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

WORK IN PROGRESS

This Version: January 31, 2015

Guillaume Chevillon[a], Alain Hecq[b], Sébastien Laurent[c]

[a] *ESSEC Business School & CREST, France*
[b] *Department of Quantitative Economics, Maastricht University, The Netherlands*
[c] *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 cointegrated vector autoregression (VAR) models of finite orders can generate long memory in the marginalized univariate series. We use a final equation representation of a VAR(1) of large dimension n to show that as n tends to infinity while the proportion of stochastic trends remains constant, individual variables may tend to fractional white noises (whose degree corresponds to the fraction of unit roots in the system). Fractional integration may also appear in the presence of a unique stochastic trend, in which case the order of integration can be identical for all process. We consider the implications of our findings for the volatility of asset returns where the so-called golden-rule of realized volatility states that volatility always exhibits fractional integration of degree close to 0.4.

Keywords: Long memory, Vector Autoregressive Model, Marginalization, Volatility.

JEL: C10, C58

☆We would like to thank participants at Aix-Marseille University Seminar, 6th French Econometrics Conference celebrating Christian Gouriéroux's contribution to econometrics, 2014 Econometrics Conference at the Oxford Martin School, CFE 2013, 15th OxMetrics User Conference at CASS Business School and in particular Karim Abadir, David Hendry, 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 but its origin is unclear. One reason, as shown by Müller & Watson (2008), lies in the fact that very large samples are needed to discriminate between the various models generating strong dependence at low frequencies. To the best of our knowledge there exist five sources of long range dependence in economics: (i) aggregation across heterogeneous series, frequencies or economic agents (Granger 1980, Chambers, 1998, and inter alia Abadir & Talmain, 2002, Zaffaroni, 2004, Lieberman & Phillips, 2008 and Altissimo, Mojon & Zaffaroni, 2009); (ii) via linear modeling of a nonlinear underlying process (e.g. Davidson & Sibbertsen, 2005, Miller & Park, 2010); (iii) structural changes (Parke, 1999, Diebold & Inoue, 2001, Gourieroux and Jasiak, 2001, Perron and Qu, 2007); (iv) learning (bounded rationality) by economic agents in forward looking models of expectations (Chevillon & Mavroeidis, 2013) and (v) network effects (Schennach, 2013).

The contribution of this paper is to show that long memory can be the result of the marginalization of a multivariate model (e.g. a stationary VAR(1)) and therefore caused by some hidden cross section dependence. Here, we prove an equivalent representation between long range dependence through time and cross-section dependence between an infinite number of integrated and cointegrated variables. 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 when $n \to \infty$.

Our theoretical results present several implications in addition to showing a new source of long memory. First, we provide a formal arguments why small or hidden cross-section dependencies may invalidate inference in the time domain. Indeed we show that these dependencies may imply that standard time series inference techniques, relying, say, on the central limit theorem, fail to hold. Second, long memory may be a feature of univariate analysis that vanishes in large systems. This sheds a new light on fractional cointegration (that we still need to explore further). Third, our result may explain many empirical features. In our empirical illustration, we consider the case of realized volatilities for which an established stylized feature is that they tend to be integrated of an order close to 0.4. Our result provides a possible explanation why they exhibit fractional integration (through the interdependence between market-valued assets) and why the degree of long memory is roughly the same for all assets. In subsequent analyzes, we will consider alternative applications, such as disaggregate inflation indices whose persistence has been studied by a range of authors since the seminal works of Fairfield Smith (1938), Cox and Townsend (1947) and the Nile river study of Hurst (1951).

This paper is still in progress and is organized as follow. In Section 2, we provide a brief review of required results from the three literatures we combine. Section 3 gives our main theoretical results. We then show some simulations in Section 4 and some stylized empirical features in

Section 5.

2. Review of existing results

Our theoretical argument draws upon three existing literatures: those of long memory time series process, Final Equation Representations (FER) of Zellner and Palm (1974), and large dimensional Toeplitz matrices. For this reason, we review briefly the existing theoretical elements that provide the canvas for our result.

2.1. Long Memory

Several definitions of short memory processes are available (e.g., various mixing conditions, see White, 2000). These definitions are not equivalent, but they typically imply that short memory requires the boundedness of the variance of partial sums, scaled by the sample size T.[1] If this does not hold, the process is said to exhibit long memory. This is the definition adopted by Diebold and Inoue (2001) in their study on the connection between structural changes and long memory. They define the 'degree of memory' d of a process z_t as the smallest value (when it exists) such that

$$\mathsf{var}\left(T^{-1/2}\sum_{t=1}^{T} z_t\right) = O\left(T^{2d}\right). \tag{1}$$

If $d = 0$, the process exhibits short memory, while $d > 0$ corresponds to long memory ($d < 0$ is sometimes referred to as antipersistence).[2]

The above definition applies generally to any stochastic processes having finite second moments. For a covariance stationary process, where the autocorrelation function is a common measure of persistence, short memory requires absolute summability of its autocorrelation function, or a finite spectral density at zero. Thus, long memory arises when the autocorrelation coefficients are non-summable, or the spectrum has a pole at frequency zero. This gives rise to the following definitions of d, that are equivalent to (1) for covariance stationary processes, see Beran (1994) or Baillie (1996):

$$\begin{aligned} \rho_z(k) &\sim c_\rho k^{2d-1}, \qquad \text{as } k \to \infty \\ f_z(\omega) &\sim c_f |\omega|^{-2d}, \quad \text{as } \omega \to 0, \end{aligned} \tag{2}$$

for some positive constants c_ρ, c_f, where $\rho_z(k) = \mathsf{Corr}[z_t, z_{t+k}]$ is the autocorrelation function (ACF) of a covariance stationary stochastic process z_t and $f_z(\omega)$ is its spectral density. For $d > 0$, the autocorrelation function at long lags and the spectrum at low frequencies have the familiar hyperbolic shape that has traditionally been used to define long memory.

[1] Any definition of short memory that implies an invariance principle satisfies the restriction on the variance of partial sums, e.g., Andrews and Pollard (1994), Rosenblatt (1956), or White (2000).

[2] In the context of nonlinear cointegration, Gonzalo and Pitarakis (2006) have introduced the terminology "summable of order d" for processes that satisfy the definition given in (1) above, see also Berenguer-Rio and Gonzalo (2014).

Fractional integration, denoted I(d), is a well-known example of a class of processes that exhibit long memory. 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 I(d) process is the fractional white noise

$$z_t = (1 - L)^{-d} \epsilon_t, \tag{3}$$

where L denotes the lag operator and ϵ_t is a white noise sequence with variance σ_ϵ^2. The spectral density of z_t is then $f_z(\omega) = \sigma_\epsilon^2 \left[2 \sin \frac{\omega}{2}\right]^{-2d}$. In the following, we use the spectral density characterization of long memory.

2.2. Final Equation Representation

Final equation representations (FER) were proposed by Zellner and Palm (1974, 2004) who show how the elements of vector processes can be marginalized with respect to one another 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. For simplicity, we consider a n-vector $\mathbf{x}_t = (x_{1t}, ..., x_{nt})'$ admitting a vector autoregressive, $\mathsf{VAR}(1)$, representation:

$$\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t \tag{4a}$$

$$\epsilon_t = (\epsilon_{1t}, ..., \epsilon_{nt})' \overset{i.i.d}{\sim} \mathsf{N}(\mathbf{0}, \Omega_\epsilon). \tag{4b}$$

where Ω_ϵ is a diagonal matrix with diagonal $\left(\sigma_{\epsilon_1}^2, ..., \sigma_{\epsilon_n}^2\right)$. The Final Equation Representation (FER) of model (4) is

$$\det(\mathbf{B}_n(L)) \mathbf{x}_t = \widetilde{\mathbf{B}_n(L)} \epsilon_t, \tag{5}$$

where $\det(\mathbf{B}_n(L))$ and $\widetilde{\mathbf{B}_n(L)}$ are respectively the determinant and adjugate of $\mathbf{B}_n(L) = \mathbf{I}_n - \mathbf{A}_n L$. If $\mathbf{A}_n$ admits unitary eigenvalues, we implicitly assume that $\epsilon_t = 0$ for $t < 0$.[3]

Expression (5) shows that element x_{it}, obtained by marginalizing the n-dimensional VAR(1), admits a finite ARMA($n, n-1$) representation with common AR lag polynomial and possible common root cancellation between the AR and MA lag polynomials. Hence, as $n \to \infty$, the univariate process x_{it} without root cancellation will follow an ARMA(∞, ∞).

[3]In our results below, to avoid dwelling on the issue of finite vs. infinite history (in relation to type I and type II fractional Brownian motions but that may not be accurately defined in the VAR(1) setting with unit roots), we implicity consider that the date of interest, t is always larger than n.

For clarity of the exposition, consider the following trivariate example:

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

with FER $A(L)\mathbf{x}_t = \mathbf{B}(L)\epsilon_t$, where

$$A(L) = (1 - aL)\left(1 - \left(aL + \sqrt{2}b\right)L\right)\left(1 - \left(aL - \sqrt{2}b\right)L\right),$$

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

so that $\mathbf{B}(L)\epsilon_t$ defines a VMA(2), i.e. a vector moving average (VMA) of order 2.

The jth, $1 \leq j \leq n$, row of $\widetilde{\mathbf{B}_n(L)}$, denoted $\widetilde{\mathbf{B}_n(L)}_{j.}$, is

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

so x_{jt} admits the ARMA representation

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

and the spectral density f_{x_j} of x_{jt} hence satisfies

$$f_{n,x_j}(\omega) = \sum_{k=1}^{n} \left| \frac{\widetilde{\mathbf{B}_n(e^{-i\omega})}_{jk}}{\det(\mathbf{B}_n(e^{-i\omega}))} \right|^2 \sigma^2_{\epsilon_k}. \tag{6}$$

Expression (6) constitutes the basis of our theoretical argument, which we delineate in Section 2.

2.3. Circulant and Toeplitz matrices

We now briefly review the key theoretical elements of circulant and Toeplitz matrix theory that we use in our argument. Assume $\mathbf{C}_n$ a circulant matrix, i.e.

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

Define $g_{\mathbf{C}}$, the spectral density[4] of $\mathbf{C} = \lim_{n\to\infty} \mathbf{C}_n$, with Fourier coefficients $c_k = \lim_{n\to\infty} c_k^{(n)}$:

$$g_{\mathbf{C}}(\lambda) \stackrel{def}{=} \lim_{n\to\infty}\left(c_0^{(n)} + 2\sum_{k=1}^{n-1} c_k^{(n)} e^{k\lambda}\right); \quad c_k = \frac{1}{2\pi}\int_0^{2\pi} g_{\mathbf{C}}(\lambda)\, e^{-ik\lambda} d\lambda.$$

[4] *Caveat lector:* unfortunately, the literature on Circulant and Toeplitz matrices uses the terminolgy "spectral density" of the matrix, which here is not to be confused with the spectral density of the stochastic processes x_{jt}.

