.6}$$

for all $t \in \mathbb{N} - 1$. It follows that $\lim_{t\to\infty} \|R_t\|^{1/t} = 1$. We can extend the function $R_{\text{reg}}(z)$ defined in (3.7) to an analytic function $W(z)$ on $D_{1+\epsilon}(0)$ by defining

$$W(z) = \begin{cases} R(z) - R_{\sin}(z) & \text{for } z \in D_{1+\epsilon}(0) \setminus \{1\}. \\ T_0 & \text{for } z = 1. \end{cases} \tag{4.7}$$

We can now express the extended function $R_{\text{reg}}(z) = W(z)$ as an $\mathcal{M}$-transform by writing

$$W(z) = \sum_{t\in\mathbb{N}-1} W_t z^t = \mathcal{M}[\{W_t\}_{t\in\mathbb{N}-1}](z), \tag{4.8}$$

for $z \in D_{1+\epsilon}(0)$ where $W_t \in \mathcal{B}(Y, X)$ and $\lim_{t\to\infty} \|W_t\|^{1/t} \leq 1/(1+\epsilon)$. Now (4.7) gives

$$W_t = R_t - U_t = (-1)^t(A_0^{-1}A_1)^t A_0^{-1} + (I_X - T_{-1}B_0)^{-t-1}T_{-1}, \tag{4.9}$$

for each $t \in \mathbb{N} - 1$. The inequality $\lim_{t\to\infty} \|W_t\|^{1/t} \leq 1/(1+\epsilon)$ means that $\|W_t\| \leq c/(1+\epsilon)^t$ for some $c > 0$ and all $t \in \mathbb{N} - 1$. If we write $\zeta = z - 1$ and recall that $R_{\text{reg}}(z) = W(z)$ then

$$\begin{aligned} R_{\text{reg}}(z) &= \sum_{t\in\mathbb{N}-1} W_t(1+\zeta)^t = \sum_{t\in\mathbb{N}-1} W_t \left[\sum_{\ell=0}^{t} \binom{t}{\ell} \zeta^\ell\right] \\ &= \sum_{\ell\in\mathbb{N}-1} \left[\sum_{t\in\mathbb{N}-1+\ell} \binom{t}{\ell} W_t\right] \zeta^\ell = \sum_{\ell\in\mathbb{N}-1} \left[\sum_{r\in\mathbb{N}-1} \binom{\ell+r}{\ell} W_{\ell+r}\right] \zeta^\ell, \end{aligned} \tag{4.10}$$

Journal of Time Series Analysis published by John Wiley & Sons Ltd.
DOI: 10.1111/jtsa.12766

for all $z \in D_{\epsilon}(1)$. It follows from (4.10) that we can also calculate $\{T_\ell\}_{\ell\in\mathbb{N}-1}$ using the formula

$$T_\ell = \sum_{r\in\mathbb{N}-1} \binom{\ell+r}{\ell} W_{\ell+r}, \tag{4.11}$$

for each $\ell \in \mathbb{N}-1$. To solve Example 6.1 in Section 4.2 we define $W(z) = R(z) - T_{-1}/(z-1)$ for $z \in D_{1+\epsilon}(0)$ and show that $W_t \in \mathcal{B}(Q^c(Y), P^c(X))$ for all $t \in \mathbb{N}-1$. Lemma 4.1 shows that this is true in general. The proof is given in Section B of the Appendix.

Lemma 4.1. Let $R_{\text{reg}}(z) = W(z)$ be represented by the $\mathcal{M}$-transform $W(z) = \sum_{t\in\mathbb{N}-1} W_t z^t$ for all $z \in D_{1+\epsilon}(0)$. Then $W_t = P^c W_t = W_t Q^c \in \mathcal{B}(Q^c(Y), P^c(X))$ for all $t \in \mathbb{N}-1$.

We have shown that $R_{\text{sin}}(z) = \sum_{t\in\mathbb{N}-1} U_t z^t = \mathcal{M}[\{U_t\}_{t\in\mathbb{N}-1}](z)$ and $R_{\text{reg}}(z) = \sum_{t\in\mathbb{N}-1} W_t z^t = \mathcal{M}[\{W_t\}_{t\in\mathbb{N}-1}](z)$. We can also see that $\Xi(z) - B_1\boldsymbol{c} = \mathcal{M}[\boldsymbol{\xi}(t) - B_1\boldsymbol{c}\ \delta(t)](z)$ where the series $\{\delta(t)\}_{t\in\mathbb{Z}}$ is defined by $\delta(0) = 1$ and $\delta(t) = 0$ if $t \neq 0$. It follows from (3.13) and (4.2) that

$$\boldsymbol{x}(t) = \boldsymbol{x}_{\text{sin}}(t) + \boldsymbol{x}_{\text{reg}}(t) = U_t \star [\boldsymbol{\xi}(t) - B_1\boldsymbol{c}\ \delta(t)] + W_t \star [\boldsymbol{\xi}(t) - B_1\boldsymbol{c}\ \delta(t)], \tag{4.12}$$

for all $t \in \mathbb{N}-1$. Formula (4.12) is an implicit version of the extended GJRT. It remains to show that (4.12) can be rewritten explicitly in the form (4.13) or in the modified form (4.17). The relevant proofs are given in Section B of the Appendix.

Proposition 4.2 (GJRT)**.** Under Assumption 2.1 the series $\boldsymbol{x}(t)$ in (2.1) admits the solution

$$\begin{aligned}\boldsymbol{x}(t) &= \boldsymbol{x}_{\text{sin}}(t) + \boldsymbol{x}_{\text{reg}}(t)\\ &= \left[\sum_{k\in\mathbb{N}} (-1)^k T_{-k}\Delta^{-k}\boldsymbol{\xi}_+(t) - U_t B_1\boldsymbol{c}\right] + \left[\sum_{s=0}^{t} W_s\ \boldsymbol{\xi}(t-s) - W_t B_1\boldsymbol{c}\right],\end{aligned} \tag{4.13}$$

for all $t \in \mathbb{N}-1$, where $\boldsymbol{x}(-1) = \boldsymbol{c}$, $\boldsymbol{x}_{\text{sin}}(t) = P\boldsymbol{x}(t) \in P(X)$, $\boldsymbol{x}_{\text{reg}}(t) = P^c\boldsymbol{x}(t) \in P^c(X)$, and

$$\begin{aligned}T_{-k} &= (-1)^{k-1}(T_{-1}B_0)^{k-1}T_{-1},\\ U_t &= (-1)(I_X - T_{-1}B_0)^{-t-1}T_{-1},\\ W_t &= (-1)^t(A_0^{-1}A_1)^t A_0^{-1} - U_t.\end{aligned}$$

The following consequence of the GJRT is proved in Section B of the Appendix.

Corollary 4.3 Under Assumption 2.1 the series $\boldsymbol{x} \in X^{\mathbb{Z}}$ satisfies (2.1) with $\boldsymbol{x}(-1) = \boldsymbol{c}$ if and only if $\boldsymbol{x}(t) = \boldsymbol{x}_{\text{sin}}(t) + \boldsymbol{x}_{\text{reg}}(t)$ where the singular part $\boldsymbol{x}_{\text{sin}} \in P(X)^{\mathbb{N}-1}$ satisfies

$$QA_0P\boldsymbol{x}_{\text{sin}}(t) + QA_1P\boldsymbol{x}_{\text{sin}}(t-1) = Q\boldsymbol{\xi}(t), \tag{4.14}$$

for $t \in \mathbb{N}-1$ and $\boldsymbol{x}_{\text{sin}}(-1) = P\boldsymbol{c}$ and the regular part $\boldsymbol{x}_{\text{reg}} \in P^c(X)^{\mathbb{N}-1}$ satisfies

$$Q^cA_0P^c\boldsymbol{x}_{\text{reg}}(t) + Q^cA_1P^c\boldsymbol{x}_{\text{reg}}(t-1) = Q^c\boldsymbol{\xi}(t), \tag{4.15}$$

for $t \in \mathbb{N}-1$ and $\boldsymbol{x}_{\text{reg}}(-1) = P^c\boldsymbol{c}$. Therefore an integrated AR(1) process $\boldsymbol{x}$ is the sum of separate AR(1) processes on complementary subspaces.

We define the extended singular and regular parts $\boldsymbol{x}_{\text{sin}} \in P(X)^{\mathbb{Z}}$ and $\boldsymbol{x}_{\text{reg}} \in P^c(X)^{\mathbb{Z}}$ by supposing that (4.14) and (4.15) are true for all $t \in \mathbb{Z}$. It is convenient to define $A_{i,\text{r}} = Q^cA_iP^c \in \mathcal{B}(P^c(X), Q^c(Y))$ for each $i = 0, 1$. The resolvent $R_{\text{reg}}(z) = (A_{0,\text{r}} + A_{1,\text{r}}z)^{-1} \in \mathcal{B}(Q^c(Y), P^c(X))$ is well defined and analytic for $z \in D_{1+\epsilon}(0)$. Therefore

1. $A_{0,\mathrm{r}}^{-1} \in \mathcal{B}(Q^c(Y), P^c(X))$ is well defined,
2. $\Phi = (-1)A_{0,\mathrm{r}}^{-1}A_{1,\mathrm{r}} \in \mathcal{B}(P^c(X))$ is well defined, and
3. $\|\Phi^s\| = \|\left(A_{0,\mathrm{r}}^{-1}A_{1,\mathrm{r}}\right)^s\| < c/(1+\epsilon)^s$ for some $c > 0$ and all $s \in \mathbb{N}$.

By rearranging the extended form of (4.15) it follows that $\boldsymbol{x}_{\mathrm{reg}} \in P^c(X)^{\mathbb{Z}}$ is an $AR(1)$ process defined by the equation

$$\boldsymbol{x}_{\mathrm{reg}}(t) = \Phi\boldsymbol{x}_{\mathrm{reg}}(t-1) + \boldsymbol{\eta}(t), \tag{4.16}$$

for all $t \in \mathbb{Z}$ where $\boldsymbol{x}_{\mathrm{reg}}(-1) = P^c\boldsymbol{c}$ and where the strong white noise $\boldsymbol{\eta} \in P^c(X)^{\mathbb{Z}}$ is defined by setting $\boldsymbol{\eta}(t) = A_{0,\mathrm{r}}^{-1}Q^c\boldsymbol{\xi}(t)$ for each $t \in \mathbb{Z}$. Our next result, Corollary 4.4, is an alternative version of the extended GJRT. This corollary is proved in Section B of the Appendix.

Corollary 4.4. Under Assumption 2.1 the AR(1) process $\boldsymbol{x} \in X^{\mathbb{Z}}$ satisfies

$$\begin{aligned}\boldsymbol{x}(t) &= \boldsymbol{x}_{\sin}(t) + \boldsymbol{x}_{\mathrm{reg}}(t)\\ &= \left[\sum_{k\in\mathbb{N}}(-1)^k T_{-k}\Delta^{-k}\boldsymbol{\xi}_+(t) - U_tB_1\boldsymbol{c}\right] + \left[\Phi^{t+1}z_\infty + \sum_{s\in\mathbb{N}-1}\Phi^s\boldsymbol{\eta}(t-s)\right],\end{aligned} \tag{4.17}$$

for $t \in \mathbb{N}-1$ where $z_\infty = \lim_{r\to\infty} z_r = \lim_{r\to\infty}\Phi^r\boldsymbol{x}_{\mathrm{reg}}(-1-r) = P^c\boldsymbol{c} - \sum_{s\in\mathbb{N}-1}\Phi^s\boldsymbol{\eta}(-1-s) \in P^c(X)$ almost surely. If $z_\infty = \boldsymbol{0}$ and $\boldsymbol{x} \sim I(d)$ for $d \in \mathbb{N}\cup\{\infty\}$ then $\boldsymbol{x}_{\sin} \sim I(d)$ and $\boldsymbol{x}_{\mathrm{reg}} \sim I(0)$.

If $P^c\boldsymbol{c} = \sum_{s\in\mathbb{N}-1}\Phi^s\boldsymbol{\eta}(-1-s)$ then $z_\infty = \boldsymbol{0}$ and $\boldsymbol{x}_{\mathrm{reg}}(t)$ is a strictly stationary standard linear process. If $z_\infty \neq \boldsymbol{0}$ then $\Phi^{t+1}z_\infty \to \boldsymbol{0}$ as $t \to \infty$ in which case we say that $\boldsymbol{x}_{\mathrm{reg}}(t)$ is asymptotically strictly stationary.

