



# Long Memory Through Marginalization of Large Systems and Hidden Cross-Section Dependence

Guillaume Chevillon, Alain Hecq, Sébastien Laurent

► **To cite this version:**

Guillaume Chevillon, Alain Hecq, Sébastien Laurent. Long Memory Through Marginalization of Large Systems and Hidden Cross-Section Dependence. ESSEC Working paper. Document de Recherche ESSEC / Centre de recherche de l'ESSEC. ISSN : 1291-9616. WP 1507. 2015. <hal-01158524>

**HAL Id: hal-01158524**

**<https://hal-essec.archives-ouvertes.fr/hal-01158524>**

Submitted on 1 Jun 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

---

# *Long Memory Through Marginalization of Large Systems and Hidden Cross-Section Dependence*

---

RESEARCH CENTER  
ESSEC Working Paper 1507

2015

Guillaume Chevillon

Alain Hecq

Sébastien Laurent

# Long Memory Through Marginalization of Large Systems and Hidden Cross-Section Dependence

This Version: May 7, 2015

Guillaume Chevillon<sup>a</sup>, Alain Hecq<sup>b</sup>, Sébastien Laurent<sup>c</sup>

<sup>a</sup>*ESSEC Business School & CREST, France*

<sup>b</sup>*Department of Quantitative Economics, Maastricht University, The Netherlands*

<sup>c</sup>*Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS, Aix-Marseille Graduate School of Management – IAE, France*

---

## Abstract

This paper shows that large dimensional vector autoregressive (VAR) models of finite order can generate long memory in the marginalized univariate series. We derive high-level assumptions under which the final equation representation of a VAR(1) leads to univariate fractional white noises and verify the validity of these assumptions for two specific models. We consider the implications of our findings for the variances of asset returns where the so-called golden-rule of realized variances states that they tend always to exhibit fractional integration of a degree close to 0.4.

*Keywords:* Long memory, Vector Autoregressive Model, Marginalization, Final Equation Representation, Volatility.

*JEL:* C10, C32, C58.

---

---

<sup>\*</sup>We would like to thank participants at the 6th French Econometrics Conference celebrating Christian Gouriéroux's contribution to econometrics, the 2014 Econometrics Conference at the Oxford Martin School, CFE 2013, the 15th OxMetrics User Conference at CASS Business School, the 2015 Symposium of the SNDE, Aix-Marseille University and CREATES seminars, and in particular Karim Abadir, Richard Baillie, David Hendry, Jurgen Doornik, Sophocles Mavroeidis, Bent Nielsen, Anders Rabhek and Paolo Zaffaroni.

*Email addresses:* chevillon@essec.edu (Guillaume Chevillon), a.hecq@maastrichtuniversity.nl (Alain Hecq), sebastien.laurent@univ-amu.fr (Sébastien Laurent)

## 1. Introduction and motivations

Long memory is commonly observed in economics and finance, dating back at least to Smith (1938), Cox and Townsend (1947) and Granger (1966), but its origin is unclear as argued by Cox (2014). Müller and Watson (2008) show that this is probably due to the fact that very large samples are needed to discriminate between the various models generating strong dependence at low frequencies. Hence several competing models of long range dependence have been proposed in the literature. For a covariance stationary process  $z_t$ , long memory of degree  $d$  is often defined, as in Beran (1994) or Baillie (1996), through the behavior of its spectral density  $f_z(\omega)$  about the origin:  $f_z(\omega) \sim c_f \omega^{-2d}$ , as  $\omega \rightarrow 0^+$ , for some positive  $c_f$ . Since Granger and Joyeux (1980), fractional integration of order  $d$ , denoted  $I(d)$ , has proved the most pervasive example of long memory processes in econometrics. When  $d < 1$ , the process is mean reverting (in the sense of Campbell and Mankiw, 1987, that the impulse response function to fundamental innovations converges to zero, see Cheung and Lai, 1993). Moreover,  $I(d)$  processes admit a covariance stationary representation when  $d \in (-1/2, 1/2)$ , and are non-stationary if  $d \geq 1/2$ . Long range dependence, or long memory, arises when the degree of fractional integration is positive,  $d > 0$ . When  $d \geq 1/2$ , the process is nonstationary, yet the spectral density characterization can still be used as the limit of the sample periodogram, see Solo (1992). The prototypical example of an  $I(d)$  process is the fractional white noise  $z_t = (1 - L)^{-d} \epsilon_t$ , where  $L$  denotes the lag operator and  $\epsilon_t$  is a white noise sequence. The class of fractionally integrated processes extends to cases where  $\epsilon_t$  admits a covariance stationary ARMA representation.

To the best of our knowledge five reasons have been put forward in the literature so far to explain the presence of long range dependence: (i) aggregation across heterogeneous series, frequencies or economic agents (Granger 1980, Chambers, 1998, and inter alia Abadir and Talmain, 2002, Zaffaroni, 2004, Lieberman and Phillips, 2008 and Altissimo, Mojon and Zaffaroni, 2009); (ii) linear modeling of a nonlinear underlying process (e.g. Davidson and Sibbertsen, 2005, Miller and Park, 2010); (iii) structural change (Parke, 1999, Diebold and Inoue, 2001, Gouriéroux and Jasiak, 2001, Perron and Qu, 2007); (iv) learning (bounded rationality) by economic agents in forward looking models of expectations (Chevillon and Mavroeidis, 2013) and (v) network effects (Schennach, 2013).

The contribution of this paper is to show that long memory can also be the result of the marginalization of a large dimensional multivariate system, hence caused by hidden dependence across variables within a system. More specifically, we provide an asymptotic parametric framework under which the variables entering an  $n$ -dimensional vector autoregressive process of finite order (here a VAR(1)) can be modelled individually as fractional white noises as  $n \rightarrow \infty$ . Long memory may therefore be a feature of univariate or low dimensional models that vanishes when considering larger systems. The source of long memory identified here differs distinctly from the five sources

listed above. In particular, it does not rely on any aggregation mechanism à la Granger (1980).

The intuition behind our theoretical result is the following. We consider a simple VAR(1) model  $\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t$ , where  $(\mathbf{A}_n)$  denotes a sequence of  $n$ -dimensional square Toeplitz matrices.<sup>1</sup> We use the final equation representation of this model (as proposed by Zellner and Palm, 1974, 2004) to derive the spectral density  $f_{n,x_j}(\omega)$  of any of the marginalized processes  $x_{jt}$  belonging to  $\mathbf{x}_t$  (for  $j = 1, \dots, n$ ). To prove that  $f_{n,x_j}(\omega)$  converges to the spectral density of a long memory process of order  $\delta \in (0, 1)$ , we introduce three high-level assumptions concerning  $(\mathbf{A}_n)$ . Under these assumptions,  $f_{n,x_j}(\omega)$  is asymptotically equivalent to the ratio of  $|\det(\mathbf{I}_{n-1} - \mathbf{A}_{n-1}e^{-i\omega})|^2$  over  $|\det(\mathbf{I}_n - \mathbf{A}_n e^{-i\omega})|^2$ . We parameterize  $\mathbf{A}_n$  by defining a scalar sequence  $(\delta_n)$  with  $\delta_n \in (0, 1)$  such that  $\lim_{n \rightarrow \infty} \delta_n = \delta$ , and a circulant matrix  $\mathbf{C}_n$  such that  $\det(\mathbf{I}_n - \mathbf{A}_n e^{-i\omega}) \sim \det(\mathbf{I}_n - \mathbf{C}_n e^{-i\omega})$  as  $n \rightarrow \infty$ .  $\mathbf{C}_n$  is assumed to possess about a fraction  $[n\delta_n]$  of unit eigenvalues ( $[\cdot]$  denotes the integer part) and  $n - [n\delta_n]$  zero eigenvalues. Hence, as  $n \rightarrow \infty$ ,  $\det(\mathbf{I}_n - \mathbf{C}_n e^{-i\omega}) \sim (1 - e^{-i\omega})^{[n\delta_n]}$ . We then use the *first theorem* of Szegö (1915) to prove that under our three high-level assumptions,  $f_{n,x_j}(\omega) \rightarrow |1 - e^{-i\omega}|^{-2\delta} \sigma_{\epsilon_j}^2$  for  $\omega \neq 0$ .

We then show that these high level assumptions are satisfied for at least two specific examples of VAR(1) models. In the first parameterization,  $\mathbf{A}_n$  denotes a Toeplitz matrix with diagonal elements converging to  $\delta = 1/2$  as  $n \rightarrow \infty$ , and with vanishing off-diagonal elements. Importantly, the off-diagonal elements decrease at an  $O(n^{-1})$  rate and the sum of each row equals 1 at all  $n$ . We show that as  $n \rightarrow \infty$ , each series  $x_{jt}$  of this system behaves as an ARFIMA(0, 1/2, 0). In the second example, we consider a similar setting but with limiting value  $\delta \in (0, 1)$  on the main diagonal of  $\mathbf{A}_n$  and with the additional assumption that one innovation (say  $\epsilon_{jt}$ ) dominates the others in terms of magnitude. In this case, we prove that the dominant series  $j$  follows an ARFIMA(0,  $\delta$ , 0) for  $\delta \in (0, 1)$ . Our results exemplify that vanishing interaction coefficients in a multivariate system can give rise to long memory in individual series.

The reason why we refer to this phenomena as “hidden cross-section dependence” is twofold. First, long memory appears through the marginalization mechanism and therefore in the univariate series or by extension, when estimating the model on a small subpart of a large system. The cross-section dependence appearing in the large system is therefore hidden in the univariate models. Second, because the off-diagonal elements of the VAR(1) are so small that, in finite samples, it is likely to be indistinguishable from a diagonal VAR(1) on the sole basis of the parameter estimates. The hidden dependencies induce modeling issues that were pointed out, inter alia, in Hendry (2009).

Our paper sheds some new light on the reasons why asset return variances exhibit long memory and in particular why the estimated degree of fractional integration of univariate realized variance

---

<sup>1</sup>The class of Toeplitz matrices is chosen for the technical reason that we use in our proofs the well-established First Theorem of Szegö (1915). We show in the section presenting our analytical results how this assumption can be relaxed.

series is generally about 0.4 (the so-called golden-rule of realized volatility, see Andersen et al., 2001 and Lieberman and Phillips, 2008). The presence of long memory in realized variances and its homogeneity across series is therefore likely due to the marginalization of a large system. We illustrate this finding by considering the  $\log(\text{MedRV})$  of 49 US stocks, where  $\text{MedRV}$  is a non-parametric robust to jumps estimator of the integrated variance (computed in our case on 5-minute returns), recently proposed by Andersen, Dobrev, and Schaumburg (2012).

The rest of this paper is organized as follows. Section 2 provides our main theoretical results. Section 3 presents some Monte Carlo simulations and compares them with some empirical evidences on  $\log(\text{MedRV})$ . Finally, Section 4 concludes. The appendix contains all the proofs.

In the paper, we use the following notation.  $\mathbb{R}$  and  $\mathbb{C}$  denote the sets of, respectively, real and complex scalars, and  $\mathbb{R}^* = \{x \in \mathbb{R}, x \neq 0\}$ . For any  $x \in \mathbb{R}$ ,  $\lfloor x \rfloor$  and  $\lceil x \rceil$  denote the floor and ceiling of  $x$ . For  $z \in \mathbb{C}$ ,  $|z|$  is the modulus of  $z$ ,  $\bar{z}$  its conjugate,  $\text{Re}(z)$  and  $\text{Im}(z)$  its real and imaginary parts; we often use the notation  $i = \sqrt{-1}$ . For any sequences  $a_n$ ,  $b_n$  and  $c_n$  of real-valued scalars  $a_n = O(b_n)$ ,  $b_n = o(c_n)$ , and  $a_n \sim b_n$  imply, respectively, that as  $n \rightarrow \infty$ ,  $|a_n/b_n|$  is bounded,  $b_n/c_n \rightarrow 0$ , and  $a_n/b_n \rightarrow 1$ . For any complex-valued square matrix  $\mathbf{A}$ ,  $\det(\mathbf{A})$  is the determinant of  $\mathbf{A}$ ,  $\text{tr}(\mathbf{A})$  its trace,  $\tilde{\mathbf{A}}$  its adjugate matrix,  $\bar{\mathbf{A}}'$  its conjugate transpose and  $|\mathbf{A}| = \text{tr}(\bar{\mathbf{A}}' \mathbf{A})^{1/2}$  its weak norm. For two sequences  $(\mathbf{A}_n)$  and  $(\mathbf{B}_n)$  of square matrices with bounded maximal eigenvalues,  $\mathbf{A}_n \sim \mathbf{B}_n$  means that  $|\mathbf{A}_n - \mathbf{B}_n| \rightarrow 0$  as  $n \rightarrow \infty$ .  $1_{\{\cdot\}}$  is the indicator function which takes value one if  $\{\cdot\}$  is true and zero otherwise.

## 2. Theory

In this section, we provide an analytical argument that shows that long memory can arise through the marginalization of a multivariate process. We first provide a set-up that introduces high-level assumptions and delineates the analysis that leads to our results. Our theoretical argument draws upon three existing literatures: those of long memory time series processes, Final Equation Representations (FER) of Zellner and Palm (1974), and large dimensional Toeplitz matrices (see, e.g., Gray, 2006).

In a second part, the section presents two parametric representations where the high-level assumptions are satisfied and long memory arises in the marginalized representation.

### 2.1. Set-up and main results

We first consider a simple model where the  $n$ -vector  $\mathbf{x}_t = (x_{1t}, \dots, x_{nt})'$  admits a vector autoregressive, VAR(1), representation:

$$\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t, \quad t = 1, \dots, T \quad (1a)$$

$$\epsilon_t = (\epsilon_{1t}, \dots, \epsilon_{nt})' \sim \text{IID}(\mathbf{0}, \Omega_\epsilon), \quad (1b)$$

where  $\Omega_\epsilon$  is a diagonal matrix with diagonal  $\sigma_\epsilon^2 = (\sigma_{\epsilon_1}^2, \dots, \sigma_{\epsilon_n}^2)$  such that  $\sigma_{\epsilon_j} \neq 0$  for  $j = 1, \dots, n$ . Hence shocks  $\epsilon_{jt}$  and  $\epsilon_{j't}$  are independent for  $j \neq j'$ ,  $j, j' = 1, \dots, n$ . Independence is not necessary for our results and could be relaxed, but it simplifies the exposition. We provide an extension at the end of the section.

Final equation representations (FER) were studied by Zellner and Palm (1974, 2004) to show how the elements of vector processes can be marginalized to yield univariate ARMA representations; see also Cubadda, Hecq and Palm (2009) in the context of factor models and Hecq, Laurent and Palm (2012) for an application to multivariate volatility processes. The FER of model (1) is

$$\det(\mathbf{B}_n(L)) \mathbf{x}_t = \widetilde{\mathbf{B}_n(L)} \epsilon_t, \quad (2)$$

where  $\mathbf{B}_n(L) = \mathbf{I}_n - \mathbf{A}_n L$ , with  $L$  the lag operator. If  $\mathbf{A}_n$  admits unitary eigenvalues, we implicitly assume that  $\epsilon_t = 0$  for  $t < 0$  and  $\mathbf{x}_0 = 0$ .<sup>2</sup> Expression (2) shows that element  $x_{jt}$ , obtained by marginalizing the  $n$ -dimensional VAR(1), admits a finite ARMA( $n, n-1$ ) representation with a common AR lag polynomial. Hence, as  $n \rightarrow \infty$ , the univariate process  $x_{jt}$  without roots cancellation follows an ARMA( $\infty, \infty$ ). For clarity of the exposition, consider the following trivariate example:

**Example.**  $\mathbf{x}_t$  is a trivariate VAR(1) specified as follows:

$$\begin{bmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{bmatrix} = \begin{bmatrix} a & b & 0 \\ b & a & b \\ 0 & b & a \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ x_{3t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{2t} \end{bmatrix},$$

where  $E[\epsilon_{jt}\epsilon_{kt}] = 0$  for  $j \neq k$ . The FER of  $\mathbf{x}_t$  is  $\det(\mathbf{B}_n(L)) \mathbf{x}_t = \widetilde{\mathbf{B}_n(L)} \epsilon_t$ , where

$$\det(\mathbf{B}_n(L)) = \frac{1}{(a^2 - 2b^2)} (1 - aL) \left(1 - (a - \sqrt{2}b)L\right) \left(1 - (a + \sqrt{2}b)L\right),$$

$$\widetilde{\mathbf{B}_n(L)} = \begin{bmatrix} (1 - aL)^2 - b^2L^2 & bL(1 - aL) & b^2L^2 \\ bL(1 - aL) & (1 - aL)^2 & bL(1 - aL) \\ b^2L^2 & bL(1 - aL) & (1 - aL)^2 - b^2L^2 \end{bmatrix}.$$

Hence  $x_{jt} \sim \text{ARMA}(3, 2)$  for  $j = 1, 2, 3$  when  $a \neq 0$  and  $b \neq 0$ , while  $x_{jt} \sim \text{ARMA}(2, 1)$  when  $a = 0$  and  $b \neq 0$  and  $x_{jt} \sim \text{AR}(1)$  when  $a \neq 0$  and  $b = 0$ .