Then the eigenvalues of $\mathbf{C}_n$ are given by

$$\lambda_k = g_{\mathbf{C}}\left(\frac{2\pi k}{n}\right), \quad 0 \le k < n.$$

We assume $g_{\mathbf{C}}$ real valued and symmetric so $\mathbf{C}_n$ is symmetric; the matrix is real when its spectral density is 2π-periodic. Now as $n \to \infty$, provided that $g_{\mathbf{C}}$ belongs to the Wiener class (it suffices that the sequence (c_k) is absolutely summable) then Szegö's first theorem (see Szegö, 1915, and Gray, 2006) states that

$$\lim_{n\to\infty} \frac{\det\left(\mathbf{C}_{n-1}\right)}{\det\left(\mathbf{C}_n\right)} \to \exp\left\{-\frac{1}{2\pi}\int_0^{2\pi} \log\left(g_{\mathbf{C}}\left(\lambda\right)\right) d\lambda\right\}. \tag{7}$$

Now consider the Wiener class Toeplitz matrix

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

with spectral density $g_{\mathbf{T}}$. Its determinant is asymptotically (when $n \to \infty$) equivalent to the one of the associated circulant matrix

$$\tilde{\mathbf{C}}_n = \begin{bmatrix} t_0^{(n)} & t_{-1}^{(n)} + t_{n-1}^{(n)} & \cdots & t_{-(n-1)}^{(n)} + t_1^{(n)} \\ t_1^{(n)} + t_{-(n-1)}^{(n)} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_{-1}^{(n)} + t_{n-1}^{(n)} \\ t_{n-1}^{(n)} + t_{-1}^{(n)} & \cdots & t_1^{(n)} + t_{-(n-1)}^{(n)} & t_0^{(n)} \end{bmatrix}. \tag{8}$$

When $g_{\mathbf{T}}$ is symmetric, $\tilde{\mathbf{C}}$ admits (see Gray (2006), Section 4.4) the spectral density

$$g_{\tilde{\mathbf{C}}}\left(x\right) = \begin{cases} \frac{g_{\mathbf{T}}(x)+g_{\mathbf{T}}(2\pi-x)}{2}, & x \neq 0, \\ g_{\mathbf{T}}\left(0\right) & x = 0. \end{cases} \tag{9}$$

3. Two results

In this section, we show that long memory, in the form of a fractional white noise, can appear when considering a marginalized univariate process from a multivariate system of infinite dimension. We start by considering a VAR(1) process where one of the innovations dominates (in the limit) all the others. We then consider an alternative representation where all innovations play a symmetric role.

3.1. One dominant process

Consider the VAR(1) process defined in (4a-b) where we set $\mathbf{A}_n = \mathbf{C}_n$, i.e. a circulant matrix with spectral density

$$g_d^*\left(x\right) = 1_{\{|x|<2\pi d\}}, \quad d \in \left(0, 1/2\right), \tag{10}$$

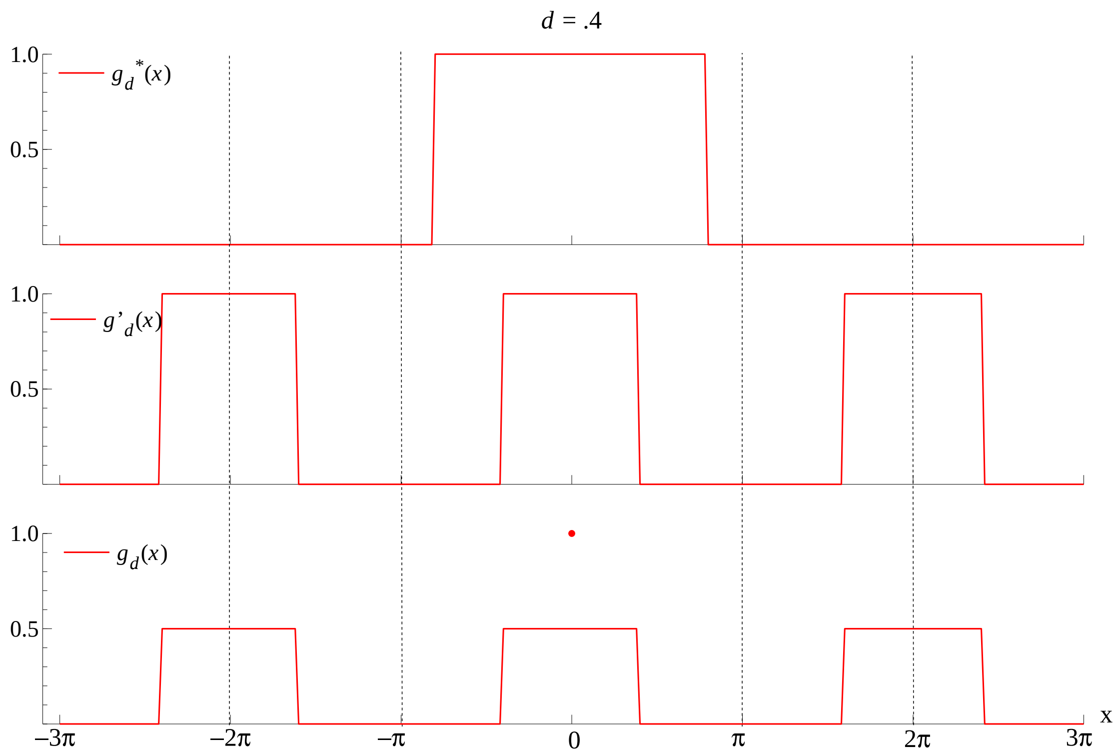

Figure 1: Spectral densities of the Circulant and Toeplitz matrices considered in this paper, i.e. $g_d^*(x)$, $g_d'(x)$ and $g_d(x)$ illustrated for $d = .4$

where $1_{\{.\}}$ denotes the indicator function. The eigenvalues of $\mathbf{C}_n$ satisfy:

$$\lambda_k = g_d^*\left(\frac{2\pi k}{n}\right) = \begin{cases} 1 & k = 0, ..., \lfloor nd \rfloor ; \\ 0 & k = \lfloor nd \rfloor + 1, ..., n-1 \end{cases}$$

so under g_d^*, $\mathbf{C}_n$ contains $\lceil nd \rceil - 1$ unit eigenvalues and $n - \lceil nd \rceil + 1$ zero eigenvalues.

The coefficients $c_k^{(n)}$ are complex and, as $n \to \infty$,

$$c_k^{(n)} \underset{n\to\infty}{\to} d\left[\cos \pi dk + i \sin \pi dk\right] \operatorname{sin}_c(\pi dk) \equiv c_k. \tag{11}$$

It is not necessary for our argument that $c_k^{(n)}$ be complex but it simplifies the exposition. The same result is obtained when considering the periodic function

$$g_d^{'}(x) = g_{d/2}^*(u), \quad \text{where } u \in [-\pi, \pi) \text{ and } x = u \bmod 2\pi \tag{12}$$

which generates a matrix $\mathbf{C}_n'$ that is real valued. For the ease of exposition, the spectral densities of all the matrices considered in this paper are reported in Figure 1.

The point of our first argument is to recognize the fact that for any element x_{jt} of $\mathbf{x}_t$, its spectral density expressed as in (6) is an infinite sum (as $n \to \infty$) of $\sigma_{\epsilon_j}^2$ times

$$\left|\frac{\widetilde{\mathbf{B}_n(e^{-i\omega})}_{jj}}{\det(\mathbf{B}_n(e^{-i\omega}))}\right|^2 \sim \left|\frac{\det\left(\mathbf{B}_{n-1}\left(e^{-i\omega}\right)\right)}{\det(\mathbf{B}_n(e^{-i\omega}))}\right|^2, \tag{13}$$

the value of (13) being given by Szegö's theorem. We now state our first theorem which considers the case where the variance of one innovation ϵ_{jt} dominates the others. For this we define $\sigma_{\backslash\epsilon_j}^2$ the vector of variances $\left(\sigma_{\epsilon_k}^2\right)$ for $k \neq j$. Our first theorem concerns the limiting behavior of x_{jt} when $n \to \infty$ and $\sigma_{\backslash\epsilon_j}^2 \to 0$.

Theorem 1. *Any element x_{jt} of the process $\mathbf{x}_t$ generated by (4a-b), where $A_n = C_n$ (i.e. a circulant matrix) with spectral density given either by (10) or its 2π-periodic extension (12), admits a spectral density f_{n,x_j} such that for $\omega > 0$, and as $\left(n, \sigma_{\backslash\epsilon_j}^2\right) \to (\infty, 0)$,*

$$f_{n,x}(\omega) \to f^{(d)}(\omega)\, \sigma_{\epsilon_j}^2,$$

where $f^{(d)}$ is the spectral density of a standard fractional white noise of degree d : $f^{(d)}(\omega) = \left|1 - e^{-i\omega}\right|^{-2d}$.

Theorem 1 shows that when the number of variables n tends to infinity and when one of the innovation processes dominates all the others, then a corresponding process entering $\mathbf{x}_t$ tends to be a fractional white noise when the modeler analyzes it separately from the others. Hence according to Theorem 1 long memory can result from the marginalization of a variable belonging to an system of infinite dimension (e.g a VAR(1)) with respect to all the other variables of this system. The vector process is the limit, as $\left(n, \sigma_{\backslash\epsilon_j}^2\right) \to (\infty, 0)$, of an integrated and cointegrated system comprising $\lceil nd \rceil$ common trends and $n - \lceil nd \rceil$ cointegration relations.

To our best 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 memory that we present, it is not aggregation that is at play here. The mechanism is closer in a sense to that which Schennach (2013) delineates of a single input that interacts within a system. But contrary to Schennach, the long memory process considered here acts as a form of common factor which drives the system.

Corollary 1 *The result of Theorem 1 extends to the cases where either*

(i) $\mathbf{A}_n = \mathbf{V}_n \mathbf{C}_n \mathbf{V}_n^{-1}$, *where* $\mathbf{C}_n$ *satisfies the conditions of the process and* $\mathbf{V}_n$ *is an orthogonal matrix of dimension* n;

or

(ii) The spectral density g^* *is replaced by*

$$g^*_{p,d}(x) = \exp\left\{-\left(\Gamma_{\frac{p+1}{p}} \frac{|x|}{2\pi d}\right)^p\right\},$$

(or a 2π*-periodic extension) and the limit is obtained as* $(d, p, n) \to \left(\frac{1}{2}, \infty, \infty\right)$.

