Granger Centre
for time series econometrics

Granger Centre Discussion Paper Series

Resuscitating the co-fractional model of Granger (1986)

by

Federico Carlini and Paolo Santucci de Magistris

Granger Centre Discussion Paper No. 19/01

University of Nottingham
UK | CHINA | MALAYSIA

Resuscitating the co-fractional model of Granger (1986)*

Federico Carlini† Paolo Santucci de Magistris ‡

January 27, 2019

Abstract

We study the theoretical properties of the model for fractional cointegration proposed by Granger (1986), namely the FVECM$_{d,b}$. First, we show that the stability of any discrete-time stochastic system of the type $\Pi(L)Y_t = \varepsilon_t$ can be assessed by means of the argument principle under mild regularity condition on $\Pi(L)$, where L is the lag operator. Second, we prove that, under stability, the FVECM$_{d,b}$ allows for a representation of the solution that demonstrates the fractional and co-fractional properties and we find a closed-form expression for the impulse response functions. Third, we prove that the model is identified for any combination of number of lags and cointegration rank, while still being able to generate polynomial co-fractionality. Finally, we show that the asymptotic properties of the maximum likelihood estimator reconcile with those of the FCVAR$_{d,b}$ model studied in Johansen and Nielsen (2012).

Keywords: Fractional cointegration, Granger representation theorem, Stability, Identification, Impulse Response Functions, Profile Maximum Likelihood

JEL Classification: C01, C02, C58, G12, G13 .

*We would like to thank Søren Johansen for his support and his precious suggestions that have improved the quality of this article. We are also grateful to Andrea Barletta, Massimo Franchi, Tobias Hartl, Morten Ørregaard Nielsen, Daniela Osterrieder and Federico Severino for their relevant remarks on our work. We would also like to thank the participants to the Long Memory Conference (Aalborg, 2018) and the seminar participants at the University of Pavia, at the University of Padua and at CREST (Paris) for useful comments. Federico Carlini gratefully acknowledges the support of the Swiss National Science Foundation for grant 105218-162633. Paolo Santucci de Magistris gratefully acknowledges the research support of CREATES, funded by the Danish National Research Foundation (DNRF78).

†Faculty of Economics, Università della Svizzera Italiana, Lugano, Switzerland. federico.carlini@usi.ch

‡Department of Economics and Finance, LUISS University. Viale Romania 32, 00197, Roma, Italy. CREATES, Aarhus University, Fuglesangs Alle 4, 8210, Aarhus, Denmark. psantucci@econ.au.dk

1 Introduction

The concept of *equilibrium* is central in many economic and financial models. In macroeconomics, equilibrium relations often originate from an economic theory linking agents' expectations to the actual outcome variables, as those behind the term structure of the interest rates. In finance, long-run equilibrium relations are often the result of no-arbitrage constraints, where deviations from the equilibrium can be interpreted as evidence against the ability of the financial markets to fully process new information and incorporate it in the asset prices. Depending on the persistence of the deviations from the no-arbitrage relation, i.e. the strength of the reversion of the system to long-run equilibrium, we might conclude on the extent of the violation of the market efficient hypothesis. For almost thirty years, the analysis of cointegrated systems has been the paradigm in the empirical investigation of equilibrium relations between economic variables. The notion of cointegration, as originally defined in Engle and Granger (1987), entails a long-run relation between variables characterized by highly persistent common stochastic trends, $I(1)$, with short-memory, $I(0)$, deviations from the equilibrium.

Unfortunately, the classification of $I(1)$ and $I(0)$ variables is very restrictive and does not accommodate the dynamic features of many economic time series. For example, the very persistent dynamics of inflation can not be described by means of integrated processes, but, consistently with the price theory of Rotemberg (1987), inflation is best described by a process with a fractional order of integration which arises from the cross-sectional aggregation of simple, possibly dependent, dynamic micro processes, see Granger (1980) and Zaffaroni (2004), and the recent contribution of Schennach (2018). In particular, fractionally integrated processes are characterized by long range dependence or *long-memory*; that is a strong relationship between observations that are distant in time, since the effects of a shock last for many periods and decay slowly and hyperbolically, see Granger (1980) and Hosking (1981). For this reason, the class of fractionally integrated processes have changed the way in which researchers describe and forecast macroeconomic and financial series, providing an elegant and parsimonious way of describing the dynamic features of economic time series with any order of integration. Evidence of long memory is found in macroeconomic aggregates, such as the consumer prices and inflation (see Geweke and Porter-Hudak, 1983), interest rates (see Shea, 1991), and in financial series as exchange rates (see Baillie and Bollerslev, 1994) and the volatility of stock prices, see, among others, Baillie et al. (1996) and Andersen and Bollerslev (1997).

In this paper, we study the properties of the multivariate model of Granger (1986) to analyze the long-run equilibrium relations between series that are integrated of a fractional order. We show that the the model of Granger (1986) is coherent with the concept of *fractional* cointegration or co-fractionality. In particular, fractional cointegration implies that linear combinations of $I(d)$ processes are $I(d-b)$, with $d, b \in \mathbb{R}_+$ and $0 < b \leq d$, see Robinson and Marinucci (2003) among others for a formal definition. In other words, the concept of fractional cointegration involves the existence of common stochastic trends integrated of order d, with short-period departures from the long-run equilibrium integrated of order $d - b$. Thus the range of applicability of the concept of cointegration is enormously extended compared to that originally defined by Engle and Granger (1987).

In his original contribution, Granger (1986, Equation 4.3) already introduces a model for co-fractionality, the fractional VECM (FVECM$_{d,b}$ henceforth). The FVECM$_{d,b}$ extends the well-known VECM to the fractional case, which is obtained by setting the parameters d and b to 1. For many years, most of the econometric analysis has been focusing to cases with d and b restricted to integers. More recently Johansen (2008b) has noted that the characteristic function of the co-fractional model of Granger (1986) involves a complicated transcendental equation, so that *it is inconvenient to analyze in the sense that the stochastic properties of the solution generated by the equations are not easily reflected in properties of the coefficients*. Hence Johansen (2008b) proposes a slightly modified version of the FVECM$_{d,b}$, namely the FCVAR$_{d,b}$, and studies the properties of the new model in terms of conditions for the stability and Granger representation theorem. The FCVAR$_{d,b}$ provides a fully parametric characterization of the long-run relations between fractional series and it encompasses the VECM analyzed in Johansen (1988), which is obtained when the parameters d and b are restricted to be equal to one. Johansen (2008b) studies the properties of the FCVAR$_{d,b}$ in terms of Granger representation, while Johansen and Nielsen (2012) derive the asymptotic properties of the profile maximum likelihood (ML) estimator of the FCVAR$_{d,b}$, see also Lasak (2010). Although alternative models for fractional cointegration can be found in Avarucci (2007) and Tschernig et al. (2013), the FCVAR$_{d,b}$ of Johansen (2008b) is probably the most commonly adopted specification in this context. Empirical applications of the FCVAR$_{d,b}$ can be found in Rossi and Santucci de Magistris (2013), Caporin et al. (2013), Bollerslev et al. (2013), Dolatabadi et al. (2015), Dolatabadi et al. (2016) and Nielsen and Shibaev (2018). Unfortunately, as noted by Johansen and Nielsen (2012) and subsequently by Carlini and Santucci de Magistris (2017), the FCVAR$_{d,b}$ is not identified when the number of lags is over-specified and the cointegration rank is also unknown. In other words, the FCVAR$_{d,b}$ can generate special cases of *polynomial* fractional cointegration analogous to those studied in Franchi (2010), when the number of lags is not correctly determined. This problem might have led to a limited use of the FCVAR$_{d,b}$ in the empirical applications. Indeed, it is often needed to impose restrictions on the coefficient d or to adopt rather computationally-intensive algorithms (such as grid-search) to study the shape of the log-likelihood function in different regions of the parameter space, see the discussion in Nielsen and Popiel (2014).

In this paper, we begin by discussing the stability properties of the FVECM$_{d,b}$ in light of the argument principle, which is a well known result in complex analysis but, to the best of our knowledge, has never been applied in the context of time-series econometrics. The application of the argument principle to determine the stability of a dynamic system is a general result that can be applied in a wide range of circumstances beyond the context of fractional cointegration. Examples of possible applications of the argument principle are in the field of rational expectation models when assessing the existence of the steady-state in reduced-form systems, see

Binder and Pesaran (1997) and Klein (2000) among others, and when dealing with non-causal processes like those introduced in Gouriéroux and Zakoïan (2017) for explosive bubbles. Under the stability condition, we derive a number of theoretical results for the FVECM$_{d,b}$ of Granger (1986). First, we show that the model of Granger (1986) admits a Granger representation in the fractional context. This makes the model suitable for analyzing equilibrium relations between fractionally integrated series. Furthermore, the impulse response functions of the FVECM$_{d,b}$ are obtained in closed-form in terms of a recursive formula built upon the type-II fractional difference operator. Second, we prove that the model is identified for any choice of the number of lags and cointegration rank. This result is expected to simplify the empirical analysis of fractionally cointegrated systems compared with the FCVAR$_{d,b}$. Third, we show that the FVECM$_{d,b}$ also allows for a Granger representation under polynomial cofractionality, which is a generalization of the I(2)-type cointegration to the fractional context. Finally, we complete the theoretical analysis by studying the asymptotic behavior of the ML estimator of the coefficients of the FVECM$_{d,b}$. We show that the conditions for applying the asymptotic results of Johansen and Nielsen (2012) hold also in the FVECM$_{d,b}$ context, such that consistency and asymptotic distribution of the ML estimator follow.

The paper is organized as follows. Section 2 presents the FVECM$_{d,b}$. Section 3 discusses the conditions for the stability of the system. Section 4 contains the theorem on the Granger representation of the FVECM$_{d,b}$ and the derivation of the impulse response functions of the FVECM$_{d,b}$. In Section 5 we prove that the FVECM$_{d,b}$ is identified for any combination of lag-length and cointegration rank. In Section 6 we show that the FVECM$_{d,b}$ allows for polynomial fractional cointegration, i.e. we provide a Granger representation theorem for $I(2)$-type fractional processes. Section 7 contains results on the consistency and asymptotic distribution of the maximum-likelihood estimator of the parameters of the FVECM$_{d,b}$. Finally, Section 8 concludes. Appendix A contains a discussion of the regularity of the characteristic polynomial, while the proofs of the theorems are in Appendix B.

2 The fractional VECM of Granger (1986)

In this section, we outline and study the properties of the FVECM$_{d,b}$ of Granger (1986), which is defined as

$$\mathcal{H}_{r,k}: \quad \Delta^d X_t = \alpha\beta' \Delta^{d-b} L_b X_t + \sum_{j=1}^{k} \Gamma_j \Delta^d X_{t-j} + \varepsilon_t, \tag{1}$$

and it is an extension of the well known VECM to the case of fractional cointegration, see also Davidson (2002). The fractional operator Δ^d in (1) is defined as

$$\Delta^d := (1-L)^d = \sum_{j=0}^{\infty} (-1)^j \binom{d}{j} L^j,$$

where L is the lag operator, such that $LX_t = X_{t-1}$ and $d \in \mathbb{R}$. The operator $\Delta^{d-b} := (1-L)^{d-b}$ is defined in an analogous way. The term $L_b := 1 - \Delta^b$ denotes the so called *fractional lag operator*. The term X_t is a p-dimensional vector, α and β are $p \times r$ matrices, where r defines the cointegration rank, ε_t is p-dimensional independent and identically distributed with mean zero and covariance matrix $\Omega > 0$, and Γ_j, $j = 1, \ldots, k$, are $p \times p$ matrices loading the short-run dynamics. The coefficient d determines the degree of fractional integration of the series X_t, while the coefficient b determines the so called *cointegration gap*, i.e. the degree of fractional integration of $\beta' X_t$ that is $d - b$. Model (1) reduces to the classic VECM when $d = b = 1$.[1] The model $\mathcal{H}_{r,k}$ in (1) has k lags and $\theta = \{d, b, \alpha, \beta, \Gamma_1, ..., \Gamma_k, \Omega\}$ is the collection of parameters. The parameter space of the model is

$$\Theta = \{\alpha \in \mathbb{R}^{p \times r}, \beta \in \mathbb{R}^{p \times r}, \Gamma_j \in \mathbb{R}^{p \times p}, j = 1, \ldots, k, d \in \mathbb{R}^+, b \in \mathbb{R}^+, d \geq b > 0, \Omega > 0 \in \mathbb{R}^{p \times p}\},$$

where r is the cointegration rank, such that $p - r$ determines the number of common stochastic trends between the series. When $r = p$, the model is

$$\mathcal{H}_{p,k}: \quad \Delta^d X_t = \Xi \Delta^{d-b} L_b X_t + \sum_{j=1}^{k} \Gamma_j \Delta^d X_{t-j} + \varepsilon_t, \tag{2}$$

where Ξ is a $p \times p$ matrix with full rank. By adopting the standard tools for the analysis of the solutions of the FVECM$_{d,b}$ in (1), Johansen (2008b) notes that it is not possible to study the stability of the system and to obtain a Granger representation for X_t. Hence, Johansen (2008b) proposes an alternative version of the FVECM$_{d,b}$, the FCVAR$_{d,b}$. The FCVAR$_{d,b}$ is defined as

$$\Delta^d X_t = \alpha \beta' \Delta^{d-b} L_b X_t + \sum_{j=1}^{k} \Gamma_j \Delta^d L_b^j X_t + \varepsilon_t, \tag{3}$$

and it replaces the usual lag operator in the autoregressive polynomial with the fractional lag operator. In other words, the FVECM$_{d,b}$ in (1) and the FCVAR$_{d,b}$ in (3) share the same cointegration component, $\alpha\beta'\Delta^{d-b} L_b X_t$, which, as noted by Johansen (2008b, p.652), arises from the formulation in terms of common trends and cofractional terms of Breitung and Hassler (2002)

[1] As also noted in Johansen (2008b), model (1) is a slightly different version of the original Granger's model in (1). Indeed, the original model reported in Granger (1986, Equation 4.3) is

$$\Delta^d X_t = \alpha \beta' \Delta^{d-b} L_b X_{t-1} + \sum_{j=1}^{k} \Gamma_j \Delta^d X_{t-j} + \varepsilon_t.$$

Imposing the restriction $d = b = 1$ leads to

$$\Delta X_t = \alpha \beta' X_{t-2} + \sum_{j=1}^{k} \Gamma_j \Delta X_{t-j} + \varepsilon_t,$$

which is not the classic VECM since the error correction term $\beta' X_t$ enters on the right-hand side of (1) lagged by two periods.

with $\beta' X_t = \Delta^{-d+b} u_{1t}$ and $\gamma' X_t = \Delta^{-d} u_{2t}$, where $u_t = (u'_{1t}, u'_{2t})' \sim iidN(0, \Sigma)$, and $(\beta', \gamma')'$ is a full rank matrix, with β being a $p \times r$ matrix and γ a $p \times (p - r)$ matrix.

The inclusion of the fractional lag operator in the short term dynamics enables Johansen (2008b) to assess the stability of the FCVAR$_{d,b}$ and to prove that the solution of the characteristic polynomial of the FCVAR$_{d,b}$ exists so that the FCVAR$_{d,b}$ admits a Granger representation. Based on this result, Johansen and Nielsen (2012) derive the asymptotic theory for the ML estimator of the parameters of the FCVAR$_{d,b}$. Recently, Carlini and Santucci de Magistris (2017) highlight the potential identification issues that emerge when the true lag structure and co-integration rank of the FCVAR$_{d,b}$ are unknown. The identification problems mostly arise as a consequence of the presence of the fractional lag operator in the autoregressive part of (3). In the following, we show that the stability conditions of the FVECM$_{d,b}$ can be studied through the argument principle and the Granger representation theorem can be obtained by the inversion of the characteristic function.

3 Stability

We first provide a number of definitions that are useful for the characterization of the properties of the FVECM$_{d,b}$.