Remark 4.5. Suppose $P^c(X) \neq \{\boldsymbol{0}\} \Leftrightarrow P^c \neq 0$. Choose $\boldsymbol{p} \in P^c(X)$ with $\boldsymbol{p} \neq \boldsymbol{0}$ and define a bounded linear functional $\boldsymbol{f}_p \neq \boldsymbol{0}$ on the subspace $\{\boldsymbol{x}|\boldsymbol{x} = \alpha\boldsymbol{p}$ for all $\alpha \in \mathbb{C}\} \subset P^c(X)$ by setting $\langle\alpha\boldsymbol{p},\boldsymbol{f}_p\rangle = \alpha$ for all $\alpha \in \mathbb{C}$. The Hahn–Banach theorem shows that we can extend $\boldsymbol{f}_p$ to a bounded linear functional $\boldsymbol{f}_p \in [P^c(X)]^*$. Now we can further extend $\boldsymbol{f}_p$ to a bounded linear functional $\boldsymbol{f}_p \in X^*$ by defining $\langle\boldsymbol{x},\boldsymbol{f}_p\rangle = \langle P^c\boldsymbol{x},\boldsymbol{f}_p\rangle$ for all $\boldsymbol{x} \in X$. Therefore $\langle P\boldsymbol{x},\boldsymbol{f}_p\rangle = 0$ for all $\boldsymbol{x} \in X$. Hence $\boldsymbol{f}_p \in P(X)^\perp$. By theorem 1 in Albrecht *et al.* (2020) one has $T_{-1}(Y) = P(X)$. Therefore $\boldsymbol{f}_p \in T_{-1}(Y)^\perp$. If $\boldsymbol{x}(-1) = \boldsymbol{c} = P\boldsymbol{d} + \sum_{s\in\mathbb{N}-1}\Phi^s\boldsymbol{\eta}(-1-s)$ for some $\boldsymbol{d} \in X$ then $P^c\boldsymbol{c} = \sum_{s\in\mathbb{N}-1}\Phi^s\boldsymbol{\eta}(-1-s)$ and so $z_\infty = \boldsymbol{0}$. Now (4.17) shows that $\langle\boldsymbol{x}(t),\boldsymbol{f}_p\rangle = \sum_{s\in\mathbb{N}-1}\Phi^s\langle\boldsymbol{\eta}(t-s),\boldsymbol{f}_p\rangle$ for all $t \in \mathbb{N}-1$. Therefore $\langle\boldsymbol{x},\boldsymbol{f}_p\rangle$ is a standard linear process and $\langle\boldsymbol{x},\boldsymbol{f}_p\rangle \sim I(0)$. We conclude that the condition $P^c \neq 0$ is sufficient to ensure that $\boldsymbol{x}$ is cointegrated.

Remark 4.6. The models (4.14) and (4.15) describe AR(1) processes on complementary subspaces. In an algebraic sense these processes are linearly independent. Any stochastic dependence is entirely due to stochastic dependence in the noise terms. If $Q\boldsymbol{\xi}$ and $Q^c\boldsymbol{\xi}$ are stochastically independent then so too are $\boldsymbol{x}_{\sin}$ and $\boldsymbol{x}_{\mathrm{reg}}$.

4.3. The attractor and cointegrating spaces

We suppose Assumption 2.1 is true. In our notation the usual definition of the attractor space is $P(X) = T_{-1}(Y) \subset X$ while the usual definition of the cointegrating space is $P(X)^\perp = T_{-1}(Y)^\perp \subset X^*$ where X^* is the dual space to X. The rationale for the definition of the attractor space is that the projected process $P^c\boldsymbol{x}$ on the subspace $P^c(X)$ is a stable process with $\|P^c\boldsymbol{x}(t)\| \to 0$ almost surely as $t \to \infty$. The rationale for the definition of the cointegrating space is that if $P^c(X) \neq \{\boldsymbol{0}\}$ then our earlier remarks show that there is some non-zero bounded linear functional $\boldsymbol{f}_0 \in T_{-1}(Y)^\perp \subset X^*$ with $\langle\boldsymbol{x},\boldsymbol{f}_0\rangle \sim I(0)$.

If $R(z)$ has a pole of order $d \in \mathbb{N}$ at $z = 1$ there is a whole hierarchy of spaces with weaker versions of the above properties. If $\boldsymbol{x} \in T_{-k}(Y)$ then we can find $\boldsymbol{y} \in Y$ such that $\boldsymbol{x} = T_{-k}\boldsymbol{y} = T_{-k+1}[(-1)B_0T_{-1}\boldsymbol{y}] = T_{-k+1}\boldsymbol{z}$

where $z = B_0 T_{-1} \boldsymbol{y} \in Y$. Therefore $\boldsymbol{x} \in T_{-k+1}(Y)$. It follows that $T_{-1}(Y) \supset T_{-2}(Y) \supset \cdots \supset T_{-d}(Y) \supset T_{-d-1}(Y) = \{\mathbf{0}\}$. The inclusions are proper as the following argument shows. Suppose $T_{-r-1}(Y) = T_{-r}(Y) \neq \{\mathbf{0}\}$ for some $r \in \mathbb{N}$. Then $(-1)(T_{-1}B_0T_{-r-1}(Y) = (-1)(T_{-1}B_0T_{-r}(Y) \Leftrightarrow T_{-r-2}(Y) = T_{-r-1}(Y)$. Therefore $T_{-r-2}(Y) = T_{-r}(Y)$. An inductive argument now shows that $T_{-k}(Y) = T_{-r}(Y)$ for all $k \in \mathbb{N} + r$ but this is a contradiction because $T_{-d-1} = \{\mathbf{0}\}$. A straightforward argument also shows that $T_{-1}(Y)^{\perp} \subset T_{-2}(Y) \subset \cdots \subset T_{-d}(Y)^{\perp} \subset X^*$. Once again the inclusions are proper. If $\boldsymbol{f} \in T_{-r}(Y)^{\perp} \setminus T_{-r+1}(Y)^{\perp}$ for some $r \in \{2, 3, \ldots, d-1\}$ then $\boldsymbol{f} \neq \mathbf{0}$ and $\langle \boldsymbol{x}, \boldsymbol{f} \rangle \sim I(r)$.

If $R(z)$ has an essential singularity at $z = 1$ the above inclusions are still valid but it is no longer necessary that the inclusions are proper. Nevertheless it is still true that if $\boldsymbol{f} \in T_{-r}(Y)^{\perp} \setminus T_{-r+1}(Y)^{\perp}$ for some $r \in \{2, 3, \ldots\}$ then $\boldsymbol{f} \neq \mathbf{0}$ and $\langle \boldsymbol{x}, \boldsymbol{f} \rangle \sim I(r)$.

5. SPECIAL CASES: THE GJRT WHEN R(Z) HAS A POLE AT Z=1

Beare and Seo (2020) consider specific results for $I(d)$ processes when $d = 1, 2$. We will show that their results are special cases of our extended results. We reiterate that our representation is also new for cointegrated AR(1) vector processes on $\mathbb{R}^n$. We suppose only that $R(z)$ is analytic on a set $[D_1(0)^a \setminus \{1\}] \cup D_{0,\delta}(1)$ for some $\delta > 0$ and has an isolated singularity at $z = 1$.

5.1. The GJRT when $R(z)$ has a simple pole at $z = 1$

If $R(z)$ has a simple pole at $z = 1$ then $T_{-k} = (-1)^{k-1}(T_{-1}B_0)^{k-1}T_{-1} = 0$ for $k > 1$ and the binomial expansion shows that

$$U_t = (-1)[I_X - T_{-1}B_0]^{-t-1}T_{-1} = (-1)[I_X + (t+1)(T_{-1}B_0) + \cdots]T_{-1} = (-1)T_{-1},$$

for all $t \in \mathbb{N} - 1$. Therefore (4.17) becomes

$$\boldsymbol{x}(t) = (-1)T_{-1}\sum_{s=0}^{t} \boldsymbol{\xi}(t-s) + T_{-1}B_1\boldsymbol{c} + \Phi^{t+1}z_\infty + \sum_{s\in\mathbb{N}-1} \Phi^s \boldsymbol{\eta}(t-s), \tag{5.1}$$

for all $t \in \mathbb{N} - 1$ where we have used the identity $\Delta^{-1}\boldsymbol{\xi}_+(t) = \sum_{s=0}^{t}\boldsymbol{\xi}(t-s)$ for all $t \in \mathbb{N} - 1$. Therefore $\boldsymbol{x} \nsim I(0)$.

Remark 5.1. If $\boldsymbol{c} = P\boldsymbol{d} + \sum_{s\in\mathbb{N}-1} \Phi^s \boldsymbol{\eta}(-1-s)$ for some $\boldsymbol{d} \in X$ then $z_\infty = \mathbf{0}$ and (5.1) implies

$$\Delta\boldsymbol{x}(t) = (-1)T_{-1}\boldsymbol{\xi}(t) + \sum_{s\in\mathbb{N}-1} \Phi^s \Delta\boldsymbol{\eta}(t-s),$$

for all $t \in \mathbb{N}$. Therefore $\Delta\boldsymbol{x} = \{\Delta\boldsymbol{x}(t)\}_{t\in\mathbb{N}} \in X^{\mathbb{N}}$ is a strictly stationary standard linear process. Thus $\Delta\boldsymbol{x} \sim I(0)$. It follows that $\boldsymbol{x} \sim I(1)$. Our results provide a natural extension of the results obtained by Beare and Seo (2020).

5.2. The GJRT when $R(z)$ has a pole of order 2 at $z = 1$

If $R(z)$ has a pole of order two at $z = 1$ then $T_{-k} = (-1)^{k-1}(T_{-1}B_0)^{k-1}T_{-1} = 0$ for $k > 2$ and the binomial expansion shows that

$$U_t = (-1)(I_X - T_{-1}B_0)^{-t-1}T_{-1} = (-1)[I_X + (t+1)(T_{-1}B_0) + \cdots]T_{-1} = -T_{-1} + (t+1)T_{-2},$$

for all $t \in \mathbb{N} - 1$. The identities $\Delta^{-1}\boldsymbol{\xi}_+(t) = \sum_{s=0}^{t} \boldsymbol{\xi}(t-s)$ and $\Delta^{-2}\boldsymbol{\xi}_+(t) = \sum_{s=0}^{t} (s+1)\boldsymbol{\xi}(t-s)$ can then be used to show that (4.17) becomes

$$\begin{aligned}\boldsymbol{x}(t) &= \left[T_{-2}\sum_{s=0}^{t}(s+1)\boldsymbol{\xi}(t-s) - T_{-1}\sum_{s=0}^{t}\boldsymbol{\xi}(t-s) - (t+1)T_{-2}B_1\boldsymbol{c} + T_{-1}B_1\boldsymbol{c}\right] \\ &\quad + \left[\Phi^{t+1}\boldsymbol{z}_\infty + \sum_{s\in\mathbb{N}-1}\Phi^s\boldsymbol{\eta}(t-s)\right], \end{aligned} \tag{5.2}$$

for all $t \in \mathbb{N} - 1$.

Remark 5.2. If $\boldsymbol{c} = P\boldsymbol{d} + \sum_{s\in\mathbb{N}-1}\Phi^s\boldsymbol{\eta}(-1-s)$ for some $\boldsymbol{d} \in X$ then $\boldsymbol{z}_\infty = \boldsymbol{0}$ and (5.2) becomes

$$\begin{aligned}\boldsymbol{x}(t) &= T_{-2}\sum_{s=0}^{t}(s+1)\boldsymbol{\xi}(t-s) - T_{-1}\sum_{s=0}^{t}\boldsymbol{\xi}(t-s) \\ &\quad -(t+1)T_{-2}B_1\boldsymbol{d} + T_{-1}B_1\boldsymbol{d} + \sum_{s\in\mathbb{N}-1}\Phi^s\boldsymbol{\eta}(t-s), \end{aligned} \tag{5.3}$$

for all $t \in \mathbb{N} - 1$. Therefore

$$\Delta\boldsymbol{x}(t) = T_{-2}\sum_{s=0}^{t}\boldsymbol{\xi}(t-s) - T_{-1}\boldsymbol{\xi}(t) - T_{-2}B_1\boldsymbol{d} + \sum_{s\in\mathbb{N}-1}\Phi^s\Delta\boldsymbol{\eta}(t-s),$$

for all $t \in \mathbb{N}$ and

$$\Delta^2\boldsymbol{x}(t) = T_{-2}\boldsymbol{\xi}(t) - T_{-1}\Delta\boldsymbol{\xi}(t) + \sum_{s\in\mathbb{N}-1}\Phi^s\Delta^2\boldsymbol{\eta}(t-s),$$

for all $t \in \mathbb{N}+1$. Therefore $\Delta^2\boldsymbol{x} = \{\Delta^2\boldsymbol{x}(t)\}_{t\in\mathbb{N}+1} \in X^{\mathbb{N}+1}$ is a strictly stationary standard linear process. Therefore $\Delta^2 \sim I(0)$. Clearly $\boldsymbol{x} \nsim I(0)$ and $\Delta\boldsymbol{x} \nsim I(0)$. It follows that $\boldsymbol{x} \sim I(2)$. Once again these results are a natural extension of the results reported in Beare and Seo (2020).

6. THREE EXAMPLES TO ILLUSTRATE THE PROPOSED REPRESENTATION

The first example is a two-dimensional time series where $R(z)$ has a simple pole at $z = 1$. We show that $\boldsymbol{x}$ is the sum of two separate AR(1) time series. The second example is an infinite-dimensional time series where $R(z)$ has an essential singularity at $z = 1$. The third example is taken from Seo (2023, Examples 3.3, pp 757–758 and 3.4, pp. 763–764) and concerns an infinite-dimensional series where $R(z)$ has a pole of order two at $z = 1$. This example is used to demonstrate the general solution procedure on separable Banach space. The general solution procedure is described in Albrecht *et al.* (2020).