Proof. (i) The proof follows from the fact that $\det(\mathbf{I}_n - \mathbf{A}_n L) = \det \mathbf{V}_n (\mathbf{I}_n - \mathbf{C}_n L) \mathbf{V}_n^{-1} = \det(\mathbf{I}_n - \mathbf{C}_n L)$ so as $n \to \infty$, the limits of $\frac{\det(\mathbf{I}_{n-1} - \mathbf{A}_{n-1} L)}{\det(\mathbf{I}_n - \mathbf{A}_n L)}$ and $\frac{\det(\mathbf{I}_{n-1} - \mathbf{C}_{n-1} L)}{\det(\mathbf{I}_n - \mathbf{C}_n L)}$ coincide.

(ii) see the appendix. ■

The above corollary shows that although Theorem 1 draws on tightly specified parametric assumptions, the results hold for a larger set of situations. Point (ii) in the corollary shows in particular that there exist processes within arbitrarily small neighborhoods of those considered in the theorem. As $g^*_{p,d}$ is continuous and so are its derivatives (if $p \in \mathbb{N}$), the off-diagonal coefficients of the resulting matrices will be much smaller than under g^*_d.

A drawback from the above theorem is that in the limit, the system is driven by a unique innovation. So instead, we consider next an alternative representation that is symmetric with respect to all innovations.

3.2. Symmetric case

We modify the previous result to consider the case where, in the VAR(1) process defined in (4), $\mathbf{A}_n = \mathbf{T}_n$ is symmetric Toeplitz with spectral density

$$g_d(x) = \begin{cases} \frac{1}{2} g^*_{d/2}(u), & u \neq 0 \\ 1 & u = 0 \end{cases} \quad \text{where } u \in [-\pi, \pi) \text{ and } x = u \bmod 2\pi. \tag{14}$$

The function $g_d(x)$ is chosen so that the Toeplitz matrix that it generates is real, symmetric, and with coefficients coinciding with the real part of the $c_k^{(n)}$ coefficients defined by g^*_d. Importantly, by construction, the determinant of $\mathbf{I}_n - z\mathbf{T}_n$ is asymptotically equivalent to that of $\mathbf{I}_n - z\tilde{\mathbf{C}}_n$ where $\tilde{\mathbf{C}}_n$ is the circulant matrix that is constructed from $\mathbf{T}_n$ as in expression (8). The spectral density of $\tilde{\mathbf{C}}_n$ is g'_d. Contrary to the case considered in the previous subsection, $\mathbf{T}_n$ itself does not contain a fixed fraction of unitary and zero eigenvalues, but $\tilde{\mathbf{C}}_n$ does. Now, as $(d, n) \to (1/2, \infty)$, the diagonal elements of $\mathbf{T}_n$ tend to d and the off-diagonal elements are of order $O\left(n^{-1} + d - 1/2\right)$, so $\mathbf{T}_n$ tends to a matrix that is close to (and probably in finite samples indistinguishable from) $d\mathbf{I}_n$.[5] We prove the following theorem in the appendix.

Theorem 2. *Any element x_{jt} of the process $\mathbf{x}_t$ generated by (4a-b), where $\mathbf{A}_n$ is Toeplitz with spectral density given by (14), admits a spectral density f_{n,x_j} such that for $\omega > 0$, and as $(d,n) \to (1/2,\infty)$, and $n(d-1/2) \to 0$,*

$$f_{n,x_j}(\omega) \to \sigma^2_{\epsilon_j} f^{(1/2)}(\omega),$$

where $f^{(1/2)}(\omega) = \left|1 - e^{-i\omega}\right|^{-1}$ is the spectral density of a standard fractional white noise of degree 1/2.

Theorem 2 shows that when the number of variables n entering in $\mathbf{x}_t$ tends to infinity and d tends to $1/2$, the spectral density of x_{jt}, for all j, is identical to the one of a fractionally integrated process of order $d \to 1/2$. The limiting ARFIMA$(0,1/2,0)$ process is often called an 1/f or flicker noise (see Mandelbrot, 1967). Theorem 2 differs from Theorem 1 in that no innovation dominates the others, so when $(d,n) \to (1/2,\infty)$ all processes x_{jt} (for $j = 1,...,n$) are independent fractional white noises.

Also, (this is still a conjecture drawn from simulations, we have not proved it formally yet) contrary to the previous theorem, Theorem 2 does not imply that $\mathbf{x}_t$ exhibits $\lceil nd \rceil$ common trends and $n - \lceil nd \rceil$ cointegration relations. Instead, it appears that the system only presents a unique stochastic trend so its rank of cointegration is $n-1$. Fractional integration can appear here via an infinite number of variables driven by a unique stochastic trend.

4. Monte Carlo

In this section, we provide some provisional simulation results based on the Toeplitz VAR(1) case considered in Section 3.2. Figures 2 and 3 report various elements regarding the simulated processes. The matrix coefficients $\mathbf{A}_n$ have spectral densities g_d with $d = 0.45$ and 0.30 respectively. Although Theorem 2 applies when $n \to \infty$ and $d \to 1/2$, we observe that the estimator biases are small despite the small sample available. The slope of the log periodogram is close to that of the corresponding fractional white noise. The correlogram exhibits the shape typical to long memory processes, namely hyperbolic decay.

The estimators reported in the figures are the MLE of d in an $I(d)$ fractional white noise (obtained by PcGive in OxMetrics 7, see Hendy and Doornik, 2001). Our results are robust to using a range of standard long-memory estimators (Geweke and Porter-Hudak, 1983, with various bandwidths, local Whittle likelihood of Robinson (1995) and Shimotsu and Phillips (2005).

In order to study the system further, Figure 4 reports some simulation results for the Toeplitz system for $n = 200$ and $T = 1000$ and leading to a theoretical value of $d = 0.45$. In this model

[5] Note however that the FER of a VAR(1) with $A_n = d\mathbf{I}_n$ leads to univariate AR(1) models with a common AR coefficient equal to d and therefore is not able to generate long memory.

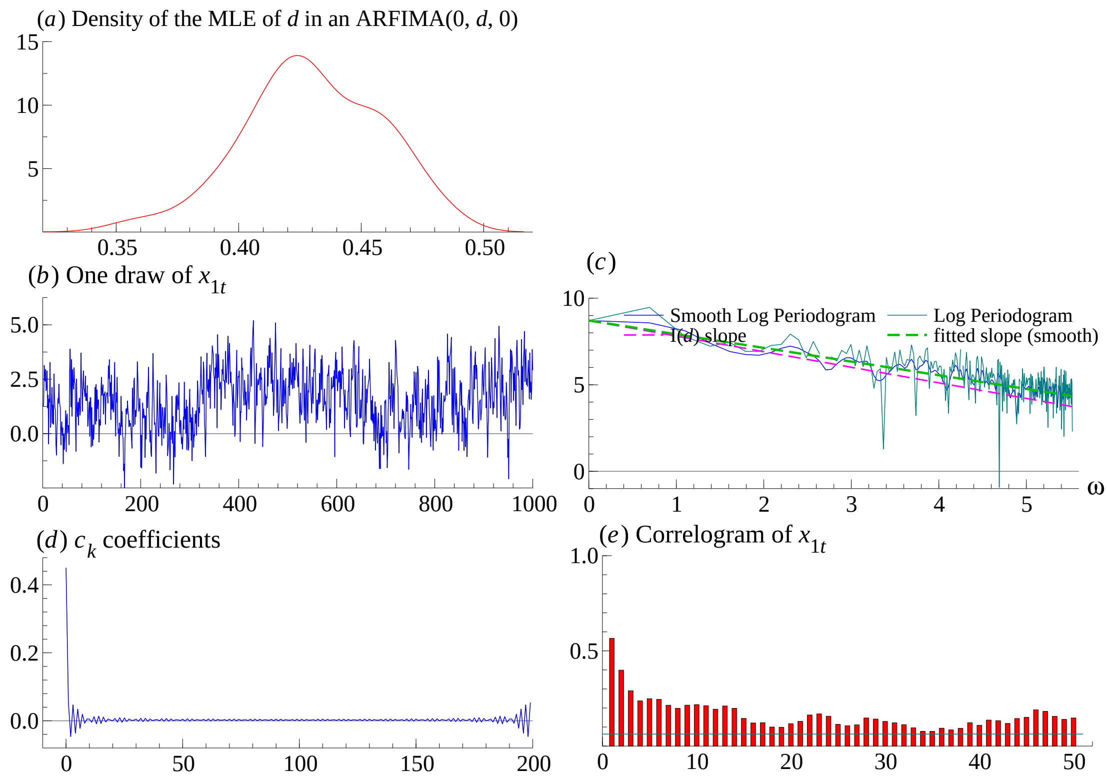

Figure 2: Simulation of stochastic processes obtained in the n-dimensional VAR(1) with Toeplitz matrix coefficient $\mathbf{A}_n$. The spectral density of $\mathbf{A}_n$ is g_d with $d = .45$. Here $n = 200$, $T = 1,000$ and the number of Monte Carlo replications is $M = 100$. The panels report, respectively, (a) the empirical density of MLE of d in an ARFIMA$(0, d, 0)$; (b) one draw of x_{1t}; (c) the Monte Carlo mean of its log (smooth or not) periodogram estimator together with the fitted slope and that of a theoretical ARFIMA$(0, 0.45, 0)$; (d) the values of the coefficients of matrix $\mathbf{A}_n$; and (e) the estimated autocorrelation function of x_{1t}.

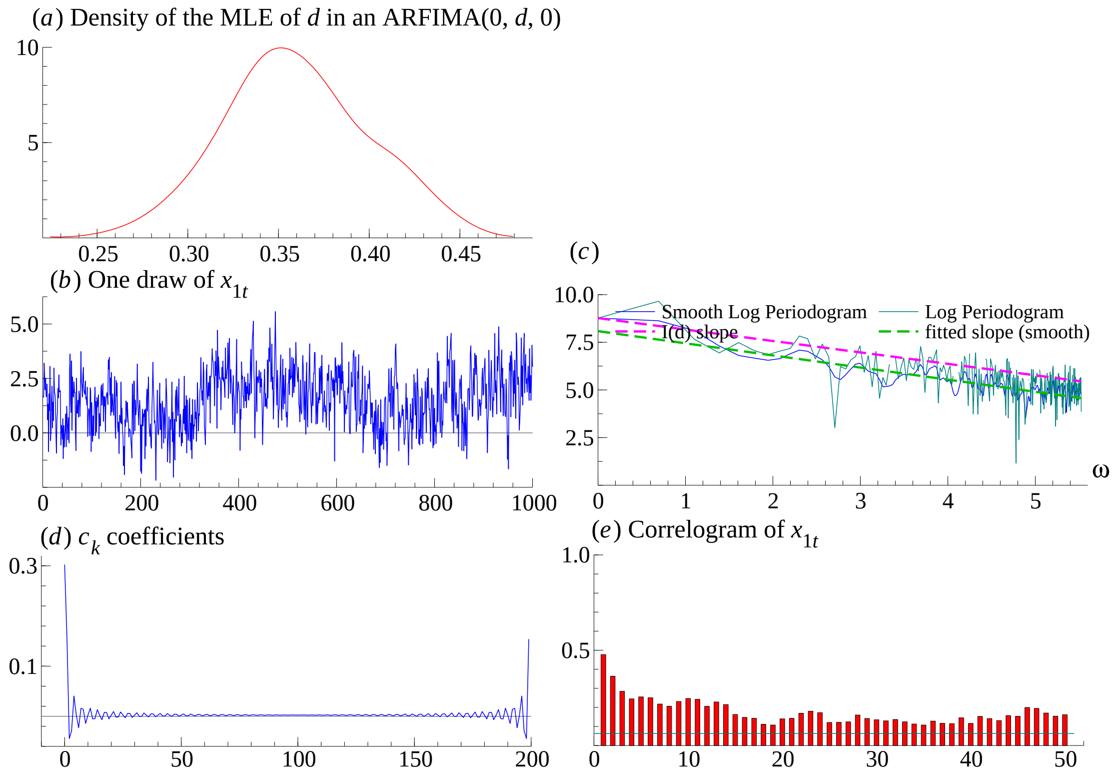

Figure 3: See note of Figure 2 but now $d = 0.3$.