Definition 3.1. Following Johansen (2008b), we define $\mathcal{F}(0)$ processes, $\mathcal{F}(d)$ processes and fractional cointegration as follows:

(i) If Ψ_j is a sequence of $p \times p$ matrices for which $\sum_{j=0}^{\infty} ||\Psi_j||^2 < \infty$ with $\Psi(z) = \sum_{j=0}^{\infty} \Psi_j z^j$. We call the stationary linear process $X_t = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}$ fractional of order zero, denoted as $X_t \sim \mathcal{F}(0)$, if the spectrum at zero $f_X(0) = \frac{1}{2\pi}\Psi(1)\Omega\Psi(1)' \neq 0$.

(ii) We denote $\mathcal{F}(0)_+$ the class of processes of the form, $X_t^+ = \Psi(L)_+ \varepsilon_t = \sum_{j=0}^{t-1} \Psi_j \varepsilon_{t-j}$.

(iii) We say that X_t is fractional of order d and write $X_t \sim \mathcal{F}(d)$, if conditionally on the past $\{X_s, s \leq 0\}$, $\Delta_+^d X_t - \mu_t \sim \mathcal{F}(0)_+$ for some function μ_t of the past where

$$\Delta_+^d X_t := (1 - L)_+^d X_t = \sum_{j=0}^{t-1} (-1)^j \binom{d}{j} L^j X_t \tag{4}$$

(iv) If $X_t \sim \mathcal{F}(d)$ and there exists a vector β so that $\beta' X_t \sim \mathcal{F}(d - b)$ for some b, $0 < b \leq d$, we call X_t co-fractional with co-fraction vector β.

For a given $r < p$ and k, the characteristic function of the FVECM$_{d,b}$ in (1) is

$$\Pi(z) = (1 - z)^d I_p - \alpha\beta'(1 - z)^{d-b}(1 - (1 - z)^b) - \sum_{j=1}^{k} \Gamma_j (1 - z)^d z^j, \tag{5}$$

or by setting $\tilde{\Pi}(z) := (1-z)^{b-d}\Pi(z)$, we have

$$\tilde{\Pi}(z) = (1-z)^b I_p - \alpha\beta'(1-(1-z)^b) - \sum_{j=1}^{k} \Gamma_j (1-z)^b z^j,$$

with I_p being the $p \times p$ identity matrix.

A crucial assumption for the stability of the FVECM$_{d,b}$ is that there are only $p-r$ roots of $|\tilde{\Pi}(z)| = 0$ in $z = 1$, while the others are outside the unit circle. While in the FCVAR$_{d,b}$ of Johansen (2008b), the trick of substituting $y = 1-(1-z)^b$ in $\tilde{\Pi}(z)$ allows to obtain a polynomial in the fractional lag operator for which the conditions of stability can be easily shown (up to a remapping to the fractional unit circle), the same can not be done for the FVECM$_{d,b}$. However, the analysis of the stability of the FVECM$_{d,b}$ can be carried out by adopting the general result in complex analysis known as the *argument principle*, see Fuchs and Shabat (1964, p.322). Let us first define the function $g(z) = |\tilde{\Pi}(z)| = 0$. Given the cointegration rank r, $g(z)$ can be further factorized as $g(z) = (1-z)^{b(p-r)} f(z)$, so that we can count the number of zeroes of $f(z)$ inside the unit circle. Provided that $f(z)$ is a holomorphic function in the unit circle, the number of zeroes is obtained through the following Cauchy integral

$$\frac{1}{2\pi i} \oint_{\mathcal{S}} \frac{f'(z)}{f(z)} dz = \mathcal{N} - \mathcal{P}, \tag{6}$$

where $\frac{f'(z)}{f(z)}$ is the logarithmic derivative of $f(z)$ in $\mathbb{C}$, and $\mathcal{N}$ and $\mathcal{P}$ are respectively the number of zeros and poles in the region $\mathcal{S} = \{z \in \mathbb{C} \text{ s.t. } |z| \leq 1\}$. In Appendix A we also show that $f(z)$ does not have poles inside the unit circle ($\mathcal{P} = 0$) nor zeros and poles on the boundary of $\mathcal{S}$, so that, by setting $z = e^{i\theta}$, the Cauchy integral becomes

$$\frac{1}{2\pi i} \int_0^{2\pi} \frac{f'(e^{i\theta})}{f(e^{i\theta})} i e^{i\theta} d\theta = \mathcal{N}. \tag{7}$$

The integral on the right-hand side admits an analytical solution, which can be approximated numerically with very high accuracy, see Delves and Lyness (1967).[2] The following lemma shows that the stability condition of the FVECM can be equivalently expressed in terms of the principle of the argument.

Lemma 3.2. *Let $f(z)$ be an holomorphic function. Then, $\mathcal{N} = 0$ if and only if $|\tilde{\Pi}(z)| = 0$ implies that either $z = 1$ or z are outside the unit circle. Hence, the FVECM$_{d,b}$ is stable.*

The lemma is a direct consequence of the Cauchy's argument principle see Ahlfors (1953). Appendix A provides a discussion on the regularity properties of $f(z) = (1-z)^{-b(p-r)} g(z)$, that is $f(z)$ is an holomorphic function in the unit circle. It should be noted that the range of applicability of the Cauchy's argument principle to assess the stability of a stochastic process extends

[2] The MATLAB code `argument_principle.m` uses the quadrature method to evaluate the integral, which is a more accurate alternative than the trapezoidal method studied in Delves and Lyness (1967).

beyond the current application to the FVECM$_{d,b}$ and it can be employed when the standard analysis of the characteristic function is complicated/unfeasible provided that $f(z)$ is a holomorphic function in the unit circle. In the following section, we show that the FVECM$_{d,b}$ admits a Granger representation given that the stability condition of the FVECM$_{d,b}$ of Granger (1986) is satisfied.

4 Granger Representation Theorem

In the following, we show that the FVECM$_{d,b}$ in (1) is coherent with the notion of fractional cointegration, as in Definition 3.1-(iv). In other words, the FVECM$_{d,b}$ admits a representation of the solution that demonstrates the fractional and co-fractional properties. In particular, Theorem 4.1 shows that the FVECM$_{d,b}$ allows for a Granger representation in the fractional context. We also introduce the variable $y = 1 - (1-z)^b$ and we define $\tilde{\Pi}(z) = \tilde{\Pi}(z, y)$ as

$$\tilde{\Pi}(z, y) = (1-y)I_p - \alpha\beta' y - \sum_{j=1}^{k} \Gamma_j (1-y) z^j.$$

Adding and subtracting $\alpha\beta' z$ from $\tilde{\Pi}(z, y)$ we obtain

$$\tilde{\Pi}(z, y) = (1-y)\left(I_p + \alpha\beta' - \sum_{j=1}^{k} \Gamma_j z^j\right) - \alpha\beta'.$$

Theorem 4.1. *If $\mathcal{N} = 0$ and α and β have rank $r < p$, and if $|\alpha'_\perp \Gamma \beta_\perp| \neq 0$ with $\Gamma = I_p - \sum_{i=1}^{k} \Gamma_i$, then*

$$X_t = C(L)\Delta_+^{-d}\varepsilon_t + \Delta_+^{-(d-b)} Y_t + \mu_t, \tag{8}$$

where $C(L) = \beta_\perp(\alpha'_\perp \Gamma(L)\beta_\perp)^{-1}\alpha'_\perp$ with $\Gamma(L) = I_p - \sum_{i=1}^{k} \Gamma_i L^i$ and $C(1) = \beta_\perp(\alpha'_\perp \Gamma(1)\beta_\perp)^{-1}\alpha'_\perp$. The term $Y_t \sim \mathcal{F}(0)$ with continuous spectrum that at zero frequency is given by $\frac{C^*\Omega C^{*'}}{2\pi} \neq 0$ and $\mu_t = -\Pi_+(L)^{-1}\Pi_-(L)X_t$ depends on the initial values. Thus, X_t is fractional of order d, whereas $\Delta^b X_t$ and $\beta' X_t$ are fractional of order $d-b$.

Proof in Appendix B.1.

Although sharing similarities with the Granger representation of the FCVAR$_{d,b}$ in Johansen (2008b), the Granger representation of the FVECM$_{d,b}$ displays one interesting difference with its predecessor. Indeed, the loading term of the common stochastic trend is not a reduced rank matrix as in Johansen (2008b), but it is reduced rank lag-polynomial matrix, $C(L)$. In particular, the leading term in (8) can be written as

$$\begin{aligned} C(L)\Delta_+^{-d}\varepsilon_t &= \beta_\perp(\alpha'_\perp(I_p - \sum_{i=1}^{k} \Gamma_i L^i)\beta_\perp)^{-1}\alpha'_\perp \Delta_+^{-d}\varepsilon_t \\ &= \sum_{j=0}^{\infty} \Delta^j \beta_\perp \Phi_j \alpha'_\perp \Delta_+^{-d}\varepsilon_t = \sum_{j=0}^{\infty} \beta_\perp \Phi_j \alpha'_\perp \Delta_+^{j-d}\varepsilon_t, \end{aligned}$$

where $\sum_{j=0}^{\infty} \Phi_j L^j = (\alpha_\perp' \Gamma(L)\beta_\perp)^{-1}$, so that

$$X_t = C(1)\Delta_+^{-d}\varepsilon_t + \sum_{j=1}^{\infty} \beta_\perp \Phi_j \alpha_\perp' \Delta_+^{j-d}\varepsilon_t + \Delta_+^{-(d-b)} Y_t + \mu_t. \tag{9}$$

Equation (9) shows that the process is composed as the sum of two *usual* terms $C(1)\Delta_+^{-d}\varepsilon_t$ and $\Delta_+^{-(d-b)} Y_t$, but the extra term $\sum_{j=1}^{\infty} \beta_\perp \Phi_j \alpha_\perp' \Delta_+^{j-d}\varepsilon_t$ is (in general) fractional of order $d-1$, but perhaps greater than the order of Y_t. In any case, we still have that

$$\beta' X_t = \beta' \sum_{j=0}^{\infty} \beta_\perp \Phi_j \alpha_\perp' \Delta_+^{j-d}\varepsilon_t + \beta' \Delta_+^{-(d-b)} Y_t + \beta' \mu_t = \beta' \Delta_+^{-(d-b)} Y_t + \beta' \mu_t,$$

that is $\beta' X_t$ is fractional of order $d-b$. This means that the FVECM reconciles with the standard notion of fractional cointegration. Furthermore, under the condition $|\alpha_\perp' \Gamma(1)\beta_\perp| \neq 0$, we cannot have polynomial fractional cointegration because $\mathrm{sp}(C(L)) = \mathrm{sp}(\beta_\perp)$, where the $\mathrm{sp}(A)$ denotes the column space of A. Section 6 discusses the case of polynomial fractional cointegration when $\alpha_\perp' \Gamma(1)\beta_\perp$ has reduced rank.

4.1 Impulse response function

The impulse response functions represent a useful tool to assess the dynamic impact of a shock of a variable on anther variable in a system. The following lemma contains the recursive formula to calculate the coefficients of the impulse response functions for the FVECM$_{d,b}$ obtained by the vector MA representation of the FVECM$_{d,b}$ arising from Theorem 4.1.

Lemma 4.2. *Consider the FVECM$_{d,b}$ with k lags defined in (1). The impulse responses $\Theta_j, j \geq 0$ are given by the following set of recursions:*

$$\begin{aligned}
&\Theta_0 = I_p, \qquad \Theta_1 = -\rho_1(d) + \alpha\beta'(\rho_1(d-b) - \rho_1(d)) + \Gamma_1, \\
&\Theta_\ell = \Theta_1 \Theta_{\ell-1} + \sum_{i=0}^{\ell-1} \Psi_i \Theta_{\ell-i-1}, \qquad \ell = 2, 3, \ldots \\
&\Psi_j = \alpha\beta'(\rho_{j+1}(d-b) - \rho_{j+1}(d)) + \sum_{i=1}^{j} \Gamma_i \rho_{j-i}(d) - I_p \rho_{j+1}(d), \qquad j = 1, \ldots, k-1 \\
&\Psi_s = \alpha\beta'(\rho_{s+1}(d-b) - \rho_{s+1}(d)) + \sum_{i=1}^{k} \Gamma_i \rho_{s-i}(d) - I_p \rho_{s+1}(d), \qquad j = k, \ldots
\end{aligned}$$

where $\rho_i(a) = (-1)^i \binom{a}{i}$, $a \in \mathbb{R}^+$.

Section B.2 in Appendix B reports the derivation of the recursive formulas for the calculation of the impulse response coefficients. Figure 1 displays an example of IRF for the FVECM$_{d,b}$ when $p = 2$, $r = 1$ and $k = 1$. The left panel displays the IRFs of a stable system, which slowly decay to zero due to the persistent nature of the variables which are fractional of order $d = 0.6$. The right panel reports the IRFs of an unstable system, which is correctly detected by computing the Cauchy integral in (6). Under an unstable setup, the IRFs explode as the horizon h increases.

5 Identification

We now study the identification property of the FVECM$_{d,b}$ for any choice of the lag, k, and cointegration rank, r. As shown in Carlini and Santucci de Magistris (2017), there exist several equivalent parametrization of the FCVAR$_{d,b}$ for different values of k and r. First, we introduce the concept of identification and equivalence between two models as in Johansen (2010).

Definition 5.1. Let $\{P_\theta, \theta \in \Theta\}$ be a family of probability measures, that is, a statistical model. We say that a parameter function $g(\theta)$ is identified if $g(\theta_1) \neq g(\theta_2)$ implies that $P_{\theta_1} \neq P_{\theta_2}$. On the other hand, if $P_{\theta_1} = P_{\theta_2}$ and $g(\theta_1) \neq g(\theta_2)$, the parameter function $g(\theta)$ is not identified. In this case, the statistical models P_{θ_1} and P_{θ_2} are equivalent.

As noted by Johansen (1995, p.177), the product $\alpha\beta'$ is identified but not the matrices α and β because if there was an invertible $r \times r$ matrix ξ, the product $\alpha\beta'$ would be equal to $\alpha_\xi\beta'_\xi$, where $\alpha_\xi = \alpha\xi$ and $\beta_\xi = \beta\xi^{-1}$. In the following, we do not discuss the identification of α and β, that is generally solved by a proper normalization of β. The following theorem states that the parameters of the FVECM$_{d,b}$ in (1) are uniquely identified.

Theorem 5.2. *For any k and r, the parameters of the FVECM$_{d,b}$ in (1) are identified, up to rotations of the vectors α and β.*

Proof in Appendix B.3.

It follows from Theorem 5.2 that the FVECM$_{d,b}$ is identified for any choice of k and r. This means that for each combination of k and r we obtain a model that is distinct from the others. Hence the following corollary highlights the nesting structure of the FVECM$_{d,b}$, that is a direct consequence of the identification property.

Corollary 5.3. *The nesting structure of the FVECM$_{d,b}$ is represented by the following scheme:*

$$\begin{array}{ccccccccc}
\mathcal{H}_{0,0} & \subset & \mathcal{H}_{0,1} & \subset & \mathcal{H}_{0,2} & \subset & \cdots & \subset & \mathcal{H}_{0,k} \\
\cap & & \cap & & \cap & & & & \cap \\
\mathcal{H}_{1,0} & \subset & \mathcal{H}_{1,1} & \subset & \mathcal{H}_{1,2} & \subset & \cdots & \subset & \mathcal{H}_{1,k} \\
\cap & & \cap & & \cap & & & & \cap \\
\vdots & & \vdots & & \vdots & & \ddots & & \vdots \\
\cap & & \cap & & \cap & & & & \cap \\
\mathcal{H}_{p,0} & \subset & \mathcal{H}_{p,1} & \subset & \mathcal{H}_{p,2} & \subset & \cdots & \subset & \mathcal{H}_{p,k}.
\end{array} \tag{10}$$

The nesting structure in (10) is a direct consequence of the identification property outlined in Theorem 5.2. In particular, row-wise we have that, for a given k, the model with full rank nests all models with reduced rank $r < p$. Column-wise, it is trivial to note that for a given r, the model with k lags nests models with $0, 1, \ldots, k-1$ lags. Finally, by Theorem 5.2, models $\mathcal{H}_{0,k}$ and $\mathcal{H}_{p,k-1}$ are distinct, and *a fortiori* $\mathcal{H}_{0,k}$ and $\mathcal{H}_{r,k-1}$ are also distinct when $r < p$. The regular nesting structure of this model facilitates the model selection in the empirical works with a general-to-specific sequence of LR tests similar to the one adopted in the standard VECM context and also discussed in Johansen and Nielsen (2012). On the contrary, the FCVAR$_{d,b}$ of Johansen (2008b) displays a non-regular nesting structure that makes the model selection more involved as a consequence of the lack of identification, see Carlini and Santucci de Magistris (2017).