Denote the  $j$ th,  $1 \leq j \leq n$ , row of  $\widetilde{\mathbf{B}_n(L)}$  by  $\widetilde{\mathbf{B}_n(L)}_j$ , such that

$$\widetilde{\mathbf{B}_n(L)}_j = \left[ \widetilde{\mathbf{B}_n(L)}_{j1} \quad \widetilde{\mathbf{B}_n(L)}_{j2} \quad \dots \quad \widetilde{\mathbf{B}_n(L)}_{jn} \right],$$

---

<sup>2</sup>In our results below, we avoid dwelling on the issue of finite vs. infinite history, in relation to type I and type II fractional Brownian motions, see Marinucci and Robinson (1999) and Davidson and Hashimzade (2009). Hence, we implicitly consider that the date of interest,  $t$  is always larger than  $n$  so the spectral density is defined.

hence  $x_{jt}$  admits the ARMA representation

$$\begin{aligned}\det(\mathbf{B}_n(L))x_{jt} &= \widetilde{\mathbf{B}}_n(L)_j \epsilon_t \\ &= \sum_{k=1}^n \widetilde{\mathbf{B}}_{jk}(L) \epsilon_{kt}.\end{aligned}$$

Therefore, the spectral density  $f_{x_j}$  of  $x_{jt}$  satisfies

$$f_{n,x_j}(\omega) = \sum_{k=1}^n \left| \frac{\widetilde{\mathbf{B}}_n(e^{-i\omega})_{jk}}{\det(\mathbf{B}_n(e^{-i\omega}))} \right|^2 \sigma_{\epsilon_k}^2 \quad (3)$$

$\forall \omega \in \mathbb{R}$  such that  $\det(\mathbf{B}_n(e^{-i\omega})) \neq 0$ .

Expression (3) constitutes the basis of our theoretical argument. In the remainder of the section, we delineate three high-level assumptions which together imply that expression (3) tends, as  $n \rightarrow \infty$ , to the spectral density of a fractional white noise.

The first assumption concerns the summation on the right-hand side of (3) defining the spectral density of  $x_{jt}$ .<sup>3</sup>

**Assumption B.** *There exists  $j \in \mathbb{N}$  for which the parameters of the VAR(1) model (1) satisfy  $\forall \omega \in \mathbb{R}^*$ , and as  $n \rightarrow \infty$ ,*

- (i)  $\sum_{k=1}^n \left| \frac{\widetilde{\mathbf{B}}_n(e^{-i\omega})_{jk}}{\det(\mathbf{B}_n(e^{-i\omega}))} \right|^2 \sigma_{\epsilon_k}^2 = \left| \frac{\widetilde{\mathbf{B}}_n(e^{-i\omega})_{jj}}{\det(\mathbf{B}_n(e^{-i\omega}))} \right|^2 \sigma_{\epsilon_j}^2 + o(1)$ ;
- (ii)  $\widetilde{\mathbf{B}}_n(e^{-i\omega})_{jj} = \det(\mathbf{B}_{n-1}(e^{-i\omega})) + o(1)$  where  $\sigma_{\epsilon_j}^2 > 0$ .

Under Assumption B(i) the summation on the right-hand side of (3) reduces to its  $j$ th element. We consider in the following how this can arise for some specific parametric assumptions for the sequence of matrices  $\mathbf{A}_n$ . In Section 2.2.2, we also consider the situation where one innovation  $\epsilon_{jt}$  dominates the others in terms of magnitude (i.e., variance). Assumption B(ii) implies in particular that, for  $\omega \neq 0$  and as  $n \rightarrow \infty$ ,

$$f_{n,x_j}(\omega) = \left| \frac{\det(\mathbf{I}_{n-1} - \mathbf{A}_{n-1}e^{-i\omega})}{\det(\mathbf{I}_n - \mathbf{A}_n e^{-i\omega})} \right|^2 \sigma_{\epsilon_j}^2 + o(1). \quad (4)$$

In (4), the spectral density of one element  $x_{jt}$  with a vector of dimension  $n$  is asymptotically expressed as a ratio of two determinants. Our main argument lies in assuming parametric expressions for  $\mathbf{A}_n$  allowing the use of a theorem concerning ratios of determinants. For this reason, we assume that  $\mathbf{A}_n$  can be expressed as functions of a sequence of Toeplitz matrices  $\mathbf{T}_n$ :

$$\mathbf{T}_n = \begin{bmatrix} t_0^{(n)} & t_1^{(n)} & \cdots & t_{n-1}^{(n)} \\ t_{-1}^{(n)} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_1^{(n)} \\ t_{-(n-1)}^{(n)} & \cdots & t_{-1}^{(n)} & t_0^{(n)} \end{bmatrix}.$$

<sup>3</sup>We denote this assumption by B to emphasize that it concerns the sequence  $\mathbf{B}_n$ . We follow the same logic for the next two assumptions.



We make the following assumption about  $\mathbf{T}_n$ .

**Assumption T.**

(i) There exists a bounded function  $g(\cdot, \cdot)$ , defined on  $(0, 1) \times \{\zeta \in \mathbb{C}, |\zeta| \leq 1\}$  which is continuous with respect to its first argument and piecewise continuous with respect to its second argument, and such that  $\forall (d, z) \in (0, 1) \times \{\zeta \in \mathbb{C}, |\zeta| \leq 1\}$ ,

$$\frac{1}{2\pi} \int_0^{2\pi} \log(1 - g(d, e^{i\omega})z) d\omega = -d \log(1 - z);$$

(ii)  $\forall d \in (0, 1)$ ,  $t_{d,k} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} g(d, e^{i\frac{2\pi j}{n}}) e^{-2i\pi jk/n}$  satisfies  $\sum_{k=-\infty}^{\infty} |t_{d,k}| < \infty$  so  $\mathbf{T}_{d,n}$ , the  $n$ -dimensional Toeplitz matrix with entries  $t_{d,j-i}$ , belongs to the Wiener class;

(iii) There exists a convergent sequence  $\delta_n \in (0, 1)^{\mathbb{N}} \rightarrow \delta$  as  $n \rightarrow \infty$  such that

$$t_k^{(n)} = \frac{1}{n} \sum_{j=0}^{n-1} g(\delta_n, e^{i\frac{2\pi j}{n}}) e^{-2i\pi jk/n}.$$

For all  $d \in (0, 1)$ , the partial function  $g_{\mathbf{T}_d}(\cdot) = g(d, \cdot)$ , such that  $g_{\mathbf{T}_d}(z) = \sum_{k=-\infty}^{\infty} t_{d,k} z^k$ , is called the *symbol* of the matrices  $(\mathbf{T}_{d,n})$ . The function  $\omega \in \mathbb{R} : \omega \rightarrow g_{\mathbf{T}_d}(e^{i\omega})$  is generally referred to as “spectral density” of  $\mathbf{T}_{d,n}$  but to avoid confusion with the spectral density of the processes  $x_{jt}$ , we do not use this terminology. Yet with a slight abuse of terminology, we refer to  $g_{\mathbf{T}_d}(e^{i\omega})$  as the function  $\omega \rightarrow g(d, e^{i\omega})$ . We discuss the details of this assumption in the proof. Notice though that, under Assumption T, the entries of the sequence  $(\mathbf{T}_{d,n})$  do not depend on  $n$ ; only the dimension of  $\mathbf{T}_{d,n}$  does. The identity matrix  $\mathbf{I}_n$  is also Toeplitz with symbol  $g_{\mathbf{I}}(\cdot) = 1$ , hence for  $|z| < 1$  the matrix  $\mathbf{I}_n - \mathbf{T}_n z$  is Toeplitz with symbol  $1 - g_{\mathbf{T}_d}(\cdot)z$ .

Assumption T ensures that the First Theorem of Szegö (1915) holds for  $\mathbf{I}_n - \mathbf{T}_{d,n}z$ . It states that as  $n \rightarrow \infty$ ,

$$\frac{\det(\mathbf{I}_{n-1} - \mathbf{T}_{d,n-1}z)}{\det(\mathbf{I}_n - \mathbf{T}_{d,n}z)} \rightarrow \exp \left\{ \frac{1}{2\pi} \int_0^{2\pi} \log(1 - g_{\mathbf{T}_d}(e^{i\omega})z) d\omega \right\}. \quad (5)$$

Under Assumption T(i) the limit above equals  $(1 - z)^{-d}$ .

We make one last high-level assumption, concerning now the parameters of the VAR(1) and how they relate to the Toeplitz structure defined above.

**Assumption A.** There exists a sequence of real matrices  $(\mathbf{A}_n)$  such that for  $\omega \in \mathbb{R}^*$  and as  $n \rightarrow \infty$ ,

$$\frac{\det(\mathbf{I}_{n-1} - \mathbf{A}_{n-1}e^{-i\omega})}{\det(\mathbf{I}_n - \mathbf{A}_n e^{-i\omega})} \sim \frac{\det(\mathbf{I}_{n-1} - \mathbf{T}_{n-1}e^{-i\omega})}{\det(\mathbf{I}_n - \mathbf{T}_n e^{-i\omega})},$$

where  $\mathbf{T}_n$  is defined in Assumption T.

The three high-level assumptions (i.e., Assumptions B, T and A) allow to prove the following theorem.

**Theorem 1.** *Let the real vector  $\mathbf{x}_t$  of dimension  $n$  be defined by the VAR(1) model (1). Under Assumptions B, T and A, there exist  $j \in \mathbb{N}$  and  $\delta \in (0, 1)$  such that the spectral density of element  $x_{jt}$  satisfies, for  $\omega \in \mathbb{R}^*$  and as  $n \rightarrow \infty$ ,*

$$f_{n,x_j}(\omega) \rightarrow |1 - e^{-i\omega}|^{-2\delta} \sigma_{\epsilon_j}^2.$$

*The convergence is uniform when  $\omega$  is restricted to closed subsets of  $\mathbb{R}^*$ .*

In Theorem 1, the spectral density of the marginalized univariate process  $x_{jt}$  tends to that of an  $I(\delta)$  fractional Brownian motion as the dimension of the system  $n \rightarrow \infty$ . Hence individual series may asymptotically (as the cross-section dimension increases) exhibit long memory although the vector process itself does not. In the theorem, the asymptotic spectral density only depends on one innovation through its variance  $\sigma_{\epsilon_j}^2$ , the others do not matter (and we show below an example where they disappear, i.e.,  $\sigma_{\epsilon_k}^2 \rightarrow 0$  for  $k \neq j$ ). Hence this is distinctly different from the example of Granger (1980, Section 4) where long memory arises in a vector process from the aggregation of moving averages; in fact our Assumption B(i) precludes it.

The set-up above can be generalized easily to transformations  $\mathbf{A}_n \rightarrow \mathbf{V}_n \mathbf{A}_n \mathbf{V}_n^{-1}$ , where  $\mathbf{V}_n$  is a non-singular  $n$ -dimensional matrix. Indeed,

$$\det(\mathbf{I}_n - \mathbf{V}_n \mathbf{A}_n \mathbf{V}_n^{-1} z) = \det[\mathbf{V}_n (\mathbf{I}_n - \mathbf{A}_n z) \mathbf{V}_n^{-1}] = \det(\mathbf{I}_n - \mathbf{A}_n z)$$

and the adjugate of  $\mathbf{I}_n - \mathbf{V}_n \mathbf{A}_n \mathbf{V}_n^{-1} z$  is  $\mathbf{V}_n (\widetilde{\mathbf{I}_n - \mathbf{A}_n z}) \mathbf{V}_n^{-1}$ . Hence Assumptions A and B apply to  $\mathbf{V}_n \mathbf{A}_n \mathbf{V}_n^{-1}$  if they hold for  $\mathbf{A}_n$ . It follows that expression (1b) can be relaxed to cases where  $\Omega_\epsilon$  is non-diagonal yet positive definite, by choosing for  $\mathbf{V}_n$  the matrix containing its orthonormal eigenvectors. Also  $\mathbf{V}_n \mathbf{A}_n \mathbf{V}_n^{-1}$  is not necessarily Toeplitz so our result is quite general.

In the following subsection, we provide examples of primitive conditions to impose on the parameters of the VAR(1) model for Theorem 1 to hold.

## 2.2. Two examples

Here we provide two parametric examples of models satisfying Assumptions B, T and A. The first example is symmetric in the sense that all processes entering the VAR are defined by the same dynamics (i.e., the results are invariant by rotations of  $\mathbf{A}_n$ ). This example stresses the fact that asymmetry, or heterogeneity, is not necessary for long memory to arise. The second example presents a heterogenous case where the results are not symmetric for all  $x_{jt}$ .

### 2.2.1. A symmetric example

The sequence of  $n$ -dimensional matrices  $\mathbf{A}_n$  is defined as

$$\mathbf{A}_n = \mathbf{T}_n^* + \eta_n \mathbf{D}_n, \tag{6}$$

where  $\mathbf{T}_n^*$  is specified below;  $\eta_n$  is a real scalar sequence that satisfies  $\eta_n = o(n^{-1})$ , and  $\mathbf{D}_n$  is a real antisymmetric Toeplitz matrix with absolutely summable entries and zero diagonal elements.

To define  $\mathbf{T}_n^*$ , we first consider  $\mathbf{T}_n$  defined as in Assumption T. We choose function  $g(\cdot, \cdot)$  such that, for  $\omega \geq 0$ ,

$$g(d, e^{i\omega}) = 1_{\{0 \leq u < \pi d\}} + 1_{\{\pi(\frac{3}{2}-d) < u \leq \frac{3\pi}{2}\}}, \quad \omega = u \bmod 2\pi, \quad (7)$$

and  $\omega \rightarrow g(d, e^{i\omega})$  is even. In the proof of Theorem 1, we show that  $\mathbf{T}_n$  is asymptotically equivalent to a circulant matrix  $\mathbf{C}_n$  defined as

$$\mathbf{C}_n = \begin{bmatrix} c_0^{(n)} & c_1^{(n)} & \cdots & c_{n-1}^{(n)} \\ c_{n-1}^{(n)} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_1^{(n)} \\ c_1^{(n)} & \cdots & c_{n-1}^{(n)} & c_0^{(n)} \end{bmatrix},$$

where  $c_k^{(n)} = t_{-k}^{(n)} + t_{n-k}^{(n)}$  for  $k \neq 0$  and  $c_0^{(n)} = t_0^{(n)}$ . Eigenvalues of circulant matrices can be expressed in terms of the associated symbol evaluated at the Fourier ordinates. Here  $\mathbf{C}_n$  is chosen so that as  $n \rightarrow \infty$ ,  $c_k^{(n)} \sim t_{\delta_n, -k} + t_{\delta_n, n-k}$ , which defines a circulant matrix with eigenvalues  $g(\delta_n, e^{2i\pi j/n})$  for  $j = 0, \dots, n-1$ , i.e., with about  $\lfloor n\delta_n \rfloor$  unit eigenvalues (the exact number is  $nt_0^{(n)}$ ) and  $n - \lfloor n\delta_n \rfloor$  zero eigenvalues.