$\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t$ so denoting $\mathbf{U}_n$ the matrix of eigenvectors of $\mathbf{A}_n$, $\mathbf{D}_n$ the diagonal matrix containing its eigenvalues, and $\mathbf{y}_t = \mathbf{U}_n^{-1}\mathbf{x}_t$, we see naturally that

$$\mathbf{y}_t = \mathbf{D}_n \mathbf{y}_{t-1} + \mathbf{U}_n^{-1}\epsilon_t,$$

and hence the element of $\mathbf{y}_t$ that corresponds to a unit eigenvalue of $\mathbf{A}_n$ follows a random walk which is an estimate of the common trend. When there is only a unique stochastic trend, it can be obtained as the cross-sectional average of the x_{jt}'s, i.e. $\overline{x}_t = n^{-1}\sum_{j=1}^{n} x_{jt}$. Figure 4 reports on Panel (a) draws of x_{jt} for $j = 1, 2$ and 3. The variables appear to be correlated and exhibit features that visually resemble to those of long memory processes. The common trend $\overline{x}_t$ is plotted in Panel (b) while Panels (c) and (d) correspond respectively to the correlograms of x_{1t} (which exhibits hyperbolic decay) and of the common trend (which exhibits a linear decay that is typical of I(1) processes). Finally, since in the Toeplitz framework considered above, as $n \to \infty$, the parameters of $\tilde{\mathbf{C}}_n$ tend to d on the diagonal and zero everywhere else, we report below a quick comparison of the GPH, Tapered (Hurvich and Chen, 2000), Whittle (Robinson, 1995) and ARFIMA estimators for x_{1t} obtained from $\mathbf{A}_n = \mathbf{T}_n$ as well as from a diagonal $d\mathbf{I}_n$ matrix.

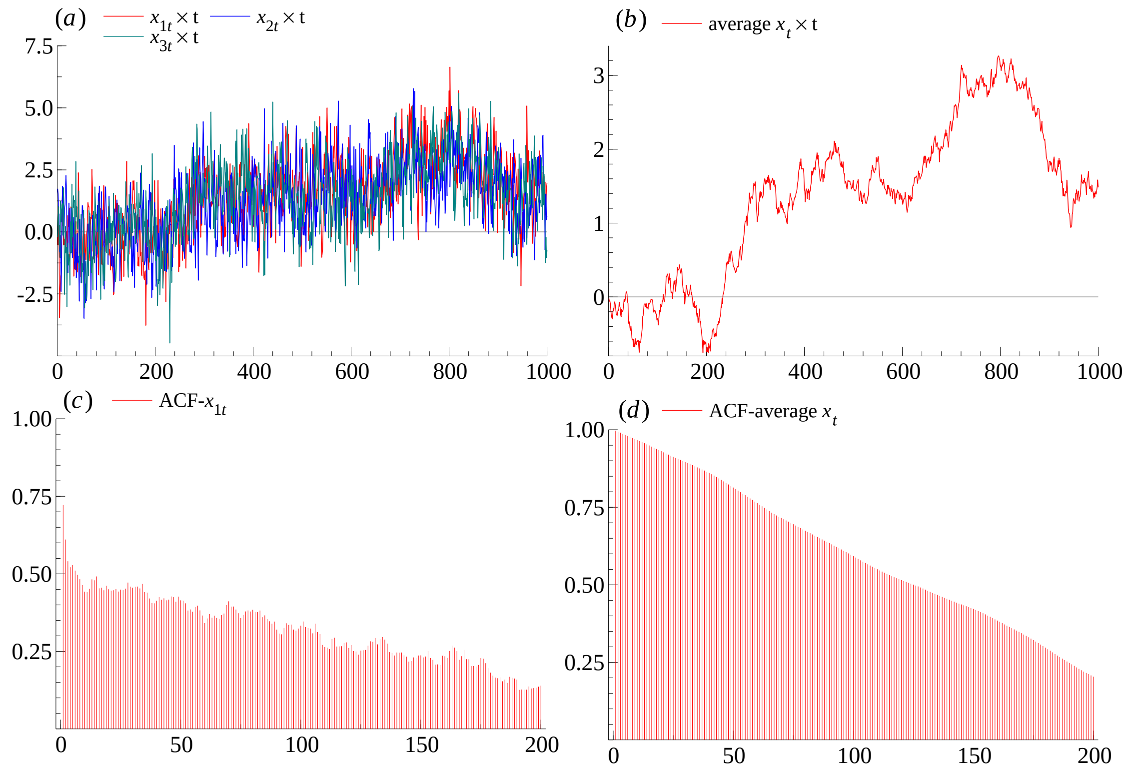

Figure 4: Simulation of stochastic processes obtained in the n-dimensional VAR(1) with Toeplitz matrix coefficient $\mathbf{A}_n$. The spectral density of $\mathbf{A}_n$ is g_d with $d = .45$. Here $n = 200$, $T = 1,000$. The panels report, respectively, (a) the first three series, i.e. x_{jt} for $j = 1, 2$ and 3; (b) the ensemble average (c) and (d) the correlograms of x_{1t} and the ensemble average respectively.

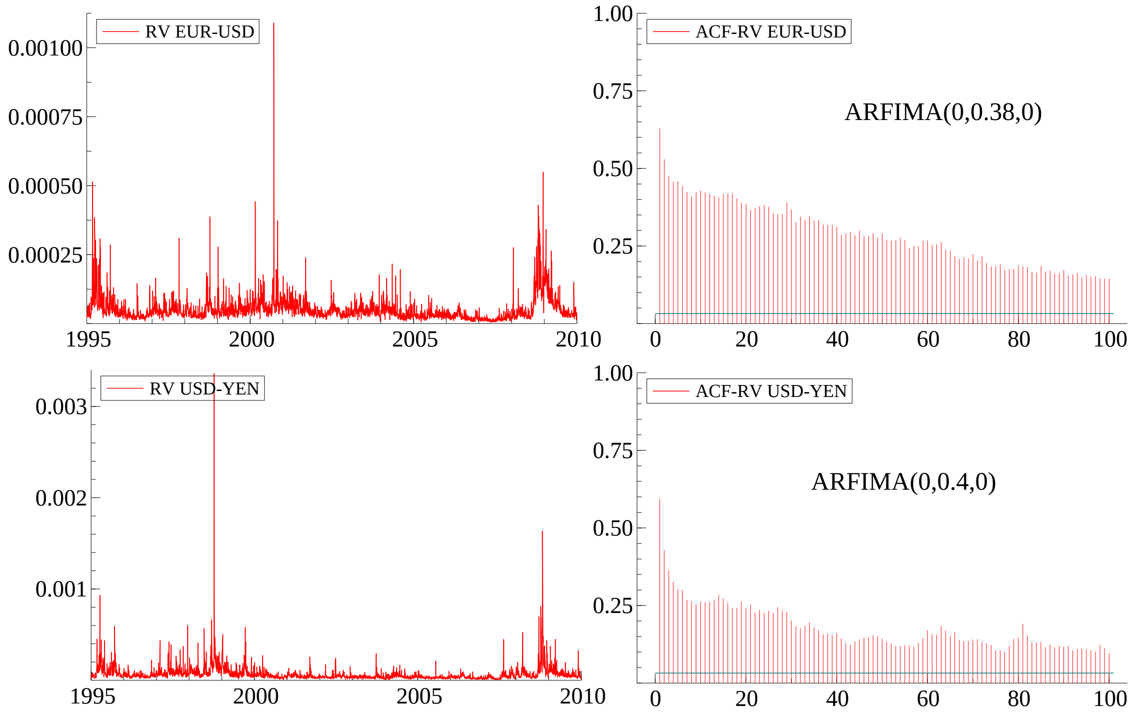

Figure 5: Daily Realized Volatility (estimated at the five minute level), Autocorrelations and fitted ARFIMA$\left(0, \widehat{d}, 0\right)$. The data are the Euro-USD and USD-JPY exchange rates.

Here $d = .4$ but similar results have been obtained for lower values $(.2, .3, \ldots)$.

$d = .4$	$\mathbf{A}_n = \mathbf{T}_n$				$\mathbf{A}_n = d\mathbf{I}_n$			
	GPH	Tapered	Whittle	ARFIMA	GPH	Tapered	Whittle	ARFIMA
$\widehat{d}$	.38	.39	.39	.41	.029	.028	.028	$-.004$
MC std dev	.025	.028	.028	.020	.023	.025	.018	.020

The table shows that although the matrix $\mathbf{T}_n$ is close to a diagonal matrix, the very small off-diagonal elements play a crucial role in the apparition of long memory.

5. Empirical stylized facts

We now present some stylized facts. 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 parameter d around 0.4 (Andersen et al., 2001; Martens et al., 2004[now 2009]).* Figure 5 presents examples of realized volatilities of exchange rates.