6 Polynomial cofractionality

In the derivation of Theorem 4.1, we assumed that $|\alpha_\perp' \Gamma(1) \beta_\perp| \neq 0$. This assumption is known as $I(1)$ condition in the classic VECM framework. In the framework of fractionally cointegrated VAR systems, Carlini and Santucci de Magistris (2017) denoted it as the "$\mathcal{F}(d)$ condition" to signal that under $|\alpha_\perp' \Gamma(1) \beta_\perp| \neq 0$ and under correct model specification, there is an unique pair of parameters d and b such that $X_t \sim \mathcal{F}(d)$ and $\beta' X_t \sim \mathcal{F}(d-b)$. Unfortunately, when the number of lags in the FCVAR$_{d,b}$ is overspecified, Carlini and Santucci de Magistris (2017) show that violations of the $\mathcal{F}(d)$ condition might arise, inducing identification problems associated with special cases of polynomial cofractionality. For example, there might exist two parameters $d_1 = d - b/2$ and $b_1 = b/2$ such that $X_t \sim \mathcal{F}(d_1 + b_1)$ and $\beta' X_t \sim \mathcal{F}(d_1 - b_1)$ when $k > k_0$. Provided that Theorem 5.2 guarantees identification of d and b for a generic lag-length in the FVECM$_{d,b}$ framework, we can now focus on the cointegration properties of X_t when imposing the restriction

$$\alpha_\perp' \left(I_p - \sum_{j=1}^{k} \Gamma_j \right) \beta_\perp = \xi \eta', \tag{11}$$

with ξ and η being $(p-r) \times s$ matrices with $\alpha_\perp$ and $\beta_\perp$ such that $\alpha' \alpha_\perp = 0$ and $\beta' \beta_\perp = 0$, and that $0 \leq b \leq d$. This is the analogous of the $I(2)$ model derived in the VECM, which is obtained when $d = 2$ and $b = 1$, see Johansen (1992). The characteristic function of the FVECM$_{d,b}$ under (11) is

$$\Lambda(z) = (1-z)^d I_p - \alpha\beta'(1-z)^{d-b}(1-(1-z)^b) - \sum_{j=1}^{k} \Gamma_j (1-z)^d z^j, \tag{12}$$

where $\Lambda(z)$ is different from $\Pi(z)$ in (5) since the restriction (11) is imposed. We can define an equivalent characteristic function as

$$\tilde{\Lambda}(z) := (1-z)^{b-d} \Lambda(z) = (1-z)^b I_p - \alpha\beta'(1-(1-z)^b) - \sum_{j=1}^{k} \Gamma_j (1-z)^b z^j.$$

The analysis of the stability of the characteristic function can be carried out again the principle of the argument as discussed above. Let us first define the function $g^*(z) = |\tilde{\Lambda}(z)| = 0$. Given the cointegration ranks r and s, $g^*(z)$ can be further factorized as $g^*(z) = (1-z)^{bs+2b(p-r-s)} f(z)$, see Johansen (1997, p.437). Hence, we can apply the argument principle as in (7) and count the number of zeroes of $f(z)$ inside the unit circle. Given the stability of the FVECM$_{d,b}$ system under the restriction (11), the following theorem provides the Granger representation of the FVECM under polynomial cofractionality.

Theorem 6.1. *If $\mathcal{N} = 0$ and α and β have rank $r < p$ with $\alpha_\perp' \left(I_p - \sum_{j=1}^{k} \Gamma_j\right) \beta_\perp$ of rank $s < p - r$ and if $\alpha_2'\Gamma(1)\bar{\beta}\bar{\alpha}'\Gamma(1)\beta_2$ is invertible with $\bar{\alpha} = \alpha(\alpha'\alpha)^{-1}$, $\bar{\beta} = \beta(\beta'\beta)^{-1}$, $\alpha_2 = \alpha_\perp\xi_\perp$ and $\beta_2 = \beta_\perp\eta_\perp$, then*

$$X_t = C_2(L)\Delta_+^{-b-d}\varepsilon_t + C_1(L)\Delta_+^{-d}\varepsilon_t + \Delta_+^{-(d-b)}Y_t^+ + \mu_t, \tag{13}$$

where $\mu_t = -\Lambda_+(L)^{-1}\Lambda_-(L)X_t$ depends on the initial values. The polynomial matrices $C_2(L)$ and $C_1(L)$ are

$$\begin{aligned} C_2(L) &= \beta_2\theta_{22}(L)^{-1}\alpha_2' \\ C_1(L) &= -\bar{\beta}_1\bar{\alpha}_1' + \left(\bar{\beta}_1\theta_{12}(L) - \bar{\beta}\bar{\alpha}'\Gamma(L)\beta_2\right)\theta_{22}(L)^{-1}\alpha_2' + \\ &\quad +\beta_2\theta_{22}(L)^{-1}\left(\theta_{21}(L)\bar{\alpha}_1' - \alpha_2'\Gamma(L)\beta_2\bar{\alpha}\right) + \beta_2\Xi(L)\alpha_2', \end{aligned}$$

where $\bar{\alpha}_1 = \alpha_1(\alpha_1'\alpha_1)^{-1}$ with $\alpha_1 = \bar{\alpha}_\perp\xi$, $\bar{\beta}_1 = \beta_1(\beta_1'\beta_1)^{-1}$ with $\beta_1 = \bar{\beta}_\perp\eta$. The process Y_t is stationary with continuous spectrum, and X_t is fractional of order $d+b$, $(\beta', \beta_1)'X_t$ is fractional of order b, and $\beta'X_t - \bar{\alpha}'\Gamma(L)\Delta_+^b X_t$ is fractional of order 0.

Proof in Appendix B.4.

In analogy with Theorem 4.1, the loadings $C_2(L)$ and $C_1(L)$ of the fractional roots of order $d+b$ and d are matrix polynomials in the lag operator.

7 Inference

As shown in Johansen and Nielsen (2012), the parameters of the FCVAR$_{d,b}$ can be estimated following a profile likelihood approach. We follow here the same approach for the estimation of the parameters of the FVECM$_{d,b}$. For fixed $\psi = (d, b)'$, the ML estimator is found by reduced rank regression of $\Delta^d X_t$ on $\Delta^{d-b} L_b X_t$ corrected for $\{\Delta^d L^i X_t\}_{i=1}^k$, see Anderson et al. (1951) or Johansen (1995). For fixed $\psi = (d, b)'$ in model $\mathcal{H}_r$, we define the residuals, $R_{it}(\psi)$ for $i = 0, 1$, of the reduced rank regression of $\Delta_+^d X_t$ on $\Delta_+^d L^j X_t$ and $\Delta_+^{d-b} L X_t$ on $\Delta_+^d L^j X_t$ for $j = 1, .., k$, respectively. We also define the product moment matrices $S_{ij}(\psi)$ for $i, j = 0, 1$, that is $S_{ij}(\psi) = T^{-1}\sum_{t=1}^T R_{it}(\psi)R_{jt}'(\psi)$. Given the product moment matrices, we can express the generalized eigenvalue problem as

$$\det\left(\omega S_{11}(\psi) - S_{10}(\psi)S_{00}^{-1}(\psi)S_{01}(\psi)\right), \tag{14}$$

whose solutions, $\omega_i(\psi)$ for $i = 1, \ldots, p$, are sorted in decreasing order. Analogously with the reduced rank regression in the VECM framework of Johansen (1991), the (profile) log-likelihood function for given fixed ψ is

$$\ell_{T,r}(\psi) = -\log\det(S_{00}(\psi)) - \sum_{i=1}^{r} \log(1-\omega_i(\psi)). \tag{15}$$

Therefore, for a given value of the cointegration rank $r = 1, \ldots, p$, ML estimates of d and b, denoted as $\hat{d}$ and $\hat{b}$, can be calculated by maximizing the profile log-likelihood function, $\ell_{T,r}$, as a function of ψ by a numerical optimization procedure, that is

$$\hat{\psi} = \arg\min_{\psi} \ell_{T,r}(\psi). \tag{16}$$

Finally, given $\hat{d}$ and $\hat{b}$, the estimates $\hat{\alpha}$, $\hat{\beta}$, $\hat{\Gamma}_j$, $j = 1, \ldots, k$, and $\hat{\Omega}$ are found by reduced rank regression as in Johansen (1991, 1995).

7.1 Asymptotic properties of the ML estimator

This section discusses the asymptotic properties (consistency and asymptotic distribution) of the ML estimator of the FVECM$_{d,b}$. The theorems outlined in this section follow Johansen and Nielsen (2012) very closely and the proofs are aimed at verifying the conditions under which the asymptotic results of Johansen and Nielsen (2012) can be extended to the FVECM$_{d,b}$ context. Similarly to Johansen and Nielsen (2012), we make the following assumptions

Assumption 7.1. We assume that:

(i) For $k \geq 0$ and $0 \leq r \leq p$, the process X_t $t = 1, 2, \ldots T$, is generated by model $\mathcal{H}_{r,k}$.

(ii) The errors ε_t are i.i.d. $(0, \Omega_0)$ with $\Omega_0 > 0$ and $E|\varepsilon_t|^8 < \infty$.

(iii) The initial values X_{-n}, $n \geq 0$ are uniformly bounded.

(iv) The true parameter value θ_0 satisfies:

1. $(d_0, b_0) \in \Psi$, with $\Psi = \{(d, b) : 0 < b \leq d \leq d_1\}$ where $d_1 > 0$ can be arbitrarily large.
2. $0 \leq d_0 - b_0 < 1/2, b_0 \neq 1/2$.[3]
3. $\Gamma_{0k} \neq 0$ (if $k > 0$), α_0 and β_0 are $p \times r$ matrices of rank r, $\alpha_0\beta_0 \neq -I_p$. Furthermore, the $\mathcal{F}(d)$ condition, $|\alpha'_{0,\perp}\Gamma_0(1)\beta_{0,\perp}| \neq 0$, with $\Gamma_0(1) = I_p - \sum_{i=1}^{k}\Gamma_{0i}$ holds.
4. If $r < p$, then $|\Pi(z)| = 0$ has $p - r$ unit roots and the remaining roots are outside the unit circle. If $k = r = 0$, only $0 < d_0 \neq 1/2$ is assumed.

[3] This assumption might be restrictive in certain macroeconomic and financial applications. In a recent contribution, Johansen and Nielsen (2018) extend the analysis of the FCVAR$_{d,b}$ to include the possibility that the cointegrating vectors are nonstationary, i.e. $d_0 - b_0 > 1/2$.

7.2 Consistency

We first have to characterize the asymptotic behavior of the profile log-likelihood function for full rank as $T \longrightarrow \infty$, that is

$$\ell_p(\psi) := \lim_{T\to\infty} \ell_{T,p}(\psi), \tag{17}$$

where

$$\ell_{T,p} = -\log\det\left(T^{-1}\sum_{t=1}^{T} R_{it}(\psi)R'_{jt}(\psi)\right) = -\log\det\left(SSR_T(\psi)\right), \tag{18}$$

so that $\ell_p(\psi)$ is the limit log-likelihood function $\ell_{T,p}(\psi)$. The following theorem states the properties of the $\ell_p(\psi)$ and the consistency of the ML estimator of ψ.

Theorem 7.2. *The function $\ell_p(\psi)$ has a strict maximum at $\psi = \psi_0$ that is,*

$$\ell_p(\psi) \le \ell_p(\psi_0) = -\log|\Omega_0|, \quad \psi \in \Psi \tag{19}$$

and equality holds if and only if $\psi = \psi_0$. Let Assumption 7.1 hold, and assuming that $(d_0, b_0) \in \Psi(\eta)$ with $\Psi(\eta) = \{(d,b) : \eta < b \le d \le d_1\} \subset \Psi$ being a family of compact sets with $\eta > 0$, then

$$\ell_{T,p}(\psi_0) \xrightarrow{p} -\log|\Omega_0|. \tag{20}$$

Finally, with probability converging to 1, $\hat{\psi}$ in model $\mathcal{H}_{r,k}$ for $r = 0, 1, \ldots, p$ exists uniquely for $\psi \in \Psi(\eta)$ and is consistent.

See proof in Appendix B.5.

The property of identification derived in Theorem 5.2 guarantees that the consistency of $\ell_{T,p}(\psi_0)$ holds true also when $k > k_0$. Figure 2 reports the surface of the expected profile log-likelihood function of the FCVAR$_{d,b}$ and FVECM$_{d,b}$ in the two-dimensional space of $(d,b) \in [0.2, 0.99]^2$ with $d \ge b$ when the DGP is a co-fractional model with $k_0 = 0$ lags. The plot clearly highlights the presence of two or three equivalent peaks for the FCVAR$_{d,b}$ log-likelihood when $k = 1$ and $k = 2$ respectively. Instead, the log-likelihood function of the FVECM$_{d,b}$ is always associated with a unique maximum for any $k \ge k_0$, as a consequence of the identification property of the FVECM$_{d,b}$. This is relevant in the empirical applications when the true value of k is unknown and it is normally selected with a general-to-specific approach.

7.3 Asymptotic distribution

Let consider again the FVECM$_{d,b}$

$$\Delta_+^d X_t = \alpha\beta'\Delta_+^{d-b} L_b X_t + \sum_{j=1}^{k} \Gamma_j \Delta_+^d X_{t-j} + \varepsilon_t,$$

where $\theta = \{d, b, \alpha, \beta, \Gamma_1, ..., \Gamma_k, \Omega\}$ is the collection of parameters and $\tilde{\theta}$ is a partition of θ such that $\theta \backslash \tilde{\theta}$ denotes all parameters but $\tilde{\theta}$. We want to find an expression for $D_{\tilde{\theta}} \varepsilon_t(\theta_0 \backslash \tilde{\theta})|_{\tilde{\theta}=\tilde{\theta}_0}$ that is the derivative of $\varepsilon_t(\theta_0 \backslash \tilde{\theta})$ with respect to $\tilde{\theta}$. Let define $\varepsilon_t(\theta)$ as

$$\varepsilon_t(\theta) = \Delta_+^d X_t - \alpha\beta' \Delta_+^{d-b} L_b X_t - \sum_{j=1}^{k} \Gamma_j \Delta_+^d X_{t-j}, \tag{21}$$

and the log-likelihood function as $-2 \log \mathcal{L}(\theta) = tr \left\{ \Omega_0^{-1} \sum_{t=1}^{T} \varepsilon_t(\theta) \varepsilon_t(\theta)' \right\}$, with $\Omega = \Omega_0$. By substituting in (21) the Granger representation of X_t evaluated in θ_0 up to the initial conditions (that asymptotically are negligible), we get

$$\begin{aligned} \varepsilon_t(\theta) =& \Delta_+^{d-d_0} (C_0 \varepsilon_t + \sum_{j=1}^{\infty} \beta_{\perp 0} \Phi_{j0} \alpha_{\perp 0} \Delta_+^j \varepsilon_t + \Delta_+^{b_0} Y_t) - \\ & -\alpha\beta' \Delta_+^{d-b-d_0} L_b (C_0 \varepsilon_t + \sum_{j=1}^{\infty} \beta_{\perp 0} \Phi_{j0} \alpha_{\perp 0} \Delta_+^j \varepsilon_t + \Delta_+^{b_0} Y_t) - \\ & - \sum_{i=1}^{k} \Gamma_i \Delta_+^{d-d_0} L^j (C_0 \varepsilon_t + \sum_{j=1}^{\infty} \beta_{\perp 0} \Phi_{j0} \alpha_{\perp 0} \Delta_+^j \varepsilon_t + \Delta_+^{b_0} Y_t). \end{aligned}$$