Expression (7) defines an even and real-valued function  $\omega \rightarrow g(d, e^{i\omega})$  so  $t_{-k}^{(n)} + t_{n-k}^{(n)} = t_k^{(n)} + t_{-k}^{(n)}$  and the entries  $c_k^{(n)}$  are real. Hence  $\mathbf{C}_n$  is also asymptotically equivalent to the matrix  $\mathbf{T}_n^* \equiv \text{Re}(\mathbf{T}_n)$  with entries  $t_k^{*(n)} = \text{Re}(t_k^{(n)})$ . Hence

$$\frac{\det(\mathbf{I}_{n-1} - \mathbf{T}_{n-1}^* z)}{\det(\mathbf{I}_n - \mathbf{T}_n^* z)} \sim \frac{\det(\mathbf{I}_{n-1} - \mathbf{C}_{n-1} z)}{\det(\mathbf{I}_n - \mathbf{C}_n z)}. \quad (8)$$

The matrix  $\mathbf{T}_n^*$  satisfies the following properties:

$$\begin{aligned} t_0^{*(n)} &= \delta_n + O(n^{-1}), \\ t_k^{*(n)} &= O(n^{-1}), \quad k \neq 0, \end{aligned}$$

where we specify

$$\delta_n = \frac{1}{2} + o(n^{-2}). \quad (9)$$

As  $n \rightarrow \infty$ , the off-diagonal entries of  $\mathbf{T}_n^*$  individually tend to zero. Yet the convergence is slow enough to ensure that for all  $n$ ,  $\mathbf{T}_n^*$  is different enough from a diagonal matrix. Indeed the off-diagonal elements of each row admit a nonzero sum:

$$\sum_{k=1}^{n-1} t_k^*(n) = 1 - g(\delta_n, 0) = 1 - \delta_n + O(n^{-1}). \quad (10)$$

To understand further the behavior of the sequence  $(\mathbf{T}_n^*)$ , consider the limiting coefficients  $t_{d,k}^* = \text{Re}(t_{d,k})$ , where  $t_{d,k}$  is defined by Assumption T:

$$t_{d,k}^* = d \sin_c \frac{\pi k d}{2} \left[ 1_{\{k \text{ odd}\}} \sin \frac{\pi k}{4} \sin \frac{\pi k (\frac{1}{2} - d)}{2} + 1_{\{k \text{ even}\}} \cos \frac{\pi k}{4} \cos \frac{\pi k (\frac{1}{2} - d)}{2} \right], \quad (11)$$

where  $\text{sin}_c(x) = \frac{\sin x}{x}$ . Hence, as  $k \rightarrow \infty$ , the coefficients  $t_{d,k}^*$  decay as  $k^{-1}$  (with a numerator bounded in absolute value between zero and one), where  $k$  measures here a distance between individual variables in the cross-section dimension.<sup>4</sup> As  $d \rightarrow 1/2$ ,  $t_{d,k}^* \rightarrow 0$ . The entries  $t_k^{*(n)}$  themselves satisfy

$$t_k^{*(n)} = t_{\delta_n, k}^* + O(n^{-1}).$$

Hence as  $n \rightarrow \infty$ , since  $\delta_n \rightarrow 1/2$ ,  $t_{\delta_n, k}^* \rightarrow 0$  for  $k \neq 0$ , and the off-diagonal elements of  $\mathbf{T}_n^*$  tend to zero although their sum does not. We verify in the appendix that Assumptions B, T and A hold for this specific  $g(\cdot, \cdot)$  function.

Theorem 1 then implies that the spectral density of  $x_{jt}$ , for all  $j \in \mathbb{N}$ , is identical to that of a fractionally integrated process of order  $1/2$ : as  $n \rightarrow \infty$ ,

$$f_{n, x_j}(\omega) \rightarrow |1 - e^{-i\omega}|^{-1} \sigma_{\epsilon_j}^2. \quad (12)$$

The limiting ARFIMA(0, 1/2, 0) process is often called an  $1/f$  or flicker noise (see Mandelbrot, 1967). Fractional integration arises here in a context where the VAR(1) matrix coefficient  $\mathbf{A}_n$  can be associated with a complex-valued circulant matrix which asymptotically presents about  $\lfloor n/2 \rfloor$  unit eigenvalues and  $\lfloor n/2 \rfloor$  zero eigenvalues.

### 2.2.2. Asymmetric example: one dominant innovation

The results presented above are not limited to the flicker noise ARFIMA(0, 1/2, 0) but can be extended to any  $I(\delta)$ ,  $\delta \in (0, 1)$ . We now give an example of sequence  $\mathbf{A}_n$  satisfying Assumptions B, T and A, but where long memory does not appear symmetrically for all  $x_{jt}$ . Consider the process where  $\mathbf{T}_n^*$  is defined as previously with  $\delta_n \equiv \delta \in (0, 1)$  and let  $\mathbf{A}_n = \mathbf{T}_n^*$ . Assumptions T and A and B(ii) are hence satisfied.

Now, assume that the variance of one innovation  $\epsilon_{jt}$  dominates the others. For this we define  $\sigma_{\epsilon_j}^2$  the vector of variances ( $\sigma_{\epsilon_k}^2$ ) for  $k \neq j$  and assume that  $\sigma_{\epsilon_j}^2 \rightarrow 0$  when  $n \rightarrow \infty$ . Assumption B(i) hence follows. Theorem 1 then implies that, as  $n \rightarrow \infty$  and for  $\omega \neq 0$

$$f_{n, x_j}(\omega) \rightarrow \sigma_{\epsilon_j}^2 |1 - e^{-i\omega}|^{-2\delta}.$$

Therefore, when the number of variables  $n$  tends to infinity and when one of the innovation processes dominates all the others, then the spectral density of the dominant process entering  $\mathbf{x}_t$  tends to that of an  $I(\delta)$  fractional white noise. Here the off-diagonal elements of  $\mathbf{A}_n$  do not tend to zero asymptotically.

To the best of our knowledge this result is new in the sense that long memory does not arise from any of the known sources. In particular, despite the multivariate nature of the source of long

---

<sup>4</sup>Although they are not comparable, we notice that the rate of decay in  $k^{-1}$  is slower than the time-dimension  $k^{-d-1}$  decay of the coefficients in the autoregressive representation of an  $I(d)$  fractional white noise (see e.g. Baillie, 1996, Table 2).

memory that we present, it is not aggregation that is at play here since only one innovation  $\epsilon_{jt}$  with nonzero variance remains in the system as  $n \rightarrow \infty$ . The mechanism is closer in a sense to that, which Schennach (2013) delineates, of a single input that transits through a network. If we interpret our VAR setting as a network, then there are  $n$  nodes,  $i = 1, \dots, n$  in the system which are in state  $x_{it-1}$  at the end of any period  $t - 1$ . At time  $t$ , each node  $i$  combines  $\mathbf{x}_{t-1}$  with an additional idiosyncratic signal  $\epsilon_{it}$  to produce  $\mathbf{x}_t$ . Since all the coefficients of  $\mathbf{A}_n$  are strictly less than unity in absolute value, signal transmission from  $x_{jt-1}$  to  $x_{it}$ , for  $i \neq j$  is attenuated in a “short memory” manner (to interpret Schennach heuristically, but she defines this precisely). In this interpretation of the VAR,  $|i - j|$  measures the distance between the two nodes  $i$  and  $j$  and the system is *homogenous* since all entries  $a_{ij} = a_{|i-j|}$ , i.e., they only depend on the distance between nodes. Contrary to Schennach, the long memory process that results in our context acts as a form of common element which drives the system when we assume that  $\sigma_{\epsilon_j}^2 \rightarrow 0$ , i.e. only one innovation process remains.

### 3. Simulation and empirical evidence

In this section, we evaluate our key theoretical results via a Monte Carlo simulation. We also show that our theoretical framework is able to replicate some stylized facts observed in the variance of US stock returns.

#### 3.1. Monte Carlo

We provide here simulations that examine the validity of our theoretical asymptotic results when the dimension of the cross-section and of the sample are finite.

An  $n$ -dimensional VAR(1), as defined in Equations (1a)-(1b), is used to generate data for different choices of  $T$  and  $n$ . To save space, we only report the results for  $n = 200$  series and  $T = 4,000$  observations.

As a benchmark, we consider in our first experiment the case of a diagonal matrix,  $\mathbf{A}_n = d\mathbf{I}_n$ , where the parameter  $d$  is set to 0.499. The first panel of Figure 1 shows the value of the elements of the first row of  $\mathbf{A}_n$ , denoted  $a_k^{(n)}$  (for  $k = 0, \dots, n - 1$ ), i.e.,  $a_k^{(n)} = 0.499$  for  $k = 0$  and 0 otherwise. In this simple setting, the derived univariate processes have short memory and follow a stationary AR(1) model with an autoregressive parameter of 0.499 for each series.

Panel 2 of Figure 1 plots the empirical distribution (over 1,000 replications) of the long memory parameter of series  $x_{1t}$  estimated using three popular estimation methods, i.e., the log periodogram regression (GPH) of Geweke and Porter-Hudak (1983), the Local Whittle Likelihood Estimator (LWLE) of Robinson (1995), both with bandwidth  $T/2$  and the MLE of an ARFIMA(1,  $d$ , 0) (see Sowell, 1992 and Doornik and Ooms, 2004).<sup>5</sup> We deliberately choose a large bandwidth, as

---

<sup>5</sup>All estimations are performed in OxMetrics 7.0 (see Doornik, 2013).

implemented by default in Doornik and Ooms (2004) to reduce the variability of the estimators. As expected the estimated long memory parameters are concentrated around 0 suggesting that there is no evidence of long memory in the individual series. This is confirmed by the third panel of Figure 1 which reports the ACF of  $x_{1t}$  for the first replication.

In the next two experiments, we consider a symmetric Toeplitz matrix  $\mathbf{A}_n = \mathbf{T}_n^*$ , under the assumptions of Section 2.2.1 (i.e., Equation (6) with  $\eta_n = 0$ ), where  $\mathbf{T}_n^*$  has symbol  $g_{\mathbf{T}_d}$ . We denote by  $d$  the value taken by  $\delta_n$ : we choose two values of  $d$  close to  $1/2$ , i.e., respectively  $d = 0.499$  in Figure 2, and  $d = 0.45$  in Figure 3. The structure of these figures is similar to that of Figure 1 except that now, since  $d$  is close to  $1/2$ , i.e., to the nonstationary region of an  $I(d)$  process, we follow the approach of Abadir, Distaso and Giraitis (2007) and apply the three long memory estimators to  $(1-L)^d x_{1t}$  (for the values we report, we have added  $d$  ex-post to the estimate). The first panel of these figures emphasizes that the diagonal elements are near  $d$  while the off-diagonal elements are small for  $d = 0.45$  and very small for  $d = 0.499$ . Recall from Equation (10) that the sum of each row of  $\mathbf{T}_n^*$  is 1 by construction and therefore although the off-diagonal elements of  $\mathbf{A}_n$  can be very small when  $d$  is close to  $1/2$ , they are nonzero. Unlike in Figure 1, long memory is detected in  $x_{1t}$ , with a Monte Carlo mean (over the 1,000 replications) of 0.444, 0.484 and 0.488 respectively for the GPH, LWLE and ARFIMA(0,  $d$ , 0) methods for  $d = 0.499$  and 0.417, 0.451 and 0.465 for  $d = 0.45$ . The ACF of  $x_{1t}$  in the first replication also suggests the presence of long memory. These figures show that although  $\mathbf{A}_n$  is near diagonal, the very small off-diagonal elements play a crucial role in the apparition of long memory.

Next, we evaluate the robustness of the previous result by using the asymmetric Toeplitz matrix given in Equation (6), i.e.,  $\mathbf{A}_n = \mathbf{T}_n^* + \eta_n \mathbf{D}_n$ , with  $d = 0.499$ ,  $\eta_n = \frac{1}{n \log(n)}$ , and where the elements of  $\mathbf{D}_n$  in the upper triangle are drawn independently from a standard normal distribution. Figure 4 suggests that results are qualitatively the same than in the case of the symmetric Toeplitz matrix in the sense that long memory is detected in  $x_{1t}$  with a parameter estimate close to  $d$ .

Theorem 1 states that, under Assumptions B, T and A, not only  $x_{1t}$  but all variables belonging to  $\mathbf{x}_t$  should be fractional white noises when  $n \rightarrow \infty$  and  $d \rightarrow 1/2$ . Our last experiment illustrates this finding for the case of a symmetric Toeplitz matrix with  $d = 0.499$ , as investigated in Figure 2. Figure 5 plots the empirical distribution of the long memory parameter estimated on all series, i.e., on  $x_{1t}, \dots, x_{200t}$ , for the three estimation methods. We only report the results for four replications, each row in the figure corresponding to a replication. Results suggest that the estimated long memory parameters do not vary much across series and are all concentrated in a region close to  $1/2$ , especially for the LWLE and MLE of the ARFIMA(0,  $d$ , 0).

### 3.2. Empirical illustration

As reported in Lieberman and Phillips (2008) “*There is an emerging consensus in empirical finance that realized volatility series typically display long range dependence with a memory*”

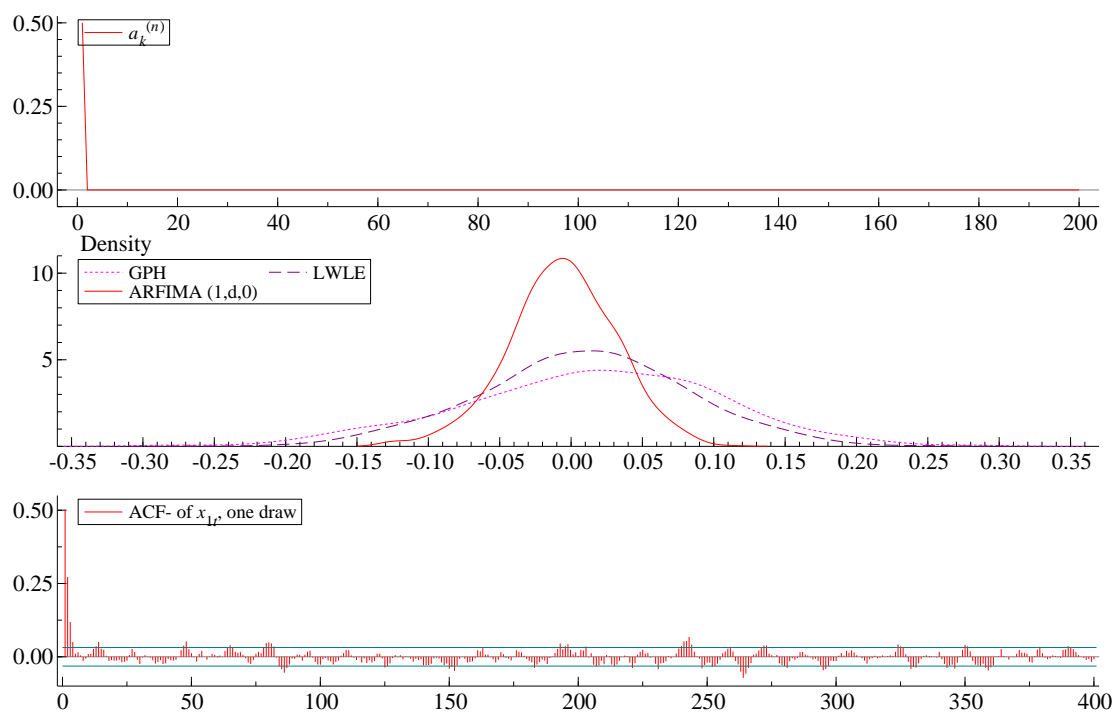


Figure 1: Simulation results for a  $n$ -dimensional diagonal VAR(1)  $\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t$ , with  $\mathbf{A}_n = d\mathbf{I}_n$ , where  $d = 0.499$ ,  $\epsilon_t \stackrel{iid}{\sim} N(\mathbf{0}, \mathbf{I}_n)$ ,  $n = 200$  and  $t = 1, \dots, 4000$ . The panels report respectively, (a) the value of the elements of the first row of  $\mathbf{A}_n$ , denoted  $a_k^{(n)}$  (for  $k = 0, \dots, n-1$ ); (b) the empirical distribution, over 1000 replications, of the estimated long memory parameter of  $x_{1t}$  obtained by the GPH, LWLE and ARFIMA(1,  $d$ , 0) methods; (c) the empirical ACF of  $x_{1t}$  for the first replication.