Example 6.1 ($\boldsymbol{x} \sim I(1)$)**.** Let $X = Y = \mathbb{R}^2$ and let $\epsilon \in (0,1)$. Let $\boldsymbol{x} \in X^{\mathbb{N}-1}$ be an AR(1) process with $A_0\boldsymbol{x}(t) + A_1\boldsymbol{x}(t-1) = \boldsymbol{\xi}(t)$ for all $t \in \mathbb{N} - 1$ where

$$A_0 = \begin{bmatrix} 0 & -1 \\ 1+\epsilon & -(2+\epsilon) \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad A(z) = \begin{bmatrix} z & -1 \\ 1+\epsilon & z-(2+\epsilon) \end{bmatrix}.$$

Journal of Time Series Analysis published by John Wiley & Sons Ltd.
DOI: 10.1111/jtsa.12766

We assume that $\boldsymbol{x}(-1) = \mathbf{0}$ and that $\mathbb{E}[\boldsymbol{\xi}(t)] = \mathbf{0} \in \mathbb{R}^2$ and $\mathbb{E}[\|\boldsymbol{\xi}(t)\|^2] = \sigma^2$ for all $t \in \mathbb{N}-1$. We define $B_0 = A_0 + A_1$ and $B_1 = A_1$. The resolvent is given by

$$R(z) = (z-1)^{-1}\begin{bmatrix} 1+\epsilon^{-1} & -\epsilon^{-1} \\ 1+\epsilon^{-1} & -\epsilon^{-1} \end{bmatrix} + (z-1-\epsilon)^{-1}\begin{bmatrix} -\epsilon^{-1} & \epsilon^{-1} \\ -(1+\epsilon^{-1}) & 1+\epsilon^{-1} \end{bmatrix},$$

for all $z \in \mathbb{C} \setminus \sigma$ where $\sigma = \{1, 1+\epsilon\}$. Therefore $R(z)$ has a pole of order one at $z = 1$ and we can write $R(z) = T_{-1}(z-1)^{-1} + \sum_{\ell \in \mathbb{N}-1} T_\ell (z-1)^\ell$ as a Laurent series for $z \in D_{0,\epsilon}(1)$ where

$$T_{-1} = \begin{bmatrix} 1+\epsilon^{-1} & -\epsilon^{-1} \\ 1+\epsilon^{-1} & -\epsilon^{-1} \end{bmatrix}, \quad T_0 = \epsilon^{-1}\begin{bmatrix} \epsilon^{-1} & -\epsilon^{-1} \\ (1+\epsilon^{-1}) & -(1+\epsilon^{-1}) \end{bmatrix} \quad \text{and} \quad T_\ell = \epsilon^{-\ell} T_0.$$

The spectral projection on X is $P = T_{-1}B_1 = T_{-1}$ with $P(X) = \{\boldsymbol{x} | P\boldsymbol{x} = \boldsymbol{x}\} = \mathrm{sp}\{\boldsymbol{e}_1 + \boldsymbol{e}_2\}$ and the complementary projection $P^c = I_X - P$ with $P^c(X) = \{\boldsymbol{x} | P^c\boldsymbol{x} = \boldsymbol{x}\} = \mathrm{sp}\{\boldsymbol{e}_1 + (1+\epsilon)\boldsymbol{e}_2\}$. The spectral projection on Y is $Q = B_1 T_{-1} = T_{-1}$ where similar remarks apply. The singular part of the resolvent $R_{\sin}(z) \in \mathcal{B}(Q(Y), P(X))$ can be written directly as an $\mathcal{M}$-transform

$$R_{\sin}(z) = T_{-1}(z-1)^{-1} = \sum_{t \in \mathbb{N}-1} (-1) T_{-1} z^t = \sum_{t \in \mathbb{N}-1} U_t z^t,$$

for $z \in D_1(0)$ where $U_t = (-1)T_{-1}$. The same direct approach for the regular part $R_{\mathrm{reg}}(z)$ gives

$$\begin{aligned} R_{\mathrm{reg}}(z) &= \sum_{\ell \in \mathbb{N}-1} \epsilon^{-\ell} T_0 (z-1)^\ell = \sum_{\ell \in \mathbb{N}-1} (-1)^\ell \epsilon^{-\ell} T_0 \sum_{t=0}^{\ell} \binom{\ell}{t} z^t \\ &= \sum_{t \in \mathbb{N}-1} \left[\sum_{\ell \in \mathbb{N}+t-1} \binom{\ell}{t} (-1)^\ell \epsilon^{-\ell} \right] T_0\, z^t, \end{aligned}$$

which is not valid because $\epsilon \in (0,1)$ and each coefficient of z^t involves the sum of a divergent series. However $R_{\mathrm{reg}}(z) \in \mathcal{B}(Q^c(Y), P^c(X))$ can be written as an $\mathcal{M}$-transform using the extended definition

$$R_{\mathrm{reg}}(z) = R(z) - R_{\sin}(z) = \left[1 - (1+\epsilon)^{-1} z\right]^{-1} \epsilon (1+\epsilon)^{-1} T_0 = \sum_{t \in \mathbb{N}-1} W_t\, z^t,$$

for $z \in D_{1+\epsilon}(0) \setminus \{1\}$ and $R_{\mathrm{reg}}(1) = T_0$ where $W_t = \epsilon(1+\epsilon)^{-(t+1)} T_0 \in \mathcal{B}(Q^c(Y), P^c(X))$. We can use the relationship $R(z) = R_{\sin}(z) + R_{\mathrm{reg}}(z)$ and the initial condition $\boldsymbol{x}(-1) = \mathbf{0}$ to rewrite (4.2) as two equations $\boldsymbol{X}_{\sin}(z) = R_{\sin}(z) Q \boldsymbol{\Xi}(z)$ and $\boldsymbol{X}_{\mathrm{reg}}(z) = R_{\mathrm{reg}}(z) Q^c \boldsymbol{\Xi}(z)$ where $\boldsymbol{X}(z) = \boldsymbol{X}_{\sin}(z) + \boldsymbol{X}_{\mathrm{reg}}(z)$. Therefore $\boldsymbol{x}(t) = \boldsymbol{x}_{\sin}(t) + \boldsymbol{x}_{\mathrm{reg}}(t) = U_t \star [Q\boldsymbol{\xi}](t) + W_t \star [Q^c \boldsymbol{\xi}](t)$ for all $t \in \mathbb{N} - 1$ where $\boldsymbol{x}_{\sin} \sim I(1)$ is an AR(1) process in $P(X)$ satisfying

$$(QA_0P)\boldsymbol{x}_{\sin}(t) + (QA_1P)\boldsymbol{x}_{\sin}(t-1) = Q\boldsymbol{\xi}(t)$$

with $\boldsymbol{x}_{\sin}(-1) = \mathbf{0}$ and $\boldsymbol{x}_{\mathrm{reg}} \sim I(0)$ is an AR(1) process in $P^c(X)$ satisfying

$$(Q^cA_0P^c)\boldsymbol{x}_{\mathrm{reg}}(t) + (Q^cA_1P^c)\boldsymbol{x}_{\mathrm{reg}}(t-1) = Q^c\boldsymbol{\xi}(t),$$

with $\boldsymbol{x}_{\mathrm{reg}}(-1) = \mathbf{0}$. The original time series $\boldsymbol{x}$ is therefore the sum of separate AR(1) time series on complementary subspaces. We show in Corollary 4.3 that this is true in general. □

Example 6.2 ($\boldsymbol{x} \sim I(\infty)$)**.** Let $\{\lambda_j\}_{j\in\mathbb{N}} \subset \mathbb{R}$ with $\lambda_j > 0$ and $\lambda_j^{1/j} \downarrow 0$ as $j \uparrow \infty$ and $\{\mu_j\}_{j\in\mathbb{N}} \subset \mathbb{R}$ with $\mu_j > 0$ and $\sum_{j\in\mathbb{N}} \mu_j = 1$. Suppose that $\boldsymbol{x}_j = \{\boldsymbol{x}_j(t)\}_{t\in\mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}$ for each $j \in \mathbb{N}$ and that $\boldsymbol{x}_j, \boldsymbol{x}_{j+1}$ satisfy the recurrence relation

$$\boldsymbol{x}_j(t) - \boldsymbol{x}_j(t-1) - \lambda_j \boldsymbol{x}_{j+1}(t) = \mu_j \boldsymbol{\xi}(t),$$

with $\boldsymbol{x}_j(-1) = 0$ for each $j \in \mathbb{N}$ where $\boldsymbol{\xi} \in \mathbb{R}^{\mathbb{Z}}$ is a strong white noise. An $\mathcal{M}$-transform gives

$$\boldsymbol{X}_j(z) - z\boldsymbol{X}_j(z) - \lambda_j \boldsymbol{X}_{j+1}(z) = \mu_j \boldsymbol{\Xi}(z),$$

for each $j \in \mathbb{N}$, where $\boldsymbol{X}_j(z) = \mathcal{M}[x_j](z)$ and $\boldsymbol{\Xi}(z) = \mathcal{M}[\boldsymbol{\xi}](z)$. In vector form we can write

$$[(1-z)I - U]\boldsymbol{X}(z) = \boldsymbol{\Xi}(z)\boldsymbol{\mu},$$

for each $z \in \mathbb{C}$ where $\boldsymbol{X}(z) = \sum_{j\in\mathbb{N}} \boldsymbol{X}_j(z)\boldsymbol{e}_j \in \mathbb{C}^\infty$, $\boldsymbol{\mu} = \sum_{j\in\mathbb{N}} \mu_j \boldsymbol{e}_j \in \ell^1$, $I = [\boldsymbol{e}_1, \boldsymbol{e}_2, \dots] \in \mathcal{B}(\ell^1)$ is the identity operator and $U = [\boldsymbol{0}, \lambda_1\boldsymbol{e}_1, \lambda_2\boldsymbol{e}_2, \lambda_3\boldsymbol{e}_3, \dots] \in \mathcal{B}(\ell^1)$ is a strictly upper triangular operator. We can see that

$$U^2 = [\boldsymbol{0}, \boldsymbol{0}, \lambda_1\lambda_2\boldsymbol{e}_1, \lambda_2\lambda_3\boldsymbol{e}_2, \lambda_3\lambda_4\boldsymbol{e}_4, \dots], \quad U^3 = [\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}, \lambda_1\lambda_2\lambda_3\boldsymbol{e}_1, \lambda_2\lambda_3\lambda_4\boldsymbol{e}_2, \lambda_3\lambda_4\lambda_5\boldsymbol{e}_3, \dots], \quad \dots$$

and so on. Thus $U^n\boldsymbol{e}_k = \boldsymbol{0}$ for $k \le n$ and $U^n\boldsymbol{e}_{n+k} = \lambda_k\lambda_{k+1} \dots \lambda_{n+k-1}\boldsymbol{e}_k$ for $n, k \in \mathbb{N}$. It follows that $\|U^n\|^{1/n} < \lambda_1\lambda_n^{1/n} \downarrow 0$ as $n \uparrow \infty$. A Neumann expansion now shows that

$$R(z) = [(1-z)I - U]^{-1} = I/(1-z) + U/(1-z)^2 + U^2/(1-z)^3 + \cdots,$$

for all $z \neq 1$. Therefore $R(z)$ has an isolated essential singularity at $z = 1$. Hence $\boldsymbol{x} \sim I(\infty)$. If we define $V = \mathcal{M}^{-1}[R]$ then

$$V(t) = \left[I + U\Delta^{-1} + U^2\Delta^{-2} + \cdots\right][\mathbf{1}](t) = I + \binom{t+1}{1} U + \binom{t+2}{2} U^2 + \cdots,$$

where $\mathbf{1}(t) = 1$ and $\binom{t+n}{n} = (t+n)(t+n-1) \dots (t+1)/n!$ for all $t \in \mathbb{N} - 1$ and $n \in \mathbb{N}$. Because $\boldsymbol{\Xi}(z) \in \mathbb{C}$ is a scalar it follows that

$$\mathcal{M}^{-1}[R(z)\boldsymbol{\Xi}(z)\boldsymbol{\mu}](t) = \mathcal{M}^{-1}[\boldsymbol{\Xi}(z)R(z)\boldsymbol{\mu}](t) = (\boldsymbol{\xi} * V)(t)\boldsymbol{\mu},$$

where $(\boldsymbol{\xi} * V)(t) = \sum_{s=0}^{t} \boldsymbol{\xi}(t-s)V(s)$ for all $t \in \mathbb{N} - 1$. Therefore

$$\boldsymbol{x}(t) = \sum_{n\in\mathbb{N}-1} \left[\sum_{s=0}^{t} \boldsymbol{\xi}(t-s)\binom{s+n}{n}\right] U^n\boldsymbol{\mu},$$

for all $t \in \mathbb{N} - 1$. We can extract the component $\boldsymbol{x}_k(t) = \langle \boldsymbol{e}_k, \boldsymbol{x}(t)\rangle$ for each $k \in \mathbb{N}$. In this regard we note that $\langle \boldsymbol{e}_k, U^n\boldsymbol{\mu}\rangle = \lambda_k \dots \lambda_{k+n-1}\mu_{k+n}$ for $n \in \mathbb{N}$ and so

$$\begin{aligned}\boldsymbol{x}_k(t) &= \left[\sum_{s=0}^{t} \boldsymbol{\xi}(t-s)\right]\mu_k + \sum_{n\in\mathbb{N}} \left[\sum_{s=0}^{t} \boldsymbol{\xi}(t-s)\binom{s+n}{n}\right]\lambda_k \dots \lambda_{k+n-1}\mu_{k+n} \\ &= \Delta^{-1}\boldsymbol{\xi}_+(t)\mu_k + \sum_{n\in\mathbb{N}} \left[\Delta^{-n-1}\boldsymbol{\xi}_+(t)\right]\lambda_k \dots \lambda_{k+n-1}\mu_{k+n},\end{aligned}$$

Journal of Time Series Analysis published by John Wiley & Sons Ltd.
DOI: 10.1111/jtsa.12766

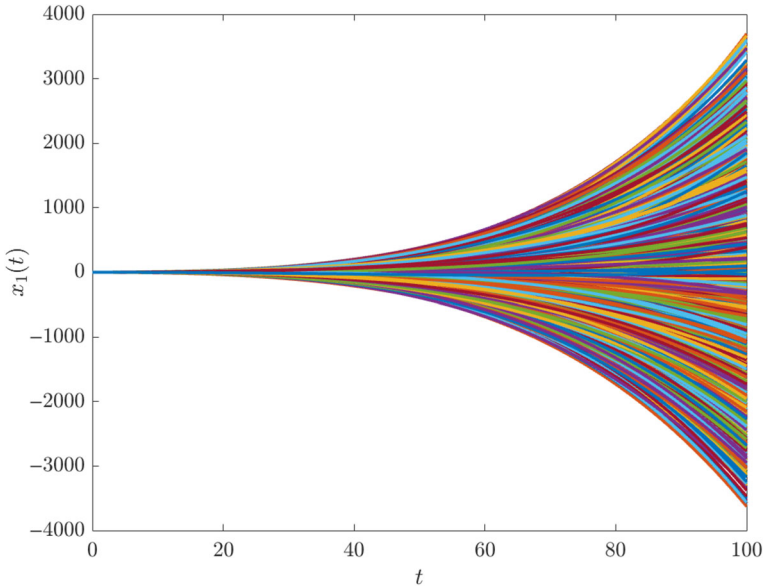

Figure 1. Composite graph for the simulated values $\{\boldsymbol{x}_1(t)\}_{t\in\{0,1,\dots,100\}}$ in Example 6.2

for all $t \in \mathbb{N} - 1$. The final form shows that no finite-order difference will eliminate all of the unbounded terms. This confirms that $\boldsymbol{x}_k \sim I(\infty)$ for each $k \in \mathbb{N}$.