To provide a first assessment of the plausibility of our explanations, we consider a dataset (provided by TickData) consisting of transaction prices at the 5-minute sampling frequency for $N = 49$ large capitalization stocks from the NYSE, AMEX NASDAQ, covering the period from

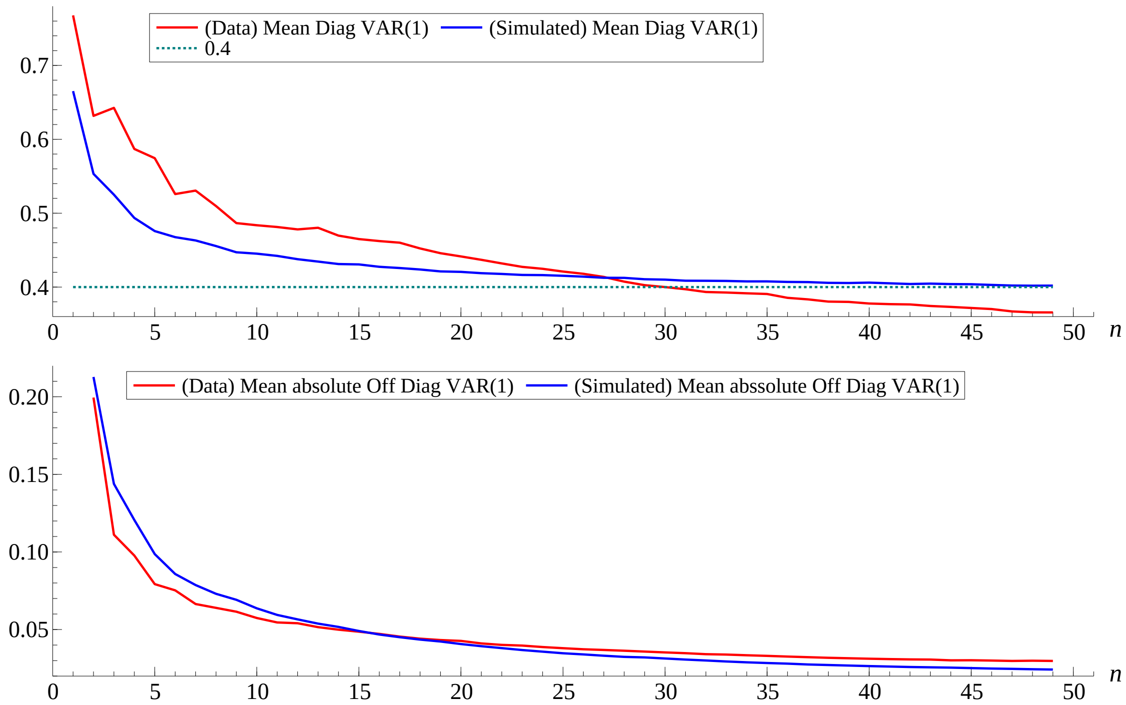

Figure 6: Average of the estimated diagonal elements (upper panel) and average of the absolute value of the off-diagonal elements (lower panels) of a VAR(1) estimated on $\log(MedRV)$ while progressively increasing the dimension of the VAR)

January 4, 1999 to December 31, 2008 (2,489 trading days). The trading session runs from 9:30 EST until 16:00 EST. The estimators of the realized variation is $MedRV$ (median RV) which is known to be robust to jumps.

A VAR(1) on $\log(MedRV)$ (progressively increasing the dimension of the VAR). Figure 6 plots the average of the estimated diagonal elements (upper panel) as well as the average of the absolute value of the off-diagonal elements (lower panels). We perform the same from the Toeplitz VAR model of Section 3.2 with $d = 0.4$ and report them in the same panels. We chose $d = 0.4$ without performing any estimation on the univariate series so the match is not perfect, but the patterns visually appear similar between the graphs.

6. Appendices

We first prove Theorem 2 as this involves the most details. Theorem 1 follows more easily.

6.1. Proof of Theorem 2

6.1.1. The coefficients

We prove here that the coefficients $t_k^{(n)}$ of $\mathbf{A}_n = \mathbf{T}_n$ satisfy, as $(d,n) \to (1/2, \infty)$

$$\begin{aligned} t_0^{(n)} &\to 1/2 \\ t_k^{(n)} &= O\left(n^{-1} + 1/2 - d\right), \quad k \neq 0. \end{aligned}$$

By construction, the coefficients $t_k^{(n)}$ coincide with the real part of those obtained by the discrete inverse Fourier transform of $g^*(\cdot) = 1_{\{|\cdot|<2\pi d\}}$; i.e. $t_k^{(n)} = \operatorname{Re} a_k^{(n)}$, where the latter satisfies

$$a_k^{(n)} = \frac{1}{n}\sum_{j=0}^{n-1} 1_{\left\{\frac{2\pi j}{n} < 2\pi d\right\}} e^{2i\pi kj/n} = \frac{1}{n}\sum_{j=0}^{\lceil nd\rceil - 1} e^{2i\pi kj/n},$$

where $\lceil\cdot\rceil$ denotes the *ceiling* of $\cdot$, i.e. the smallest integer at least as large as $\cdot$. (e.g. $\lceil 3\rceil = 3$ and $\lceil 3.1\rceil = 4$), since $\frac{2\pi j}{n} < 2\pi d$ for $j < nd$, i.e. $j \leq \lfloor nd\rfloor = \lceil nd\rceil - 1$ is $nd \notin \mathbb{N}$ and $j \leq nd - 1 = \lfloor nd\rfloor - 1 = \lceil nd\rceil - 1$ otherwise. Now

$$\begin{aligned} \frac{1}{n}\sum_{j=0}^{\lceil nd\rceil - 1} e^{2i\pi kj/n} &= \frac{1}{n}\frac{1 - e^{2i\pi k\lceil nd\rceil/n}}{1 - e^{2i\pi k/n}} = \frac{1}{n}\frac{e^{i\pi k\lceil nd\rceil/n}}{e^{i\pi k/n}}\frac{e^{-i\pi k\lceil nd\rceil/n} - e^{i\pi k\lceil nd\rceil/n}}{e^{-i\pi k/n} - e^{i\pi k/n}} \\ &= \frac{1}{n}\frac{e^{i\pi k\lceil nd\rceil/n}}{e^{i\pi k/n}}\frac{-2i\sin\left(\pi k\lceil nd\rceil/n\right)}{-2i\sin\left(\pi k/n\right)} = \frac{\lceil nd\rceil}{n}e^{i\pi k(\lceil nd\rceil - 1)/n}\frac{\operatorname{sin}_c\left(\pi k\lceil nd\rceil/n\right)}{\operatorname{sin}_c\left(\pi k/n\right)} \\ &= \frac{\lceil nd\rceil}{n}\left(\cos\frac{\pi k\left(\lceil nd\rceil - 1\right)}{n} + i\sin\frac{\pi k\left(\lceil nd\rceil - 1\right)}{n}\right)\frac{\operatorname{sin}_c\left(\pi k\lceil nd\rceil/n\right)}{\operatorname{sin}_c\left(\pi k/n\right)}. \end{aligned}$$

Hence

$$\begin{aligned} \operatorname{Re} a_k^{(n)} &= \frac{\lceil nd\rceil}{n}\cos\frac{\pi k\left(\lceil nd\rceil - 1\right)}{n}\frac{\operatorname{sin}_c\left(\pi k\lceil nd\rceil/n\right)}{\operatorname{sin}_c\left(\pi k/n\right)} \\ &= \frac{\lceil nd\rceil}{n}\left[\cos\frac{\pi k\lceil nd\rceil}{n}\cos\frac{\pi k}{n} + \sin\frac{\pi k\lceil nd\rceil}{n}\sin\frac{\pi k}{n}\right]\frac{\operatorname{sin}_c\left(\pi k\lceil nd\rceil/n\right)}{\operatorname{sin}_c\left(\pi k/n\right)} \\ \operatorname{Im} a_k^{(n)} &= \frac{\lceil nd\rceil}{n}\sin\frac{\pi k\left(\lceil nd\rceil - 1\right)}{n}\frac{\operatorname{sin}_c\left(\pi k\lceil nd\rceil/n\right)}{\operatorname{sin}_c\left(\pi k/n\right)} \\ &= \frac{\lceil nd\rceil}{n}\left[\sin\frac{\pi k\lceil nd\rceil}{n}\cos\frac{\pi k}{n} - \cos\frac{\pi k\lceil nd\rceil}{n}\sin\frac{\pi k}{n}\right]\frac{\operatorname{sin}_c\left(\pi k\lceil nd\rceil/n\right)}{\operatorname{sin}_c\left(\pi k/n\right)}. \end{aligned}$$

Note that as $(d,n) \to (1/2^-, \infty)$, $\operatorname{sin}_c(\pi k/n) = 1 + O\left((k/n)^{-4}\right)$ so $a_0^{(n)} \to d$ and, for $k > 0$,

$$\begin{aligned} a_k^{(n)} &= \frac{\lceil nd\rceil}{n}\cos\frac{\pi k\left(\lceil nd\rceil - 1\right)}{n}\frac{\operatorname{sin}_c\left(\pi k\lceil nd\rceil/n\right)}{\operatorname{sin}_c\left(\pi k/n\right)} \\ &= \left(d + \frac{\lceil nd\rceil - nd}{n}\right)\cos\left(\frac{\pi k}{2} + \pi k\frac{\left(\lceil nd\rceil - n/2 - 1\right)}{n}\right)\frac{\operatorname{sin}_c\left(\frac{\pi k}{2} + \pi k\frac{\lceil nd\rceil - n/2}{n}\right)}{\operatorname{sin}_c\left(\pi k/n\right)} \\ &= \left(d + \frac{\lceil nd\rceil - nd}{n}\right)\cos\left(\frac{\pi k}{2} + \pi k\frac{\left(\lceil nd\rceil - n/2 - 1\right)}{n}\right) \\ &\times \frac{\sin\left(\frac{\pi k}{2} + \pi k\frac{\lceil nd\rceil - n/2}{n}\right)}{\sin\left(\pi k/n\right)}\frac{\pi k/n}{\frac{\pi k}{2} + \pi k\frac{\lceil nd\rceil - n/2}{n}}. \end{aligned}$$

Now, using $\cos a \sin b = \frac{\sin(a+b)+\sin(a-b)}{2}$

$$
\begin{aligned}
a_k^{(n)} &= \left(d + \frac{\lceil nd \rceil - nd}{n}\right) \\
&\times \frac{\sin\left(\frac{\pi k}{2} + \pi k \frac{(\lceil nd \rceil - n/2 - 1)}{n} + \frac{\pi k}{2} + \pi k \frac{\lceil nd \rceil - n/2}{n}\right) + \sin\left(\frac{\pi k}{2} + \pi k \frac{\lceil nd \rceil - n/2}{n} - \frac{\pi k}{2} - \pi k \frac{(\lceil nd \rceil - n/2 - 1)}{n}\right)}{2} \\
&\times \frac{1}{\sin(\pi k/n)} \frac{\pi k/n}{\frac{\pi k}{2} + \pi k \frac{\lceil nd \rceil - n/2}{n}} \\
&= \left(d + \frac{\lceil nd \rceil - nd}{n}\right) \frac{\sin\left(\pi k + 2\pi k \frac{(\lceil nd \rceil - n/2 - 1/2)}{n}\right) + \sin\left(\frac{\pi k}{n}\right)}{2} \frac{1}{\sin(\pi k/n)} \frac{\pi k/n}{\frac{\pi k}{2} + \pi k \frac{\lceil nd \rceil - n/2}{n}}.
\end{aligned}
$$

Now $\sin(\pi k + a) = (-1)^k \sin a$

$$
\begin{aligned}
a_k^{(n)} &= \left(d + \frac{\lceil nd \rceil - nd}{n}\right) \left[(-1)^k \sin\left(2\pi k \frac{(\lceil nd \rceil - n/2 - 1/2)}{n}\right) + \sin\left(\frac{\pi k}{n}\right)\right] \\
&\times \frac{1}{\sin(\pi k/n)} \frac{2/n}{1 + 2\frac{\lceil nd \rceil - n/2}{n}} \\
&= \frac{2/n}{1 + 2\frac{\lceil nd \rceil - n/2}{n}} \left[\frac{1}{2} + \frac{\lceil nd \rceil - n/2}{n}\right] \left[1 + (-1)^k \frac{\sin\left(2\pi k \frac{(\lceil nd \rceil - n/2 - 1/2)}{n}\right)}{\sin(\pi k/n)}\right]
\end{aligned}
$$

i.e. denoting $d = 1/2 - \epsilon_d$, then

$$
\begin{aligned}
\lceil nd \rceil - n/2 = \lceil n(1/2 - \epsilon_d) \rceil - n/2 &= (\lceil n/2 - n\epsilon_d \rceil - \lceil n/2 \rceil) - (\lceil n/2 \rceil - n/2) \\
&= O(n\epsilon_d) + 1_{\{n \text{ odd}\}}/2
\end{aligned}
$$

so

$$
a_k^{(n)} = \frac{1}{n}\left[1 + (-1)^k \frac{\sin\left(2\pi k O(\epsilon_d) - 1_{\{n \text{ even}\}} \pi k/n\right)}{\sin(\pi k/n)}\right],
$$

where either $k = O(n)$ or $k = o(n)$.