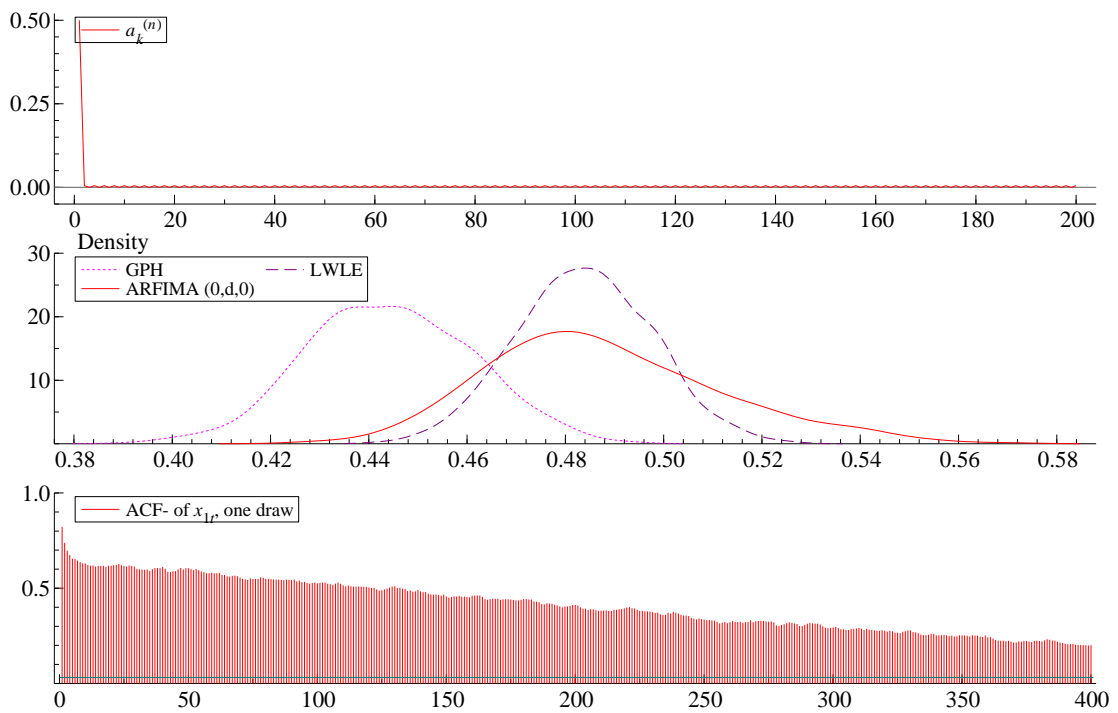


Figure 2: Simulation results for a  $n$ -dimensional diagonal VAR(1)  $\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t$ , with  $\mathbf{A}_n = \mathbf{T}_n^*$ , where  $\mathbf{T}_n^* \equiv \text{Re}(\mathbf{T}_n)$ ,  $\mathbf{T}_n$  has symbol defined by (7),  $d = 0.499$ ,  $\epsilon_t \stackrel{iid}{\sim} N(\mathbf{0}, \mathbf{I}_n)$ ,  $n = 200$  and  $t = 1, \dots, 4000$ . The panels report respectively, (a) the value of the elements of the first row of  $\mathbf{A}_n$ , denoted  $a_k^{(n)}$  (for  $k = 0, \dots, n-1$ ); (b) the empirical distribution, over 1000 replications, of the estimated long memory parameter of  $x_{1t}$  obtained by the GPH, LWLE and ARFIMA(0,  $d$ , 0) methods; (c) the empirical ACF of  $x_{1t}$  for the first replication.



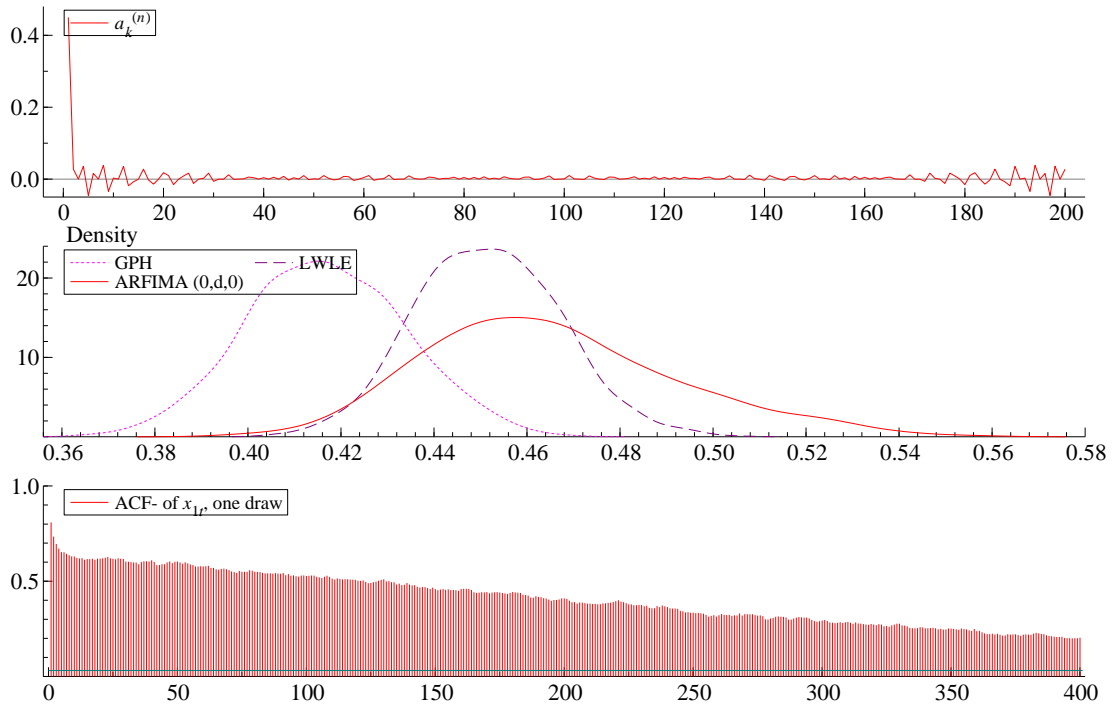


Figure 3: Simulation results for a  $n$ -dimensional diagonal VAR(1)  $\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t$ , with  $\mathbf{A}_n = \mathbf{T}_n^*$ , where  $\mathbf{T}_n^* \equiv \text{Re}(\mathbf{T}_n)$ ,  $\mathbf{T}_n$  has symbol defined by (7),  $d = 0.45$ ,  $\epsilon_t \stackrel{iid}{\sim} N(\mathbf{0}, \mathbf{I}_n)$ ,  $n = 200$  and  $t = 1, \dots, 4000$ . The panels report respectively, (a) the value of the elements of the first row of  $\mathbf{A}_n$ , denoted  $a_k^{(n)}$  (for  $k = 0, \dots, n-1$ ); (b) the empirical distribution, over 1000 replications, of the estimated long memory parameter of  $x_{1t}$  obtained by the GPH, LWLE and ARFIMA(0,  $d$ , 0) methods; (c) the empirical ACF of  $x_{1t}$  for the first replication.

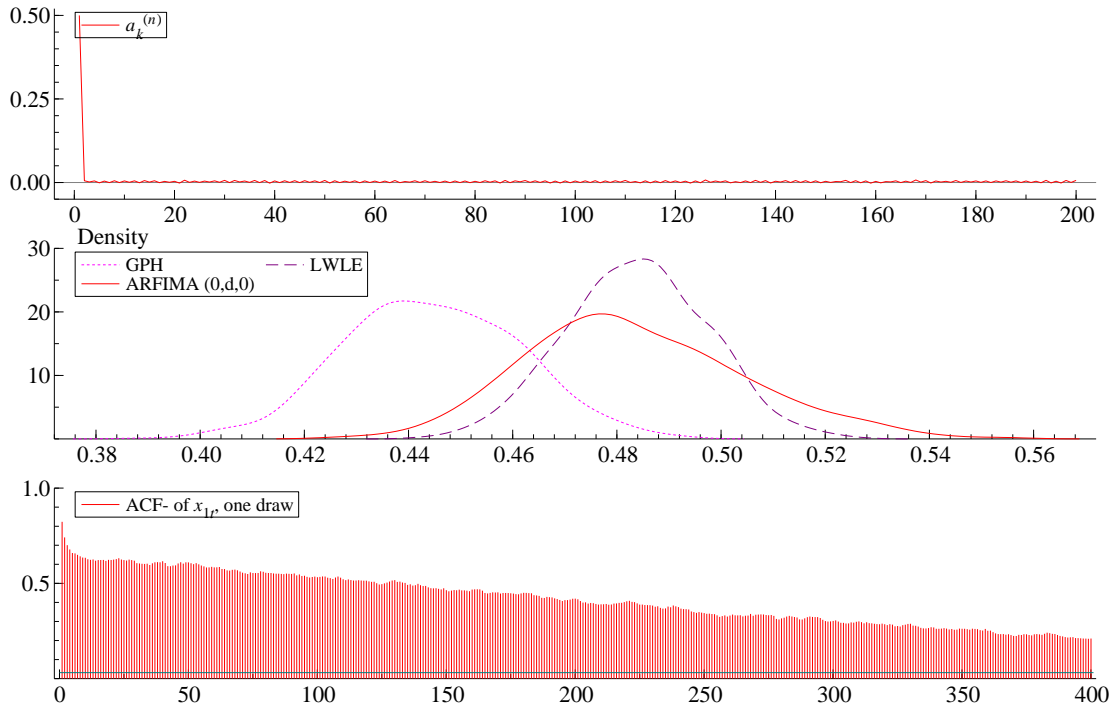


Figure 4: Simulation results for a  $n$ -dimensional diagonal VAR(1)  $\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t$ , with  $\mathbf{A}_n = \mathbf{T}_n^* + \eta_n \mathbf{D}_n$ , where  $\mathbf{T}_n^* \equiv \text{Re}(\mathbf{T}_n)$ ,  $\mathbf{T}_n$  has symbol defined by (7),  $\eta_n = \frac{1}{n \log(n)}$ ,  $\mathbf{D}_n$  is an antisymmetric Toeplitz matrix with above diagonal elements drawn from a standard normal distribution,  $d = 0.499$ ,  $\epsilon_t \stackrel{iid}{\sim} N(\mathbf{0}, \mathbf{I}_n)$ ,  $n = 200$  and  $t = 1, \dots, 4000$ . The panels report respectively, (a) the value of the elements of the first row of  $\mathbf{A}_n$ , denoted  $a_k^{(n)}$  (for  $k = 0, \dots, n-1$ ); (b) the empirical distribution, over 1000 replications, of the estimated long memory parameter of  $x_{1t}$  obtained by the GPH, LWLE and ARFIMA(0,  $d$ , 0) methods; (c) the empirical ACF of  $x_{1t}$  for the first replication.

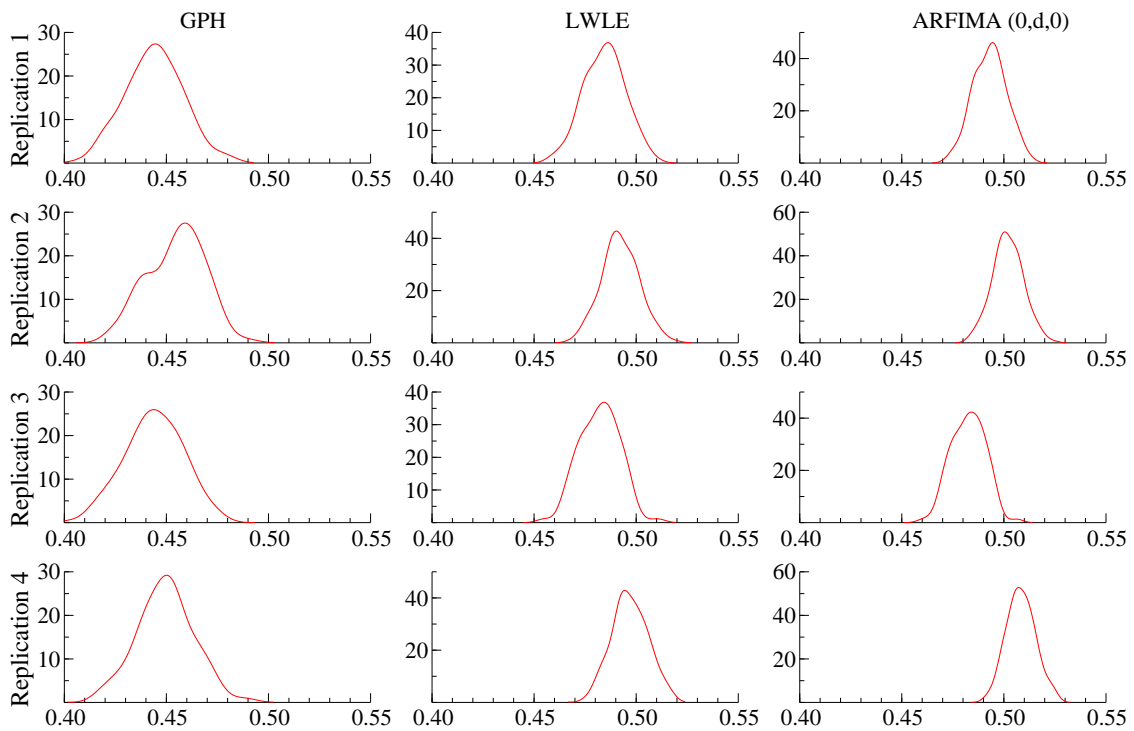


Figure 5: Simulation results for a  $n$ -dimensional diagonal VAR(1)  $\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t$ , with  $\mathbf{A}_n = \mathbf{T}_n^*$ , where  $\mathbf{T}_n^* \equiv \text{Re}(\mathbf{T}_n)$ ,  $\mathbf{T}_n$  has symbol defined by (7),  $d = 0.499$ ,  $\epsilon_t \stackrel{iid}{\sim} N(\mathbf{0}, \mathbf{I}_n)$ ,  $n = 200$  and  $t = 1, \dots, 4000$ . The figure plots the empirical distribution of the long memory parameter estimated on all series, i.e., on  $x_{1t}, \dots, x_{200t}$ , using GPH, LWLE and ARFIMA(0,  $d$ , 0). Each row corresponds to a separate replication.

parameter  $d$  around 0.4 (Andersen et al., 2001; Martens et al., 2004[now 2009]).”

To illustrate this claim and also to provide a first assessment of the plausibility of our explanation for the origin of long memory, we consider a dataset (provided by TickData) consisting of transaction prices at the 5-minute sampling frequency for 49 large capitalization stocks from the NYSE, AMEX NASDAQ, covering the period from January 4, 1999 to December 31, 2008 (2,489 trading days).<sup>6</sup> The trading session runs from 9:30 EST until 16:00 EST. Using 5-minute returns, we computed the MedRV estimator of Andersen, Dobrev, and Schaumburg (2012), a non-parametric robust to jumps estimator of the integrated variance.<sup>7</sup>

Figure 6 plots the long memory parameter estimated using an ARFIMA model on  $\log(\text{MedRV}_{it})$  for  $i = 1, \dots, 49$ .<sup>8</sup> The estimated long memory parameters fluctuate around 0.45, with a minimum of about 0.40 and a maximum of about 0.53.

VAR models for the logarithm of realized variances have been used for instance by Anderson and Vahid (2007). Figure 7 plots some summary statistics on the estimated parameters of a VAR(1) model estimated on  $\log(\text{MedRV}_{it})$ , by progressively increasing the dimension of the VAR (i.e., by adding one variable at a time to the system, following the alphabetical order of the tickers).

The solid lines correspond to the average of the diagonal elements (upper panel) and the average of the absolute value of the off-diagonal elements (lower panel). For instance, the average of the diagonal elements is about 0.63 for the VAR(1) of dimension 2 (i.e., series AAPL and ABT) and the absolute value of the off-diagonal element is about 0.2. Figure 7 suggests that the average of the diagonal elements converges to about 0.4 when the dimension of the system increases while the off-diagonal elements converge to a very small value. This is in agreement with our theoretical model for which the diagonal elements correspond roughly to  $d$  and the off-diagonal elements are small.