We used MATLAB to simulate the evolution of the component $\boldsymbol{x}_1 \sim I(\infty)$ setting $\lambda_j = 1/j!$ and $\mu_j = 1/2^{j+1}$ with noise defined by a pseudo-random variable with a uniform distribution on the interval $(-1/2, 1/2)$. We performed 20,000 independent trials for the series $\{\boldsymbol{x}_1(t)\}_{t\in\{0,1,\dots,100\}}$. The composite graph is shown in Figure 1. The histogram for the frequency distribution of the normalized values $[\boldsymbol{x}_1(t) - m(\boldsymbol{x}_1(t)]/s(\boldsymbol{x}_1(t)$ at $t = 100$ is shown in Figure 2. The mean and SD for the trials were $m[x_1(100)] \approx -12.8811$ and $s[\boldsymbol{x}_1(100)] \approx 1.0051 \times 10^3$. Although the composite graph shows that the range of values for $\boldsymbol{x}_1(t)$ increases rapidly as t increases the histogram shows that the maximum frequency occurs when $\boldsymbol{x}_1(t) \approx 0$. □

Example 6.3 (**The general solution process**). Let $X = Y = c_0$ be the space of real-valued vectors $\boldsymbol{x} = [\boldsymbol{x}_n]_{n\in\mathbb{N}}$ with $\lim_{n\to\infty} \boldsymbol{x}_n = 0$ and norm $\|\boldsymbol{x}\|_\infty = \max_{n\in\mathbb{N}} |\boldsymbol{x}_n|$. The space c_0 is a closed separable subspace of ℓ^∞ with basis $\{\boldsymbol{e}_n\}_{n\in\mathbb{N}}$. Let $\lambda \in (0, 1)$ and suppose the time series $\boldsymbol{x} \in X^{\mathbb{Z}}$ satisfies $A_0\boldsymbol{x}(t) + A_1\boldsymbol{x}(t-1) = \boldsymbol{\xi}(t)$ for all $t \in \mathbb{Z}$ with $\boldsymbol{x}(-1) = \boldsymbol{c} \in X$ where

$$A_0 = I = [\boldsymbol{e}_1, \boldsymbol{e}_2, \dots] \in \mathcal{B}(X, Y) \quad \text{and} \quad A_1 = (-1)[\boldsymbol{e}_1, \boldsymbol{e}_1 + \boldsymbol{e}_2, \lambda\boldsymbol{e}_3, \lambda^2\boldsymbol{e}_4, \lambda^3\boldsymbol{e}_5, \dots] \in \mathcal{B}(X, Y).$$

The autoregressive polynomial is $A(z) = A_0 + A_1 z \in \mathcal{B}(X, Y)$ for all $z \in \mathbb{C}$. The operator $A(1) = A_0 + A_1$ is singular and so we write $A(z) = B_0 + B_1(z-1)$ where

$$B_0 = A_0 + A_1 = [\boldsymbol{0}, -\boldsymbol{e}_1, (1-\lambda)\boldsymbol{e}_3, (1-\lambda^2)\boldsymbol{e}_4, (1-\lambda^3)\boldsymbol{e}_5, \dots] \quad \text{and} \quad B_1 = A_1.$$

We wish to find a Laurent series $R(z) = \sum_{j\in\mathbb{Z}} T_j(z-1)^j$ which converges on some region $z \in D_{0,\epsilon}(1)$. Thus we need to find $\{T_j\}_{j\in\mathbb{Z}} \in \mathcal{B}(Y, X)$ satisfying the fundamental equations (3.1) and (3.2) and the magnitude constraints (3.3). We follow Albrecht *et al.* (2020) and begin by finding the infinite-length Jordan chains at $z = 1$.

The singular Jordan chains $\{\boldsymbol{u}_{-n}\}_{n\in\mathbb{N}}$ satisfy

$$B_0\boldsymbol{u}_{-n} + B_1\boldsymbol{u}_{-n-1} = \boldsymbol{0}, \tag{6.1}$$

DOI: 10.1111/jtsa.12766

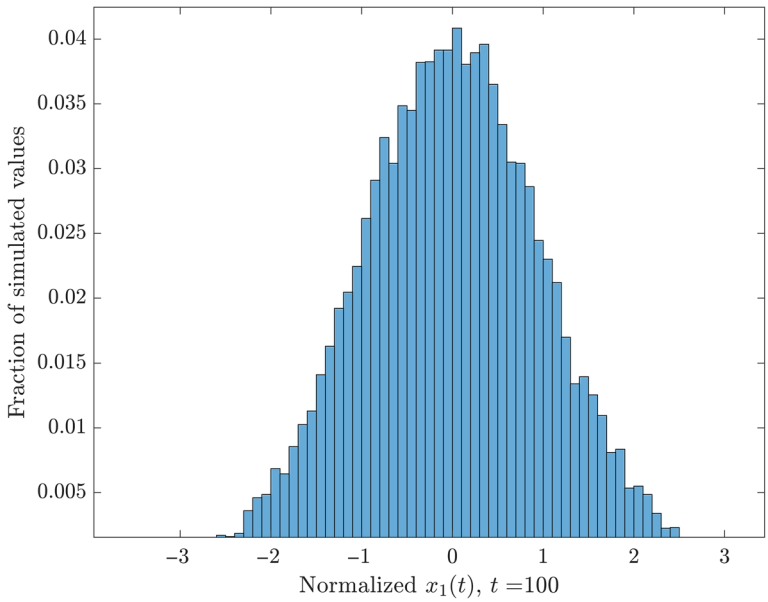

Figure 2. Histogram showing the frequency distributed of simulated values for the normalized variable $[\boldsymbol{x}_1(t) - m(\boldsymbol{x}_1(t))]/s(\boldsymbol{x}_1(t))$ at $t = 100$ in Example 6.2

for all $n \in \mathbb{N}$ with $\|\boldsymbol{u}_{-n}\|_\infty^{1/n} \to 0$ as $n \to \infty$. If we set $\boldsymbol{u}_{-1} = \boldsymbol{t} = \boldsymbol{t}_1\boldsymbol{e}_1 + \boldsymbol{t}_2\boldsymbol{e}_2 + \boldsymbol{t}_3\boldsymbol{e}_3 + \cdots$ then (6.1) shows that $\boldsymbol{u}_{-n,2} = 0$ for $n \geq 2$, $\boldsymbol{u}_{-n,1} = -\boldsymbol{u}_{-n+1,2} = 0$ for $n \geq 3$ and $\boldsymbol{u}_{-n,j} = (1 - 1/\lambda^{j-2})^{n-1}\boldsymbol{t}_j$ for all $n \in \mathbb{N}$ and all $j \geq 3$. If $\boldsymbol{t}_j \neq 0$ for $j \geq 3$ then

$$\|\boldsymbol{u}_{-n}\|_\infty^{1/n} \geq \left(1/\lambda^{j-2} - 1\right)^{1-1/n}|\boldsymbol{t}_j|^{1/n} \to \left(1/\lambda^{j-2} - 1\right) > 0,$$

for each $j \geq 3$ as $n \to \infty$. Thus we must choose $\boldsymbol{t}_j = 0$ for $j \geq 3$. Therefore $\boldsymbol{u}_{-1} = \boldsymbol{t}_1\boldsymbol{e}_1 + \boldsymbol{t}_2\boldsymbol{e}_2$, $\boldsymbol{u}_{-2} = \alpha\boldsymbol{e}_1$ for some $\alpha \in \mathbb{R}$ and $\boldsymbol{u}_{-n} = \boldsymbol{0}$ for $n \geq 3$. This means the subspace $X_{\sin} \subset X$ is the closed subspace $\mathrm{sp}(\{\boldsymbol{e}_1, \boldsymbol{e}_2\}) \subset c_0$.

The regular Jordan chains $\{\boldsymbol{u}_n\}_{n\in\mathbb{N}}$ satisfy

$$B_0\boldsymbol{u}_{n+1} + B_1\boldsymbol{u}_n = \boldsymbol{0}, \tag{6.2}$$

for all $n \in \mathbb{N}$ with $\|\boldsymbol{u}_n\|_\infty^{1/n} \to a$ for some $a > 0$. If we set $\boldsymbol{u}_1 = \boldsymbol{s} = \boldsymbol{s}_1\boldsymbol{e}_1 + \boldsymbol{s}_2\boldsymbol{e}_2 + \boldsymbol{s}_3\boldsymbol{e}_3 + \cdots$ then (6.2) shows that $\boldsymbol{u}_{1,2} = \boldsymbol{s}_2 = 0$. We also have $\boldsymbol{u}_{2,2} = 0$ and $\boldsymbol{u}_{2,2} = -\boldsymbol{u}_{1,1} = -\boldsymbol{s}_1$. Therefore $\boldsymbol{s}_1 = 0$ as well. The remaining equations show that $\boldsymbol{u}_{n,j} = [1 - 1/(1 - \lambda^{j-2})]^{n-1}\boldsymbol{s}_j$ for all $n \in \mathbb{N}$ and all $j \geq 3$. If we choose $\boldsymbol{u}_1 = \boldsymbol{e}_k$ for some $k \geq 3$ then $\boldsymbol{u}_n = [1 - 1/(1 - \lambda^{k-2})]^{n-1}\boldsymbol{e}_k$ and hence

$$\|\boldsymbol{u}_n\|_\infty^{1/n} = \left[1/\left(1 - \lambda^{k-2}\right) - 1\right]^{1-1/n} \to \left[1/\left(1 - \lambda^{k-2}\right) - 1\right] = \lambda^{k-2}/\left(1 - \lambda^{k-2}\right) \leq \lambda/(1 - \lambda)$$

as $n \to \infty$. It follows that $X_{\mathrm{reg}} \subset X$ is the closed subspace $c_0 \cap \mathrm{sp}(\{\boldsymbol{e}_3, \boldsymbol{e}_4, \ldots\})^a \subset c_0$.

Therefore $X = X_{\sin} \oplus X_{\mathrm{reg}} \cong X_{\sin} \times X_{\mathrm{reg}}$ and $Y = Y_{\sin} \oplus Y_{\mathrm{reg}} \cong Y_{\sin} \times Y_{\mathrm{reg}}$ where $Y_{\sin} = B_1(X_{\sin}) = \mathrm{sp}(\{\boldsymbol{e}_1, \boldsymbol{e}_2\})$ and $Y_{\mathrm{reg}} = B_0(X_{\mathrm{reg}}) = c_0 \cap \mathrm{sp}(\{\boldsymbol{e}_3, \boldsymbol{e}_4, \ldots\})^a$. The corresponding key projections are $P = Q = [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{0}, \boldsymbol{0}, \ldots] \in \mathcal{B}(c_0)$ and the basic solution is found by solving $T_{-1}B_1 = P$ and $B_1T_{-1} = Q$ to find $T_{-1} \in \mathcal{B}(Y, X)$ and by solving $T_0B_0 = P^c$ and $B_0T_0 = Q^c$ to find $T_0 \in \mathcal{B}(Y, X)$. Elementary algebra gives

$$T_{-1} = [-\boldsymbol{e}_1, \boldsymbol{e}_1 - \boldsymbol{e}_2, \boldsymbol{0}, \boldsymbol{0}, \ldots] \quad \text{and} \quad T_0 = [\boldsymbol{0}, \boldsymbol{0}, \{1/(1 - \lambda)\}\boldsymbol{e}_3, \{1/(1 - \lambda^2)\}\boldsymbol{e}_4, \ldots].$$

Journal of Time Series Analysis published by John Wiley & Sons Ltd.
DOI: 10.1111/jtsa.12766

The algebra is simply Gaussian elimination but the calculation of T_0 requires consideration of both $T_0B_0 = P^c$ and $B_0T_0 = Q^c$. Now we can calculate the entire Laurent series using $T_{-k} = (-1)^{k-1}(T_{-1}B_0)^{k-1}T_{-1} \in \mathcal{B}(Y,X)$ and $T_\ell = (-1)^\ell(T_0B_1)^\ell T_0 \in \mathcal{B}(Y,X)$ which gives

$$T_{-2} = [\mathbf{0}, \boldsymbol{e}_1, \mathbf{0}, \mathbf{0}, \dots] \quad \text{and} \quad T_{-k} = 0 \text{ for } k \geq 3,$$

and

$$T_\ell = [\mathbf{0}, \mathbf{0}, \{\lambda^\ell/(1-\lambda)^{\ell+1}\}\boldsymbol{e}_3, \{\lambda^{2\ell}/(1-\lambda^2)^{\ell+1}\}\boldsymbol{e}_4, \dots] \text{ for } \ell \in \mathbb{N}.$$

Therefore $R(z)$ has a pole of order 2 at $z = 1$. The singular part is $R_{\sin}(z) = T_{-2}/(z-1)^2 + T_{-1}/(z-1)$ for $z \neq 1$ and the regular part is $R_{\text{reg}}(z) = T_0 + T_1(z-1) + T_2(z-1)^2 + \cdots$ for $z \in D_{1/(1-\lambda)}(1)$. Thus the Laurent series converges for $z \in D_{0,1/(1-\lambda)}(1)$. □

7. AN APPLICATION TO THE CONTROL OF ELECTRICAL POWER GENERATION

We consider an application to the dynamic control of electrical power generation in a wind farm. The proposed feedback control is designed to mitigate the high volatility of power generation in wind farms observed by Agrawal *et al.* (2013). The time series $\boldsymbol{x}$ is a non-autonomous AR(1) process. Thus – strictly speaking – this example lies beyond the theoretical scope of the article. However, we will show that the new methods can be used to find $\boldsymbol{x}(t)$.