To derive the asymptotic distribution of θ it is necessary to characterize the asymptotic behavior of the product moments needed to calculate the log-likelihood function. For this purpose, it is useful to use a local parametrization of the FVECM$_{d,b}$. We define the following quantities

$$X_{-1,t} = (\Delta^{d-b} - \Delta^d) X_t \quad X_{it} = (\Delta^{d+i} - \Delta^{d+k}) X_t \quad X_{kt} = \Delta^{d+k} X_t,$$

where $i = 0, \dots, k-1$ and the errors as

$$\varepsilon_t(\lambda) = X_{kt} - \alpha\beta' X_{-1,t} + \sum_{i=0}^{k-1} \Psi_i X_{it},$$

where $\lambda = (d, b, \alpha, \beta, \Psi_*)$ with $\Psi_* = (\Psi_0, \dots, \Psi_{k-1})$. As in Johansen and Nielsen (2012) we locally parametrize the likelihood with the following formulation $\beta = \beta_0 + \beta_{0\perp}(\bar{\beta}_{0\perp}' \beta) = \beta_0 + \beta_{0\perp} \vartheta$. Let $\mathcal{N}(\psi_0, \epsilon) = \{\psi : |\psi - \psi_0| < \epsilon\}$. Then for $(d, b) \in \mathcal{N}(\psi_0, \epsilon), \epsilon < 1/2$ with $\delta_{-1} = d - b - d_0 < -1/2$ and $d + i - d_0 \geq -\epsilon$ for $i \geq 0$. the process $\beta_{0\perp}' X_{-1,t}$ is the only non-stationary process in $\varepsilon_t(\lambda)$. We also introduce the normalized parameter $\zeta = \bar{\beta}_{0\perp}' (\beta - \beta_0) T^{-(\delta_{-1}+1/2)} = \vartheta T^{-(\delta_{-1}+1/2)}$, such that $\beta = \beta_0 + \beta_{0\perp} \zeta T^{\delta_{-1}+1/2}$. Let us define $V_t = (X_{-1,t}' \beta_0, \{X_{it}'\}_{i=0}^{k-1}, X_{kt}')'$ and $\phi = (d, b, \alpha, \Psi_*)$ such that $\lambda = (\phi, \zeta)$. We can write the error as

$$\varepsilon_t(\lambda) = -\alpha T^{\delta_{-1}+1/2} \zeta' \beta_{0\perp}' X_{-1,t} + (-\alpha, \Psi_*, I_p) V_t.$$

When $b_0 > 1/2$, the product moments in the conditional likelihood function $-2T^{-1} \log L_T(\phi, \zeta) =$

$\log|\Omega| + tr\left(\Omega^{-1}T^{-1}\sum_{t=1}^{T}\varepsilon_t(\lambda)\varepsilon_t(\lambda)'\right)$ are

$$\begin{pmatrix} \mathcal{A}_T(\psi) & C_T(\psi) \\ C_T(\psi)' & \mathcal{B}_T(\psi) \end{pmatrix} = T^{-1}\sum_{t=1}^{T}\begin{pmatrix} T^{\delta_{-1}+1/2}\beta_{0\perp}'X_{-1,t} \\ V_t \end{pmatrix}\begin{pmatrix} T^{\delta_{-1}+1/2}\beta_{0\perp}'X_{-1,t} \\ V_t \end{pmatrix}'.$$

Finally we define

$$C_{\varepsilon T}^0 = T^{-1/2}\sum_{t=1}^{T}T^{1/2-b_0}\beta_{0\perp}'X_{-1,t}^0\varepsilon_t',$$

where $X_{-1,t}^0$ is $X_{-1,t}$ with $\lambda = \lambda_0$. When $b_0 < 1/2$, we replace $\delta_{-1}+1/2$ by zero in the definition of $\mathcal{A}_t(\psi), B_t(\psi), C_t(\psi)$ and $C_{\varepsilon T}^0$. The asymptotic behavior of $\mathcal{A}_T(\psi), \mathcal{B}_T(\psi), C_T(\psi)$ and their derivatives when $1/2 < b_0 < d_0$ and $0 < b_0 < 1/2$ is derived in Theorem 6 in Johansen and Nielsen (2012).

We can now outline the following theorem, which is analogous to Theorem 10 in Johansen and Nielsen (2012).

Theorem 7.3. *Under Assumption 7.1, with $X_{-n} = 0$ for $n \geq T^{\nu}$ for some $\nu < 1/2$, the asymptotic distribution of the ML estimator of the FVECM$_{d,b}$ is as follows:*

- *If $b_0 > 1/2$ and $E|\varepsilon_t|^q < \infty$ for some $q > (b_0 - 1/2)^{-1}$, the asymptotic distribution of the ML estimator $\hat{\phi} = (\hat{d}, \hat{b}, \hat{\alpha}, \hat{\Gamma}_j)$ and $\hat{\beta}$ is given by*

$$\begin{pmatrix} T^{1/2}vec(\hat{\phi}-\phi_0) \\ T^{b_0}\bar{\beta}_{0\perp}'(\hat{\beta}-\beta_0) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} N(0,\Sigma_0) \\ \left(\int_0^1 F_0F_0'\right)^{-1}\int_0^1 F_0(dG_0)'(\alpha_0\Omega_0^{-1}\alpha_0)^{-1} \end{pmatrix},$$

where $\Sigma_0 > 0$, $F_0 = \beta_{0\perp}'C_0W_{b_0-1}$ with W_{b_0-1} is the (non-standardized) type II fractional Brownian motion of order $b_0 - 1$, and $G_0 = \alpha_0'\Omega_0^{-1}W$ are independent with $W := W_0$ denoting the Brownian motion generated by ε_t. The two components of the asymptotic distribution are independent (see Lemma 10 in Johansen and Nielsen, 2010). It follows that the asymptotic distribution of $vec(T^{b_0}\bar{\beta}_{0\perp}'(\hat{\beta}-\beta_0))$ is mixed Gaussian with conditional variance given by

$$\mathcal{V} = (\alpha_0'\Omega_0^{-1}\alpha_0)^{-1}\otimes\left(\int_0^1 F_0F_0'du\right)^{-1}.$$

- *If $0 < b_0 < 1/2$, the estimators $(\hat{d}, \hat{b}, \hat{\alpha}, \hat{\Gamma}_j, \hat{\beta})$ are asymptotically Gaussian.*

- *If $k = r = 0$, and $d = b$ the model is $\Delta^d X_t = \varepsilon_t$, and $\hat{d}$ is asymptotically Gaussian.*

Proof. See the proof in Appendix B.7.

7.4 Testing for the cointegration rank

We now focus on the likelihood ratio test for the determination of the co-fractional rank and we rely on the results of Johansen and Nielsen (2012) to prove its asymptotic distribution. Let us first define the model $\mathcal{H}_{p,k}$ as

$$\mathcal{H}_{p,k}: \quad \Delta^d X_t = \Pi \Delta^{d-b} L_b X_t + \sum_{i=1}^{k} \Gamma_i \Delta^d L_b^i X_t + \varepsilon_t,$$

where the following analysis holds for any given $k = k_0$. We consider the test for the null hypothesis $\mathcal{H}_r : rank(\Pi) \leq r$ against the alternative $\mathcal{H}_p : rank(\Pi) \leq p$. We define the LR statistic as

$$-2\log LR(\mathcal{H}_r|\mathcal{H}_p) = T\log\frac{|S_{00}(\hat{\psi}_r)|\prod_{i=1}^{r}(1-\hat{\omega}_i(\hat{\psi}_r))}{|S_{00}(\hat{\psi}_p)|\prod_{i=1}^{p}(1-\hat{\omega}_i(\hat{\psi}_p))} = T(\ell_{T,r}(\hat{\psi}_r) - \ell_{T,p}(\hat{\psi}_p)). \tag{22}$$

The following theorem presents the asymptotic distribution of the LR test.

Theorem 7.4. *Under Assumption 7.1, with $X_{-n} = 0$ for $n \geq T^{\nu}$ for some $\nu < 1/2$, the asymptotic distribution of the LR test in (22) is:*

- *If $b_0 > 1/2$,*

$$-2\log LR(\mathcal{H}_r|\mathcal{H}_p) \xrightarrow{d} tr\left(\int_0^1 (dB)B'_{b_0-1}\left(\int_0^1 B_{b_0-1}B'_{b_0-1}du\right)^{-1}\int_0^1 B_{b_0-1}(dB)'\right)$$

  where $B(u)$ is a $(p-r)-$dimensional standard Brownian motion and $B_{b_0-1}(u)$ is the corresponding standardized type II fractional Brownian motion. The limit distribution is continuous in b_0.

- *If $0 < b_0 < 1/2$,*

$$-2\log LR(\mathcal{H}_r|\mathcal{H}_p) \xrightarrow{d} \chi^2\left((p-r)^2\right).$$

- *Let $P_{\mathcal{H}_1}$ the probability measure under the alternative $\Pi_1 = \alpha_1\beta_1' = \alpha\beta' + \alpha^*\beta^{*\prime}$, where $\alpha_1 = (\alpha, \alpha^*)$ and $\beta_1 = (\beta, \beta^*)$ are $p \times (r + r^*)$ matrices of rank $r_1 = r + r^* > r$, and hence $rank(\Pi_1) > r$. Under the Assumption that X_t is generated by model $\mathcal{H}_r$, then*

$$-2\log LR(\mathcal{H}_r|\mathcal{H}_p) \xrightarrow{P_{\mathcal{H}_1}} \infty,$$

  under the alternative.

Proof. See the proof of Theorem 11 in Johansen and Nielsen (2012).

In the framework of the $\text{FCVAR}_{d,b}$, the parameter b is not identified when $k = 0$ and we are testing $r = 0$ (i.e. $\Pi = 0$). Johansen and Nielsen (2012) suggest to follow the approach of Lasak (2010) and to adopt a sup-type test, $\sup_b LR(b)$, where $LR(b) = -2\log LR(\Pi = 0|b)$, where the supremum is taken over the values of the index b.[4] In the $\text{FVECM}_{d,b}$, the parameter b is not identified for any $k = 0, 1, \ldots$ when testing $r = 0$. Hence, the $\sup_b LR(b)$ statistic should be

[4]Alternatively, Lasak and Velasco (2015) propose a two-step procedure to determine the cointegration rank.

computed for any choice of k under $r = 0$. For a given k, the co-fractional rank can be determined with a sequence of tests for a given nominal size $\varsigma \in (0, 1)$. The sequence of tests is performed by considering the null hypothesis $\mathcal{H}_r$, for $r = 0, 1, \ldots$ in sequence until rejection, and the estimated co-fractional rank $\hat{r}$ is the last non-rejected value of r. The consistency of the test guarantees that any test with $r < r_0$, where r_0 is the true cointegrating rank, will reject with probability 1 as $T \to \infty$. Finally, if the asymptotic size is ς, then $P(\hat{r} < r_0) \to \varsigma$, so that $P(\hat{r} = r_0) \to 1 - \varsigma$. Similarly to MacKinnon and Nielsen (2014), the critical values of the limiting distribution need to be tabulated.

8 Conclusion

In this paper, we have shown that the multivariate co-fractional model of Granger (1986) is suitable to carry out inference on the long-run equilibrium relations between series that are integrated of a fractional order. Indeed, we have proved that the $\text{FVECM}_{d,b}$ allows for a Granger representation theorem and its stability conditions can be studied through the argument principle. Notably, the model is always identified for any combination of number of lags and cointegration rank. Finally, the parameters $\text{FVECM}_{d,b}$ can be estimated by ML in a similar fashion as in Johansen and Nielsen (2012) and they are associated with the same asymptotic behavior as those of the $\text{FCVAR}_{d,b}$.

References

Ahlfors, L. V. (1953). Complex analysis: an introduction to the theory of analytic functions of one complex variable. *New York, London*, page 177.

Andersen, T. G. and Bollerslev, T. (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *Journal of Finance*, 52(3):975–1005.

Anderson, T. W. et al. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. *The Annals of Mathematical Statistics*, 22(3):327–351.

Avarucci, M. (2007). *Three essays on fractional cointegration*. PhD thesis, University of Rome Tor Vergata.

Baillie, R. T. and Bollerslev, T. (1994). Cointegration, fractional cointegration, and exchange rate dynamics. *Journal of Finance*, 49(2):737–45.

Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3–30.

Binder, M. and Pesaran, M. H. (1997). Multivariate linear rational expectations models: characterization of the nature of the solutions and their fully recursive computation. *Econometric Theory*, 13(6):877–888.

Bollerslev, T., Osterrieder, D., Sizova, N., and Tauchen, G. (2013). Risk and return: Long-run relations, fractional cointegration, and return predictability. *Journal of Financial Economics*, 108(2):409–424.

Breitung, J. and Hassler, U. (2002). Inference on the cointegration rank in fractionally integrated processes. *Journal of Econometrics*, 110(2):167–185.

Caporin, M., Ranaldo, A., and Santucci de Magistris, P. (2013). On the predictability of stock prices: A case for high and low prices. *Journal of Banking & Finance*, 37(12):5132–5146.

Carlini, F. and Santucci de Magistris, P. (2017). On the identification of fractionally cointegrated VAR models with the $\mathcal{F}$(d) condition. *Journal of Business & Economic Statistics*, pages 1–13.

Davidson, J. (2002). A model of fractional cointegration, and tests for cointegration using the bootstrap. *Journal of Econometrics*, 110(2):187 – 212.

Delves, L. and Lyness, J. (1967). A numerical method for locating the zeros of an analytic function. *Mathematics of Computation*, 21(100):543–560.

Dolatabadi, S., Nielsen, M. Ø., and Xu, K. (2015). A fractionally cointegrated VAR analysis of price discovery in commodity futures markets. *Journal of Futures Markets*, 35(4):339–356.

Dolatabadi, S., Nielsen, M. Ø., and Xu, K. (2016). A fractionally cointegrated VAR model with deterministic trends and application to commodity futures markets. *Journal of Empirical Finance*, 38:623 – 639.

Engle, R. and Granger, C. W. J. (1987). Cointegration and error correction: representation estimation, and testing. *Econometrica*, 55:251–276.

Franchi, M. (2010). A representation theory for polynomial cofractionality in vector autoregressive models. *Econometric Theory*, 26(04):1201–1217.

Fuchs, B. A. and Shabat, B. V. (1964). *Functions of a complex variable and some of their applications*, volume 1. Pergamon Press.

Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4:221–238.

Gouriéroux, C. and Zakoïan, J.-M. (2017). Local explosion modelling by non-causal process. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(3):737–756.

Granger, C. W. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of econometrics*, 14(2):227–238.

Granger, C. W. J. (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics*, 48(3):213–28.

Hamilton, J. D. (1994). *Time series analysis*. Princeton university press, Princeton.

Hosking, J. (1981). Fractional differencing. *Biometrika*, 68:165–76.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12:231–254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59(6):1551–80.

Johansen, S. (1992). A representation of vector autoregressive processes integrated of order 2. *Econometric Theory*, 8(2):188–202.

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

Johansen, S. (1997). Likelihood analysis of the I(2) model. *Scandinavian Journal of Statistics*, 24(4):433–462.

Johansen, S. (2008a). Representation of cointegrated autoregressive processes with application to fractional processes. *Econometric Reviews*, 28(1-3):121–145.

Johansen, S. (2008b). A representation theory for a class of vector autoregressive models for fractional processes. *Econometric Theory*, 24(3):651–676.

Johansen, S. (2010). Some identification problems in the cointegrated vector autoregressive model. *Journal of Econometrics*, 158(2):262–273.

Johansen, S. and Nielsen, M. Ø. (2010). Likelihood inference for a nonstationary fractional autoregressive model. *Journal of Econometrics*, 158(1):51–66.

Johansen, S. and Nielsen, M. Ø. (2012). Likelihood inference for a fractionally cointegrated vector autoregressive model. *Econometrica*, 80(6):2667–2732.

Johansen, S. and Nielsen, M. Ø. (2018). Nonstationary cointegration in the fractionally cointegrated VAR model. *Forthcoming on the Journal of Time Series Analysis*.