Figure 7 (dotted lines) also reports similar quantities but for simulated data following a VAR(1) with a symmetric Toeplitz matrix  $\mathbf{A}_n = \mathbf{T}_n^*$ , where  $\mathbf{T}_n^*$  has symbol  $g_{\mathbf{T}_d}$  given in (7),  $n = 200$  and  $d = 0.4$ . While the true dimension of the system is  $n = 200$ , the VAR is estimated on a smaller system whose dimension progressively increases up to 49 series. A similar pattern is observed both for real and simulated data. Indeed, the average of the diagonal of the VAR(1) estimated on simulated data decreases with the size of the system and converges to 0.4 while the average of the off-diagonal elements converges to a very small value.

---

<sup>6</sup>To save space, we do not report company names but only the ticker symbols. There are AAPL, ABT, AXP, BA, BAC, BMY, BP, C, CAT, CL, CSCO, CVX, DELL, DIS, EK, EXC, F, FDX, GE, GM, HD, HNZ, HON, IBM, INTC, JNJ, KO, LLY, MCD, MMM, MOT, MRK, MS, MSFT, ORCL, PEP, PFE, PG, QCOM, SLB, T, TWX, UN, VZ, WFC, WMT, WYE, XOM, XRX.

<sup>7</sup>If  $r_{t,i}$  is the  $i$ th 5-minutes return of a day  $t$  containing  $M$  of such returns, the MedRV of day  $t$  is computed as  $\text{MedRV}_t = \frac{\pi}{6-4\sqrt{3}+\pi} \frac{M}{M-2} \sum_{i=3}^M \text{med}(|r_{t,i}|, |r_{t,i-1}|, |r_{t,i-2}|)^2$ , where  $\text{med}(\cdot)$  denotes the median.

<sup>8</sup>Similar to the previous section, the ARFIMA model is estimated on  $(1-L)^{1/2} \log(\text{MedRV}_{it})$  and 1/2 is added ex-post to the estimated value to ensure the estimated  $d$  to lie in the covariance stationarity region.

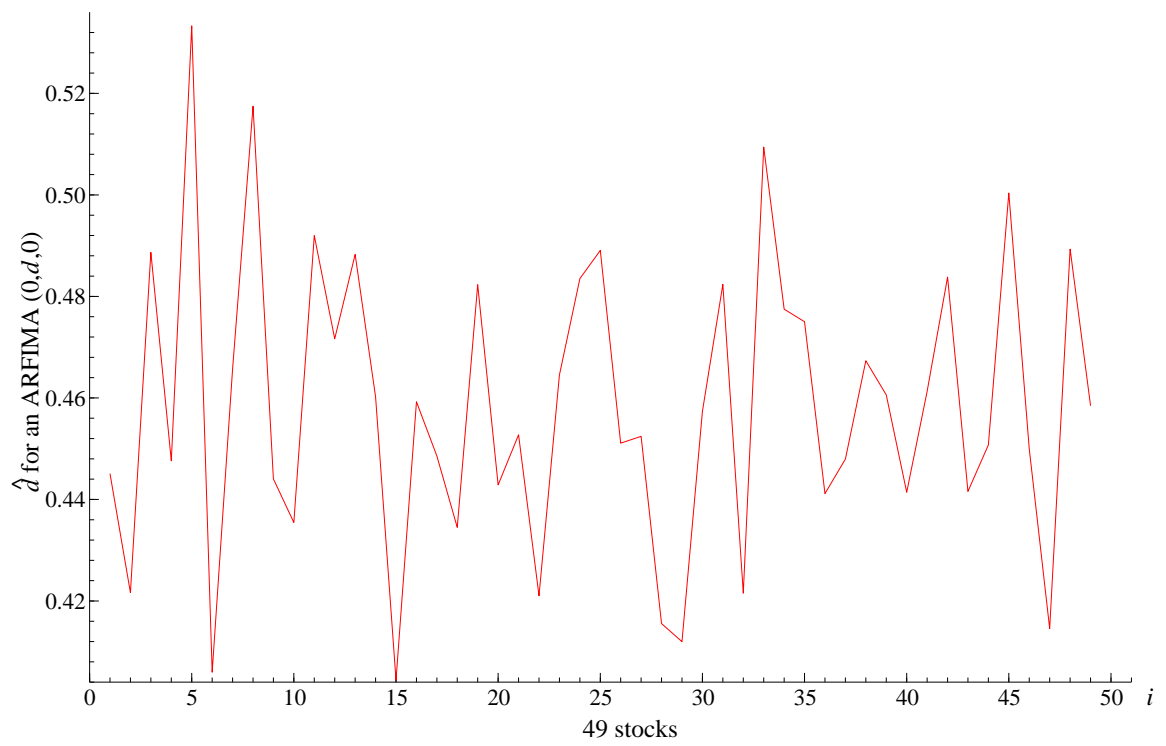


Figure 6: Long memory parameter of an ARFIMA(0,  $d$ , 0) model estimated by maximum likelihood on  $\log(\text{Med}RV_{it})$  for the 49 stocks  $i = 1, \dots, 49$ .

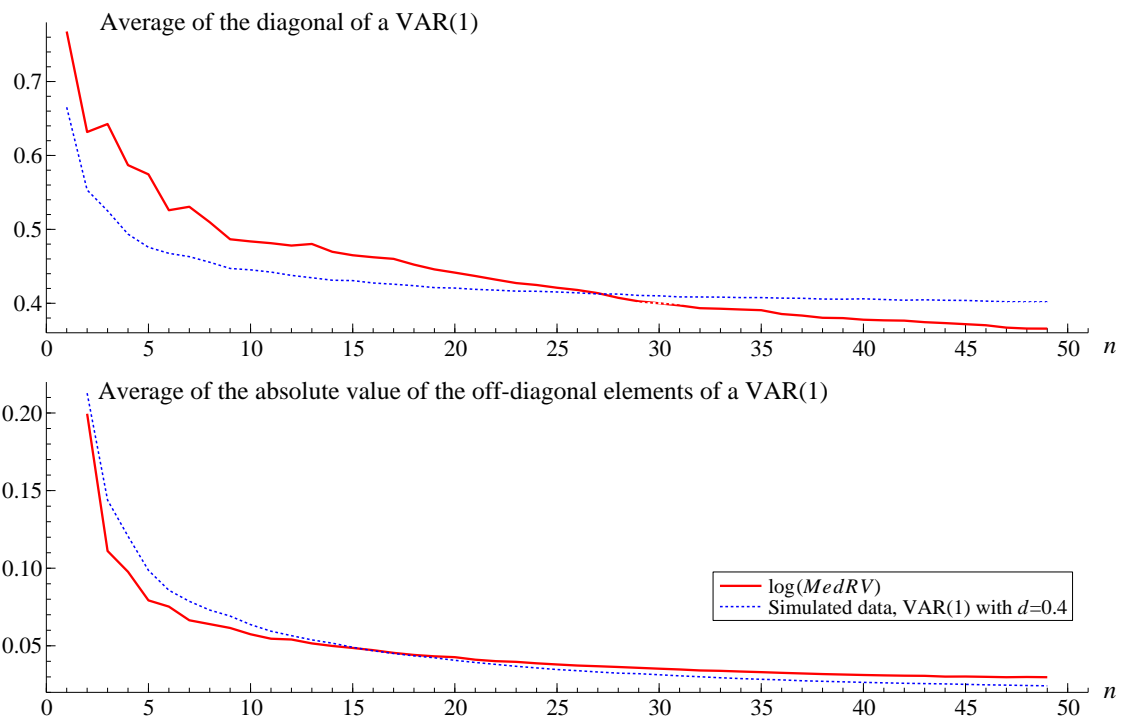


Figure 7: Average of the diagonal elements (upper panel) and average of the absolute value of the off-diagonal elements (lower panels) of a VAR(1) estimated on  $\log(MedRV_{it})$  while progressively increasing the dimension of the VAR)

## 4. Conclusion

Our paper contributes to the time series literature investigating the mechanisms generating slowly decaying autocorrelations and low frequency variability, in particular those leading to long memory processes. We show that an  $n$ -dimensional vector autoregressive model of order 1, can generate long memory in the marginalized univariate series. To achieve this goal, we consider the final equation representation of this model and obtain the  $n$  univariate implied ARMA( $n, n - 1$ ) models. The spectral density of each univariate series is expressed as the sum of  $n$  ratios that are derived from the determinant and the adjugate of the matrix lag polynomial of the VAR. We then develop three high-level assumptions ensuring the spectral density of each series converges to that of a long memory process of order  $\delta \in (0, 1)$ . We show that these assumptions are satisfied for two specific examples of an  $n$ -dimensional VAR(1) model where either (i) all univariate processes appear  $I(\frac{1}{2})$  as  $n \rightarrow \infty$ , or (ii) the spectral density of one univariate process tends to that of an  $I(\delta)$  fractional white noise.

We consider the implications of our findings for the variance of asset returns where the so-called golden-rule of realized variance states that they always exhibit fractional integration of degree close to 0.4. The assumption of a “quasi-diagonal” multivariate time series model is motivated by the fact that it is common to see in empirical works parameter values of large dimensional VAR, VEC or BEKK models such that each series is strongly explained by its own lags and that cross-correlation or contagion parameters (i.e., off-diagonal elements) are individually small, weakly significant (if not insignificant) but jointly highly significant.

Our approach is general enough to allow extending it to groups of time series sharing within each group the properties we study in this paper and where each group is orthogonal to others. This would be the case in a large dimensional block-diagonal VAR or in a GVAR for instance.

## 5. Appendix

### 5.1. Proof of Theorem 1

Together, Assumptions B and A imply that

$$f_{n,x_j}(\omega) \sim \left| \frac{\det(\mathbf{I}_{n-1} - \mathbf{T}_{n-1}e^{-i\omega})}{\det(\mathbf{I}_n - \mathbf{T}_n e^{-i\omega})} \right|^2 \sigma_{\epsilon_j}^2.$$

Hence the only element we need to consider is the convergence, as  $n \rightarrow \infty$ ,

$$\frac{\det(\mathbf{I}_{n-1} - \mathbf{T}_{n-1}e^{-i\omega})}{\det(\mathbf{I}_n - \mathbf{T}_n e^{-i\omega})} \rightarrow (1 - z)^{-\delta}.$$

Assumption T(i) states that  $g \in \mathbb{R}$  which implies that  $t_{d,k} = \int_0^{2\pi} g(d, e^{i\omega}) e^{-ik\omega} d\omega = \overline{t_{d,-k}}$ , i.e.,  $\mathbf{T}_{d,n}$  is Hermitian. This entails in particular that  $t_k^{(n)} + t_{n-k}^{(n)} = t_k^{(n)} + t_{-k}^{(n)} \in \mathbb{R}$ . Also  $g_{\mathbf{T}_d}$  being bounded ensures  $(\mathbf{T}_{d,n})$  and the associated matrices below are uniformly bounded in strong

norm. Assumption T(ii) ensures that  $\sum_{k=-(n-1)}^{n-1} t_{d,k} e^{ik\omega}$  converges uniformly to  $g_{\mathbf{T}_d}(e^{i\omega})$  which is required for applying Szegő's theorem. Assumption T(iii) ensures that  $\det(\mathbf{T}_n - \mathbf{T}_{\delta_n, n}) \rightarrow 0$  as  $n \rightarrow \infty$  since the determinant is a continuous function of the matrix entries.  $\mathbf{T}_n$  and  $\mathbf{T}_{\delta_n, n}$  are said to be asymptotically equivalent, which is denoted  $\mathbf{T}_n \sim \mathbf{T}_{\delta_n, n}$  (see Gray, 2006, Section 2.3 for the definition of equivalent matrices).

To any Toeplitz matrix  $\mathbf{T}_{d,n}$  within the Wiener class, we can associate a Circulant matrix  $\mathbf{C}_{d,n}$  such that  $\mathbf{C}_{d,n} \sim \mathbf{T}_{d,n}$  defined as

$$\mathbf{C}_{d,n} = \begin{bmatrix} c_{d,0}^{(n)} & c_{d,1}^{(n)} & \cdots & c_{d,n-1}^{(n)} \\ c_{d,n-1}^{(n)} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_{d,1}^{(n)} \\ c_{d,1}^{(n)} & \cdots & c_{d,n-1}^{(n)} & c_{d,0}^{(n)} \end{bmatrix}. \quad (13)$$

The sequence  $(\mathbf{C}_{d,n})$  is not uniquely defined, see Grenander and Szegő (1958, Section 7.6). Gray (2006, Lemma 4.6) shows in particular that choosing  $c_{d,k}^{(n)} = t_{d,-k} + t_{d,n-k}$  yields a matrix which is asymptotically equivalent to  $\mathbf{T}_{d,n}$ .

Here we define  $t_{\delta,k}^{(n)} = \frac{1}{n} \sum_{j=0}^{n-1} g\left(\delta, e^{i\frac{2\pi j}{n}}\right) e^{-2i\pi jk/n}$  with  $\delta = \lim_{n \rightarrow \infty} \delta_n$  and

$$\begin{aligned} c_{\delta,k}^{(n)} &= t_{\delta,-k}^{(n)} + t_{\delta,n-k}^{(n)} \\ c_k^{(n)} &= t_{-k}^{(n)} + t_{n-k}^{(n)}. \end{aligned}$$

$\mathbf{C}_{\delta,n}$  and  $\mathbf{C}_n$  are therefore circulant matrices with entries  $c_{\delta,k}^{(n)}$  and  $c_k^{(n)}$ . Since  $t_k^{(n)}$  converges to  $t_{\delta,k}$ , the matrix  $\mathbf{C}_n$  is asymptotically equivalent to that with entries  $t_{\delta,-k} + t_{\delta,n-k}$ . Since asymptotic equivalence is transitive (see Gray, 2006, Theorem 2.1), it follows that  $\mathbf{C}_n \sim \mathbf{T}_{\delta_n, n}$  and  $\mathbf{C}_n \sim \mathbf{T}_n$ . Hence

$$\frac{\det(\mathbf{I}_{n-1} - z\mathbf{T}_{n-1})}{\det(\mathbf{I}_n - z\mathbf{T}_n)} \sim \frac{\det(\mathbf{I}_{n-1} - z\mathbf{C}_{n-1})}{\det(\mathbf{I}_n - z\mathbf{C}_n)}. \quad (14)$$

The symbol of  $\mathbf{I}_n - \mathbf{T}_{d,n}z$  is  $1 - g_{\mathbf{T}_d}(\cdot)z$  for all  $z \in \mathbb{C}$ ,  $|z| < 1$ . Hence Gray (2006, Lemma 4.6) implies that:

$$\frac{\det(\mathbf{I}_{n-1} - z\mathbf{C}_{\delta,n-1})}{\det(\mathbf{I}_n - z\mathbf{C}_{\delta,n})} \rightarrow \exp \left\{ \frac{1}{2\pi} \int_0^{2\pi} \log(1 - g_{\mathbf{T}_\delta}(e^{i\omega})z) d\omega \right\}.$$

The determinants can be decomposed as

$$\begin{aligned} \det(\mathbf{I}_n - z\mathbf{C}_n) &= \det(\mathbf{I}_n - z\mathbf{C}_{\delta,n} + z(\mathbf{C}_{\delta,n} - \mathbf{C}_n)) \\ &= \det(\mathbf{I}_n - z\mathbf{C}_{\delta,n}) \det\left(\mathbf{I}_n + z(\mathbf{I}_n - z\mathbf{C}_{\delta,n})^{-1}(\mathbf{C}_{\delta,n} - \mathbf{C}_n)\right) \end{aligned}$$

and

$$\begin{aligned} \frac{\det(\mathbf{I}_{n-1} - z\mathbf{C}_{n-1})}{\det(\mathbf{I}_n - z\mathbf{C}_n)} &= \frac{\det(\mathbf{I}_{n-1} - z\mathbf{C}_{\delta,n-1})}{\det(\mathbf{I}_n - z\mathbf{C}_{\delta,n})} \\ &\quad \times \frac{\det\left(\mathbf{I}_{n-1} + z(\mathbf{I}_{n-1} - z\mathbf{C}_{\delta,n-1})^{-1}(\mathbf{C}_{\delta,n-1} - \mathbf{C}_{n-1})\right)}{\det\left(\mathbf{I}_n + z(\mathbf{I}_n - z\mathbf{C}_{\delta,n})^{-1}(\mathbf{C}_{\delta,n} - \mathbf{C}_n)\right)}, \end{aligned}$$



where  $\mathbf{C}_{\delta,n-1} - \mathbf{C}_{n-1} \rightarrow 0$  so for  $|z| < 1$ ,

$$\det \left( \mathbf{I}_{n-1} + z (\mathbf{I}_{n-1} - z \mathbf{C}_{\delta,n-1})^{-1} (\mathbf{C}_{\delta,n-1} - \mathbf{C}_{n-1}) \right) \rightarrow 1.$$