Example 7.1 (Power generation)**.** Business Managers of Australian wind-farms sign short-term contracts with the Australian Energy Market Operator to supply an agreed number $H = H_{\mathcal{T}}$ of kilowatt-hours of electrical energy during the next trading interval $t \in \mathcal{T} = \{1, \dots, N\}$. Let $\boldsymbol{w} = \{\boldsymbol{w}(t)\}_{t\in\mathbb{N}_{-1}} \in \mathbb{R}^{\mathbb{N}_{-1}}$ be the number of kilowatt-hours of electrical energy collected on the interval $[0,t]$. To moderate the observed high volatility we seek a suitable feedback mechanism that encourages an appropriate target for the energy $\Delta\boldsymbol{w}(t) = \boldsymbol{w}(t) - \boldsymbol{w}(t-1)$ collected during each subinterval $[t-1,t]$ for $t \in \mathcal{T}$. Our aim will be to ensure that

$$\boldsymbol{x}(t) = \boldsymbol{w}(t) - Ht/N \to 0$$

as $t \to N$ where $\boldsymbol{x}(t)$ is the notional cumulative error at time $t \in \mathcal{T}$. Intuitively we could manage the error by setting $\boldsymbol{x}(t) - \boldsymbol{x}(t-1) = -p\boldsymbol{x}(t-1)/(t+q)$ for some $p, q \in \mathbb{N}$ with $p \leq q$ and all $t \in \mathcal{T}$. Thus we would set $\boldsymbol{x}(t) = [1 - p/(t+q)]\boldsymbol{x}(t-1)$. Now $1 - p/(t+q) \in (0,1)$ for all $t \in \mathcal{T}$ and so either $\boldsymbol{x}(t-1) \leq \boldsymbol{x}(t) \leq 0$ or $0 \leq \boldsymbol{x}(t) \leq \boldsymbol{x}(t-1)$. Thus the notional error must decrease in magnitude. In practice, we observe $\boldsymbol{x}_\xi(t) = \boldsymbol{x}(t) + \boldsymbol{\xi}(t)$ where $\boldsymbol{\xi}(t)$ is an inherent stochastic noise. Hence our target becomes

$$\boldsymbol{x}(t) = [1 - p/(t+q)]\boldsymbol{x}_\xi(t-1) \Leftrightarrow \boldsymbol{x}_\xi(t) = [1 - p/(t+q)]\boldsymbol{x}_\xi(t-1) + \boldsymbol{\xi}(t), \tag{7.1}$$

for all $t \in \mathcal{T}$. Suppose (7.1) is true for all $t \in \mathbb{N}$ and assume $\boldsymbol{x}_\xi(0) = 0$. It follows that

$$\sum_{t\in\mathbb{N}} \boldsymbol{x}_\xi(t)z^t = z\sum_{t\in\mathbb{N}} \boldsymbol{x}_\xi(t-1)z^{t-1} - (p/z^q)\left[\sum_{t\in\mathbb{N}} \boldsymbol{x}_\xi(t-1)z^{t+q}/(t+q)\right] + \sum_{t\in\mathbb{N}} \boldsymbol{\xi}(t)z^t,$$

which we can rewrite in terms of the $\mathcal{M}$-transform as

$$\boldsymbol{X}_\xi(z) = z\boldsymbol{X}_\xi(z) - W\boldsymbol{X}_\xi(z) + \boldsymbol{\Xi}(z) - \boldsymbol{\xi}(0) \Leftrightarrow [(1-z)I + W]\boldsymbol{X}_\xi(z) = \boldsymbol{\Xi}(z) - \boldsymbol{\xi}(0), \tag{7.2}$$

where W is a weighted Volterra operator defined by the formula $W\boldsymbol{F}(z) = (p/z^q)\int_0^z F(\zeta)\zeta^q d\zeta$ for all $\boldsymbol{F}(z) = \sum_{t\in\mathbb{N}_{-1}} \boldsymbol{f}(t)z^t$ which are analytic when $z \in D_r(0)$ for some $r > 0$. An intuitive application of the Neumann expansion suggests that we may have

$$[I(z-1) + W]^{-1} = \left\{I - (1-z)^{-1}W + [(1-z)^{-1}W]^2 - [(1-z)^{-1}W]^3 + \cdots\right\}(1-z)^{-1}, \tag{7.3}$$

DOI: 10.1111/jtsa.12766 *Journal of Time Series Analysis* published by John Wiley & Sons Ltd.

for $z \neq 1$. This is not entirely true but a detailed justification of (7.3) for $z \in D_1(0)$ is given in section C of the Appendix. Order is important because $(1-z)^{-1}W \neq W(1-z)^{-1}$. The formula (7.3) suggests that $[I(z-1)+W]^{-1}$ has an essential singularity at $z=1$ and that $\boldsymbol{x}_\xi \sim I(\infty)$. If $z \in D_1(0)$ and we write $(1-z)^{-1} = 1+z+z^2+\cdots$ then straightforward calculations give

$$(1-z)^{-1}[\Xi(z)-\xi(0)] = \xi(1)z + [\xi(1)+\xi(2)]z^2 + [\xi(1)+\xi(2)+\xi(3)]z^3 + \cdots,$$

$$\begin{aligned}&(1-z)^{-1}W(1-z)^{-1}[\Xi(z)-\xi(0)]\\ &\quad = \{p\xi(1)/(q+2)\}z^2 + \{p\xi(1)/(q+2) + p[\xi(1)+\xi(2)]/(q+3)\}z^3 + \cdots,\end{aligned}$$

$$\begin{aligned}&[(1-z)^{-1}W]^2(1-z)^{-1}[\Xi(z)-\xi(0)] = \{p^2\xi(1)/[(q+2)(q+3)]\}z^3\\ &\quad + \{p^2\xi(1)/[(q+2)(q+3)] + \left[\{p\xi(1)/(q+2) + p[\xi(1)+\xi(2)]/(q+3)\}/(q+4)\right]\}z^4 + \cdots\end{aligned}$$

and so on. By collecting the coefficients of the various powers we have

$$\begin{aligned}&[(1-z)I+W]^{-1}[\Xi(z)-\xi(0)]\\ &\quad = \xi(1)z + \{[1-p/(q+2)]\xi(1) + \xi(2)\}z^2\\ &\qquad + \{[1-p/(q+3)][1-p/(q+2)]\xi(1) + [1-p/(q+3)]\xi(2) + \xi(3)\}z^3 + \cdots\\ &\quad = \boldsymbol{x}_\xi(1)z + \boldsymbol{x}_\xi(2)z^2 + \boldsymbol{x}_\xi(3)z^3 + \cdots.\end{aligned}$$

In general we can show that

$$\boldsymbol{x}_\xi(t) = \begin{cases} \xi(1) & \text{for } t=1\\ \xi(t) + \sum_{s=1}^{t-1}\left\{\prod_{r=t-s+1}^{t}[1-p/(q+r)]\right\}\xi(t-s) & \text{for } t \in \mathbb{N}+1.\end{cases}$$

Therefore $\boldsymbol{x}_\xi(t)$ may grow without bound as $t \to \infty$. It can be shown that $\Delta^k \boldsymbol{x}_\xi(t)$ may also grow without bound for all $k \in \mathbb{N}$ as $t \to \infty$. Thus $\boldsymbol{x}_\xi \sim I(\infty)$. The corresponding increment is $\Delta \boldsymbol{w}_\xi(t) = -p\boldsymbol{x}_\xi(t-1)/(t+q) + H/N$ for all $t \in \mathcal{T}$. In practice p, q could be adjusted to get the best results. The detailed calculations are justified in section C of the Appendix. □

Remark 7.2. In retrospect we could change the dependent variable to $\boldsymbol{y}_\xi(t) = (t+q)^{(q-p)}\boldsymbol{x}_\xi(t)$ in Example 7.1. Now $\boldsymbol{y}_\xi(t) - \boldsymbol{y}_\xi(t-1) = (t+q)^{(q-p)}\xi(t)$ for all $t \in \mathbb{N}$ with $\boldsymbol{y}_\xi(0) = 0$ where the factorial power is defined by $(t+r)^{(s)} = (t+r)(t+r-1)\cdots(t+r-s+1)$ for all $t, r, s \in \mathbb{N}-1$ with $s \in [r, t+r]$. Hence $\boldsymbol{y}_\xi$ is an autonomous AR(1) process with an unbounded $I(\infty)$ noise term. A simple recursion gives $\boldsymbol{y}_\xi(t) = (t+q)^{(q-p)}\xi(t) + (t-1+q)^{(q-p)}\xi(t-1) + \cdots (1+q)^{(q-p)}\xi(1)$ for all $t \in \mathbb{N}$. Hence $\boldsymbol{y}_\xi(t)$ may grow without bound as $t \to \infty$. It can be shown that $\Delta^k \boldsymbol{y}_\xi(t)$ may also grow without bound for each $k \in \mathbb{N}$ as $t \to \infty$. Therefore $\boldsymbol{y}_\xi \sim I(\infty)$.

8. CONCLUSIONS

We have used the new formula $R(z) = [(z-1)I_X + T_{-1}B_0]^{-1}T_{-1} + [I_X + T_0B_1(z-1)]^{-1}T_0$ for the resolvent of the autoregressive polynomial to establish an extended GJRT that applies to all integrated autoregressive processes irrespective of the dimension of the space, the number of stochastic trends and cointegrating relations, and the order of integration, all of which may be finite or infinite. The present results are a substantial extension of those reported in Beare *et al.* (2017), Beare and Seo (2020), Franchi and Paruolo (2020), and Seo (2023). Standard linear algebra can be used to find the key spectral separation projections in finite-dimensional problems but it may be necessary to use infinite-length Jordan chains in more general problems.

Journal of Time Series Analysis published by John Wiley & Sons Ltd. DOI: 10.1111/jtsa.12766

ACKNOWLEDGEMENTS

Massimo Franchi gratefully acknowledges partial financial assistance from Ministero dell' Università e della Ricerca for the project PRIN 2022372SWE *Methodological and computational issues in large-scale time series models for economics and finance*. Phil Howlett and Brendan Beare wish to thank the organizers of ANZESG 2023 and the participants for critiquing the initial presentation of this research. Open access publishing facilitated by University of South Australia, as part of the Wiley - University of South Australia agreement via the Council of Australian University Librarians.

DATA AVAILABILITY STATEMENT

There is no data used in this document.

REFERENCES

Agrawal MR, Boland J, Ridley B. 2013. Analysis of wind farm output: estimation of volatility using high-frequency data. *Environmental Modeling and Assessment* **18**:481–492. https://doi.org/10.1007/s10666-013-9357-3.

Albrecht A, Howlett P, Pearce C. 2014. The fundamental equations for inversion of operator pencils on Banach space. *Journal of Mathematical Analysis and Applications* **413**:411–421. https://doi.org/10.1016/j.jmaa.2013.11.060 See also *arXiv*, arXiv.2202.12410 for a corrected version.

Albrecht A, Howlett P, Verma G. 2020. Inversion of operator pencils on Banach space using Jordan chains when the generalized resolvent has an isolated essential singularity. *Linear Algebra and Its Applications* **595**:33–62. https://doi.org/10.1016/j.laa.2020.02.030.

Aue A, van Delft A. 2020. Testing for stationarity of functional time series in the frequency domain. *The Annals of Statistics* **48**(5):2505–2547. https://doi.org/10.1214/19-AOS1895.

Avrachenkov KE, Haviv M, Howlett PG. 2001. The inversion of singularly perturbed matrices. *SIAM Journal on Matrix Analysis and Applications* **22**(4):1175–1189. https://doi.org/10.1137/S0895479898337555.

Avrachenkov KE, Filar JA, Howlett PG. 2013. *Analytic Perturbation Theory and Its Applications*, Vol. OT 135. Philadelphia: SIAM.

Beare BK, Seo W-K. 2020. Representation of I(1) and I(2) autoregressive Hilbertian processes. *Econometric Theory* **36**(5):773–802. https://doi.org/10.1017/S0266466619000276.

Beare BK, Seo J, Seo W-K. 2017. Cointegrated linear processes in Hilbert space. *Journal of Time Series Analysis* **38**(6):1010–1027. https://doi.org/10.1111/jtsa.12251.

Bosq D. 2000. *Linear Processes in Function Spaces*. New York: Springer-Verlag.

Chang Y, Kim CS, Park JY. 2016. Nonstationarity in time series of state densities. *Journal of Econometrics* **192**:152–167. https://doi.org/10.1016/j.jeconom.2015.06.025.

Chang Y, Kaufmann RK, Kim CS, Isaac Miller J, Park JY, Park S. 2020. Evaluating trends in time series of distributions: a spatial fingerprint of human effects on climate. *Journal of Econometrics* **214**:274–294. https://doi.org/10.1016/j.jeconom.2019.05.014.

Doob JL. 1953. *Stochastic Processes*. New York: John Wiley & Sons.