First if $n/k = O(1)$, then $\sin(\pi k/n)$ does not tend to zero and $a_k^{(n)} = O\left(n^{-1}\right)$, now if $k = o(n)$, we assume $k\epsilon_d = o(1)$ then

$$
\begin{aligned}
a_k^{(n)} &\sim \frac{1}{n}\left[1 + (-1)^k \left[O(n\epsilon_d) - 1_{\{n \text{ even}\}}\right]\right] \\
&= O(\epsilon_d) + O\left(n^{-1}\right).
\end{aligned}
$$

This can be formalized as

$$
a_k^{(n)} = O\left(n^{-1} + 1/2 - d\right). \tag{15}
$$

QED.

6.1.2. Using Final Equations

The point of the proof is to show that

$$
\sum_{j=1}^{n} \left| \frac{\widetilde{\mathbf{B}_n\left(e^{-i\omega}\right)}_{1j}}{\det\left(\mathbf{B}_n\left(e^{-i\omega}\right)\right)} \right|^2 \tag{16}
$$

amounts to its first element $\left|\frac{\widetilde{\mathbf{B}_n(e^{-i\omega})}_{11}}{\det(\mathbf{B}_n(e^{-i\omega}))}\right|^2$ as $(d,n)\to\left(\frac{1}{2},\infty\right)$ and also to characterize $\widetilde{\mathbf{B}_n\left(e^{-i\omega}\right)}_{11}$ in terms of $\det\left(\mathbf{B}_{n-1}\left(e^{-i\omega}\right)\right)$.

We start with the latter, i.e. by showing that $\widetilde{\mathbf{B}_n\left(e^{-i\omega}\right)}_{11}$ is asymptotically equivalent to $\det\left(\mathbf{B}_{n-1}\left(e^{-i\omega}\right)\right)$.

We first consider the elements $\widetilde{\mathbf{B}_n\left(z\right)}_{1j}$ of the first row of $\widetilde{\mathbf{B}_n\left(z\right)}$, i.e.

$$\widetilde{\mathbf{B}_n\left(z\right)}_{11}=\det\left(\begin{bmatrix}1-t_0^{(n)}z & -t_1^{(n)}z & \cdots & -t_{n-2}^{(n)}z\\ -t_1^{(n)}z & 1-t_0^{(n)}z & \ddots & \vdots\\ \vdots & \ddots & \ddots & -t_1^{(n)}z\\ -t_{n-2}^{(n)}z & \cdots & -t_1^{(n)}z & 1-t_0^{(n)}z\end{bmatrix}\right).$$

This is a matrix of dimension $n-1$ which resembles $\mathbf{B}_{n-1}\left(z\right)$ except that the coefficients are the $t_k^{(n)}$ where as the $t_k^{(n-1)}$ enter $\mathbf{B}_{n-1}\left(z\right)$. Hence we rewrite the matrix as:

$$\widetilde{\mathbf{B}_n\left(z\right)}_{11}=\det\left(\mathbf{B}_{n-1}\left(z\right)+\begin{bmatrix}t_0^{(n-1)}-t_0^{(n)} & t_1^{(n-1)}-t_1^{(n)} & \cdots & t_{n-2}^{(n-1)}-t_{n-2}^{(n)}\\ t_1^{(n-1)}-t_1^{(n)} & t_0^{(n-1)}-t_0^{(n)} & \ddots & \vdots\\ \vdots & \ddots & \ddots & t_1^{(n-1)}-t_1^{(n)}\\ t_{n-2}^{(n-1)}-t_{n-2}^{(n)} & \cdots & t_1^{(n-1)}-t_1^{(n)} & t_0^{(n-1)}-t_0^{(n)}\end{bmatrix}z\right).$$

Denote by $\mathbf{T}_n^{(1)}$ the symmetric submatrix of $\mathbf{T}_n$ of dimension $n-1$ obtained by removing the last row and column. Then

$$\begin{aligned}\widetilde{\mathbf{B}_n\left(z\right)}_{11}&=\det\left(\mathbf{B}_{n-1}\left(z\right)+\left(\mathbf{T}_{n-1}-\mathbf{T}_n^{(1)}\right)z\right)\\&=\det\left(\mathbf{B}_{n-1}\left(z\right)\right)\det\left(\mathbf{I}_{n-1}+\left[\mathbf{B}_{n-1}\left(z\right)\right]^{-1}\left(\mathbf{T}_{n-1}-\mathbf{T}_n^{(1)}\right)z\right).\end{aligned}$$

Consider $\mathbf{T}_{n-1}-\mathbf{T}_n^{(1)}$. It consists of the deviations $t_k^{(n-1)}-t_k^{(n)}$. For ease of the proof, we consider the coefficients obtained by the inverse Fourier transform of the function $g^*\left(\cdot\right)=1_{\{|\cdot|<2\pi d\}}$ since these are easier to compute; recall that the real part of the latter coincide with the coefficients of interest (which are real). Since $n\geq\lceil nd\rceil-1$, for $d\leq 1/2$, we express differences as

$$\begin{aligned}a_k^{(n+1)}-a_k^{(n)}&=\frac{1}{n+1}\sum_{j=0}^{n}g_d\left(\frac{2\pi j}{n+1}\right)e^{2i\pi kj/(n+1)}-\frac{1}{n}\sum_{j=0}^{n-1}g_d\left(\frac{2\pi j}{n}\right)\\&=\frac{1}{n+1}g_d\left(\frac{2\pi n}{n+1}\right)e^{2i\pi kn/(n+1)}\\&+\frac{1}{n}\sum_{j=0}^{n-1}\left[\frac{n}{n+1}g_d\left(\frac{2\pi j}{n+1}\right)e^{2i\pi kj\left[\frac{1}{n+1}-\frac{1}{n}\right]}-g_d\left(\frac{2\pi j}{n}\right)\right]e^{2i\pi kj/n}.\end{aligned}$$

Now since $\frac{2\pi j}{n} > \frac{2\pi j}{n+1}$, $g_d\left(\frac{2\pi j}{n}\right) = 1$ implies $g_d\left(\frac{2\pi j}{n+1}\right) = 1$, so

$$\begin{aligned}
&\sum_{j=0}^{n-1}\left[\frac{n}{n+1}g_d\left(\frac{2\pi j}{n+1}\right)e^{2i\pi kj\left[\frac{1}{n+1}-\frac{1}{n}\right]} - g_d\left(\frac{2\pi j}{n}\right)\right]e^{2i\pi kj/n} \\
&= \sum_{j=0}^{\lceil nd\rceil-1}\left[\frac{n}{n+1}g_d\left(\frac{2\pi j}{n+1}\right)e^{-\frac{2i\pi kj}{n(n+1)}} - g_d\left(\frac{2\pi j}{n}\right)\right]e^{2i\pi kj/n} \\
&+ \sum_{j=\lceil nd\rceil}^{n-1}\frac{n}{n+1}g_d\left(\frac{2\pi j}{n+1}\right)e^{-\frac{2i\pi kj}{n(n+1)}} \\
&= \sum_{j=0}^{\lceil nd\rceil-1}\left[\frac{n}{n+1}e^{-\frac{2i\pi kj}{n(n+1)}} - 1\right]e^{2i\pi kj/n} + \sum_{j=\lceil nd\rceil}^{\lceil (n+1)d\rceil-1}\frac{n}{n+1}g_d\left(\frac{2\pi j}{n+1}\right)e^{-\frac{2i\pi kj}{n(n+1)}}. \\
&= \sum_{j=0}^{\lceil nd\rceil-1}\left[\frac{n}{n+1}e^{-\frac{2i\pi kj}{n(n+1)}} - 1\right]e^{2i\pi kj/n} + \frac{n}{n+1}g_d\left(\frac{2\pi\lceil nd\rceil}{n+1}\right)e^{-\frac{2i\pi kj}{n(n+1)}}.
\end{aligned}$$

Now $n \geq d/(1-d)$, so $g_d\left(\frac{2\pi n}{n+1}\right) = 0$ then

$$\begin{aligned}
a_k^{(n+1)} - a_k^{(n)} &= \frac{1}{n}\sum_{j=0}^{\lceil nd\rceil-1}\left[\frac{n}{n+1}e^{-\frac{2i\pi kj}{n(n+1)}} - 1\right]e^{2i\pi kj/n} \\
&+ \frac{1}{n+1}g_d\left(\frac{2\pi\lceil nd\rceil}{n+1}\right)e^{-\frac{2i\pi k\lceil nd\rceil}{n(n+1)}}
\end{aligned}$$

with $\frac{n}{n+1}e^{-2i\pi kj\frac{1}{n(n+1)}} - 1 = O(k/n)$. Hence

$$\frac{1}{n}\sum_{j=0}^{\lceil nd\rceil-1}\left[\frac{n}{n+1}e^{-\frac{2i\pi kj}{n(n+1)}} - 1\right]e^{2i\pi kj/n} = O\left(\frac{k}{n}a_k^{(n)}\right)$$

and $e^{-\frac{2i\pi k\lceil nd\rceil}{n(n+1)}} \to 1$ as $n \to \infty$ and $g_d\left(\frac{2\pi\lceil nd\rceil}{n+1}\right) = 1_{\{n \text{ even}\}}$ as $d \to 1/2$. For notational ease we write

$$1_n^* \equiv 1_{\{n \text{ even}\}}$$

Hence

$$t_k^{(n+1)} - t_k^{(n)} = O\left(\frac{kt_k^{(n)} + 1_n^*}{n}\right). \tag{17}$$

and

$$\det\left(\mathbf{T}_{n-1} - \mathbf{T}_n^{(1)}\right) = O\left(\left[\frac{\max_{0\leq k<n} kt_k^{(n)} + 1_n^*}{n}\right]^n\right) \tag{18}$$

and since $\max_{0\leq k<n} t_k^{(n)} \sim t_0^{(n)} \to d$ as $n \to \infty$

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

and

$$\widetilde{\mathbf{B}_n(z)_{11}} = \det\left(\mathbf{B}_{n-1}(z)\right)\det\left(\mathbf{I}_{n-1} + O\left(e^{-n\log|d|}\left[\mathbf{B}_{n-1}(z)\right]^{-1}z\right)\right). \tag{19}$$