Hence for all  $\eta > 0$ , there exists a value  $N$  such that for  $n \geq N$ ,

$$\left| \frac{\det \left( \mathbf{I}_{n-1} + z (\mathbf{I}_{n-1} - z \mathbf{C}_{\delta,n-1})^{-1} (\mathbf{C}_{\delta,n-1} - \mathbf{C}_{n-1}) \right)}{\det \left( \mathbf{I}_n + z (\mathbf{I}_n - z \mathbf{C}_{\delta,n})^{-1} (\mathbf{C}_{\delta,n} - \mathbf{C}_n) \right)} - 1 \right| < \eta$$

implying that

$$\left| \frac{\det(\mathbf{I}_{n-1} - z \mathbf{C}_{n-1})}{\det(\mathbf{I}_n - z \mathbf{C}_n)} - \frac{\det(\mathbf{I}_{n-1} - z \mathbf{C}_{\delta,n-1})}{\det(\mathbf{I}_n - z \mathbf{C}_{\delta,n})} \right| < \left| \frac{\det(\mathbf{I}_{n-1} - z \mathbf{C}_{\delta,n-1})}{\det(\mathbf{I}_n - z \mathbf{C}_{\delta,n})} \right| \eta.$$

Since  $\lim_{n \rightarrow \infty} \frac{\det(\mathbf{I}_{n-1} - z \mathbf{C}_{\delta,n-1})}{\det(\mathbf{I}_n - z \mathbf{C}_{\delta,n})} = (1-z)^{-\delta}$  it follows that  $\forall \varepsilon > 0$  and  $z \in (0, 1)$ , there exists  $\eta = \varepsilon (1-z)^\delta$  and  $N$  such that for  $n > N$

$$\left| \frac{\det(\mathbf{I}_{n-1} - z \mathbf{C}_{n-1})}{\det(\mathbf{I}_n - z \mathbf{C}_n)} - \frac{\det(\mathbf{I}_{n-1} - z \mathbf{C}_{\delta,n-1})}{\det(\mathbf{I}_n - z \mathbf{C}_{\delta,n})} \right| < \varepsilon.$$

Uniform convergence of  $\frac{\det(\mathbf{I}_{n-1} - z \mathbf{C}_{n-1})}{\det(\mathbf{I}_n - z \mathbf{C}_n)}$  to  $(1-z)^{-\delta}$  also holds on any set  $(0, z_0)$ , where  $|z_0| < 1$ , by choosing  $\forall \varepsilon > 0$ ,  $\eta = \varepsilon (1-z_0)^\delta$  and associated  $N$ .

It follows from Equation (14) that

$$\frac{\det(\mathbf{I}_{n-1} - \mathbf{T}_{n-1}z)}{\det(\mathbf{I}_n - \mathbf{T}_nz)} \rightarrow (1-z)^{-\delta},$$

with the same uniform convergence property.

## 5.2. Proofs relative to Section 2.2.1

We collect here the proofs related to Section 2.2.1 that show that Assumptions B, T and A are satisfied for  $\mathbf{A}_n = \mathbf{T}_n^* + \eta_n \mathbf{D}_n$ , where  $\mathbf{T}_n^*$ ,  $\eta_n$  and  $\mathbf{D}_n$  are specified as in Section 2.2.1. We start by proving the following lemma.

**Lemma 1.** *Under the assumptions of Section 2.2.1, and as  $n \rightarrow \infty$ ,*

*L(i) The coefficients of  $\mathbf{C}_n$  satisfy,*

$$\begin{aligned} t_0^{*(n)} &= 1/2 + O(n^{-1}), \\ t_k^{*(n)} &= O(n^{-1}), \quad 0 < |k| < n. \end{aligned}$$

*L(ii)  $\forall k \neq 0$  and  $t_k^{*(n)} \neq 0$ ,  $\frac{t_k^{*(n+1)} - t_k^{*(n)}}{t_k^{*(n)}} = O\left(\frac{k}{n}\right)$ .*

*L(iii)  $\exists N$  such that  $\forall n \geq N$ ,  $\sup_{0 \leq k \leq n} t_k^{*(n)} < 1$ .*

**Proof.** We start with the proof of L(i).

Since  $g(d, x) = 1_{\{0 \leq x < \pi d\}} + 1_{\{\pi(\frac{3}{2}-d) < x \leq \frac{3\pi}{2}\}}$  for  $d \in (0, 1)$  and  $x \in [0, 2\pi]$ , the coefficients of  $\mathbf{T}_n$

satisfy:

$$\begin{aligned} t_{-k}^{(n)} &= \frac{1}{n} \sum_{j=0}^{n-1} 1_{\{j < \frac{n\delta_n}{2}\}} e^{2i\pi k j/n} + \frac{1}{n} \sum_{j=0}^{n-1} 1_{\{(\frac{3n}{4} - \frac{n\delta_n}{2}) < j \leq \frac{3n}{4}\}} e^{2i\pi k j/n} \\ &= \frac{1}{n} \sum_{j=0}^{\lceil n\delta_n/2 \rceil - 1} e^{2i\pi k j/n} + \frac{1}{n} \sum_{j=\lfloor \frac{3n}{4} - \frac{n\delta_n}{2} \rfloor + 1}^{\lfloor \frac{3n}{4} \rfloor} e^{2i\pi k j/n}. \end{aligned}$$

Hence for  $k = 0$ ,

$$\begin{aligned} t_0^{(n)} &= t_0^{*(n)} = \frac{\lceil n\delta_n/2 \rceil + \lfloor \frac{3n}{4} \rfloor - \lfloor \frac{3n}{4} - \frac{n\delta_n}{2} \rfloor}{n} \\ &= \frac{1}{2} + \left( \delta_n - \frac{1}{2} \right) + \frac{\lceil n\delta_n/2 \rceil - n\delta_n/2 + \lfloor \frac{3n}{4} \rfloor - \frac{n\delta_n}{2} - \lfloor \frac{3n}{4} - \frac{n\delta_n}{2} \rfloor}{n} \\ &= \frac{1}{2} + O(n^{-1} + 1/2 - \delta_n), \end{aligned}$$

and therefore when  $n^2(1/2 - \delta_n) \rightarrow 0$ ,  $t_0^{*(n)} = \frac{1}{2} + O(n^{-1})$ .

Now, when  $k \neq 0$ ,

$$\begin{aligned} t_{-k}^{(n)} &= \frac{1}{n} \frac{1 - e^{2i\pi k \lceil n\delta_n/2 \rceil/n} + e^{2i\pi k (\lfloor \frac{3n}{4} - \frac{n\delta_n}{2} \rfloor + 1)/n} - e^{2i\pi k (\lfloor \frac{3n}{4} \rfloor + 1)/n}}{1 - e^{2i\pi k/n}} \\ &= \frac{1}{n} \frac{e^{\frac{i\pi k (2\lceil n\delta_n/2 \rceil)}{2n}} \left( e^{-\frac{i\pi k (2\lceil n\delta_n/2 \rceil)}{2n}} - e^{\frac{i\pi k (2\lceil n\delta_n/2 \rceil)}{2n}} \right)}{e^{\frac{i\pi k}{n}} \left( e^{-\frac{i\pi k}{n}} - e^{\frac{i\pi k}{n}} \right)} \\ &\quad - \frac{1}{n} \frac{e^{2i\pi k (\lfloor \frac{3n}{4} \rfloor + 1)/n} \left[ 1 - e^{-2i\pi k (\frac{n\delta_n}{2} + (\lfloor \frac{3n}{4} \rfloor - \frac{n\delta_n}{2} - \lfloor \frac{3n}{4} - \frac{n\delta_n}{2} \rfloor))/n} \right]}{e^{\frac{i\pi k}{n}} \left( e^{-\frac{i\pi k}{n}} - e^{\frac{i\pi k}{n}} \right)} \\ &= \frac{1}{n} \frac{e^{\frac{i\pi k \delta_n (\lceil n\delta_n/2 \rceil)}{\delta_n n/2}} \sin \left\{ \frac{\pi k \delta_n \lceil n\delta_n/2 \rceil}{2 \delta_n n/2} \right\}}{e^{\frac{i\pi k}{n}} \sin \frac{\pi k}{n}} \\ &\quad + \frac{1}{n} \frac{e^{\frac{3i\pi k}{2} (1 + \frac{\lfloor 3n/4 \rfloor - 3n/4 + 1}{3n/4})} e^{-\frac{i\pi k \delta_n}{2} \left( 1 + \frac{\lfloor \frac{3n}{4} \rfloor - \frac{n\delta_n}{2} - \lfloor \frac{3n}{4} - \frac{n\delta_n}{2} \rfloor}{n\delta_n/2} \right)} \sin \left\{ \frac{\pi k \delta_n}{2} \left( 1 + \frac{\lfloor \frac{3n}{4} \rfloor - \frac{n\delta_n}{2} - \lfloor \frac{3n}{4} - \frac{n\delta_n}{2} \rfloor}{n\delta_n/2} \right) \right\}}{e^{\frac{i\pi k}{n}} \sin \frac{\pi k}{n}} \\ &= \frac{e^{\frac{i\pi k \delta_n (\lceil n\delta_n/2 \rceil)}{\delta_n n/2}} \left( \sin \left\{ \frac{\pi k \delta_n \lceil n\delta_n/2 \rceil}{2 \delta_n n/2} \right\} \right)}{n e^{\frac{i\pi k}{n}} \sin \frac{\pi k}{n}} \\ &\quad + \frac{e^{\frac{i\pi k (3-\delta_n)}{2}} e^{\frac{3i\pi k}{2} \frac{\lfloor 3n/4 \rfloor - 3n/4 + 1}{3n/4} - \frac{i\pi k \delta_n}{2} \frac{\lfloor \frac{3n}{4} \rfloor - \frac{n\delta_n}{2} - \lfloor \frac{3n}{4} - \frac{n\delta_n}{2} \rfloor}{n\delta_n/2}} \left[ \sin \left\{ \frac{\pi k \delta_n}{2} \left( 1 + \frac{\lfloor \frac{3n}{4} \rfloor - \frac{n\delta_n}{2} - \lfloor \frac{3n}{4} - \frac{n\delta_n}{2} \rfloor}{n\delta_n/2} \right) \right\} \right]}{n e^{\frac{i\pi k}{n}} \sin \frac{\pi k}{n}} \\ &= \frac{\sin \frac{\pi k \delta_n}{2}}{n e^{\frac{i\pi k}{n}} \sin \frac{\pi k}{n}} \left[ e^{\frac{i\pi k \delta_n (\lceil n\delta_n/2 \rceil)}{\delta_n n/2}} \frac{\sin \left\{ \frac{\pi k \delta_n \lceil n\delta_n/2 \rceil}{2 \delta_n n/2} \right\}}{\sin \frac{\pi k \delta_n}{2}} + e^{\frac{i\pi k \delta_n (\lceil n\delta_n/2 \rceil)}{\delta_n n/2}} \frac{\sin \left\{ \frac{\pi k \delta_n \lceil n\delta_n/2 \rceil}{2 \delta_n n/2} \right\}}{\sin \frac{\pi k \delta_n}{2}} \right] \\ &\quad + e^{\frac{i\pi k (3-\delta_n)}{2}} e^{\frac{3i\pi k}{2} \frac{\lfloor 3n/4 \rfloor - 3n/4 + 1}{3n/4} - \frac{i\pi k \delta_n}{2} \frac{\lfloor \frac{3n}{4} \rfloor - \frac{n\delta_n}{2} - \lfloor \frac{3n}{4} - \frac{n\delta_n}{2} \rfloor}{n\delta_n/2}} \frac{\sin \left\{ \frac{\pi k \delta_n}{2} \left( 1 + \frac{\lfloor \frac{3n}{4} \rfloor - \frac{n\delta_n}{2} - \lfloor \frac{3n}{4} - \frac{n\delta_n}{2} \rfloor}{n\delta_n/2} \right) \right\}}{\sin \frac{\pi k \delta_n}{2}} \end{aligned}$$

Using the fact that  $e^{i(\frac{3\pi k}{2} - x)} = (-1)^k e^{i(\frac{\pi k}{2} - x)}$ , the previous expression can be rewritten as follows:

$$\begin{aligned}
t_{-k}^{(n)} &= \frac{\sin \frac{\pi k \delta_n}{2}}{n e^{\frac{i\pi k}{n}} \sin \frac{\pi k}{n}} \left[ e^{\frac{i\pi k \delta_n}{2}} + (-1)^k e^{\frac{i\pi k(1-\delta_n)}{2}} \right] + \zeta_k(\delta_n, n) \\
&= \frac{\delta_n}{2} \frac{\sin_c \frac{\pi k \delta_n}{2}}{e^{\frac{i\pi k}{n}} \sin_c \frac{\pi k}{n}} \left[ e^{\frac{i\pi k \delta_n}{2}} + (-1)^k e^{\frac{i\pi k(1-\delta_n)}{2}} \right] + \zeta_k(\delta_n, n),
\end{aligned}$$

where

$$\begin{aligned}
\zeta_k(\delta_n, n) &= \left( e^{\frac{i\pi k \delta_n}{2} \left( \frac{[n\delta_n/2] - n\delta_n/2}{n\delta_n/2} \right)} \right) \left( \frac{\sin \left\{ \frac{\pi k \delta_n}{2} \frac{[n\delta_n/2]}{n\delta_n/2} \right\}}{\sin \frac{\pi k \delta_n}{2}} - 1 \right) \\
&\quad - e^{\frac{i\pi k(1-\delta_n)}{2}} \left( e^{\frac{3i\pi k}{2} \frac{[3n/4] - 3n/4 + 1}{3n/4}} - \frac{i\pi k \delta_n}{2} \frac{[3n/4] - \frac{n\delta_n}{2} - \frac{[3n/4] - n\delta_n/2}{n\delta_n/2}}{n\delta_n/2} \right) \frac{\sin \left\{ \frac{\pi k \delta_n}{2} \left( 1 + \frac{[3n/4] - \frac{n\delta_n}{2} - \frac{[3n/4] - n\delta_n/2}{n\delta_n/2} \right) \right\}}{\sin \frac{\pi k \delta_n}{2}} - 1 \right).
\end{aligned}$$