Engle RF, Granger CWJ. 1987. Co-integration and error correction: representation, estimation, and testing. *Econometrica* **55**:251–276. https://doi.org/10.2307/1913236.

Franchi M, Paruolo P. 2016. Inverting a matrix function around a singularity via local rank factorization. *SIAM Journal on Matrix Analysis and Applications* **37**:774–797. https://doi.org/10.1137/140999839.

Franchi M, Paruolo P. 2020. Cointegration in functional autoregressive processes. *Econometric Theory* **36**(5):803–839. https://doi.org/10.1017/S0266466619000306.

Gabrys R, Hörmann S, Kokoszka P. 2013. Monitoring the intraday volatility pattern. *Journal of Time Series Econometrics* **5**:87–116. https://doi.org/10.1515/jtse-2012-0006.

Hansen PR. 2005. Granger's representation theorem: a closed-form expression for $I(1)$ processes. *The Econometrics Journal* **8**(1):23–28. https://doi.org/10.1111/j.1368-423X.2005.00149.x.

Hörmann S, Horváth L, Reeder R. 2013. A functional version of the ARCH model. *Econometric Theory* **29**:267–288. https://doi.org/10.1017/S0266466612000345.

Horváth L, Kokoszka P. 2012. *Inference for Functional Data with Applications*, Vol. 2012. New York: Springer.

Horváth L, Kokoszka P, Rice G. 2014. Testing stationarity of functional time series. *Journal of Econometrics* **179**:66–82. https://doi.org/10.1016/j.jeconom.2013.11.002.

Howlett PG. 1982. Input retrieval in finite dimensional linear systems. *ANZIAM Journal* **23**:357–382. https://doi.org/10.1017/S033427000000031X.

Howlett P, Avrachenkov K, Pearce C, Ejov V. 2009. Inversion of analytically perturbed linear operators that are singular at the origin. *Journal of Mathematical Analysis and Applications* **353**:68–84. https://doi.org/10.1016/j.jmaa.2008.11.074.

Hu B, Park JY. 2016. *Econometric Analysis of Functional Dynamics in the Presence of Persistence* Indiana University; 1–70 cireqmontreal.com/wp-content/uploads/2017/03/park1.pdf.
Johansen S. 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* **59**:1551–1580. https://doi.org/10.2307/2938278.
Johansen S. 1992. A representation of vector autoregressive processes integrated of order 2. *Econometric Theory* **8**(2):188–202. https://doi.org/10.1017/S0266466600012755.
Kargin V, Onatski A. 2008. Curve forecasting by functional autoregression. *Journal of Multivariate Analysis* **99**:2508–2526. https://doi.org/10.1016/j.jmva.2008.03.001.
Kokoszka P, Reimherr M. 2017. *Introduction to Functional Data Analysis*. Boca Raton: Chapman and Hall.
Kokoszka P, Young G. 2016. KPSS test for functional time series. *Statistics* **50**:957–973. https://doi.org/10.1080/02331888.2015.1128937.
Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* **54**:159–178. https://doi.org/10.1016/0304-4076(92)90104-Y.
Luenberger DG. 1969. *Optimization by Vector Space Methods* John Wiley & Sons, New York.
Nielsen MØ, Seo W-K, Seong D. 2022. Inference on the dimension of the nonstationary subspace in functional time series. *Econometric Theory* **39**:443–480. https://doi.org/10.1017/S0266466622000111.
Seo W-K. 2023. Cointegration and representation of cointegrated autoregressive processes in Banach spaces. *Econometric Theory* **39**(4):737–788. https://doi.org/10.1017/S0266466622000172.
Yosida K. 1978. *Functional Analysis*. Classics in Mathematics, 5th ed. Springer-Verlag, New York.

APPENDIX A: NOTATION AND TERMINOLOGY

We use the following notation throughout. In general X, Y will denote complex Banach spaces and (Ω, Σ, μ) will denote a probability space.

- $\mathbb{N} = \{1, 2, \dots\}$ is the set of natural numbers, $\mathbb{N} - 1$ is the set of non-negative integers, $\mathbb{Z} = (-\mathbb{N}) \cup (\mathbb{N} - 1)$ is the set of integers, $\mathbb{N} + m$ is the set $\{1 + m, 2 + m, \dots\}$ for each $m \in \mathbb{Z}$, $\mathbb{R}$ is the set of real numbers, and $\mathbb{C}$ is the set of complex numbers. If $z \in \mathbb{C}$ then $\bar{z} \in \mathbb{C}$ denotes the complex conjugate.
- If $D \subset \mathbb{C}$ we follow Yosida (1978, p. 3) and write D^a to denote the closure of D.
- For $z_0 \in \mathbb{C}$ and $s, r \in [0, \infty] \subset \mathbb{R} \cup \{\infty\}$ with $s < r$ we write $D_{\ r}(z_0) = \{z \in \mathbb{C} | |z - z_0| < r\}$ and $D_{\ s,r}(z_0) = \{z \in \mathbb{C} | s < |z - z_0| < r\}$.
- A linear operator $A : X \to Y$ is bounded if $\|A\boldsymbol{x}\|_Y \leq c\|\boldsymbol{x}\|_X$ for some $c \in \mathbb{R}$ with $c > 0$. The set of all such bounded linear operators is denoted by $\mathcal{B}(X, Y)$. If $A \in \mathcal{B}(X, Y)$ and $T \subset Y$ we write $A^{-1}(T) = \{\boldsymbol{x} \in X | A\boldsymbol{x} \in T\}$ whether or not A is invertible. The set $\mathcal{B}(X, X)$ is denoted by $\mathcal{B}(X)$. The identity operator $I_X \in \mathcal{B}(X)$ is defined by $I_X\boldsymbol{x} = \boldsymbol{x}$ for all $\boldsymbol{x} \in X$. If $P \in \mathcal{B}(X)$ is a projection we write $P^c = I_X - P$ to denote the complementary projection. We may write I instead of I_X if X is clear from the context.
- We follow Luenberger (1969, sections 5.6, 5.7, pp. 115–118) and use the scalar product notation $\langle \boldsymbol{x}, \boldsymbol{f} \rangle \in \mathbb{C}$ to denote the value of the bounded linear functional $\boldsymbol{f} \in X^*$ at the point $\boldsymbol{x} \in X$. If $S \subset X$ is a subset of X then $S^\perp \subset X^*$ is the subspace defined by $S^\perp = \{\boldsymbol{f} \in X^* | \langle \boldsymbol{s}, \boldsymbol{f} \rangle = 0 \text{ for all } \boldsymbol{s} \in S\}$.
- A stochastic time series on X is a sequence of strongly measurable maps from Ω to X. The stochastic time series $\boldsymbol{x} = \{\boldsymbol{x}(t)\}_{t \in \mathbb{Z}} : \Omega \to X^{\mathbb{Z}}$ is a series with value $\boldsymbol{x}(\boldsymbol{\omega}) = \{\boldsymbol{x}(t)\}_{t \in \mathbb{Z}}(\boldsymbol{\omega}) = \{\boldsymbol{x}(\boldsymbol{\omega}, t)\}_{t \in \mathbb{Z}} \in X^{\mathbb{Z}}$ that depends on the outcome $\boldsymbol{\omega} \in \Omega$ of some random process in the probability space (Ω, Σ, μ). Thus the value of the series is simply defined as the corresponding series of values. If $A \in \mathcal{B}(X, Y)$ is a bounded linear operator then the stochastic time series $\boldsymbol{y} = A\boldsymbol{x}$ is defined by $\boldsymbol{y}(\boldsymbol{\omega}) = [A\{\boldsymbol{x}(t)\}_{t \in \mathbb{Z}}](\boldsymbol{\omega}) = \{A\boldsymbol{x}(\boldsymbol{\omega}, t)\}_{t \in \mathbb{Z}} \in Y^{\mathbb{Z}}$ for each $\boldsymbol{\omega} \in \Omega$. The dependence on $\boldsymbol{\omega}$ is normally suppressed unless it is directly relevant to the issue at hand.
- According to Doob (1953, p 94) a stochastic time series $\boldsymbol{x} = \{\boldsymbol{x}(t)\}_{t \in \mathbb{Z}} \in X^{\mathbb{Z}}$ is *strictly stationary* if the multi-variate joint distribution of the variables $\boldsymbol{x}(t_1 + q), \dots, \boldsymbol{x}(t_p + q)$ for each finite collection of times $\{t_1, \dots, t_p\} \in \mathbb{Z}^p$ does not depend on $q \in \mathbb{Z}$.
- For each stochastic time series $\boldsymbol{u} \in X^{\mathbb{Z}}$ we define the associated observable time series $\boldsymbol{u}_+ = \{\boldsymbol{u}_+(t)\}_{t \in \mathbb{Z}} \in X^{\mathbb{Z}}$ by setting $\boldsymbol{u}_+(t) = \boldsymbol{u}(t)$ for $t \in \mathbb{N} - 1$ and $\boldsymbol{u}_+(t) = \boldsymbol{0}$ otherwise.
- For each $\boldsymbol{u} \in X^{\mathbb{Z}}$ the time series $\Delta\boldsymbol{u} \in X^{\mathbb{Z}}$ of first differences is defined using the (backward) difference operator $\Delta\boldsymbol{u}(t) = \boldsymbol{u}(t) - \boldsymbol{u}(t - 1)$ and the corresponding series delayed by one unit of time $L\boldsymbol{u} \in X^{\mathbb{Z}}$ is defined using the lag operator $L\boldsymbol{u}(t) = \boldsymbol{u}(t - 1)$ for each $t \in \mathbb{Z}$. The operator identity $\Delta = I_X - L$ shows that

$$\Delta^{-k}\boldsymbol{u}(t) = (I_X - L)^{-k}\boldsymbol{u}(t) = \sum_{s\in\mathbb{N}-1}\binom{k+s-1}{s}L^s\boldsymbol{u}(t) = \sum_{s\in\mathbb{N}-1}\binom{k+s-1}{s}\boldsymbol{u}(t-s), \tag{A1}$$

for all $k \in \mathbb{N}$, provided $\mathbb{E}[\|\boldsymbol{u}(t-s)\|] \le c(t)/r^s$ for some $c(t), r > 0$ and all sufficiently large $s \in \mathbb{N}$ and that

$$\Delta^{\ell}\boldsymbol{u}(t) = (I_X - L)^{\ell}\boldsymbol{u}(t) = \sum_{s=0}^{\ell}\binom{\ell}{s}(-1)^s L^s\boldsymbol{u}(t) = \sum_{s=0}^{\ell}\binom{\ell}{s}(-1)^s\boldsymbol{u}(t-s), \tag{A2}$$

for all $\ell \in \mathbb{N} - 1$. In each case $\binom{p}{q}$ is the binomial coefficient for $p, q \in \mathbb{N} - 1$ with $p \ge q$.

- A stochastic time series $\boldsymbol{\xi} = \{\boldsymbol{\xi}(\boldsymbol{\omega}, t)\}_{t\in\mathbb{Z}} \in Y^{\mathbb{Z}}$ is called a *strong white noise* process if it is an i.i.d. process with $0 < \sigma^2 = \mathbb{E}\left[\|\boldsymbol{\xi}\|^2\right] = \int_\Omega \|\boldsymbol{\xi}(\boldsymbol{\omega}, t)\|^2 d\mu(\boldsymbol{\omega}) < \infty$ and $\boldsymbol{\mu} = \mathbb{E}[\boldsymbol{\xi}] = \int_\Omega \boldsymbol{\xi}(\boldsymbol{\omega}, t)d\mu(\boldsymbol{\omega}) = \mathbf{0}$ for each $t \in \mathbb{Z}$, where the latter integral is a Bochner integral (Yosida, 1978, chapter V, sections 4 and 5, pp. 130–134). The noise process is not necessarily Gaussian. A strong white noise process is strictly stationary.
- A stochastic time series $\boldsymbol{x} = \{\boldsymbol{x}(\boldsymbol{\omega}, t)\}_{t\in\mathbb{Z}} \in X^{\mathbb{Z}}$ is said to be *integrated of order zero*, written $\boldsymbol{x} \sim I(0)$, if there exists a strong white noise process $\{\boldsymbol{\xi}(\boldsymbol{\omega}, t)\}_{t\in\mathbb{Z}} \in Y^{\mathbb{Z}}$ and a sequence of bounded linear operators $\{A_s\}_{s\in\mathbb{N}-1} \in \mathcal{B}(Y, X)^{\mathbb{N}-1}$ with $\sum_{s\in\mathbb{N}-1}\|A_s\| < \infty$ and $\sum_{s\in\mathbb{N}-1} A_s \neq 0$ so that $\boldsymbol{x}(\boldsymbol{\omega}, t) = \sum_{s\in\mathbb{N}-1} A_s\boldsymbol{\xi}(\boldsymbol{\omega}, t-s)$ for all $t \in \mathbb{Z}$. In Bosq (2000, p. 183) a series in this form is called a *standard linear process*. A standard linear process is strictly stationary and converges almost surely. See Bosq (2000, lemma 7.1, pp. 182–183).
- A stochastic linear process $\boldsymbol{x}$ is said to be *integrated of order d* for some $d \in \mathbb{N}$, written $\boldsymbol{x} \sim I(d)$, if d is the smallest natural number such that $\Delta^d\boldsymbol{x} \sim I(0)$. A stochastic linear process $\boldsymbol{x} \in X^{\mathbb{Z}}$ is said to be *integrated of order infinity*, written $\boldsymbol{x} \sim I(\infty)$ if there is no $d \in \mathbb{N}$ such that $\Delta^d\boldsymbol{x} \sim I(0)$.