Notice that $\mathbf{B}_{n-1}(z)^{-1}$ is finite for all $|z| < 1$ since its determinant is equivalent to that of $\left(\mathbf{I}_{n-1} - \tilde{\mathbf{C}}_{n-1} z\right)^{-1}$. The latter is finite since $\tilde{\mathbf{C}}_{n-1}$ contains $\lceil (n-1) d \rceil - 1$ unit eigenvalues and the rest are zero eigenvalues, so

$$\det \mathbf{B}_{n-1}(z)^{-1} \underset{n\to\infty}{\sim} (1-z)^{-(\lceil (n-1)d \rceil - 1)}. \tag{20}$$

Hence for all $|z| < 1$,

$$\widetilde{\mathbf{B}_n(z)}_{11} \underset{n\to\infty}{\sim} \det(\mathbf{B}_{n-1}(z)). \tag{21}$$

This constitutes the first part of the proof.

6.1.3. Using properties of Toeplitz matrices

We now turn to showing that expression (16) amounts to its first element ($j = 1$). To show this, we show that the $\widetilde{\mathbf{B}_n(z)}_{1j}$ all tend to zero for $|z| < 1$ as $(n, d) \to (\infty, 1/2)$. By symmetry of the system, we can in fact focus the proof on $\widetilde{\mathbf{B}_n(z)}_{12}$. This is equal to:

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

The key feature that is shared by all the $\widetilde{\mathbf{B}_n(z)}_{1j}$, $j \neq 1$ 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

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

Without loss of generality, we assume for instance that the maximum is $\left|t_1^{(n)}\right|$. We have shown before that as $(d, n) \to (1/2, \infty)$,

$$t_k^{(n)} = O\left(n^{-1} + 1/2 - d\right) = O\left(n^{-1} + \epsilon_d\right).$$

Hence as $(d, n) \to (1/2, \infty)$,

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

The formula for partitioned matrices is

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

where $\det\left(-1-O\left(n^{-1}+\epsilon_d\right)\mathbf{B}_{n-2}\left(z\right)^{-1}O\left(1\right)\right)\to -1$ as $n\to\infty$ since $\mathbf{B}_{n-2}\left(z\right)^{-1}$ is finite for $|z|<1$.

It follows that

$$\widetilde{\mathbf{B}_n\left(z\right)}_{12}=O\left(\left[n^{-1}+\epsilon_d\right]\det\mathbf{B}_{n-2}\left(z\right)\right).$$

Now

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

Hence as $(d,n)\to(1/2,\infty)$

$$f_{n,x}\left(\omega\right)=\frac{\sigma_\epsilon^2}{2\pi}\frac{\left|\det\left(\mathbf{B}_{n-1}\left(e^{-i\omega}\right)\right)\right|^2}{\left|\det\left(\mathbf{B}_n\left(e^{-i\omega}\right)\right)\right|^2}+O\left(\frac{\left(1+n\epsilon_d\right)^2}{n}\frac{\left|\det\left(\mathbf{B}_{n-2}\left(e^{-i\omega}\right)\right)\right|^2}{\left|\det\left(\mathbf{B}_n\left(e^{-i\omega}\right)\right)\right|^2}\right). \tag{22}$$

The rest of the proof relies on letting $(d,n)\to(1/2,\infty)$. We notice that $\mathbf{B}_{n-1}\left(z\right)$ is Toeplitz with density $g_{B(z)}\left(\cdot\right)=1-g_d\left(\cdot\right)z$ (since $\frac{1}{n}\sum_{j=0}^{n-1}e^{2i\pi jk/n}=1_{\{k=0\}}$). The circulant associated to $\mathbf{B}_n\left(z\right)$ is $\mathbf{I}_n-z\tilde{\mathbf{C}}_n$ whose density is $1-g_d^*\left(\cdot\right)z$. Hence, as $n\to\infty$

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

The densities g_d and g_d^* belong to the Wiener class and we can use Szegö's first theorem (see Szegö, 1915, and Gray, 2006) which states that

$$\lim_{n\to\infty}\frac{\det\left(\mathbf{I}_{n-1}-z\tilde{\mathbf{C}}_{n-1}\right)}{\det\left(\mathbf{I}_n-z\tilde{\mathbf{C}}_n\right)}\to\exp\left\{-\frac{1}{2\pi}\int_0^{2\pi}\log\left(1-g_d^*\left(\lambda\right)z\right)d\lambda\right\}, \tag{23}$$

where

$$\begin{aligned}\exp\left\{-\frac{1}{2\pi}\int_0^{2\pi}\log\left(1-g_d^*\left(\lambda\right)z\right)d\lambda\right\}&=\exp\left\{-\frac{\log\left(1-z\right)}{2\pi}\int_0^{2\pi d}d\lambda\right\}\\ &=\left(1-z\right)^{-d}.\end{aligned}$$

This expression finite for $|z|<1$ so

$$\frac{\left|\det\left(\mathbf{B}_{n-2}\left(z\right)\right)\right|^2}{\left|\det\left(\mathbf{B}_n\left(z\right)\right)\right|^2}=\frac{\left|\det\left(\mathbf{B}_{n-2}\left(z\right)\right)\right|^2}{\left|\det\left(\mathbf{B}_{n-1}\left(z\right)\right)\right|^2}\frac{\left|\det\left(\mathbf{B}_{n-1}\left(z\right)\right)\right|^2}{\left|\det\left(\mathbf{B}_n\left(z\right)\right)\right|^2}\underset{n\to\infty}{\to}\left|1-z\right|^{-4d}$$

hence in expression (22) $f_{n,x}\left(\omega\right)=\frac{\sigma_\epsilon^2}{2\pi}\frac{\left|\det\left(\mathbf{B}_{n-1}\left(e^{-i\omega}\right)\right)\right|^2}{\left|\det\left(\mathbf{B}_n\left(e^{-i\omega}\right)\right)\right|^2}+O\left(\frac{\left(1+n\epsilon_d\right)^2}{n}\right)$ and provided that $n\epsilon_d^2\to 0$

$$\lim_{(d,n)\to(1/2,\infty)}f_{n,x}\left(\omega\right)=\frac{\sigma_\epsilon^2}{2\pi}\left(1-e^{-i\omega}\right)^{-1}.$$

6.2. Proof of Theorem 1

To prove the theorem, it suffices to notice that the proof follows the same lines except that the coefficients of the circulant matrix $\mathbf{C}_n$ are $a_k^{(n)}$ which are now complex. Because $\operatorname{Im} a_k^{(n)}$ does not tend to zero as $(d,n) \to (1/2,\infty)$ the main difference is that $\sum_{j=2}^{n} \left| \frac{\widetilde{\mathbf{B}_n(z)}_{1j}}{\det(\mathbf{B}_n(z))} \right|^2$ does not vanish asymptotically. Yet, when $\sigma^2_{\setminus \epsilon_j} \to 0$, the latter tend to zero (when $|z| < 1$, the ratios do not diverge as $n \to \infty$). Hence for all d,

$$\sigma^2_{\epsilon_j} \left| \frac{\widetilde{\mathbf{B}_n(z)}_{11}}{\det(\mathbf{B}_n(z))} \right|^2 \sim \sigma^2_{\epsilon_j} \left| \frac{\det(\mathbf{B}_{n-1}(z))}{\det(\mathbf{B}_n(z))} \right|^2$$

so Szegö's theorem applies and the theorem follows.

6.3. Proof of Corollary 1 (ii)

We show below that convergence results hold when considering Toeplitz matrices with spectral density

$$g_{d,p}(x) = \begin{cases} 1 & u = 0 \\ \frac{1}{2} \exp\left\{ -\left(\Gamma_{\frac{p+1}{p}} \frac{|u|}{2\pi d} \right)^p \right\} & u \neq 0, \end{cases}$$

where $x = u \bmod 2\pi$ and $u \in (-\pi, \pi)$.

The function $g_{d,p}$ yields the $a_k^{(p,d,n)}$ coefficients which are real and coincide with the real part of those generated from:

$$g^*_{d,p}(x) = \exp\left\{ -\left(\Gamma_{\frac{p+1}{p}} \frac{|x|}{2\pi d} \right)^p \right\}.$$

Hence, for simplicity we use the latter function in the proof. The proof follows the same lines as with g_d and g^*_d. We need to show three points in the previous proofs: (i) the limit and behavior of $a_k^{(n)}$ as $(d,n,p) \to (1/2,\infty,\infty)$, (ii) that of $a_k^{(n+1)} - a_k^{(n)}$ and (iii) the limit of $\frac{\det(\mathbf{B}_{n-1}(z))}{\det(\mathbf{B}_n(z))}$.