Therefore

$$\begin{aligned}
t_{-k}^{(n)} &= \frac{\delta_n}{2} \frac{\sin_c \frac{\pi k \delta_n}{2}}{e^{\frac{i\pi k}{n}} \sin_c \frac{\pi k}{n}} e^{\frac{i\pi k}{4}} \left[ e^{-\frac{i\pi k(1/2-\delta_n)}{2}} + (-1)^k e^{\frac{i\pi k(1/2-\delta_n)}{2}} \right] + \zeta_k(\delta_n, n) \\
&= \delta_n \frac{\sin_c \frac{\pi k \delta_n}{2}}{\sin_c \frac{\pi k}{n}} e^{\frac{i\pi}{2} \left( \left( \frac{1}{2} - \frac{\delta_n}{n} \right) k \right)} \left[ 1_{\{k \text{ odd}\}} e^{-i\frac{\pi}{2}} \sin \frac{\pi k \left( \frac{1}{2} - \delta_n \right)}{2} + 1_{\{k \text{ even}\}} \cos \frac{\pi k \left( \frac{1}{2} - \delta_n \right)}{2} \right] + \zeta_k(\delta_n, n) \\
&= \delta_n e^{i\frac{\pi k}{4}} \sin_c \frac{\pi k \delta_n}{2} \left[ 1_{\{k \text{ odd}\}} e^{-i\frac{\pi}{2}} \sin \frac{\pi k \left( \frac{1}{2} - \delta_n \right)}{2} + 1_{\{k \text{ even}\}} \cos \frac{\pi k \left( \frac{1}{2} - \delta_n \right)}{2} \right] \left( \frac{e^{-i\pi \frac{k}{n}}}{\sin_c \frac{\pi k}{n}} \right) + \zeta_k(\delta_n, n) \\
&= \delta_n e^{i\frac{\pi k}{4}} \sin_c \frac{\pi k \delta_n}{2} \left[ 1_{\{k \text{ odd}\}} e^{-i\frac{\pi}{2}} \sin \frac{\pi k \left( \frac{1}{2} - \delta_n \right)}{2} + 1_{\{k \text{ even}\}} \cos \frac{\pi k \left( \frac{1}{2} - \delta_n \right)}{2} \right] + \xi_k(\delta_n, n) + \zeta_k(\delta_n, n),
\end{aligned}$$

where

$$\xi_k(\delta_n, n) = \delta_n e^{i\frac{\pi k}{4}} \sin_c \frac{\pi k \delta_n}{2} \left[ 1_{\{k \text{ odd}\}} e^{-i\frac{\pi}{2}} \sin \frac{\pi k \left( \frac{1}{2} - \delta_n \right)}{2} + 1_{\{k \text{ even}\}} \cos \frac{\pi k \left( \frac{1}{2} - \delta_n \right)}{2} \right] \left( \frac{e^{-i\pi \frac{k}{n}}}{\sin_c \frac{\pi k}{n}} - 1 \right).$$

It remains to be shown that both  $\xi_k(\delta_n, n)$  and  $\zeta_k(\delta_n, n)$  are  $O(n^{-1})$ . We use the fact that, as  $x \rightarrow 0$ ,  $\sin x = x + O(x^3)$ ,  $\sin_c x = 1 + O(x^2)$  and, when  $\sin a \neq 0$ ,  $\sin(a+x) = \sin a + x \cos a + O(x^2)$  and  $\sin_c(a+x) = \sin_c a + O(x)$ . Hence

$$\begin{aligned}
\xi_k(\delta_n, n) &= \left( \frac{1}{2} + O\left(\frac{1}{2} - \delta_n\right) \right) e^{i\frac{\pi(k-2)}{4}} \left( \sin_c \left( \frac{\pi k}{4} \right) + O\left(k \left( \frac{1}{2} - \delta_n \right)\right) \right) \\
&\quad \times \left[ \frac{\pi k \left( \frac{1}{2} - \delta_n \right)}{2} + O\left(k \left( \frac{1}{2} - \delta_n \right)^3\right) \right] \left( \frac{1 + O\left(\frac{k}{n}\right)}{1 + O\left(\frac{k}{n}\right)} - 1 \right) \\
&= \frac{1}{2} e^{i\frac{\pi(k-2)}{4}} \sin_c \left( \frac{\pi k}{4} \right) \left( 1 + O\left(\frac{1}{2} - \delta_n\right) \right) \\
&\quad \times \left[ \frac{\pi k \left( \frac{1}{2} - \delta_n \right)}{2} + O\left(k \left( \frac{1}{2} - \delta_n \right)^3\right) \right] \left( O\left(\frac{k}{n}\right) \right) \\
&= O\left(\frac{k^2}{n} \left( \frac{1}{2} - \delta_n \right)\right) = O\left(n \left( \frac{1}{2} - \delta_n \right)\right)
\end{aligned}$$

and therefore when  $n^2(1/2 - \delta_n) \rightarrow 0$ ,  $\xi_k(\delta_n, n) = O(n^{-1})$  while

$$\begin{aligned}
\zeta_k(\delta_n, n) &= (1 + O(n^{-1})) \left( \frac{\sin \frac{\pi k \delta_n}{2} + O(n^{-1})}{\sin \frac{\pi k \delta_n}{2}} - 1 \right) - e^{\frac{i\pi k(1-\delta_n)}{2}} \left( [1 + O(n^{-1})] \frac{\sin \frac{\pi k \delta_n}{2} + O(n^{-1})}{\sin \frac{\pi k \delta_n}{2}} - 1 \right) \\
&= O(n^{-1}).
\end{aligned}$$

Now

$$\begin{aligned}
t_{-k}^{*(n)} &= \operatorname{Re} \left( t_{-k}^{(n)} \right) \\
&= \delta_n \sin_c \frac{\pi k \delta_n}{2} \left[ 1_{\{k \text{ odd}\}} \sin \frac{\pi k}{4} \sin \frac{\pi k (\frac{1}{2} - \delta_n)}{2} + 1_{\{k \text{ even}\}} \cos \frac{\pi k}{4} \cos \frac{\pi k (\frac{1}{2} - \delta_n)}{2} \right] \\
&\quad + O(n^{-1}).
\end{aligned}$$

Notice that  $k(\frac{1}{2} - \delta_n) = o(n^{-1}) \forall k < n$  hence for  $k$  odd  $t_{-k}^{*(n)} = O(n^{-1})$ . When  $k$  is even, we need to consider the cases where there exists an odd integer  $m$  such that  $k = 4m$  or  $k = 4m + 2$ . First if  $k = 4m$  then  $\sin \frac{\pi k \delta_n}{2} = \sin 2\pi m \delta_n = O(m(\frac{1}{2} - \delta_n)) = o(n^{-1})$  and if  $k = 4m + 2$ , then  $\cos \frac{\pi k}{4} = \cos(m\pi + \frac{\pi}{2}) = 0$ . Hence for all  $k$  such that  $0 < |k| < n$

$$t_{-k}^{*(n)} = O(n^{-1}),$$

which concludes the proof of L(i).

Now we prove L(ii). We consider  $t_{-k}^{(n+1)} - t_{-k}^{(n)}$  for  $k \neq 0$  and assume  $t_{-k}^{(n)} \neq 0$ .

$$\begin{aligned}
t_{-k}^{(n+1)} &= \frac{1}{n} \left( 1 - \frac{1}{n+1} \right) \left[ \frac{1 - e^{2i\pi k \frac{[n\delta_n/2]}{n}} \left( 1 + \frac{n \frac{[(n+1)\delta_n/2] - (n+1)[n\delta_n/2]}{(n+1)[n\delta_n/2]} \right)}{1 - e^{2i\pi \frac{k}{n} \left( 1 - \frac{1}{(n+1)} \right)}} + \right. \\
&\quad \left. \frac{e^{2i\pi k \frac{[\frac{3n}{4} - \frac{n\delta_n}{2}] + 1}{n}} \left( 1 + \frac{n \left( [\frac{3(n+1)}{4} - \frac{(n+1)\delta_n}{2}] + 1 \right) - (n+1) \left( [\frac{3n}{4} - \frac{n\delta_n}{2}] + 1 \right)}{(n+1) \left( [\frac{3n}{4} - \frac{n\delta_n}{2}] + 1 \right)} \right)}{1 - e^{2i\pi \frac{k}{n} \left( 1 - \frac{1}{(n+1)} \right)}} \right. \\
&\quad \left. - \frac{e^{2i\pi k \frac{[\frac{3n}{4}] + 1}{n}} \left( 1 + \frac{n \left( [\frac{3(n+1)}{4}] + 1 \right) - (n+1) \left( [\frac{3n}{4}] + 1 \right)}{(n+1) \left( [\frac{3n}{4}] + 1 \right)} \right)}{1 - e^{2i\pi \frac{k}{n} \left( 1 - \frac{1}{(n+1)} \right)}} \right] \\
&= \frac{1}{n} (1 + O(n^{-1})) \frac{\left[ 1 - e^{2i\pi k \frac{[n\delta_n/2]}{n} + O(kn^{-1})} + e^{2i\pi k \frac{[\frac{3n}{4} - \frac{n\delta_n}{2}] + 1}{n} + O(kn^{-1})} - e^{2i\pi k \frac{[\frac{3n}{4}] + 1}{n} + O(kn^{-1})} \right]}{\left[ 1 - e^{2i\pi \frac{k}{n} + O(kn^{-2})} \right]} \\
&= t_{-k}^{(n)} (1 + O(kn^{-1}))
\end{aligned}$$

hence for  $t_{-k}^{(n)} \neq 0$ ,

$$\frac{t_{-k}^{*(n+1)} - t_{-k}^{*(n)}}{t_{-k}^{*(n)}} = O\left(\frac{k}{n}\right).$$

Now the proof of L(iii) follows from L(i) and the convergence of  $t_k^{(n)}$ :  $\forall \varepsilon > 0, \exists N$  such that for  $n \geq N$

$$\sup_{0 \leq |k| < n} t_k^{(n)} \leq \frac{1}{2} + \varepsilon.$$

hence choosing  $\varepsilon = 1/4$  is sufficient.

### 5.2.1. Proof of the validity of Assumption T

Assumptions T(i) and T(iii) follow from the definitions of  $g$  and  $\delta_n$ . To prove that Assumption T(ii) holds, we need to show that  $\mathbf{T}_{d,n}$  belongs to the Wiener class for all  $d \in (0, 1)$ . This follows

from the fact that the derivative

$$\frac{\partial}{\partial \omega} g(d, e^{i\omega})$$

is continuous at  $\omega = 0$ . Hence, the Fourier series of  $g(d, e^{i\omega})$  is absolutely summable at  $\omega = 0$  (see Whittaker, 1930-31), i.e.,

$$\lim_{n \rightarrow \infty} \sum_{k=-(n-1)}^{n-1} |c_{d,k}^{(n)}| < \infty.$$

Hence  $\sum_{k=-\infty}^{\infty} |t_{d,k}| < \infty$  and  $\sum_{k=-\infty}^{\infty} |t_{d,k}^*| < \infty$ .

### 5.2.2. Proof of the validity of Assumption B

We let  $j = 1$  without loss of generality. The assumption B(i) is that  $\sum_{k=1}^n \left| \frac{\widetilde{\mathbf{B}_n(e^{-i\omega})_{1k}}}{\det(\widetilde{\mathbf{B}_n(e^{-i\omega}))} \right|^2$  amounts to its first element  $\left| \frac{\widetilde{\mathbf{B}_n(e^{-i\omega})_{11}}}{\det(\widetilde{\mathbf{B}_n(e^{-i\omega}))} \right|^2$  as  $n \rightarrow \infty$  and B(ii) characterizes  $\widetilde{\mathbf{B}_n(e^{-i\omega})_{11}}$ .

We start by showing point B(ii) holds.

Elements  $\widetilde{\mathbf{B}_n(z)_{1k}}$ , for  $k = 1, \dots, n$ , of the first row of  $\widetilde{\mathbf{B}_n(z)}$ , satisfy  $\widetilde{\mathbf{B}_n(z)_{1k}} = (-1)^{k+1} \det(\mathbf{CoB}_n(z)_{1k})$ , where  $\mathbf{CoB}_n(z)_{ik}$  is the  $i, j$  entry of the matrix of cofactors of  $\mathbf{B}_n(z)$ . We consider first  $\mathbf{CoB}_n(z)_{11}$  which is

$$\mathbf{CoB}_n(z)_{11} = \begin{bmatrix} 1 - t_0^{*(n)} z & -\left(t_1^{*(n)} + \eta_n \gamma_{23}^{(n)}\right) z & \cdots & -\left(t_{n-2}^{*(n)} + \eta_n \gamma_{2n}^{(n)}\right) z \\ -\left(t_1^{*(n)} + \eta_n \gamma_{32}^{(n)}\right) z & 1 - t_0^{*(n)} z & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\left(t_1^{*(n)} + \eta_n \gamma_{(n-1)n}^{(n)}\right) z \\ -\left(t_{n-2}^{*(n)} + \eta_n \gamma_{n2}^{(n)}\right) z & \cdots & -\left(t_1^{*(n)} + \eta_n \gamma_{n(n-1)}^{(n)}\right) z & 1 - t_0^{*(n)} z \end{bmatrix},$$

where  $\gamma_{ik}$  denotes the  $i, j$  entry of  $\mathbf{D}_n$ .

Denoting respectively by  $\mathbf{T}_n^{*(1)}$  and  $\mathbf{D}_n^{(1)}$  the submatrices of  $\mathbf{T}_n^*$  and  $\mathbf{D}_n$  of dimension  $n-1$  obtained by removing their first row and first column,  $\mathbf{CoB}_n(z)_{11}$  can be written in a matrix form as

$$\mathbf{CoB}_n(z)_{11} = \mathbf{B}_{n-1}(z) + \left(\mathbf{T}_{n-1}^* - \mathbf{T}_n^{*(1)} + \eta_{n-1} \mathbf{D}_{n-1} - \eta_n \mathbf{D}_n^{(1)}\right) z$$

and therefore

$$\widetilde{\mathbf{B}_n(z)_{11}} = \det(\mathbf{B}_{n-1}(z)) \det\left(\mathbf{I}_{n-1} + [\mathbf{B}_{n-1}(z)]^{-1} \left(\mathbf{T}_{n-1}^* - \mathbf{T}_n^{*(1)} + \eta_{n-1} \mathbf{D}_{n-1} - \eta_n \mathbf{D}_n^{(1)}\right) z\right).$$

Consider first  $\mathbf{T}_{n-1}^* - \mathbf{T}_n^{*(1)} + \eta_{n-1} \mathbf{D}_{n-1} - \eta_n \mathbf{D}_n^{(1)}$ . Since  $\det(\mathbf{M}) \leq (n^{-1} \text{tr}(\mathbf{M}))^n$  for any symmetric matrix  $\mathbf{M}$  of dimension  $n$ , we can bound

$$\det\left(\mathbf{T}_{n-1}^* - \mathbf{T}_n^{*(1)} + \eta_{n-1} \mathbf{D}_{n-1} - \eta_n \mathbf{D}_n^{(1)}\right)$$

as follows

$$\begin{aligned} & \det\left(\mathbf{T}_{n-1}^* - \mathbf{T}_n^{*(1)} + \eta_{n-1} \mathbf{D}_{n-1} - \eta_n \mathbf{D}_n^{(1)}\right) \\ &= O\left(\left[\max\left(t_k^{*(n+1)} - t_k^{*(n)}\right) + \eta_{n-1} \max\left|\gamma_{ij}^{(n-1)}\right| + \eta_n \left|\max\gamma_{ij}^{(n)}\right|\right]^n\right). \end{aligned}$$

Lemma L(ii) provides the order of magnitude of the difference  $t_k^{*(n-1)} - t_k^{*(n)}$ , and we have assumed  $\eta_n \max_{ij} |\gamma_{ij}^{(n)}| = o(n^{-1})$ . Hence

$$\det \left( \mathbf{T}_{n-1}^* - \mathbf{T}_n^{*(1)} + \eta_{n-1} \mathbf{D}_{n-1} - \eta_n \mathbf{D}_n^{(1)} \right) = O \left( \left[ \frac{\max_{0 \leq k < n} k t_k^{(n)} + 1}{n} \right]^n \right)$$

and by Lemma L(iii), as  $n \rightarrow \infty$  :

$$\det \left( \mathbf{T}_{n-1}^* - \mathbf{T}_n^{*(1)} + \eta_{n-1} \mathbf{D}_{n-1} - \eta_n \mathbf{D}_n^{(1)} \right) = O \left( \left[ \delta_n + \frac{1}{n} + O \left( \frac{1}{n} \right) \right]^n \right).$$

We notice that for  $d \in (0, 1)$ ,  $[d + \frac{1}{n} + O(\frac{1}{n})]^n \sim \exp \{-n |\log d|\}$  as  $n \rightarrow \infty$ . Consequently,

$$\widetilde{\mathbf{B}}_n(z)_{11} = \det(\mathbf{B}_{n-1}(z)) \det \left( \mathbf{I}_{n-1} + O \left( e^{-n |\log \delta_n|} [\mathbf{B}_{n-1}(z)]^{-1} z \right) \right).$$

By construction, the determinant of  $\mathbf{B}_{n-1}(z)^{-1}$  is equivalent to that of  $(\mathbf{I}_{n-1} - \mathbf{C}_{n-1}z)^{-1}$ . For  $n$  large enough, the latter determinant is finite for all  $|z| < 1$  by Assumption T. Indeed the matrix  $\mathbf{C}_n$  is asymptotically equivalent to that with symbol  $g(\delta, \cdot)$ , whose eigenvalues are  $g \left( \delta, e^{i \frac{2\pi k}{n}} \right) \in [0, 1]$  for  $k = 0, \dots, n-1$ . Hence  $\forall z \in \mathbb{C}$ ,  $|z| < 1$ ,  $|\det(\mathbf{I}_{n-1} - \mathbf{C}_{n-1}z)| > 0$  so  $|\det(\mathbf{I}_{n-1} - \mathbf{C}_{n-1}z)^{-1}| < \infty$ . Hence  $\forall |z| < 1$ ,  $\det \mathbf{B}_{n-1}(z)^{-1}$  is finite and  $\det \left( \mathbf{I}_{n-1} + O \left( e^{-n |\log \delta_n|} [\mathbf{B}_{n-1}(z)]^{-1} z \right) \right) \rightarrow 1$ . It follows that

$$\widetilde{\mathbf{B}}_n(z)_{11} \underset{n \rightarrow \infty}{\sim} \det(\mathbf{B}_{n-1}(z)).$$

This constitutes the first part of the proof.