APPENDIX B: PROOFS OF THE MAIN RESULTS

Proof of Lemma 3.1. Suppose $R(z)$ is analytic on $D_{1+\epsilon}(0) \setminus \{1\}$ for some $\epsilon > 0$. Therefore $R(z)$ is analytic on $[\,D_{\,1}(0)^a \setminus \{1\}\,] \cup D_{\,0,\,\epsilon}(1) \subset D_{1+\epsilon}(0) \setminus \{1\}$. Now suppose $R(z)$ is analytic on $[\,D_{\,1}(0)^a \setminus \{1\}\,] \cup D_{\,0,\,\delta}(1)$ for some $\delta > 0$. Define the compact set $\mathcal{C} = \{z \in \mathbb{C} \,|\, |z| = 1 \text{ and } z \notin D_{\,\delta/2}(1)\}$. The resolvent $R(z)$ is analytic on $\mathcal{C}$ and so for each $w \in \mathcal{C}$ we can find $\epsilon(w) > 0$ such that $R(z)$ is analytic on the open neighborhood $D_{\,\epsilon(w)}(w)$. Therefore $\mathcal{C} \subset \bigcup_{w\in\mathcal{C}} D_{\,\epsilon(w)}(w)$. The set $\mathcal{C}$ is compact so there exists a finite subcollection of open neighbourhoods $\{D_{\epsilon_i}(w_i)\}_{i=1}^n$ with $w_i \in \mathcal{C}$ and $\epsilon_i > 0$ for each $i = 1, \ldots, n$ such that $\mathcal{C} \subset \bigcup_{i=1}^n D_{\,\epsilon_i}(w_i)$. The set $\bigcup_{i=1}^n D_{\,\epsilon_i}(w_i)$ is open so there is some $\epsilon \in (0, \delta/2)$ such that $N(\mathcal{C}, \epsilon) = \{z \in \mathbb{C} \,|\, |z - w| < \epsilon \text{ for some } w \in \mathcal{C}\} \subset \bigcup_{i=1}^n D_{\,\epsilon_i}(w_i)$. Thus $R(z)$ is analytic for $z \in N(\mathcal{C}, \epsilon)$. Now $D_{1+\epsilon}(0) \setminus \{1\} \subset N(\mathcal{C}, \epsilon) \cup [\,D_{\,1}(0)^a \setminus \{1\}\,] \cup D_{\,0,\,\delta}(1)$. Consequently $R(z)$ is analytic for $D_{1+\epsilon}(0) \setminus \{1\}$. ■

Proof of Lemma 4.1. We observe that $PW(z) = \sum_{t\in\mathbb{N}-1} PW_t z^t$ for all $z \in D_{1+\epsilon}(0)$. However we also know that $PT_\ell = 0$ for all $\ell \in \mathbb{N} - 1$ and so $PW(z) = PR_{\,\text{reg}}(z) = \sum_{\ell\in\mathbb{N}-1} PT_\ell(z-1)^\ell = 0$ for all $z \in D_{\,\epsilon}(1)$. Since $PW(z) \equiv 0$ on a non-trivial open set it follows by analytic continuation that $PW(z) \equiv 0$ for all $z \in D_{1+\epsilon}(0)$ and hence $PW_t = 0$ for all $t \in \mathbb{N} - 1$. A similar argument shows that $W(z)Q = 0$ for all $z \in D_{1+\epsilon}(0)$. Therefore $W_t Q = 0$ for all $t \in \mathbb{N} - 1$. ■

Proof of Proposition 4.2. Consider the term $\boldsymbol{X}_{\,\sin}(z)$. It follows from (3.11) that

$$\begin{aligned}\boldsymbol{X}_{\sin}(z) &= R_{\sin}(z)[\boldsymbol{\Xi}(z) - B_1\boldsymbol{c}] = \sum_{k\in\mathbb{N}} T_{-k}(z-1)^{-k}\left[\boldsymbol{\Xi}(z) - B_1\boldsymbol{c}\right] \\ &= \sum_{k\in\mathbb{N}} T_{-k}\,(-1)^k\mathcal{M}\left[\{\Delta^{-k}(\boldsymbol{\xi} - B_1\boldsymbol{c}\,\boldsymbol{\delta})_+(t)\}_{t\in\mathbb{N}-1}\right](z).\end{aligned} \tag{B1}$$

Now we can apply (3.4), (3.9) and (4.4) to deduce that

$$\begin{aligned}\boldsymbol{X}_{\sin}(z) &= \mathcal{M}\left[\left\{\sum_{k\in\mathbb{N}}(-1)^k T_{-k}\Delta^{-k}\boldsymbol{\xi}_+(t)\right\}_{t\in\mathbb{N}-1}\right](z)\\ &\quad -\mathcal{M}\left[\left\{\sum_{k\in\mathbb{N}}\binom{t+k-1}{k-1}(T_{-1}B_0)^{k-1}T_{-1}B_1\boldsymbol{c}\right\}_{t\in\mathbb{N}-1}\right](z)\\ &= \mathcal{M}\left[\sum_{k\in\mathbb{N}}(-1)^k T_{-k}\Delta^{-k}\boldsymbol{\xi}_+(t) - U_t B_1\boldsymbol{c}\right](z). \end{aligned} \tag{B2}$$

We justify convergence of (B2) as follows. For each $\eta > 0$, we can find $c_\eta > 0$ such that $\|T_{-k}\| \le c_\eta \eta^k$ for all $k \in \mathbb{N}$. If we choose $\eta \in (0,1)$ then

$$\begin{aligned}\mathbb{E}\left[\|\sum_{k\in\mathbb{N}}(-1)^k T_{-k}\Delta^{-k}\boldsymbol{\xi}_+(t)\|\right] &= \mathbb{E}\left[\|\sum_{k\in\mathbb{N}}(-1)^k T_{-k}\sum_{s=0}^{t}\binom{s+k-1}{k-1}\boldsymbol{\xi}(t-s)\|\right]\\ &\le \mathbb{E}\left[\|\boldsymbol{\xi}\|\right]\cdot\sum_{k\in\mathbb{N}} c_\eta\eta^k\cdot\sum_{s=0}^{t}\binom{s+k-1}{k-1}\\ &\le c_\eta\,\mathbb{E}\left[\|\boldsymbol{\xi}\|\right]\cdot\sum_{k\in\mathbb{N}}\binom{t+k}{k}\eta^k\\ &= c_\eta\,\mathbb{E}\left[\|\boldsymbol{\xi}\|\right]\cdot\left[1/(1-\eta)^{t+1}-1\right]\end{aligned}$$

is well defined and finite. Therefore $\boldsymbol{X}_{\sin}(z) = \mathcal{M}\left[\{\boldsymbol{x}_{\sin}(t)\}_{t\in\mathbb{N}-1}\right](z)$ where

$$\boldsymbol{x}_{\sin}(t) = \sum_{k\in\mathbb{N}}(-1)^k T_{-k}\Delta^{-k}\boldsymbol{\xi}_+(t) - U_t B_1\boldsymbol{c} \in P(X) \tag{B3}$$

for each $t \in \mathbb{N}-1$. Now consider the term $\boldsymbol{X}_{\text{reg}}(z)$. We have

$$\begin{aligned}\boldsymbol{X}_{\text{reg}}(z) &= W(z)[G(z) - B_1\boldsymbol{c}]\\ &= \sum_{s\in\mathbb{N}-1} W_s z^s \sum_{r\in\mathbb{N}-1}(\boldsymbol{\xi} - B_1\boldsymbol{c}\,\boldsymbol{\delta})_+(r)z^r\\ &= \sum_{t\in\mathbb{N}-1}\left[\sum_{s\in\mathbb{N}-1} W_s\,(\boldsymbol{\xi} - B_1\boldsymbol{c}\,\boldsymbol{\delta})_+(t-s)\right]z^t\\ &= \sum_{t\in\mathbb{N}-1}\left[\sum_{s=0}^{t} W_s\,\boldsymbol{\xi}(t-s) - W_t\,B_1\boldsymbol{c}\right]z^t\end{aligned} \tag{B4}$$

for all $z \in D_{1+\epsilon}(0)$. Therefore $\boldsymbol{X}_{\text{reg}}(z) = \mathcal{M}\left[\boldsymbol{x}_{\text{reg}}(t)\right](z)$ where

$$\boldsymbol{x}_{\text{reg}}(t) = \sum_{s=0}^{t} W_s\,\boldsymbol{\xi}(t-s) - W_t\,B_1\boldsymbol{c} \in P^c(X), \tag{B5}$$

Journal of Time Series Analysis published by John Wiley & Sons Ltd.
DOI: 10.1111/jtsa.12766

for all $t \in \mathbb{N} - 1$. We know that $\|W_s\| < c/(1+\epsilon)^s$ and so convergence is straightforward with

$$\mathbb{E}\left[\|\sum_{s=0}^{t} W_s\ \xi(t-s)\|\right] \leq \mathbb{E}[\|\xi\|]\sum_{s=0}^{t} c/(1+\epsilon)^s \leq c\ \mathbb{E}[\|\xi\|](1+\epsilon)/\epsilon,$$

for all $t \in \mathbb{N} - 1$. Finally we set $\boldsymbol{x}(t) = \boldsymbol{x}_{\ \sin}(t) + \boldsymbol{x}_{\text{reg}}(t)$ for all $t \in \mathbb{N} - 1$. ■

Proof of Corollary 4.3. We have $R_{\sin}(z) = [(z-1)I_X + T_{-1}B_0]^{-1}T_{-1} = \sum_{t\in\mathbb{N}-1} U_t z^t$ for $z \in D_1(0)$ where $PU_t = U_t$ for all $t \in \mathbb{N} - 1$. Therefore $PR_{\ \sin}(z) = R_{\sin}(z)$. The equation

$$\boldsymbol{X}_{\sin}(z) = R_{\sin}(z)[\boldsymbol{\Xi}(z) - B_1\boldsymbol{c}], \tag{B6}$$

shows that we also have $P\boldsymbol{X}_{\sin}(z) = \boldsymbol{X}_{\sin}(z)$. If we multiply both sides of (B6) on the left by $B_1[(z-1)I_X + T_{-1}B_0]$ and use the identity $T_0B_0 + T_{-1}B_1 = I_X$ we get

$$B_1[(z-1)I_X + T_{-1}B_1][T_0B_0 + T_{-1}B_1]\boldsymbol{X}_{\ \sin}(z) = B_1T_{-1}\boldsymbol{\Xi}(z) - B_1T_{-1}\boldsymbol{c},$$

for all $z \in D_1(0)$. Now we can use the relationships $B_0 = A_0 + A_1$ and $B_1 = A_1$, the identity $T_{-1}B_0T_0 = 0$ and the definitions $P = T_{-1}B_1$ and $Q = B_1T_{-1}$ to deduce that

$$[QA_0P + QA_1Pz]\boldsymbol{X}_{\sin}(z) = Q\boldsymbol{\Xi}(z) - QB_1\boldsymbol{c},$$

for all $z \in D_1(0)$. Finally we have $QB_1 = B_1T_{-1}[B_1T_{-1} + B_0T_0]B_1 = B_1P^2 = A_1P$ from which it follows that $QB_1 = Q^2B_1 = QA_1P$. Thus we obtain

$$[QA_0P + QA_1Pz]\boldsymbol{X}_{\sin}(z) = Q\boldsymbol{\Xi}(z) - QA_1P\boldsymbol{c}, \tag{B7}$$

for all $z \in D_1(0)$. This means that $\boldsymbol{x}_{\sin}(t) \in P(X)$ satisfies the AR(1) equation

$$QA_0P\boldsymbol{x}_{\sin}(t) + QA_1P\boldsymbol{x}_{\sin}(t-1) = Q\boldsymbol{\xi}(t), \tag{B8}$$

for all $t \in \mathbb{N} - 1$ with $\boldsymbol{x}_{\sin}(-1) = QA_1P\boldsymbol{c}$. We know that $QB_iP + Q^cB_iP^c = B_i$ for each $i = 0, 1$. It follows that $QA_iP + Q^cA_iP^c = A_i$ for each $i = 0, 1$. Therefore

$$A_0\boldsymbol{x}(t) + A_1\boldsymbol{x}(t-1) - [QA_0P\boldsymbol{x}_{\sin}(t)(t) + QA_1P\boldsymbol{x}_{\sin}(t-1) = \boldsymbol{\xi}(t) - Q\boldsymbol{\xi}(t),$$

for all $t \in \mathbb{N}-1$. We have $A_i\boldsymbol{x}(t) = (QA_iP + Q^cA_iP^c)[\boldsymbol{x}_{\sin}(t) + \boldsymbol{x}_{\text{reg}}(t)] = QA_iP\boldsymbol{x}_{\sin}(t) + Q^cA_iP^c\boldsymbol{x}_{\text{reg}}(t)$ for each $i = 0, 1$. Therefore the previous equation becomes

$$Q^cA_0P^c\boldsymbol{x}_{\text{reg}}(t) + Q^cA_1P^c\boldsymbol{x}_{\text{reg}}(t-1) = Q^c\boldsymbol{\xi}(t), \tag{B9}$$

for all $t \in \mathbb{N} - 1$ with $\boldsymbol{x}_{\text{reg}}(-1) = \boldsymbol{c} - Q\boldsymbol{c} = Q^c\boldsymbol{c}$. If $\boldsymbol{x} \in X^{\mathbb{N}-1}$ and Assumption 2.1 is true we have shown that $\boldsymbol{x} = \boldsymbol{x}_{\sin} + \boldsymbol{x}_{\text{reg}}$ where $\boldsymbol{x}_{\sin} \in P(X)^{\mathbb{N}-1}$ and $\boldsymbol{x}_{\text{reg}} \in P^c(X)^{\mathbb{N}-1}$. Similar algebraic manipulations show that the reverse implication is also true. ■