Klein, P. (2000). Using the generalized Schur form to solve a multivariate linear rational expectations model. *Journal of Economic Dynamics and Control*, 24(10):1405–1423.

Lasak, K. (2010). Likelihood based testing for no fractional cointegration. *Journal of Econometrics*, 158(1):67–77.

Lasak, K. and Velasco, C. (2015). Fractional cointegration rank estimation. *Journal of Business & Economic Statistics*, 33(2):241–254.

MacKinnon, J. G. and Nielsen, M. Ø. (2014). Numerical distribution functions of fractional unit root and cointegration tests. *Journal of Applied Econometrics*, 29(1):161–171.

Neusser, K. et al. (2016). *Time Series Econometrics.* Springer.

Nielsen, M. Ø. and Popiel, M. K. (2014). A MATLAB program and user's guide for the fractionally cointegrated VAR model. Technical report, Queen's Economics Department Working Paper.

Nielsen, M. Ø. and Shibaev, S. S. (2018). Forecasting daily political opinion polls using the fractionally cointegrated vector auto-regressive model. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 181(1):3–33.

Robinson, P. M. and Marinucci, D. (2003). Semiparametric frequency domain analysis of fractional cointegration. In Robinson, P. M., editor, *Time Series with Long Memory*, pages 334–373. Oxford University Press.

Rossi, E. and Santucci de Magistris, P. (2013). A no-arbitrage fractional cointegration model for futures and spot daily ranges. *Journal of Futures Markets*, 33(1):77–102.

Rotemberg, J. J. (1987). The new Keynesian microfoundations. *Macroeconomics Annual*, 2:69–104.

Schennach, S. M. (2018). Long memory via networking. *Econometrica*, 86(6):2221–2248.

Shea, G. S. (1991). Uncertainty and implied variance bounds in long-memory models of the interest rate term structure. *Empirical Economics*, 16(3):287–312.

Tschernig, R., Weber, E., and Weigand, R. (2013). Long-run identification in a fractionally integrated system. *Journal of Business & Economic Statistics*, 31(4):438–450.

Zaffaroni, P. (2004). Contemporaneous aggregation of linear dynamic models in large economies. *Journal of Econometrics*, 120(1):75–102.

A Regularity of $f(z)$

In this Appendix, we discuss the regularity properties of $f(z) = (1-z)^{-b(p-r)}g(z)$ such that the argument principle can be adopted to count the number of zeroes inside the unit circle. In particular, we have to show that $f(z)$ is an holomorphic function on the unit circle and it does not have poles inside. An holomorphic function is defined as a complex-valued differentiable function on an open set $\mathbb{D}$ of the $\mathbb{C}$. For instance, the functions $h_1(x) = 1 - (1-z)^b$ and $h_2(x) = (1-z)^b$ are holomorphic in the unit circle for any $b \in \mathbb{R}^+$, see Johansen (2008b). A useful property of holomorphic functions is that the composition of two holomorphic functions is also an holomorphic function. It follows from this property that $\tilde{\Pi}(z)$ is an holomorphic matrix function. Analogously, the determinant $g(z) = |\tilde{\Pi}(z)|$ is holomorphic since the determinant is a continuous function. Hence, $f(z)$ is holomorphic in the unit circle and it does not have any zero on the contour $|z| = 1$. Moreover, the function $f(z)$ does not have any pole inside the unit circle because $g(z)$ does not involve any inverse function of z.

B Proofs

B.1 Proof of Theorem 4.1

To ease the exposition of the proof, we first derive the Granger representation of the model

$$\Delta_+^d X_t = \alpha\beta' L_d X_t + \sum_{j=1}^{k} \Gamma_j \Delta_+^d X_{t-j} + \varepsilon_t,$$

where $d = b$. First of all, let us write the characteristic polynomial as

$$\Pi_d(z) = (1-z)^d (I_p - \sum_{j=1}^{k} \Gamma_j z^j) - \alpha\beta'(1-(1-z)^d). \tag{23}$$

We introduce the variable $y = 1-(1-z)^d$ and we write $\Pi(z) = \Pi^*(z,y)$ as

$$\Pi_d^*(z,y) = (1-y)(I_p - \sum_{j=1}^{k} \Gamma_j z^j) - \alpha\beta' y.$$

Following the proof of Theorem 3 of Johansen (2008a) we calculate $A'\Pi_d^*(z,y)B$ with $A = (\bar{\alpha}, \alpha_\perp)$ and $B = (\bar{\beta}, \beta_\perp)$, with $\bar{\alpha} = \alpha(\alpha'\alpha)^{-1}$ and $\bar{\beta} = \beta(\beta'\beta)^{-1}$. We compute the Taylor expansion of $\Pi_d^*(z,y)$ in $y = 1$ (with $y = 1 \iff z = 1$) and we get

$$A'\Pi_d^*(z,y)B = \begin{pmatrix} -I_r & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \bar{\alpha}'(\Gamma(z) + \alpha\beta')\bar{\beta} & \bar{\alpha}'\Gamma(z)\beta_\perp \\ \alpha_\perp'\Gamma(z)\bar{\beta} & \alpha_\perp'\Gamma(z)\beta_\perp \end{pmatrix}(1-y),$$

where $\Gamma(z) = I_p - \sum_{j=1}^{k} \Gamma_j z^j$. Now, we calculate $A'\Pi_d^*(z,y)BF(y)$ where

$$F(y) = \begin{pmatrix} I_r & 0 \\ 0 & (1-y)^{-1} I_{p-r} \end{pmatrix},$$

and we get

$$K(z,y) = A'\Pi_d^*(z,y)BF(y) = \underbrace{\begin{pmatrix} -I_r & \bar{\alpha}'\Gamma(z)\beta_\perp \\ 0 & \alpha_\perp'\Gamma(z)\beta_\perp \end{pmatrix}}_{K(z)} + \underbrace{\begin{pmatrix} \bar{\alpha}'(\Gamma(z)+\alpha\beta')\bar{\beta} & 0 \\ \alpha_\perp'\Gamma(z)\bar{\beta} & 0 \end{pmatrix}}_{\dot{K}(z)}(1-y).$$

Then

$$K(z,y)^{-1} = (A'\Pi_d^*(z,y)BF(y))^{-1} = K^{-1}(z) + K^{-1}(z)\dot{K}(z)K^{-1}(z)\cdot(1-y) + (1-y)^2 H_1(z,y),$$

$H_1(z,y)$ is the remainder term of the infinite series $K(z,y)^{-1}$ in $y=1$, and

$$K^{-1}(z) = \begin{pmatrix} -I_r & (\bar{\alpha}'\Gamma(z)\beta_\perp)(\alpha_\perp'\Gamma(z)\beta_\perp)^{-1} \\ 0 & (\alpha_\perp'\Gamma(z)\beta_\perp)^{-1} \end{pmatrix},$$

which is computed with the formula of the partitioned inverse. We now calculate

$$F(y)K(z,y)^{-1} = (1-y)^{-1}M_{-1}(z) + M_0(z) + (1-y)H_2(z,y),$$

with

$$M_{-1}(z) = \begin{pmatrix} 0 & 0 \\ 0 & (\alpha_\perp'\Gamma(z)\beta_\perp)^{-1} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & (\alpha_\perp'\Gamma\beta_\perp)^{-1} \end{pmatrix} + (1-z)H_3(z),$$

where $\Gamma = I_p - \sum_{j=1}^k \Gamma_j$ and $|\alpha_\perp'\Gamma\beta_\perp| \neq 0$ and $M_0(z)$ contains term of degree 0 in $(1-y)$. Therefore, by pre-multiplying by B and post-multiplying by A', we find that the inverse of $\Pi_d^*(z,y)$ with respect to y is

$$\begin{aligned} \Pi_d^*(z,y)^{-1} &= BF(y)(A'\Pi_d^*(z,y)BF(y))^{-1}A' = \\ & (1-y)^{-1}\beta_\perp(\alpha_\perp'\Gamma(z)\beta_\perp)^{-1}\alpha_\perp' + C^*(z) + (1-y)H(z,y), \end{aligned} \tag{24}$$

and the only pole of (24) is $(1-y)$ and $H(z,y)$ has zeros in $z=1$ and $y=1$. The function $\tilde{H}(z,y) = C^*(z) + (1-y)H(z,y)$ is regular[5] in the complex circle with no singularity at $y=z=1$. When $b>0$, the function $y = 1-(1-z)^d$ is regular for $|z|<1$ and continuous for $|z|\leq 1$. Hence,

$$F(z) = \tilde{H}(1-(1-z)^d, z), \quad |z| \leq 1,$$

is continuous for $|z|\leq 1$ and regular without singularities on the open unit disk $|z|<1$. Hence, the expansion $F(z) = \sum_{n=0}^\infty F_n z^n$, $|z|<1$ is defined with $\sum_{n=0}^\infty \|F_n\|^2 < \infty$. We define now $Y_t = F(L)\varepsilon_t = \sum_{n=0}^\infty F_n\varepsilon_{t-n}$ as a stationary process with mean zero, finite variance and continuous

[5] A regular (or holomorphic) function is defined to be a complex-valued differentiable function on an open (and arc connected) set $\mathbb{D}$ of $\mathbb{C}$, where $\mathbb{C}$ denotes the set of complex numbers. For further details see Johansen (2008b).

spectral density given by

$$f_Y(\lambda) = \frac{1}{2\pi}F(e^{-i\lambda})\Omega F(e^{i\lambda})' = \frac{1}{2\pi}\tilde{H}(1-(1-e^{-i\lambda})^d, e^{-i\lambda})\Omega\tilde{H}(1-(1-e^{i\lambda})^d, e^{i\lambda})',$$

and for $\lambda = 0$ we get

$$\frac{1}{2\pi}F(1)\Omega F(1)' = \frac{1}{2\pi}\tilde{H}(1,1)\Omega\tilde{H}(1,1) = \frac{1}{2\pi}C^*(1)\Omega C^*(1)'.$$

Given the inequality

$$\Omega - \alpha(\alpha'\Omega\alpha)^{-1}\alpha' = \Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\alpha_\perp'\Omega \geq 0,$$

then it follows that

$$\beta' C^*(1)\Omega C^{*\prime}(1)\beta \geq 0,$$

because $\beta' C^*(1)\alpha = -I_r$. Hence, we have shown that $f_Y(0) \neq 0$, hence $Y_t \sim \mathcal{F}(0)$. Now, we know that

$$\Pi_d^{-1}(z) = C(z)(1-z)^{-d} + F(z),$$

and applying the operator $\Pi_{d,+}^{-1}(L)$ (defined analogously to the truncated filter in (4)) to the equation $\Pi_d(L)X_t = \varepsilon_t$ we find the solution

$$X_t = C(L)(1-z)_+^{-d} + Y_t^+ - \Pi_{d,+}^{-1}(L)\Pi_{d,-}(L)X_t.$$

This means that $X_t \sim \mathcal{F}(d)$ because $C(1) \neq 0$ and that $\beta' X_t = \beta' Y_t^+ \sim \mathcal{F}(0)_+$ because $Y_t \sim \mathcal{F}(0)$. The case $d > b$ can be solved in a similar way by noting that

$$\Delta_+^d X_t = \alpha\beta'\Delta_+^{d-b}L_b X_t + \sum_{j=1}^{k}\Delta_+^d\Gamma_j X_{t-j} + \varepsilon_t,$$

has the characteristic polynomial given by

$$\Pi(z) = (1-z)^d I_p - \alpha\beta'(1-z)^{d-b}(1-(1-z)^b) - \sum_{j=1}^{k}\Gamma_j(1-z)^d z^j.$$

that can be written as

$$\Pi(z) = (1-z)^{d-b}\left[(1-z)^b I_p - \alpha\beta'(1-(1-z)^b) - \sum_{j=1}^{k}\Gamma_j(1-z)^b z^j\right].$$

The polynomial $(1-z)^{d-b}$ is trivially invertible and the polynomial $[(1-z)^b I_p - \alpha\beta'(1-(1-z)^b) - \sum_{j=1}^{k}\Gamma_j(1-z)^b z^j]$ is the same as in (23) where $d = b$ and we proved is invertible.

B.2 Proof of Lemma 4.2

To illustrate the steps to obtain the recursion to compute the IRFs, we first consider the following $\text{FVECM}_{d,b}$ with one lag,

$$\Delta_+^d X_t = \alpha\beta' \Delta_+^{d-b} L_b X_t + \Gamma_1 \Delta_+^d X_{t-1} + \varepsilon_t,$$

which can be written as

$$\Delta_+^d X_t = \alpha\beta'(\Delta_+^{d-b} - \Delta_+^d) X_t + \Gamma_1 \Delta_+^d X_{t-1} + \varepsilon_t.$$

Now, let us write explicitly $X_t, t = 1, \ldots, T$ as a function of ε_1. The first term is $X_1 = \varepsilon_1$ and the second is given by

$$X_2 - dX_1 = \alpha\beta'(-(d-b) + d)X_1 + \Gamma_1 X_1 + \varepsilon_2,$$

so that

$$X_2 = (d + b\alpha\beta' + \Gamma_1)\varepsilon_1 + \varepsilon_2.$$

Let us define $\Theta_1 := d + b\alpha\beta' + \Gamma_1$, the third recursion is given by

$$X_3 - dX_2 + \frac{d(d-1)}{2} X_1 = b\alpha\beta' X_2 + \alpha\beta'[(d-b)(d-b-1)/2 - d(d-1)/2]X_1 + \Gamma_1 X_2 - d \cdot \Gamma_1 X_1 + \varepsilon_3,$$

and rearranging the terms we get

$$X_3 = d\Theta_1\varepsilon_1 - \frac{d(d-1)}{2}\varepsilon_1 + b\alpha\beta'\Theta_1\varepsilon_1 + \alpha\beta'[(d-b)(d-b-1)/2 - d(d-1)/2]\varepsilon_1 + \Gamma_1\Theta_1\varepsilon_1 - d\Gamma_1\varepsilon_1 + \varepsilon_3$$

Hence we can define

$$\Theta_2 = [\Theta_1\Theta_1 + \alpha\beta'[(d-b)(d-b-1)/2 - b(b-1)/2] - d\Gamma_1 - d(d-1)/2]\,\varepsilon_1.$$

Iterating this process, we can get the impulse response coefficients, Θ_j $j = 1, 2, \ldots$, for the $\text{FVECM}_{d,b}$.

B.3 Proof of Theorem 5.2

We have to show that

$$P_{\theta_0} = P_{\theta_1} \implies \theta_0 = \theta_1,$$

under the condition $\varepsilon_t \sim N(0, \Omega)$, so that the conditional variance of X_t is $\text{Var}(X_t|\mathcal{I}_{t-1}) = \Omega$, where the filtration is the σ-field generated as $\mathcal{I}_{t-1} = \{\mu_0, X_0, X_1, \ldots, X_{t-1}\}$. Hence, the matrix $\Omega = \text{Var}(\varepsilon_t)$ is identified, so that $\Omega = \Omega_0$. We now show that the conditional mean of the process X_t is identified for given k and r, i.e. that the characteristic polynomial is uniquely determined as a function of the parameters, θ_0.

Identification when both k and r are known

Let us consider the two characteristic polynomials

$$\Pi_0(z) = (1-z)^{d_0} I_p - \alpha_0 \beta_0' (1-z)^{d_0-b_0} (1-(1-z)^{b_0}) - \sum_{j=1}^{k} \Gamma_{j,0} (1-z)^{d_0} z^j,$$

and

$$\Pi_1(z) = (1-z)^{d_1} I_p - \alpha_1 \beta_1' (1-z)^{d_1-b_1} (1-(1-z)^{b_1}) - \sum_{j=1}^{k} \Gamma_{j,1} (1-z)^{d_1} z^j.$$

We identify the parameters of the model when $\Pi_0(z) = \Pi_1(z)$ if and only if $\theta_0 = \theta_1$. The following set of equalities holds under the FVECM$_{d,b}$ when k and r are known and fixed

$$\begin{aligned}
(1-z)^{d_0} I_p = (1-z)^{d_1} I_p \quad &\Longleftrightarrow \quad d_0 = d_1 \\
\alpha_0 \beta_0' (1-z)^{d_0-b_0} (1-(1-z)^{b_0}) = \alpha_1 \beta_1' (1-z)^{d_1-b_1} (1-(1-z)^{b_1}) \quad &\Longleftrightarrow \quad b_0 = b_1 \\
\Gamma_{j,0} (1-z)^{d_0} z^j = \Gamma_{j,1} (1-z)^{d_1} z^j, j = 1, \ldots, k \quad &\Longleftrightarrow \quad \Gamma_{j,0} = \Gamma_{j,1},
\end{aligned}$$

with $\alpha_1 = \alpha_0 \xi$ and $\beta_1 = \beta_0 \xi^{-1}$. Hence, $d, b, \Gamma_j, j = 1, \ldots, k$ are identified as well as α and β up to rotations, ξ.