We now turn to showing point (i). We first show that  $\widetilde{\mathbf{B}}_n(z)_{1j} \rightarrow 0 \forall j \neq 1$  when  $|z| < 1$  as  $n \rightarrow \infty$ . By symmetry of the system, we can in fact focus the proof on  $\widetilde{\mathbf{B}}_n(z)_{12}$ . Ignoring  $\eta_n \mathbf{D}_n$  which is of lower order, as  $n \rightarrow \infty$  :

$$\widetilde{\mathbf{B}}_n(z)_{12} \sim - \det \left( \begin{bmatrix} -t_1^{*(n)} z & -t_1^{*(n)} z & \cdots & -t_{n-2}^{*(n)} z \\ -t_2^{*(n)} z & 1 - t_0^{*(n)} z & \ddots & \vdots \\ \vdots & \ddots & \ddots & -t_1^{*(n)} z \\ -t_{n-1}^{*(n)} z & -t_{n-3}^{*(n)} z & \cdots & 1 - t_0^{*(n)} z \end{bmatrix} \right).$$

The key feature that is shared by all the  $\widetilde{\mathbf{B}}_n(z)_{1j}$ , for  $j \neq 1$ , is that one of their columns (here the first) contains no element from the diagonal of  $\mathbf{B}_n(L)$  (where a 1 appears). Hence

$$\widetilde{\mathbf{B}}_n(z)_{12} \sim - \left( \max_{0 < k < n} |t_k^{*(n)}| z \right) \det \left( \begin{bmatrix} \frac{-t_1^{*(n)}}{\max_{0 < k < n} |t_k^{*(n)}|} & -t_2^{*(n)} z & \cdots & -t_{n-2}^{*(n)} z \\ \frac{-t_1^{*(n)}}{\max_{0 < k < n} |t_k^{*(n)}|} & 1 - t_0^{*(n)} z & \ddots & \vdots \\ \vdots & \vdots & \ddots & -t_1^{*(n)} z \\ \frac{-t_{n-1}^{*(n)}}{\max_{0 < k < n} |t_k^{*(n)}|} & -t_{n-3}^{*(n)} z & \cdots & 1 - t_0^{*(n)} z \end{bmatrix} \right).$$

Without loss of generality, we assume for instance that the maximum is  $|t_1^{*(n)}|$ . We have shown before that as  $n \rightarrow \infty$ ,

$$t_k^{*(n)} = O(n^{-1}).$$

Hence,

$$\widetilde{\mathbf{B}_n(z)}_{12} \sim O\left(\frac{z}{n}\right) \det \left( \begin{bmatrix} -1 & O(n^{-1}) \\ O(1) & \mathbf{B}_{n-2}(z) \end{bmatrix} \right).$$

The formula for the determinant of partitioned matrices is

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(D) \det(A - BD^{-1}C),$$

where in our case the second determinant on the right, i.e.,  $\det(-1 - O(n^{-1})\mathbf{B}_{n-2}(z)^{-1}O(1))$ , converges to  $-1$  as  $n \rightarrow \infty$  since  $\mathbf{B}_{n-2}(z)^{-1}$  is finite for  $|z| < 1$ . It follows that

$$\widetilde{\mathbf{B}_n(z)}_{12} = O(n^{-1} \det \mathbf{B}_{n-2}(z)).$$

Now,

$$\begin{aligned} \sum_{j=2}^n \left| \frac{\widetilde{\mathbf{B}_n(z)}_{1j}}{\det(\mathbf{B}_n(z))} \right|^2 &= \sum_{j=2}^n O\left( [n^{-1}]^2 \frac{|\det(\mathbf{B}_{n-2}(z))|^2}{|\det(\mathbf{B}_n(z))|^2} \right) \\ &= O\left( n^{-1} \frac{|\det(\mathbf{B}_{n-2}(z))|^2}{|\det(\mathbf{B}_n(z))|^2} \right). \end{aligned}$$

Hence as  $n \rightarrow \infty$ ,

$$\sum_{j=1}^n \left| \frac{\widetilde{\mathbf{B}_n(z)}_{1j}}{\det(\mathbf{B}_n(z))} \right|^2 = \frac{\sigma_\epsilon^2 |\det(\mathbf{B}_{n-1}(e^{-i\omega}))|^2}{2\pi |\det(\mathbf{B}_n(e^{-i\omega}))|^2} + O\left( \frac{1}{n} \frac{|\det(\mathbf{B}_{n-2}(e^{-i\omega}))|^2}{|\det(\mathbf{B}_n(e^{-i\omega}))|^2} \right). \quad (15)$$

The circulant matrix associated to  $\mathbf{B}_n(z)$  has symbol  $1 - g(\delta_n, \cdot)z$  since  $\mathbf{D}_n$  is antisymmetric. Hence, as  $n \rightarrow \infty$ , under assumption T (using the same argument than to prove Theorem 1),

$$\frac{\det(\mathbf{B}_{n-1}(z))}{\det(\mathbf{B}_n(z))} \sim \frac{\det(\mathbf{I}_{n-1} - z\mathbf{C}_{n-1})}{\det(\mathbf{I}_n - z\mathbf{C}_n)}. \quad (16)$$

The limit  $(1-z)^{-1/2}$  is finite for  $|z| < 1$  so

$$\frac{|\det(\mathbf{B}_{n-2}(z))|^2}{|\det(\mathbf{B}_n(z))|^2} = \frac{|\det(\mathbf{B}_{n-2}(z))|^2 |\det(\mathbf{B}_{n-1}(z))|^2}{|\det(\mathbf{B}_{n-1}(z))|^2 |\det(\mathbf{B}_n(z))|^2} \xrightarrow{(d,n) \rightarrow (1/2, \infty)} |1-z|^{-2}$$

hence in expression (15) the second term on the right-hand side is  $O(n^{-1})$ . Then together with (16), (15) implies that  $f_{n,x}(\omega)$  converges pointwise to  $\frac{\sigma_\epsilon^2}{2\pi} (1 - e^{-i\omega})^{-1}$  for all  $\omega \neq 0$  and uniformly on sets such that  $|\omega| > \omega_0$  for all  $\omega_0 > 0$ .

### 5.2.3. Proof of the validity of Assumption A

The assumption follows from Assumption T. By construction,  $\mathbf{T}_n^*$  is real valued and bounded. By transitivity of asymptotic equivalence (see Gray, 2006, Theorem 2.1),

$$\frac{\det(\mathbf{I}_{n-1} - \mathbf{T}_{n-1}^* z)}{\det(\mathbf{I}_n - \mathbf{T}_n^* z)} \sim \frac{\det(\mathbf{I}_{n-1} - \mathbf{C}_{n-1} z)}{\det(\mathbf{I}_n - \mathbf{C}_n z)} \sim \frac{\det(\mathbf{I}_{n-1} - \mathbf{T}_{n-1} z)}{\det(\mathbf{I}_n - \mathbf{T}_n z)}.$$

Now the circulant associated with  $\eta_n \mathbf{D}_n$  has zero asymptotic entries so  $\mathbf{A}_n \sim \mathbf{C}_n$  and the result follows.

## 6. References

- Abadir, K., Distaso, W., Giraitis, L., 2007. Nonstationarity-extended local Whittle estimation. *Journal of Econometrics* 141 (2), 1353–1384.
- Abadir, K. M., Talmain, G., 2002. Aggregation, persistence and volatility in a macro model. *Review of Economic Studies* 69 (4), 749–79.
- Altissimo, F., Mojon, B., Zaffaroni, P., 2009. Can aggregation explain the persistence of inflation? *Journal of Monetary Economics* 56 (2), 231–241.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2001. The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* 96 (453), 42–55.
- Andersen, T. G., Dobrev, D., Schaumburg, E., 2012. Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics* 169 (1), 75–93.
- Anderson, H. M., Vahid, F., 2007. Forecasting the volatility of Australian stock returns: Do common factors help? *Journal of Business & Economic Statistics* 25 (1), 76–90.
- Baillie, R. T., 1996. Long memory processes and fractional integration in econometrics. *Journal of Econometrics* 73, 5–59.
- Beran, J., 1994. *Statistics for Long-Memory Processes*. Chapman & Hall.
- Campbell, J. Y., Mankiw, N. G., 1987. Are output fluctuations transitory? *Quarterly Journal of Economics* 102 (4), 857–880.
- Chambers, M. J., 1998. Long memory and aggregation in macroeconomic time series. *International Economic Review* 39 (4), pp. 1053–1072.
- Cheung, Y.-W., Lai, K. S., 1993. A fractional cointegration analysis of purchasing power parity. *Journal of Business and Economic Statistics* 11 (1), 103–112.
- Chevillon, G., Mavroeidis, S., 2013. Learning can generate long memory. Working paper, ESSEC Business School.
- Cox, D. R., 2014. Scaling. Symposium presentation and video, Nuffield College, Oxford.
- Cox, D. R., Townsend, M. W. H., 1947. The use of the correlogram in measuring yarn irregularities. *Proceedings of the Royal Society of Edinburgh, Section A* 63, 290–311.
- Cubadda, G., Hecq, A., Palm, F. C., 2009. Studying co-movements in large multivariate data prior to multivariate modelling. *Journal of Econometrics* 148 (1), 25–35.
- Davidson, J., Hashimzade, N., 2009. Type I and type II fractional brownian motions: A reconsideration. *Computational Statistics & Data Analysis* 53 (6), 2089–2106.
- Davidson, J., Sibbertsen, P., 2005. Generating schemes for long memory processes: regimes, aggregation and linearity. *Journal of Econometrics* 128 (2), 253–82.
- Diebold, F. X., Inoue, A., 2001. Long memory and regime switching. *Journal of Econometrics* 105 (1), 131–159.
- Doornik, J., 2013. *Object-Oriented Matrix Programming Using Ox* (7th ed.). Timberlake Consultants Press.
- Doornik, J. A., Ooms, M., 2004. Inference and forecasting for ARFIMA models, with an application to US and UK inflation. *Studies in Nonlinear Dynamics and Econometrics* 8.
- Geweke, J., Porter-Hudak, S., 1983. The estimation and application of long memory time series models. *Journal of Time Series Analysis* 4, 221–38.
- Gourieroux, C., Jasiak, J., 2001. Memory and infrequent breaks. *Economics Letters* 70, 29–41.
- Granger, C. W. J., 1966. The typical spectral shape of an economic variable. *Econometrica: Journal of the Econometric Society*, 150–161.
- Granger, C. W. J., 1980. Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics* 14 (2), 227–238.
- Granger, C. W. J., Joyeux, R., 1980. An introduction to long-memory time series models and fractional differencing. *Journal of time series analysis* 1 (1), 15–29.
- Gray, R. M., 2006. *Toeplitz and circulant matrices: A review*. NOW Publishers Inc.



- Grenander, U., Szegő, G., 1958. Toeplitz forms and their applications. Vol. 321. University of California Press.
- Hecq, A., Laurent, S., Palm, F. C., 2012. On the univariate representation of BEKK models with common factors. Tech. rep., RM/12/018, Maastricht University.
- Hendry, D. F., 2009. The methodology of empirical econometric modeling: Applied econometrics through the looking-glass. In: Mills, T. C., Patterson, K. D. (Eds.), *Palgrave Handbook of Econometrics, Volume 2: Applied Econometrics*. Palgrave MacMillan, Basingstoke, Ch. 1, pp. 3–67.
- Lieberman, O., Phillips, P. C., 2008. Refined inference on long memory in realized volatility. *Econometric Reviews* 27 (1-3), 254–267.
- Mandelbrot, B., 1967. Some noises with  $1/f$  spectrum, a bridge between direct current and white noise. *Information Theory, IEEE Transactions on* 13 (2), 289–298.
- Marinucci, D., Robinson, P. M., 1999. Alternative forms of fractional Brownian motion. *Journal of Statistical Planning and Inference* 80 (1), 111–122.
- Martens, M., van Dijk, D., de Pooter, M., 2009. Modeling and forecasting S&P 500 volatility: Long memory, structural breaks and nonlinearity. *International Journal of Forecasting* 27, 282–303.
- Miller, J. I., Park, J. Y., 2010. Nonlinearity, nonstationarity, and thick tails: How they interact to generate persistence in memory. *Journal of Econometrics* 155 (1), 83 – 89.
- Müller, U. K., Watson, M. W., 2008. Testing models of low-frequency variability. *Econometrica* 76 (5), 979–1016.
- Parke, W. R., 1999. What is fractional integration? *Review of Economics and Statistics* 81 (4), 632–638.
- Perron, P., Qu, Z., 2007. An analytical evaluation of the log-periodogram estimate in the presence of level shifts. Working paper, Boston University.
- Robinson, P. M., 1995. Gaussian semiparametric estimation of long range dependence. *Annals of Statistics* 23, 1630–61.
- Schennach, S., 2013. Long memory via networking. Working paper CWP13/13, CEMMAP.
- Smith, H. F., 1938. An empirical law describing heterogeneity in the yields of agricultural crops. *The Journal of Agricultural Science* 28 (01), 1–23.
- Solo, V., 1992. Intrinsic random functions and the paradox of  $1/f$  noise. *SIAM Journal on Applied Mathematics* 52 (1), 270–291.
- Sowell, F., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics* 53, 165–188.
- Szegő, G., 1915. Ein Grenzwertsatz über die Toeplitzschen Determinanten einer reellen positiven Funktion. *Mathematische Annalen* 76 (4), 490–503.
- Whittaker, J. M., 1930-31. The absolute summability of Fourier series. *Proceedings of the Edinburgh Mathematical Society* 2 (2), 1–5.
- Zaffaroni, P., 2004. Contemporaneous aggregation of linear dynamic models in large economies. *Journal of Econometrics* 120 (1), 75–102.
- Zellner, A., Palm, F., 1974. Time series analysis and simultaneous equation econometric models. *Journal of Econometrics* 2 (1), 17–54.
- Zellner, A., Palm, F., 2004. *The structural econometric time series analysis approach*. Cambridge University Press.

---

**ESSEC Business School**

Avenue Bernard Hirsch  
CS 50105 Cergy  
95021 Cergy-Pontoise Cedex  
France  
Tél. +33 (0)1 34 43 30 00  
Fax +33 (0)1 34 43 30 01  
[www.essec.fr](http://www.essec.fr)

**ESSEC Executive Education**

CNIT BP 230  
92053 Paris-La Défense  
France  
Tél. +33 (0)1 46 92 49 00  
Fax +33 (0)1 46 92 49 90  
[www.executive-education.essec.fr](http://www.executive-education.essec.fr)

**ESSEC Asia-Pacific**

100 Victoria Street  
National Library # 13-02  
Singapore 188064  
Tél. +65 6884 9780  
Fax +65 6884 9781  
[www.essec.edu/asia](http://www.essec.edu/asia)

**Contact:**

Research Center  
+33 (0)1 34 43 33 58  
[research.center@essec.fr](mailto:research.center@essec.fr)

ISSN 1291-9616