Proof of Corollary 4.4. By recursive application of (4.16) we obtain

$$\boldsymbol{x}_{\text{reg}}(t) = \Phi^{t+r+1}\boldsymbol{x}_{\text{reg}}(-1-r) + \sum_{s=0}^{t+r}\Phi^s\boldsymbol{\eta}(t-s), \tag{B10}$$

for all $r \in \mathbb{N} - 1$. In particular for $t = -1$ we have

$$P^c \boldsymbol{c} = \Phi^r \boldsymbol{x}_{\text{reg}}(-1-r) + \sum_{s=0}^{r-1} \Phi^s \boldsymbol{\eta}(-1-s),$$

for all $r \in \mathbb{N} - 1$. We know that

$$\mathbb{E}\left[\| \sum_{s=0}^{r-1} \Phi^s \boldsymbol{\eta}(-1-s) \| \right] \leq \sum_{s=0}^{r-1} k/(1+\epsilon)^s \leq k(1+1/\epsilon),$$

for some $k > 0$ and all $r \in \mathbb{N}$. It follows that $\sum_{s=0}^{r-1} \Phi^s \boldsymbol{\eta}(t-s)$ converges almost surely as $r \to \infty$. Therefore $z_r = \Phi^r \boldsymbol{x}_{\text{reg}}(-1-r)$ also converges almost surely as $r \to \infty$. Thus we obtain

$$z_\infty = \lim_{r\to\infty} z_r = \lim_{r\to\infty} \Phi^r \boldsymbol{x}_{\text{reg}}(-1-r) = P^c \boldsymbol{c} - \sum_{s\in\mathbb{N}-1} \Phi^s \boldsymbol{\eta}(-1-s) \in P^c(X).$$

If we return to (B10) we now have

$$\boldsymbol{x}_{\text{reg}}(t) = \Phi^{t+1} \cdot \Phi^r \boldsymbol{x}_{\text{reg}}(-1-r) + \sum_{s=0}^{t+r} \Phi^s \boldsymbol{\eta}(t-s), \tag{B11}$$

for all $r \in \mathbb{N} - 1$. By taking the limit in (B11) as $r \to \infty$ we get

$$\boldsymbol{x}_{\text{reg}}(t) = \Phi^{t+1} z_\infty + \sum_{s\in\mathbb{N}-1} \Phi^s \boldsymbol{\eta}(t-s),$$

for all $t \in \mathbb{Z}$. If $R(z)$ has a pole of order d at $z = 1$ then

$$\begin{aligned} U_t &= (-1)(I_X - T_{-1}B_0)^{-t-1} \\ &= \sum_{k\in\mathbb{N}-1} \binom{t+k}{k} (T_{-1}B_0)^k T_{-1} \\ &= \sum_{k=0}^{d-1} \binom{t+k}{k} (-1)^k T_{-k-1}, \end{aligned}$$

because $T_{-k-1} = 0$ for $k > d-1$. If $z_\infty = \mathbf{0}$ it follows from (4.13) that

$$\boldsymbol{x}(t) = \sum_{k=0}^{d} (-1)^k T_{-k} \Delta^{-k} \boldsymbol{\xi}_+(t) + \sum_{k=0}^{d-1} (-1)^{k+1} T_{-k-1} \binom{t+k}{k} B_1 \boldsymbol{c} + \sum_{s\in\mathbb{N}-1} \Phi^s \boldsymbol{\eta}(t-s),$$

for all $t \in \mathbb{N} - 1$. Now $\Delta \binom{t+k}{k} = \binom{t+k}{k} - \binom{t-1+k}{k} = \binom{t+k-1}{k-1}$. An inductive argument shows that $\Delta^k \binom{t+k}{k} = 1$. Therefore $\Delta^d \binom{t+k}{k} = 0$ for all $k < d$. Some elementary algebra gives

$$\Delta^d \boldsymbol{x}(t) = T_{-d} \boldsymbol{\xi}(t) + \sum_{s\in\mathbb{N}-1} \Phi^s \Delta^d \boldsymbol{\eta}(t-s),$$

Journal of Time Series Analysis published by John Wiley & Sons Ltd.
DOI: 10.1111/jtsa.12766

for all $t \in \mathbb{N} - 1 + d$. It follows that $\boldsymbol{x} \sim I(d)$, $\boldsymbol{x}_{\sin} \sim I(d)$ and $\boldsymbol{x}_{\text{reg}} \sim I(0)$. If $\boldsymbol{z}_\infty = \boldsymbol{0}$ but $R(z)$ has an isolated essential singularity at $z = 1$ then no finite number of differences will reduce $\boldsymbol{x}_{\sin}(t)$ to a strictly stationary standard linear process. Therefore $\boldsymbol{x} \sim I(\infty)$, $\boldsymbol{x}_{\sin} \sim I(\infty)$ and $\boldsymbol{x}_{\text{reg}} \sim I(0)$. ■

APPENDIX C: THE WEIGHTED VOLTERRA OPERATOR

The weighted Volterra operator W is defined by $W[\boldsymbol{F}(z)] = (p/z^q)\int_0^z \boldsymbol{F}(\zeta)\zeta^q d\zeta$ for all functions $\boldsymbol{F}(z) = \boldsymbol{f}(0) + \boldsymbol{f}(1)z + \boldsymbol{f}(2)z^2 + \cdots$ which are analytic on some region $D_r(0)$ where $r \in (0, 1]$. We wish to justify the formula $S = [(1-z)I + W]^{-1}$ where S is defined by

$$S\boldsymbol{G}(z) := \left[(1-z)^{-1} - [(1-z)^{-1}W](1-z)^{-1} + [(1-z)^{-1}W]^2(1-z)^{-1} - \cdots\right]\boldsymbol{G}(z), \tag{C1}$$

for all $\boldsymbol{G}(z) = \boldsymbol{g}(0) + \boldsymbol{g}(1)z + \boldsymbol{g}(2)z^2 + \cdots$ which are analytic for $z \in D_r(0)$. Consider the operator $[(1-z)I + W]$. Direct calculations show that $[(1-z)I + W]\boldsymbol{F}(z) = \boldsymbol{G}(z)$ where

$$\boldsymbol{g}(0) = \boldsymbol{f}(0),\ \boldsymbol{g}(1) = \boldsymbol{f}(1) - [1 - p/(q+1)]\boldsymbol{f}(0),\ \boldsymbol{g}(2) = \boldsymbol{f}(2) - [1 - p/(q+2)]\boldsymbol{f}(1), \cdots, \tag{C2}$$

and so on. If $\boldsymbol{F}(z)$ is analytic for $D_r(0)$ it follows that $\boldsymbol{G}(z) = \boldsymbol{g}(0) + \boldsymbol{g}(1)z + \boldsymbol{g}(2)z^2 + \cdots$ is also analytic for $z \in D_r(0)$. Solving (C2) gives

$$\begin{aligned} \boldsymbol{f}(0) &= \boldsymbol{g}(0), \quad \boldsymbol{f}(1) = \boldsymbol{g}(1) + [1 - p/(q+1)]\boldsymbol{g}(0), \\ \boldsymbol{f}(2) &= \boldsymbol{g}(2) + [1 - p/(q+2)]\boldsymbol{g}(1) + [1 - p/(q+2)][1 - p/(q+1)]\boldsymbol{g}(0), \cdots \end{aligned} \tag{C3}$$

and so on. If $\boldsymbol{G}(z)$ is analytic for $z \in D_r(0)$ then $\boldsymbol{F}(z) = \boldsymbol{f}(0) + \boldsymbol{f}(1)z + \boldsymbol{f}(2)z^2 + \cdots$ is also analytic for $z \in D_r(0)$. In general we have

$$\begin{aligned} \boldsymbol{f}(t) &= \boldsymbol{g}(t) + [1 - p/(q+t)]\boldsymbol{g}(t-1) + \cdots + [1 - p/(q+t)]\cdots[1 - p/(q+1)]\boldsymbol{g}(0) \\ &= \boldsymbol{g}(t) + \sum_{s=1}^{t}[1 - p/(q+t)]\cdots[1 - p/(q+t-s+2)][1 - p/(q+t-s+1)]\boldsymbol{g}(t-s), \end{aligned}$$

for all $t \in \mathbb{N} - 1$. Thus we now have a general formula for the desired inverse operator. To justify (C1) we argue as follows. If $z \in D_1(0)$ and we write $(1-z)^{-1} = 1 + z + z^2 + \cdots$ then for each $k \in \mathbb{N} - 1$ direct calculation gives $(1-z)^{-1}[z^k] = \sum_{s \in \mathbb{N}-1} z^{s+k}$ from which it follows that

$$[(1-z)^{-1}W](1-z)^{-1}[z^k] = p \cdot \sum_{s \in \mathbb{N}-1}\left[\sum_{1 \le i_1 \le s+1}\prod_{j=1}^{1}(q+k+i_j)^{-1}\right]z^{s+1+k},$$

$$[(1-z)^{-1}W]^2(1-z)^{-1}[z^k] = p^2 \cdot \sum_{s \in \mathbb{N}-1}\left[\sum_{1 \le i_1 < i_2 \le s+2}\prod_{j=1}^{2}(q+k+i_j)^{-1}\right]z^{s+2+k},$$

and so on. An elementary induction can now be used to show that

$$[(1-z)^{-1}W]^n(1-z)^{-1}[z^k] = p^n \cdot \sum_{s \in \mathbb{N}-1}\left[\sum_{1 \le i_1 < \cdots < i_n \le s+n}\prod_{j=1}^{n}(q+k+i_j)^{-1}\right]z^{s+n+k},$$

for all $n \in \mathbb{N}$ provided $z \in D_1(0)$. If we calculate $[(1-z)I + W]^{-1}[z^k]$ but extract only the coefficients of z^t for some fixed value of $t \in \mathbb{N} - 1$ we need only consider values of $k \leq t$. Thus we have

$$\left((1-z)^{-1}[z^k]\right)_t = 1,$$

$$\left((-1)[(1-z)^{-1}W](1-z)^{-1}[z^k]\right)_t = (-1)p \cdot \sum_{1 \leq i_1 \leq t-k} \prod_{j=1}^{1} (q+k+i_j)^{-1},$$

$$\left((-1)^2[(1-z)^{-1}W]^2(1-z)^{-1}[z^k]\right)_t = (-1)^2 p^2 \cdot \sum_{1 \leq i_1 < i_2 \leq t-k} \prod_{j=1}^{2} (q+k+i_j)^{-1},$$

$$\left((-1)^3[(1-z)^{-1}W]^3(1-z)^{-1}[z^k]\right)_t = (-1)^3 p^3 \cdot \sum_{1 \leq i_1 < i_2 < i_3 \leq t-k} \prod_{j=1}^{3} (q+k+i_j)^{-1},$$

and so on. In general

$$\left((-1)^{t-k}[(1-z)^{-1}W]^{t-k}[z^k]\right)_t = (-1)^{t-k} p^{t-k} \cdot \sum_{1 \leq i_1 < \cdots < i_{t-k} \leq t-k} \prod_{j=1}^{t-k} (q+k+i_j)^{-1}.$$

Finally $\left((-1)^n[(1-z)^{-1}W]^n[z^k]\right)_t = 0$ for $n > t-k$. Collecting all of the terms shows that

$$\begin{aligned}\left(S[z^k]\right)_t &= [1 - p/(q+k+1)][1 - p/(q+k+2)] \cdots [1 - p/(q+t)] \\ &= \prod_{s=1}^{t-k} [1 - p/(q+k+s)],\end{aligned}$$

for each $t, k \in \mathbb{N}$ with $k \leq t$. It follows that

$$\begin{aligned}&S[\boldsymbol{g}(0) + \boldsymbol{g}(1)z + \boldsymbol{g}(2)z^2 + \cdots + \boldsymbol{g}(t)z^t] \\ &\to \left\{ \boldsymbol{g}(t) + \sum_{s=1}^{t} [1 - p/(q+t)] \cdots [1 - p/(q+t-s+2)][1 - p/(q+t-s+1)]\boldsymbol{g}(t-s) \right\} \\ &= \boldsymbol{f}(t),\end{aligned}$$

for all $t \in \mathbb{N} - 1$. Therefore $S\boldsymbol{G}(z) = \boldsymbol{F}(z)$. It follows that $S \equiv [(1-z)I + W]^{-1}$ for $z \in D_1(0)$. If $\boldsymbol{F}(z)$ is analytic for $z \in D_{1+\epsilon}(0)$ where $\epsilon > 0$ the definition of $W\boldsymbol{F}(z)$ as an analytic function remains valid on this region. The difficulty with the definition of $S = [(1-z)I + W]^{-1}$ arises because if we define $\boldsymbol{H}(z) = \boldsymbol{F}(z)z^q$ and write

$$H(z) = \boldsymbol{H}(1) + \boldsymbol{H}'(1)(z-1) + \cdots$$

for $z \in D_{1+\epsilon}(0)$ then

$$W(1-z)^{-1}\boldsymbol{F}(z) = (p/z^q) \int_0^z \left[\boldsymbol{H}(1)/(1-\zeta) + \cdots\right] d\zeta = (-1)(p/z^q)\boldsymbol{H}(1) \log_e(1-z) + \cdots$$

contains the term $\log_e(1-z)$ which can only be defined as a single-valued analytic function if a suitable cut terminating at $z = 1$ is inserted into the complex plane. For instance the cut $\{z \in \mathbb{C} | \Re(z) \in [1, \infty), \Im(z) = 0\}$ is suitable. Our earlier intuitive analysis suggested that $[(1-z)I + W]^{-1}$ has an essential singularity at $z = 1$. We can now see that it is not an isolated singularity.

Journal of Time Series Analysis published by John Wiley & Sons Ltd.
DOI: 10.1111/jtsa.12766