Identification of $\mathcal{H}_{k_0}$ when $k > k_0$

Let us consider the following two models

$$\mathcal{H}_{k_0} : \Delta_+^{d_0} X_t = \alpha_0 \beta_0' \Delta_+^{d_0-b_0} L_{b_0} X_t + \Gamma_{1,0} \Delta_+^{d_0} X_{t-1} + \cdots + \Gamma_{k_0,0} \Delta_+^{d_0} X_{t-k_0} + \varepsilon_t,$$

and

$$\mathcal{H}_k : \Delta_+^{d} X_t = \alpha \beta' \Delta_+^{d-b} L_b X_t + \Gamma_1 \Delta_+^{d} X_{t-1} + \cdots + \Gamma_{k_0} \Delta_+^{d} X_{t-k} + \varepsilon_t,$$

where k is such that $k \geq k_0$ and the rank, r, is known and fixed. The characteristic polynomials of $\mathcal{H}_{k_0}$ and $\mathcal{H}_k$ are

$$\Pi_{k_0}(z) = (1-z)^{d_0} I_p - \alpha_0 \beta_0' (1-z)^{d_0-b_0} (1-(1-z)^{b_0}) - \sum_{i=1}^{k_0} \Gamma_{i,0} (1-z)^{d_0} z^i,$$

and

$$\Pi_k(z) = (1-z)^{d} I_p - \alpha \beta' (1-z)^{d-b} (1-(1-z)^{b}) - \sum_{i=1}^{k} \Gamma_i (1-z)^{d} z^i.$$

By equating $\Pi_{k_0}(z)$ and $\Pi_k(z)$ we get the following set of conditions

$$\begin{aligned}
(1-z)^{d_0} I_p = (1-z)^d I_p &\iff d = d_0 \\
\alpha_0 \beta_0'(1-z)^{d_0-b_0}(1-(1-z)^{b_0}) = \alpha\beta'(1-z)^{d-b}(1-(1-z)^b) &\iff b = b_0 \\
\Gamma_{i,0}(1-z)^{d_0} z^i = \Gamma_i (1-z)^d z^i, \quad i = 1, \dots, k_0 &\iff \Gamma_{i,0} = \Gamma_i \\
0 = \Gamma_i (1-z)^d z^i, \quad i = k_0+1, \dots, k &\iff \Gamma_i = 0,
\end{aligned}$$

with $\alpha_0 = \alpha\xi$ and $\beta_0 = \beta\xi^{-1}$. Hence, the model $\mathcal{H}_{k_0}$ is always uniquely identified as a subset of model $\mathcal{H}_k$ associated with the restriction $\Gamma_i = 0$ for $i = k_0+1, \dots, k$ (up to rotations ξ of α and β).

Identification when rank and lags are unknown

Let us consider the following two models

$$\mathcal{H}_{0,k} : \Delta_+^{d_{0,k}} X_t = \sum_{j=1}^{k} \Gamma_{j,(0,k)} \Delta_+^{d_{0,k}} X_{t-j} + \varepsilon_t,$$

$$\mathcal{H}_{p,k-1} : \Delta_+^{d_{p,k-1}} X_t = \Xi_{p,k-1} \Delta_+^{d_{p,k-1}-b_{p,k-1}} L_{b_{p,k-1}} X_t + \sum_{j=1}^{k-1} \Gamma_{j,(p,k-1)} \Delta_+^{d_{p,k-1}} X_{t-j} + \varepsilon_t,$$

The goal is to prove that $\mathcal{H}_{0,k} \neq \mathcal{H}_{p,k-1}$. The characteristic polynomials are

$$\Pi_{0,k}(z) = (1-z)^{d_{0,k}} I_p - \sum_{j=1}^{k} \Gamma_{j,(0,k)} (1-z)^{d_{0,k}} z^j,$$

and

$$\Pi_{p,k-1}(z) = (1-z)^{d_{p,k-1}} I_p - \Xi_{p,k-1}(1-z)^{d_{p,k-1}-b_{p,k-1}}(1-(1-z)^{b_{p,k-1}}) + \sum_{j=1}^{k-1} \Gamma_{j,(p,k-1)}(1-z)^{d_{p,k-1}} z^j.$$

The polynomial $\Pi_{p,k-1}(z)$ contains the term $(1-z)^{d_{p,k-1}-b_{p,k-1}}(1-(1-z)^{b_{p,k-1}})$ that does not appear in $\Pi_{0,k}(z)$ and there are no restrictions on $d_{p,k-1}$, $b_{p,k-1}$, $\Gamma_{j,(p,k-1)}$ such that $\mathcal{H}_{0,k} = \mathcal{H}_{p,k-1}$. Hence $\mathcal{H}_{0,k} \neq \mathcal{H}_{p,k-1}$. ■

B.4 Proof of Theorem 6.1

To ease the exposition of the proof, we first derive the Granger representation of the FVECM$_{d,b}$ under (11) of

$$\Delta_+^d X_t = \alpha\beta' L_d X_t + \sum_{j=1}^{k} \Gamma_j \Delta_+^d X_{t-j} + \varepsilon_t, \tag{25}$$

where $d = b$ and $\alpha_\perp' \left(I_p + \alpha'\beta' - \sum_{j=1}^k \Gamma_j \right) \beta_\perp = \xi\eta'$ with ξ and η being $p - r \times s$ matrices with $\alpha_\perp$ and $\beta_\perp$ such that $\alpha'\alpha_\perp = 0$ and $\beta'\beta_\perp = 0$. The characteristic polynomial of (25) is

$$\Lambda_d(z) = (1-z)^d I_p - \alpha\beta'(1-(1-z)^d) - \sum_{j=1}^k \Gamma_j (1-z)^d z^j,$$

which can be written as

$$\Lambda_d^*(z,y) = (1-y) I_p - \alpha\beta' y - \sum_{j=1}^k \Gamma_j (1-y) z^j,$$

where $y = 1 - (1-z)^d$. Hence

$$\Lambda_d^*(z,y) = (1-y) \underbrace{\left(I_p + \alpha\beta' - \sum_{j=1}^k \Gamma_j (1-y) z^j \right)}_{\Gamma(z)} - \alpha\beta'.$$

Let us define $A = (\bar{\alpha}, \bar{\alpha}_1, \alpha_2)$ and $B = (\bar{\beta}, \bar{\beta}_1, \beta_2)$, where $\bar{\alpha} = \alpha(\alpha'\alpha)^{-1}$, $\bar{\beta} = \beta(\beta'\beta)^{-1}$, $\bar{\alpha}_1 = \alpha_1(\alpha_1'\alpha_1)^{-1}$ with $\alpha_1 = \bar{\alpha}_\perp \xi$, $\bar{\beta}_1 = \beta_1(\beta_1'\beta_1)^{-1}$ with $\beta_1 = \bar{\beta}_\perp \eta$, $\alpha_2 = \bar{\alpha}_\perp \xi_\perp$ and $\beta_1 = \bar{\beta}_\perp \eta_\perp$. We can compute the Taylor expansion of $A'\Lambda_d^*(z,y)B$ in $y = 1$ (with $y = 1 \iff z = 1$) as

$$A'\Lambda_d^*(z,y)B = \begin{pmatrix} -I_r + (1-y)\bar{\alpha}'\Gamma(z)\bar{\beta} & \bar{\alpha}'\Gamma(z)\bar{\beta}_1(1-y) & \bar{\alpha}'\Gamma(z)\beta_2(1-y) \\ (1-y)\bar{\alpha}_1'\Gamma(z)\bar{\beta} & (1-y)I_s & 0 \\ (1-y)\bar{\alpha}_2'\Gamma(z)\bar{\beta} & 0 & 0 \end{pmatrix}.$$

Let us now define

$$F(y) = \begin{pmatrix} I_r & 0 & (1-y)^{-1}\bar{\alpha}'\Gamma(z)\beta_2 \\ 0 & (1-y)^{-1}I_s & 0 \\ 0 & 0 & (1-y)^{-2}I_{p-r-s} \end{pmatrix},$$

and calculate $K(z,y) = A'\Lambda_d^*(z,y)BF(z) = K(z) + (1-y)\dot{K}(z)$ where

$$K(z) = \begin{pmatrix} -I_r & \bar{\alpha}'\Gamma(z)\bar{\beta}_1 & \bar{\alpha}'\Gamma(z)\bar{\beta}\bar{\alpha}'\Gamma(z)\beta_2 \\ 0 & I_s & \bar{\alpha}_1'\Gamma(z)\bar{\beta}\bar{\alpha}'\Gamma(z)\beta_2 \\ 0 & 0 & \alpha_2'\Gamma(z)\bar{\beta}\bar{\alpha}'\Gamma(z)\beta_2 \end{pmatrix},$$

and

$$\dot{K}(z) = \begin{pmatrix} \bar{\alpha}'\Gamma(z)\bar{\beta} & 0 & 0 \\ \bar{\alpha}_1'\Gamma(z)\bar{\beta} & 0 & 0 \\ \alpha_2'\Gamma(z)\bar{\beta} & 0 & 0 \end{pmatrix}.$$

Then, to guarantee that $K(z)$ is invertible, we have to impose that

$$|\alpha_2'\Gamma(1)\bar{\beta}\bar{\alpha}'\Gamma(1)\beta_2| \neq 0, \tag{26}$$

which we name $\mathcal{F}(2b)$ condition. A necessary condition for (26) to hold is that $p < 2r + s$. By inversion of $K(z, y)$, we get

$$K(z,y)^{-1} = (A'\Lambda_d^*(z,y)BF(y))^{-1} = K^{-1}(z) + (1-y)K^{-1}(z)\dot{K}(z)K^{-1}(z) + (1-y)^2 H_1(z,y)$$

where $H_1(z, y)$ is the remainder term of the infinite series $K(z, y)^{-1}$ in $y = 1$. Assuming that a $\delta > 0$ exists, such that $0 < |z - 1| < \delta$, $H_1(z, y)$ is regular for $|1 - y| < \delta$. Hence, by the formula of the partitioned inverse, we get

$$K^{-1}(z) = \begin{pmatrix} -I_r & \bar{\alpha}'\Gamma(z)\beta_\perp & \left(\theta_{02}(z) - \bar{\alpha}'\Gamma(z)\bar{\beta}_1\theta_{12}(z)\right)\theta_{22}(z)^{-1} \\ 0 & I_s & -\theta_{12}(z)\theta_{22}(z)^{-1} \\ 0 & 0 & \theta_{22}(z)^{-1} \end{pmatrix},$$

where $\theta_{ij}(z) = A_{i+1}'\Gamma(z)\bar{\beta}\bar{\alpha}'\Gamma(z)B_{j+1}$ for $i, j = 0, 1, 2$. It follows that

$$F(y)^{-1}K(z,y)^{-1} = (1-y)^{-2}M_{-2}(z) + (1-y)^{-1}M_{-1}(z) + M_0(z) + (1-y)H_2(z,y),$$

with

$$M_{-2}(z) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \theta_{22}(z)^{-1} \end{pmatrix},$$

and

$$M_{-1}(z) = \begin{pmatrix} 0 & 0 & -\bar{\alpha}'\Gamma(z)\beta_2\theta_{22}(z)^{-1} \\ 0 & -I_s & \theta_{12}(z)\theta_{22}(z)^{-1} \\ -\theta_{22}^{-1}\alpha_2'\Gamma(z)\beta_2 & \theta_{22}(z)^{-1}\theta_{21}(z) & \Xi(z) \end{pmatrix},$$

with

$$\Xi(z) = \theta_{22}(z)^{-1}\left[\alpha_2'\Gamma(z)\bar{\beta}\bar{\alpha}'\Gamma(z)\bar{\beta}\bar{\alpha}'\Gamma(z)\beta_2 - \theta_{21}(z)\theta_{12}(z)\right]\theta_{22}(z)^{-1}.$$

The matrix $M_0(z)$ is very involved but it has the following form

$$M_0(z) = \begin{pmatrix} -I_r + \bar{\alpha}'\Gamma(z)\beta_2\theta_{22}(z)^{-1}\alpha_2'\Gamma(z)\bar{\beta} & * & * \\ * & * & * \\ * & * & * \end{pmatrix}.$$

Finally, we use

$$\begin{aligned}\Lambda_d^*(z,y)^{-1} &= BF(y)(A'\Lambda_d^*(z,y)BF(y))^{-1}A' = BF(y)K(z)^{-1}A'\\ &= C_2(z)\frac{1}{(1-y)^2}+C_1(z)\frac{1}{1-y}+C_0(z)+(1-y)H(z,y),\end{aligned}$$

where $H(z,y)$ is regular for $|z-1|<\delta$, and $C_0(z)$ and $C_1(z)$ and $C_2(z)$ are

$$\begin{aligned}C_2(z) &= \beta_2\theta_{22}(z)^{-1}\alpha_2'\\ C_1(z) &= -\bar\beta_1\bar\alpha_1' + \left(\bar\beta_1\theta_{12}(z) - \bar\beta\bar\alpha'\Gamma(z)\beta_2\right)\theta_{22}(z)^{-1}\alpha_2' +\\ &\quad +\beta_2\theta_{22}(z)^{-1}\left(\theta_{21}(z)\bar\alpha_1' - \alpha_2'\Gamma(z)\beta_2\bar\alpha\right) + \beta_2\Xi(z)\alpha_2'\\ \beta'C_0(z)\alpha &= -I_r + \bar\alpha'\Gamma(z)C_2\Gamma(z)\bar\beta.\end{aligned}$$

The function $\Lambda^*(z,y) = C_0(z)+(1-y)H(z,y)$ under the condition that the roots of $|\Lambda(z,1-(1-z)^b)|=0$ are outside the unit circle is regular without singularities inside the unit circle. We define $F(z)=\Lambda^*(z,1-(1-z)^b)$ for $|z|\le 1$. By Lemma A.1 in Johansen (2008b) $F(z)$ is regular for $|z|<1$ so that $Y_t=\sum_{n=0}^{\infty}F_n\varepsilon_{t-n}$ is a stationary process with continuous spectrum, where $F(z)=\sum_{n=0}^{\infty}F_nz^n$, $|z|<1$. We find then

$$\Lambda_d^*(z,y)^{-1} = C_2(z)(1-z)_+^{2b} + C_1(z)(1-z)^b + F(z). \tag{27}$$

The solution of the equation $\Lambda(L)X_t=\varepsilon_t$ is obtained by taking $\Lambda_+^{-1}(L)$ and find

$$X_t = C_2(L)\Delta_+^{2b} + \varepsilon_t + C_1(L)\Delta_+^{b} + \varepsilon_t + Y_t^+ - \Lambda_+(L)^{-1}\Lambda_-(L)X_t. \tag{28}$$

It is seen that $X_t\sim\mathcal{F}(2b)$ because $C_2(L)\neq 0$ that $(\beta,\beta_1)'X_t\sim\mathcal{F}(b)$. Instead the polynomial co-fractionality can be obtained by taking $\beta'X_t-\bar\alpha'\Gamma(L)\Delta_+^bX_t\sim\mathcal{F}(0)$. To extend to the case $d\ge b>0$, it is sufficient to consider the case

$$\Delta_+^{d-b}[\Delta_+^bX_t - \alpha\beta'L_bX_t - \sum_{j=1}^{k}\Gamma_j\Delta_+^bLX_t] = \varepsilon_t,$$

with characteristic polynomial given by

$$\Lambda(z) = (1-z)^{d-b}\left[(1-z)^bI_p - \alpha\beta'(1-(1-z)^b) - \sum_{j=1}^{k}\Gamma_j(1-z)^bz^j\right].$$

Based on the previous results, this implies that

$$\Delta_+^{d-b}X_t = \frac{1}{\Delta_+^{2b}}C_2(L)\varepsilon_t + \frac{1}{\Delta_+^{b}}C_1(L)\varepsilon_t + Y_t^+ + \psi_t,$$

where $\psi_t = \Delta_+^{d-b}\mu_t$, so that

$$X_t = \Delta_+^{-b-d}C_2(L)\varepsilon_t + \Delta_+^{-d}C_1(L)\varepsilon_t + \Delta_+^{-(d-b)}Y_t^+ + \mu_t. \qquad \blacksquare$$

B.5 Proof of Theorem 7.2

The proof of Theorem 7.2 consists of reconciling with the convergence results of the product moments, $S_{ij,t}(\psi)$, as outlined in Appendix A in Johansen and Nielsen (2012). In particular, we have to prove that the stochastic properties of X_t and of the stationary process $U_t = C_0(L)\varepsilon_t + \Delta_+^{b_0}Y_t$ for the FVECM$_{d,b}$ are the same as for the FCVAR$_{d,b}$. In particular, we can define the following quantities

$$\begin{aligned}
X_{-1,t} &= (\Delta_+^{d-b} - \Delta_+^{d})X_t, \quad X_{k,t} = \Delta_+^{d+k}X_t,\\
X_{i,t} &= (\Delta_+^{d+i} - \Delta_+^{d+k})X_t, \quad i = 0,\dots,k-1\\
U_{-1,t} &= (\Delta_+^{d-b-d_0} - \Delta_+^{d-d_0})U_t, \quad U_{k,t} = \Delta_+^{d+k-d_0}X_t,\\
U_{i,t} &= (\Delta_+^{d+i} - \Delta_+^{d+k})\Delta_+^{-d_0}U_t, \quad i = 0,\dots,k-1
\end{aligned}$$

such that we can determine the class of stationary processes for a given ψ as

$$\mathcal{F}_{stat}(\psi) = \left\{\beta_0' U_{jt} \text{ for all j, and } U_{it} \text{ for } d - d_0 > -1/2\right\}. \tag{29}$$

For $d_0 < 1/2$, $d - d_0 \geq -d_0 > -1/2$, the set $\mathcal{F}_{stat}(\psi)$ contains $U_{i,t}$ for all i. We next want to define the probability limit, $\ell_p(\psi)$, of the profile likelihood function $\ell_{T,p}(\psi)$. The limit of $\log\det(SSR_T(\psi))$ is infinite if $X_{k,t}$ is non-stationary and is finite if $X_{k,t}$ is (asymptotically) stationary. Let us now focus on the stochastic properties of $\Delta_+^d X_t = C(L)\varepsilon_t + \Delta_+^b Y_t$, up to the initial conditions that are asymptotically negligible by assumption. We first define an analogous of the Beveridge-Nelson decomposition for fractional processes similar to that of Definition 2 in Johansen and Nielsen (2012, p. 2673). In particular, the polynomial $C(z) = \sum_{j=0}^{\infty} A_j(1-z)^j$ can be factorized as

$$C(z) = C(1) + (1-z)C^*(z), \tag{30}$$

with $C^*(z) = \sum_{j=0}^{\infty} \varphi_j^* z$ and φ_j^* defining an absolute summable sequence by the classic Beveridge-Nelson decomposition. It follows that the process $\Delta_+^d X_t$ can be written as

$$\Delta_+^d X_t \quad = \quad C\varepsilon_t + \Delta_+ Y_t^* + \Delta_+^b Y_t, \tag{31}$$

where $\tilde{Y}_t = \Delta_+^{\lfloor b \rfloor} Y_t$ and with $Y_t^* = C^*(L)\varepsilon_t$. As shown in Lemma B.2 below, the process $\Delta_+^d X_t$ belongs to the $\mathcal{Z}_b$ class. This means that the limit theory for product moments of the stochastic terms in (31) is the same as Johansen and Nielsen (2012), and that Lemma A.9 and Corollary A.10 in Johansen and Nielsen (2012) hold also for the FVECM$_{d,b}$. Therefore, the concentrated log-likelihood function $\ell_{T,p}(\psi) = -\log|SSR_T(\psi)|$ has the same limit as in Johansen and Nielsen

(2012) for the set of intervals for the parameters d and b given in (29). Hence, consistency follows.

B.6 The $\mathcal{Z}_b$ class

To characterize the asymptotic behaviour of the product moments in the log-likelihood function, we follow Johansen and Nielsen (2012) and introduce the class of processes $\mathcal{Z}_b$, as defined below.

Definition B.1. Following Johansen and Nielsen (2012, p. 2673), we define the class $\mathcal{Z}_b$ as the set of stationary processes Z_t that can be represented as

$$Z_t = \varphi\varepsilon_t + \Delta_+^b \sum_{n=0}^{\infty} \varphi_n^* \varepsilon_{t-n}, \tag{32}$$

where $\sum_{n=0}^{\infty} |\varphi_n^*| < \infty$.

In the following, we show that X_t generated by the FCVECM$_{d,b}$ belongs to the class $\mathcal{Z}_b$.

Lemma B.2. *The process*

$$Z_t := \Delta_+^d X_t = C\varepsilon_t + \Delta_+ Y_t^* + \Delta_+^b Y_t, \tag{33}$$

belongs to the class $\mathcal{Z}_b$ specified in Definition B.1.

The proof of Lemma B.2 proceeds as follows. Let us define $B(z) := \alpha_\perp' \Gamma(z)\beta_\perp$. $B(z)$ is a stationary process because $\alpha_\perp' \Pi(z)\beta_\perp = \alpha_\perp' \Gamma(z)\beta_\perp(1-z)^b$ and $\Pi(z) = \Gamma(z)(1-z)^b - \alpha\beta'(1-(1-z)^b)$ has roots in 1 or outside the unit circle. Given that the $\mathcal{F}(d)$ condition holds, $B(z)$ has roots outside the unit circle and it is an autoregressive process. We want to study the behaviour of $B(z)^{-1} = C(z) = \sum_{i=0}^{\infty} C_i z^i$. It follows from Hamilton (1994, p.263) that the (ℓ, k) elements $(C_{\ell k})_i$ of the matrix C_i are such that $|(C_{\ell k})_i| \leq M_1|\lambda|^i$, where $|\lambda| < 1$ where M is an universal constant that bounds $|(C_{\ell k})_i|$ for any $i = 1, 2, \ldots$ This means that $||C_i|| \leq M_2|\lambda|^i$, where $|\lambda| < 1$, where $|| \cdot ||$ denotes a norm defined on the space of matrices. Let us focus on the expansion $C(z) = C(1) + (1-z)C^*(z)$. Then $C^*(z) = \frac{C(z)-C(1)}{(1-z)} = \sum_{i=0}^{\infty} \frac{C_i(z^i-1)}{1-z} = \sum_{i=0}^{\infty} C_i \sum_{j=0}^{i} z^j = \sum_{i=0}^{\infty} C_i^* z^i$ where $C_i^* = \sum_{j \geq i} C_j$. Let us prove that the power series $C^*(z)$ is absolutely summable. It follows that

$$\begin{aligned}\sum_{i=0}^{\infty} \sum_{j \geq i} ||C_j|| &= M \textstyle\sum_{i=0}^{\infty} \sum_{j \geq i} |\lambda|^j = M \sum_{i=0}^{\infty} \frac{1}{1-|\lambda|} - \frac{1}{1-|\lambda|}(1 + |\lambda| + \ldots + |\lambda|^{i-1}) \\ &= M \textstyle\sum_{i=0}^{\infty} \frac{1}{1-|\lambda|} - \frac{1-|\lambda|^i}{1-|\lambda|} = \frac{M}{1-|\lambda|} \sum_{i=0}^{\infty} |\lambda|^i < \infty.\end{aligned}$$

Using the fact that $\sum_{i=0}^{\infty} |C_i| < \infty$ if and only if $\sum_{i=0}^{\infty} ||C_i|| < \infty$, see Neusser et al. (2016, p.206), $C^*(z)$ is absolute summable. Hence, $Y_t^* = \sum_{j=0}^{\infty} C_j^* \varepsilon_{t-j}$ with $\sum_{j=0}^{\infty} |C_j^*| < \infty$. We now turn our attention to the term $\Delta_+^b Y_t^*$ for $b > 1$, which can be written as $\Delta_+^b Y_t^* = \Delta_+^{\{b\}} Y_t^{**}$, where $Y_t^{**} = \sum_{j=0}^{\infty} C_j^{**} \varepsilon_{t-j}$ with $\sum_{j=0}^{\infty} |C_j^{**}| < \infty$, and $\{b\}$ is defined as $\{b\} = b - \lfloor b \rfloor$, where $\lfloor b \rfloor$ denotes the greatest integer less than b. Hence, if $b > 1$, the process $\Delta^{\lfloor b \rfloor} Z_t$ is in the class $\mathcal{Z}_{\{b\}}$, a subset of the class $\mathcal{Z}_b$. ■

B.7 Proof of Theorem 7.3

B.7.1 The asymptotic distribution of $\hat{\beta}$

Let us first assume that $d_0, b_0 > 1/2$, so that we are in the non-stationary region and normalize β as $\beta = \beta_0 + \beta_{0\perp}\vartheta$. Let now set all the other parameters with the exception of ϑ to their true values. We obtain

$$\begin{aligned}\varepsilon_t(\theta_0\backslash\vartheta) =&(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^j\varepsilon_t + \Delta_+^{b_0}Y_t)-\\ &-\alpha_0(\beta_0' + \vartheta'\beta_{\perp 0}')\Delta_+^{-b_0}L_{b_0}(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^j\varepsilon_t + \Delta_+^{b_0}Y_t)-\\ &-\sum_{i=1}^{k}\Gamma_{0,i}L^j(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^j\varepsilon_t + \Delta_+^{b_0}Y_t).\end{aligned}$$

Differentiating with respect to ϑ, we find

$$D_\vartheta\varepsilon_t(\theta_0\backslash\vartheta) = -\alpha_0(d\theta)'\beta_{\perp 0}'\Delta_+^{-b_0}L_{b_0}(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^j\varepsilon_t + \Delta_+^{b_0}Y_t). \tag{34}$$

In this expression we keep the non-stationary fractional terms of higher order, which determine the asymptotic behavior of the score function, and find

$$D_\theta\varepsilon_t(\theta_0\backslash\vartheta)|_{\vartheta=\vartheta_0} = -\alpha_0(d\vartheta)'\beta_{\perp 0}'(\Delta_+^{-b_0} - 1)C_0\varepsilon_t,$$

where $d\vartheta$ denotes the increment on the coefficients ϑ. The score function then becomes

$$\begin{aligned}-2T^{-b_0-1/2}D_\vartheta\log\mathcal{L}(\theta_0) = \quad & tr\{(d\vartheta)'\beta_{\perp 0}'C_0T^{-b_0-1/2}\sum_{t=1}^{T}(\Delta_+^{-b_0} - 1)\varepsilon_t\varepsilon_t'\Omega_0^{-1}\alpha_0\}\\ \xrightarrow{d} \quad & tr\{(d\vartheta)'\beta_{\perp 0}'C_0\int_0^1 W_{b_0-1}(dW)'\Omega_0^{-1}\alpha_0\},\end{aligned}$$

where

$$S_{T,t} = T^{-b_0+1/2}(\Delta_+^{-b_0} - 1)\varepsilon_t \xrightarrow{d} W_{b_0-1}(u),$$

$$T^{-1}\sum_{t=1}^{T}S_{T,t}\varepsilon_t' = T^{-b_0-1/2}\sum_{t=1}^{T}(\Delta_+^{-b_0} - 1)\varepsilon_t\varepsilon_t' \xrightarrow{d} \int_0^1 W_{b_0-1}(dW)',$$

$$T^{-1}\sum_{t=1}^{T}S_{T,t}S_{T,t}' = T^{-2b_0}\sum_{t=1}^{T}\{(\Delta_+^{-b_0} - 1)\varepsilon_t\}\{(\Delta_+^{-b_0} - 1)\varepsilon_t\}' \xrightarrow{d} \int_0^1 W_{b_0-1}W_{b_0-1}'du.$$

The information matrix is found as the limit

$$T^{-2b_0} tr\{\Omega_0^{-1} \sum_{t=1}^{T} D_\vartheta \varepsilon_t(\theta_0) D_\vartheta \varepsilon_t(\theta_0)'\} \xrightarrow{d} tr\{\Omega_0^{-1}\alpha_0 (d\vartheta)'\beta_{\perp 0}' C_0 \int_0^1 W_{b_0-1} W_{b_0-1}' du \beta_{\perp 0}(d\vartheta)\alpha_0'\}.$$

Given that the estimator is consistent, we find that for all matrices $d\vartheta$

$$tr\{(d\vartheta)'\beta_{\perp 0}' C_0 T^{-1} \sum_t S_{T,t}\varepsilon_t' \Omega_0^{-1}\alpha_0'\} \approx -tr\{(d\vartheta)'\beta_{\perp 0}' C_0 T^{-1} \sum_t S_{T,t} S_{T,t}' C_0' \beta_{\perp 0}(\hat{\theta} - \theta_0)(\alpha_0' \Omega_0^{-1}\alpha_0)\}.$$

Hence

$$\begin{aligned}
T^{b_0}(\hat{\vartheta} - \vartheta_0) \simeq & [\beta_{\perp 0}' C_0 T^{-1} \sum_{t=1}^{T} S_{T,t} S_{T,t}' C_0' \beta_{\perp 0}]^{-1} \beta_{\perp 0}' C T^{-1} \sum_{t=1}^{T} S_{T,t}\varepsilon_t' \Omega_0^{-1}\alpha_0 (\alpha_0 \Omega_0^{-1}\alpha_0)^{-1} = \\
&= \left[\beta_{\perp 0}' C_0 \left(\int_0^1 W_{b_0-1} W_{b_0-1}' du\right) C_0' \beta_{\perp 0}\right]^{-1} \beta_{\perp 0}' C \int_0^1 W_{b_0-1}(dW)' \Omega_0^{-1}\alpha_0(\alpha_0 \Omega_0^{-1}\alpha_0)^{-1} = \\
&= \left[\int_0^1 F_0 F_0' du\right]^{-1} \int_0^1 F_0 (dG_0)' (\alpha_0' \Omega_0^{-1}\alpha_0)^{-1}
\end{aligned}$$

where $F_0 = \beta_{0\perp}' C_0 W_{b_0-1}$ and $G_0 = \alpha_0' \Omega_0^{-1} W$. When $b_0 < 1/2$, the right hand side of (34) is a stationary process because Δ^{-b_0} is applied to an $I(0)$ process. Hence, standard asymptotics applies in this case.

B.7.2 The asymptotic distribution of $\hat{d}$

Let now assume that all the parameters are set to their DGP values, with the exception of d. The error term is

$$\begin{aligned}
\varepsilon_t(\theta_0 \backslash d) = & \Delta_+^{d-d_0}(C_0\varepsilon_t + \sum_{j=1}^{\infty} \beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^j \varepsilon_t + \Delta_+^{b_0} Y_t) - \\
& -\alpha_0\beta_0' \Delta_+^{d-d_0}\Delta_+^{-b_0} L_{b_0}(C_0\varepsilon_t + \sum_{j=1}^{\infty} \beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^j \varepsilon_t + \Delta_+^{b_0} Y_t) - \\
& -\sum_{i=1}^{k} \Gamma_{i,0}\Delta_+^{d-d_0} L^j (C_0\varepsilon_t + \sum_{j=1}^{\infty} \beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^j \varepsilon_t + \Delta_+^{b_0} Y_t).
\end{aligned}$$

Exploiting that $\beta_0' C_0 = 0$, then it follows that

$$\begin{aligned}
\varepsilon_t(\theta_0 \backslash d) = & \Delta_+^{d-d_0}(C_0\varepsilon_t + \sum_{j=1}^{\infty} \beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^j \varepsilon_t + \Delta_+^{b_0} Y_t) - \\
& -\alpha_0\beta_0' \Delta_+^{d-d_0} L_{b_0}(Y_t) - \sum_{i=1}^{k} \Gamma_{i,0}\Delta_+^{d-d_0} L^j (C_0\varepsilon_t + \sum_{j=1}^{\infty} \beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^j \varepsilon_t + \Delta_+^{b_0} Y_t),
\end{aligned}$$

so that the non-stationary fractional terms disappear and the derivative $D_d\varepsilon_t(\theta_0)$ is stationary. By the martingale CLT the score $T^{-\frac{1}{2}}D_d \log \mathcal{L}(\theta_0) = T^{-\frac{1}{2}} tr\left\{\sum_{t=1}^{T} D_d\varepsilon_t(\theta_0)\varepsilon_t(\theta_0)'\Omega_0^{-1}\right\}$ is asymptotically Gaussian, and the information matrix is found as the limit of the outer product of the gradients, that is $T^{-1}tr\left\{\sum_{t=1}^{T} D_d\varepsilon_t(\theta_0)D_d\varepsilon_t(\theta_0)'\Omega_0^{-1}\right\}$. Thus the asymptotic distribution of $T^{\frac{1}{2}}(\hat{d}-d_0)$ is Gaussian.

B.7.3 The asymptotic distribution of $\hat{b}$

Let now assume that all the parameters are set to their DGP values, with the exception of b. The error term is

$$\begin{aligned}\varepsilon_t(\theta_0\backslash b) =&(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^{j}\varepsilon_t + \Delta_+^{b_0}Y_t)-\\ &-\alpha_0\beta_0'\Delta_+^{-b}L_b(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^{j}\varepsilon_t + \Delta_+^{b_0}Y_t)-\\ &-\sum_{i=1}^{k}\Gamma_{i,0}L^j(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^{j}\varepsilon_t + \Delta_+^{b_0}Y_t).\end{aligned}$$

Again, we exploit the fact that $\beta_0'C_0 = \beta_0'\beta_{\perp 0} = 0$ and we get

$$\begin{aligned}\varepsilon_t(\theta_0\backslash b) =&(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^{j}\varepsilon_t + \Delta_+^{b_0}Y_t)-\\ &-\alpha_0\beta_0'\Delta_+^{b_0-b}L_b(Y_t) - \sum_{i=1}^{k}\Gamma_{i,0}L^j(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^{j}\varepsilon_t + \Delta_+^{b_0}Y_t).\end{aligned}$$

Taking the derivative with respect to b, we find $D_b\varepsilon_t(\theta_0\backslash b)|_{b=b_0} = -\alpha_0\beta_0'D_b(\Delta_+^{-b-b_0})|_{b=b_0}Y_t$, so that $D_b\varepsilon_t(\theta_0\backslash b)$ is stationary and the asymptotic distribution of $\hat{b}$ is Gaussian. The information is found as the limit of $T^{-1}tr\left\{\sum_{t=1}^{T} D_b\varepsilon_t(\theta_0)D_b\varepsilon_t(\theta_0)'\Omega_0^{-1}\right\}$.

B.7.4 The asymptotic distribution of $\hat{\Gamma}_i, i = 1, \ldots, k$

Let now assume that all the parameters are set to their DGP values, with the exception of Γ_i. The error term is

$$\begin{aligned}\varepsilon_t(\theta_0\backslash b) =&(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^{j}\varepsilon_t + \Delta_+^{b_0}Y_t) - \alpha_0\beta_0'L_{b_0}(Y_t)-\\ &-\sum_{j\neq i}\Gamma_{j,0}L^j(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^{j}\varepsilon_t + \Delta_+^{b_0}Y_t)\\ &-\Gamma_iL^i(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^{j}\varepsilon_t + \Delta_+^{b_0}Y_t).\end{aligned}$$

Taking the derivative with respect to Γ_i we get

$$D_{\Gamma_i}\varepsilon_t(\theta_0\backslash\Gamma_i) = -(d\Gamma_i)(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^j\varepsilon_t + \Delta_+^{b_0}Y_t),$$

that is stationary and hence the asymptotic distribution of $\hat{\Gamma}_i$ is Gaussian. The score $T^{-\frac{1}{2}}D_{\Gamma_i}\log\mathcal{L}(\theta_0)$ is asymptotically Gaussian and the information is found as the limit of $T^{-1}tr\{\sum_{t=1}^{T}D_{\Gamma_i}\varepsilon_t(\theta_0)D_{\Gamma_i}\varepsilon_t(\theta_0)'\Omega_0^{-1}\}$.

B.7.5 The asymptotic distribution of $\hat{\alpha}$

Let now assume that all the parameters are set to their DGP values, with the exception of α. The error term is

$$\begin{aligned}\varepsilon_t(\theta_0\backslash\alpha) =&(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^j\varepsilon_t + \Delta_+^{b_0}Y_t)- \\ &-\alpha\beta_0'L_{b_0}Y_t - \sum_{j=1}^{k}\Gamma_{j,0}L^j(C_0\varepsilon_t + \sum_{j=1}^{\infty}\beta_{\perp 0}\Phi_{j0}\alpha_{\perp 0}\Delta_+^j\varepsilon_t + \Delta_+^{b_0}Y_t).\end{aligned}$$

Taking the derivative with respect to α we get

$$D_\alpha\varepsilon_t(\theta_0\backslash\alpha) = -(d\alpha)\beta_0'LY_t.$$

Hence $D_\alpha\varepsilon_t(\theta_0\backslash\alpha)$ is stationary and the asymptotic distribution of $\hat{\alpha}$ is therefore Gaussian. The score $T^{-\frac{1}{2}}D_\alpha\log\mathcal{L}(\theta_0)$ is asymptotically Gaussian and the information matrix is found as the limit of $T^{-1}tr\{T^{-1}\sum_{t=1}^{T}D_\alpha\varepsilon_t(\theta_0)D_\alpha\varepsilon_t(\theta_0)'\Omega_0^{-1}\}$.

B.7.6 Asymptotic covariance of $\hat{\theta}\backslash\hat{\beta}$

The off diagonal elements of the asymptotic information matrix of $\hat{\theta}\backslash\hat{\beta}$ is given by

$$tr\{T^{-1}\sum_{t=1}^{T}D_{\Gamma_i}(\theta_0)\varepsilon_t D_{\Gamma_j}\varepsilon_t(\theta_0)\Omega_0^{-1}\}, tr\{T^{-1}\sum_{t=1}^{T}D_\alpha(\theta_0)\varepsilon_t D_{\Gamma_i}\varepsilon_t(\theta_0)\Omega_0^{-1}\},$$

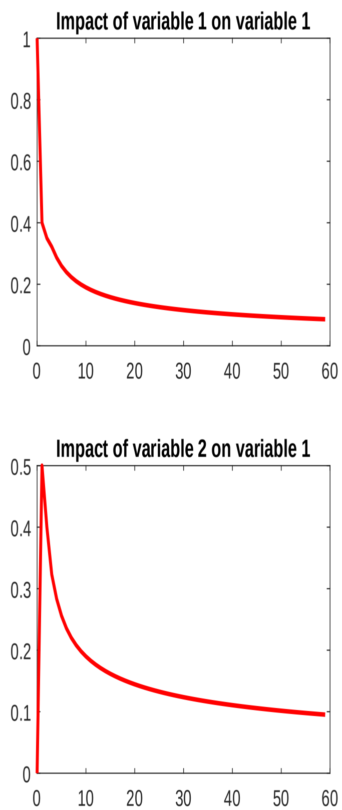
$$tr\{T^{-1}\sum_{t=1}^{T}D_d(\theta_0)\varepsilon_t D_{\Gamma_i}\varepsilon_t(\theta_0)\Omega_0^{-1}\}, tr\{T^{-1}\sum_{t=1}^{T}D_b(\theta_0)\varepsilon_t D_{\Gamma_i}\varepsilon_t(\theta_0)\Omega_0^{-1}\},$$

$$tr\{T^{-1}\sum_{t=1}^{T}D_\alpha(\theta_0)\varepsilon_t D_d\varepsilon_t(\theta_0)\Omega_0^{-1}\}, tr\{T^{-1}\sum_{t=1}^{T}D_\alpha(\theta_0)\varepsilon_t D_b\varepsilon_t(\theta_0)\Omega_0^{-1}\},$$

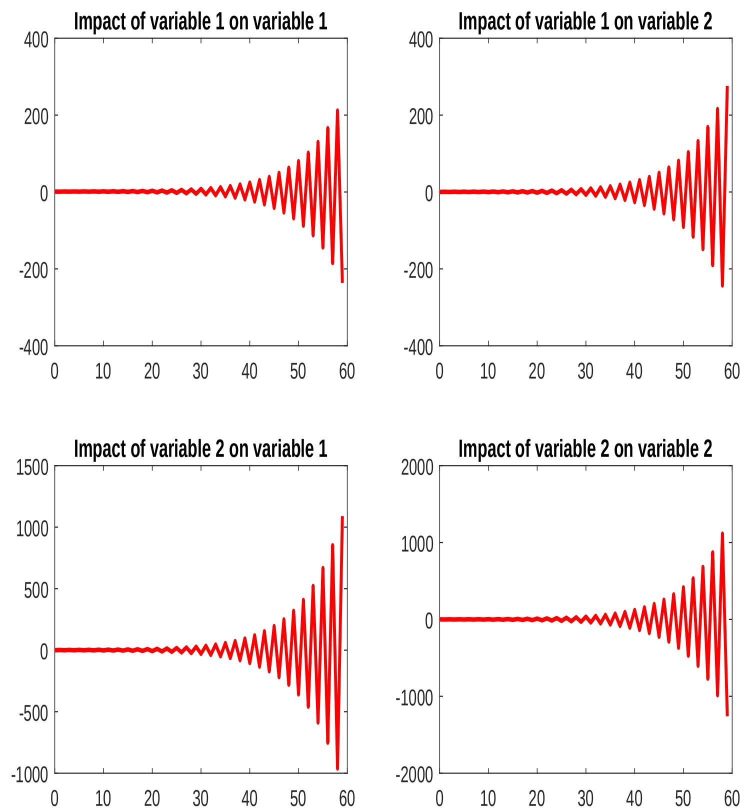
which are product of stationary components and have a finite limit. Hence the asymptotic distribution of

$$T^{\frac{1}{2}}vec(\hat{d}-d_0, \hat{b}-b_0, \hat{\Gamma}-\Gamma_0, \hat{\alpha}-\alpha_0),$$

where $\Gamma = [\Gamma_1 : \ldots : \Gamma_k]$ is multivariate Gaussian and it is independent with respect to $\hat{\beta}$, see Lemma 10 in Johansen and Nielsen (2010). ■

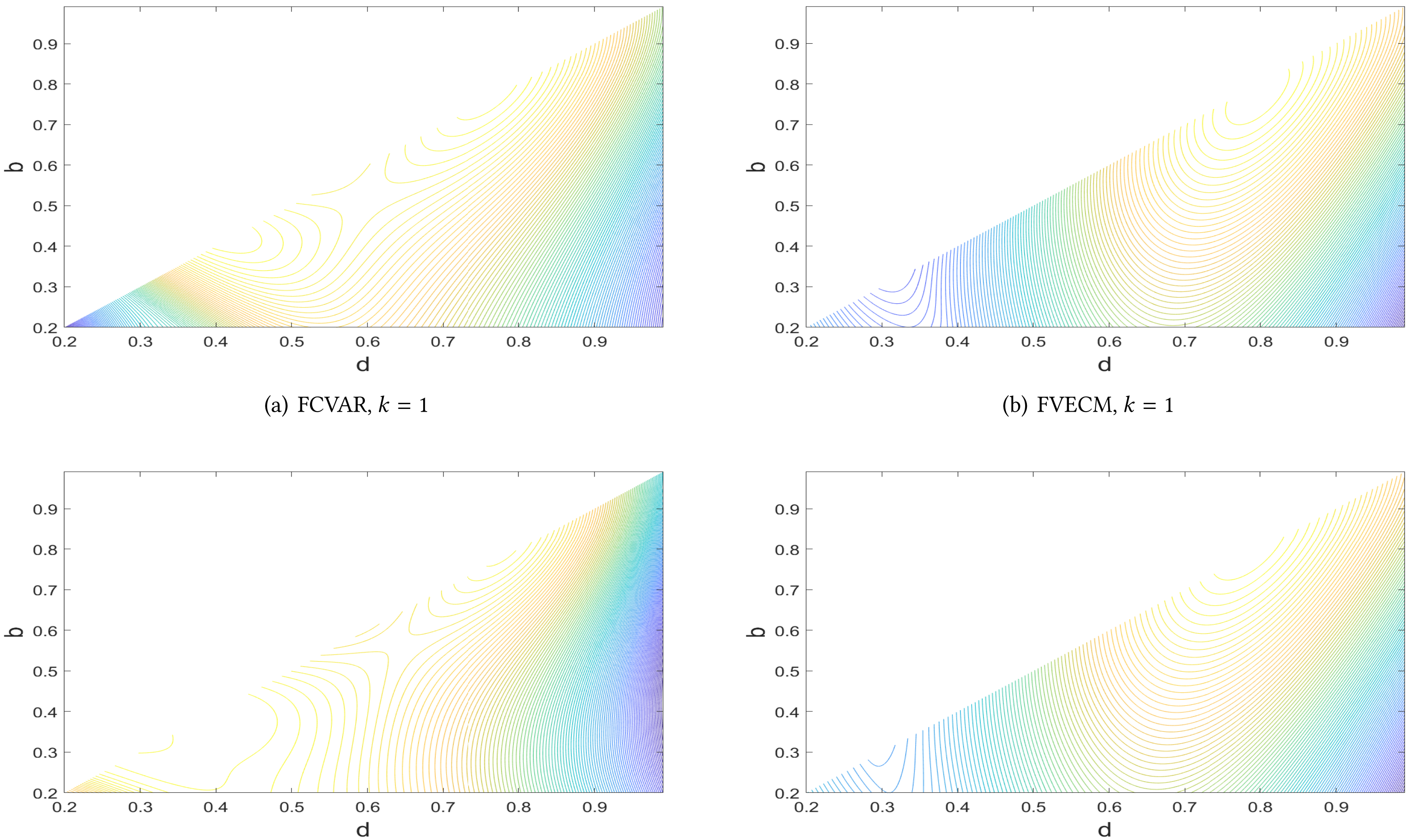


(a) Stable



(b) Unstable

Figure 1: Impulse response function for the FVECM$_{d,b}$ when $p = 2$, $r = 1$ and $k = 1$. The left panel is generated with $d = 0.6$, $b = 0.4$, $\beta = [1, -0.8]'$, $\alpha = [-0.4; 0.3]$, $\Gamma_1 = \begin{bmatrix} 0.2 & -0.1 \\ 0.2 & 0.4 \end{bmatrix}$ with $\mathcal{N} = 0$. The right panel is generated with $d = 1.1$, $b = 0.8$, $\beta = [1, -1.2]'$, $\alpha = [-0.6, 1.7]'$, $\Gamma_1 = \begin{bmatrix} 0.3 & -0.2 \\ -0.1 & 0.3 \end{bmatrix}$ with $\mathcal{N} = 1$.



(a) FCVAR, $k = 1$

(b) FVECM, $k = 1$

(c) FCVAR, $k = 2$

(d) FVECM, $k = 2$

Figure 2: The figure reports the contour plot of the values of the function $\ell(\psi)$ for different combinations of $d \in [0.2, 0.99]$ (x-axis) and $b \in [0.2, 0.99]$ (y-axis). The observations from the DGP are generated with $k_0 = 0$ lags and both the $\text{FCVAR}_{d,b}$ and $\text{FVECM}_{d,b}$ with $k = 1$ and $k = 2$ lags are estimated. The parameters of the DGP are $d_0 = b_0 = 0.8$, $\beta_0 = [1, -1]'$, $\alpha_0 = [-0.5, 0.5]'$. The empty area is associated with values of $b > d$ which are ruled out by assumption.