The function $g^*_{d,p}$ satisfies, for $x \neq \pm 2\pi d$,

$$\lim_{p\to\infty} g^*_{d,p}(x) = g^*_d(x)$$

since

$$\left(\frac{|x|}{2\pi d} \right)^p \underset{p\to\infty}{\to} \begin{matrix} 0 & \text{if } |x| < 2\pi d \\ 1 & \text{if } |x| = 2\pi d \\ +\infty & \text{if } |x| > 2\pi d \end{matrix}$$

and

$$\left[\Gamma\left(\frac{p+1}{p} \right) \right]^p \underset{p\to\infty}{\sim} \left(1 - \frac{\gamma}{p} \right)^p \to e^{-\gamma},$$

where γ is the Euler-Mascheroni constant. We denote the coefficients $\alpha_{\cdot}^{(n,p,d)}$ instead of $a_{\cdot}^{(\cdot)}$ to avoid confusion with those defined previously, and we remove indices in (n,p,d) to indicate that we have taken one of the limits. The coefficients satisfy:

$$\alpha_k^{(p,d,n)} = \frac{1}{n}\sum_{j=0}^{n-1} g_{p,d}^{*}\left(\frac{2\pi j}{n}\right) e^{-2i\pi kj/n}$$

$$\underset{n\to\infty}{\to} \frac{1}{2\pi}\int_0^{2\pi} \exp\left\{-\left(\Gamma\left(1+1/p\right)\frac{|u|}{2\pi d}\right)^p - iku\right\} du = \alpha_k^{(p,d)}$$

with

$$\begin{aligned}
\alpha_0^{(p,d)} &= \frac{1}{2\pi}\int_0^{2\pi} \exp\left\{-\left(\Gamma\left(1+1/p\right)\frac{|u|}{2\pi d}\right)^p\right\} du \\
&= \underset{p\to\infty}{\sim} \frac{1}{2\pi}\int_0^{2\pi d} \exp\left\{-\left(\Gamma\left(1+1/p\right)\frac{|u|}{2\pi d}\right)^p\right\} du \\
&+ \frac{1}{2\pi}\int_{2\pi d}^{2\pi} \exp\left\{-\left(\Gamma\left(1+1/p\right)\frac{|u|}{2\pi d}\right)^p\right\} du \\
&\underset{p\to\infty}{\sim} \frac{1}{2\pi}\int_0^{2\pi d} \exp\left\{-\left(\Gamma\left(1+1/p\right)\frac{|u|}{2\pi d}\right)^p\right\} du.
\end{aligned}$$

Hence

$$\alpha_0^{(p,d)} \underset{p\to\infty}{\to} d.$$

Also

$$\begin{aligned}
\alpha_k^{(p,d,n)} &\underset{n\to\infty}{\to} \frac{1}{2\pi}\int_0^{2\pi} \exp\left\{-\left(\Gamma\left(1+1/p\right)\frac{|u|}{2\pi d}\right)^p - iku\right\} du \\
&\underset{p\to\infty}{\sim} \frac{1}{2\pi}\int_0^{2\pi d} \exp\left\{-iku\right\} du \\
&= \frac{e^{-2i\pi kd}-1}{2i\pi k} \\
&= \frac{1}{2i\pi k} e^{-i\pi kd}\left(e^{-i\pi kd} - e^{i\pi kd}\right) \\
&= -\frac{1}{\pi k}\left(\cos\pi kd - i\sin\pi kd\right)\sin\left(\pi kd\right) \\
\operatorname{Re}\alpha_k^{(p,d)} &= d\cos\pi kd \operatorname{sin_c}\pi kd.
\end{aligned}$$

Hence as $(p,n)\to(\infty,\infty)$

$$\alpha_k^{(p,d,n)} \sim a_k^{(d)}.$$

And the points (i) and (ii) of the proof follow. Now we need to show point (iii), the integral that appears in Szegö's theorem:

$$\exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log\left(1 - g_{p,d}\left(u\right)z\right) du\right\}.$$

Using a integer series expansion for the logarithm:

$$\begin{aligned}
\exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log\left(1 - g_{p,d}\left(u\right)z\right) du\right\} &= \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log\left(1 - \exp\left\{-\left(\Gamma_{\frac{p+1}{p}}\frac{u}{2\pi d}\right)^p\right\} z\right) du\right\} \\
&= \exp\left\{-\frac{1}{2\pi}\int_0^{2\pi}\sum_{j=1}^{\infty} \exp\left\{-j\left(\Gamma_{\frac{p+1}{p}}\frac{u}{2\pi d}\right)^p\right\}\frac{z^j}{j} du\right\}.
\end{aligned}$$

Now, the integer series summation and integral can be swapped so

$$\exp\left\{-\frac{1}{2\pi}\int_0^{2\pi}\sum_{j=1}^{\infty}\exp\left\{-j\left(\Gamma_{\frac{p+1}{p}}\frac{u}{2\pi d}\right)^p\right\}\frac{z^j}{j}du\right\}$$
$$=\exp\left\{-\frac{1}{2\pi}\sum_{j=1}^{\infty}\frac{z^j}{j}\int_0^{2\pi}\exp\left\{-j\left(\frac{\Gamma_{\frac{p+1}{p}}}{2\pi d}\right)^p u^p\right\}du\right\}. \quad (24)$$

We now use the change of variable formula, for some b, and where we let $v = bu^p$ so $dv = bpu^{p-1}du$:

$$\int_0^{2\pi}\exp(-bu^p)\,du = (bp)^{-1}\int_0^{b(2\pi)^p}\left(\frac{v}{b}\right)^{1/p-1}\exp(-v)\,dv$$
$$=\frac{1}{b^{1/p}}\frac{1}{p}\int_0^{b(2\pi)^p}v^{1/p-1}\exp(-v)\,dv.$$

Now, we use the definition of the incomplete Gamma function $\Gamma_x(\cdot) = \int_x^\infty z^{\cdot-1}\exp(-z)\,dz$, so

$$\int_0^{2\pi}\exp(-bu^p)\,du = \frac{1}{b^{1/p}}\frac{1}{p}\left(\Gamma\left(\frac{1}{p}\right) - \Gamma_{b(2\pi)^p}\left(\frac{1}{p}\right)\right).$$

Plugging this into expression (24) yields:

$$\exp\left\{-\frac{1}{2\pi}\sum_{j=1}^{\infty}\frac{z^j}{j}\int_0^{2\pi}\exp\left\{-j\left(\frac{\Gamma\left(\frac{p+1}{p}\right)}{2\pi d}\right)^p u^p\right\}du\right\}$$
$$=\exp\left\{-\frac{1}{2\pi}\sum_{j=1}^{\infty}\frac{z^j}{j}\left[\frac{2\pi d}{j^{1/p}\Gamma\left(\frac{p+1}{p}\right)}\frac{1}{p}\left(\Gamma\left(\frac{1}{p}\right)-\Gamma_{j\left(\frac{1}{d}\Gamma_{\frac{p+1}{p}}\right)^p}\left(\frac{1}{p}\right)\right)\right]\right\}.$$

Now as $p \to \infty$, letting as before γ denote the Euler-Mascheroni constant,

$$\Gamma\left(\frac{p+1}{p}\right) \sim 1 - \frac{\gamma}{p} \text{ and } \Gamma\left(\frac{1}{p}\right) \sim p - \gamma$$

so as $p \to \infty$

$$\exp\left\{-\frac{1}{2\pi}\sum_{j=1}^{\infty}\frac{z^j}{j}\left[\frac{2\pi d}{j^{1/p}\Gamma\left(\frac{p+1}{p}\right)}\frac{1}{p}\left(\Gamma\left(\frac{1}{p}\right)-\Gamma_{j\left(\frac{1}{d}\Gamma_{\frac{p+1}{p}}\right)^p}\left(\frac{1}{p}\right)\right)\right]\right\}$$
$$\sim\exp\left\{-\frac{1}{2\pi}\sum_{j=1}^{\infty}\frac{z^j}{j}\left[\frac{2\pi d}{j^{1/p}\left(1-\frac{\gamma}{p}\right)}\frac{1}{p}\left(p-\gamma-\Gamma_{j\left(\frac{1}{d}\Gamma_{\frac{p+1}{p}}\right)^p}\left(\frac{1}{p}\right)\right)\right]\right\}$$
$$\sim\exp\left\{-d\sum_{j=1}^{\infty}\frac{z^j}{j^{1+1/p}}\right\} = \exp\left\{-d\mathsf{Li}_{1+1/p}(z)\right\}.$$

Where Li denotes the polylogarithm. As $p \to \infty$, $\mathsf{Li}_{(p+1)/p}(z) \to -\log(1-z)$ so

$$\exp\left\{\frac{1}{2\pi}\int_0^{2\pi}\log\left(1-g_{p,d}(u)\,z\right)du\right\}\underset{p\to\infty}{\to}(1-z)^d. \quad (25)$$

7. References

Abadir, K. M. and G. Talmain (2002). Aggregation, persistence and volatility in a macro model. *Review of Economic Studies 69*(4), 749–79.

Altissimo, F., B. Mojon, and P. Zaffaroni (2009). Can aggregation explain the persistence of inflation? *Journal of Monetary Economics 56*(2), 231–241.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American statistical association 96*(453), 42–55.

Andrews, D. W. K. and D. Pollard (1994). An introduction to functional central limit theorems for dependent stochastic processes. *International Statistical Review / Revue Internationale de Statistique 62*(1), pp. 119–132.

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.

Berenguer-Rico, V. and J. Gonzalo (2014). Summability of stochastic processesa generalization of integration for non-linear processes. *Journal of Econometrics 178*, 331–341.

Campbell, J. Y. and N. G. Mankiw (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. and K. S. Lai (1993). A fractional cointegration analysis of purchasing power parity. *Journal of Business and Economic Statistics 11*(1), 103–112.

Chevillon, G. and S. Mavroeidis (2013). Learning can generate long memory. working paper, ESSEC Business School.

Cox, D. R. (2014). Scaling. Sir david cox 90th birthday symposium, Nuffield College, Oxford.

Cox, D. R. and M. W. H. Townsend (1947). The use of the correlogram in measuring yarn irregularities. *Proc. Roy. Soc. Edinburgh Sec. A 63*, 290–311.

Cubadda, G., A. Hecq, and F. C. Palm (2009). Studying co-movements in large multivariate data prior to multivariate modelling. *Journal of Econometrics 148*(1), 25–35.

Davidson, J. and P. Sibbertsen (2005). Generating schemes for long memory processes: regimes, aggregation and linearity. *Journal of Econometrics 128*(2), 253–82.

Diebold, F. X. and A. Inoue (2001). Long memory and regime switching. *Journal of Econometrics 105*(1), 131–159.

Geweke, J. and S. Porter-Hudak (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis 4*, 221–38.

Gonzalo, J. and J.-Y. Pitarakis (2006). Threshold effects in cointegrating relationships. *Oxford Bulletin of Economics and Statistics 68*, 813–833.

Gourieroux, C. and J. Jasiak (2001). Memory and infrequent breaks. *Economics Letters 70*, 29–41.

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

Gray, R. M. (2006). *Toeplitz and circulant matrices: A review.* NOW Publishers Inc.

Hecq, A., S. Laurent, and F. C. Palm (2012). Common intraday periodicity. *Journal of Financial Econometrics 10*(2), 325–353.

Hendry, D. F. and J. A. Doornik (2001). *PcGive 10: Empirical econometric modelling using.* Timberlake Consultants.

Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng. 116*, 770–808.

Hurvich, C. M. and W. W. Chen (2000). An efficient taper for potentially overdifferenced long-memory time series. *Journal of Time Series Analysis 21*(2), 155–180.

Lieberman, O. and P. C. Phillips (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.

Martens, M., D. van Dijk, and M. de Pooter (2009). Modeling and forecasting s&p 500 volatility: Long memory, structural breaks and nonlinearity. *International Journal of Forecasting 27*, 282–303.

Miller, J. I. and J. Y. Park (2010). Nonlinearity, nonstationarity, and thick tails: How they interact to generate persistence in memory. *Journal of Econometrics 155*(1), 83 – 89.

Parke, W. R. (1999). What is fractional integration? *Review of Economics and Statistics 81*(4), 632–638.

Perron, P. and Z. Qu (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.

Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. 42*(1), 43–47.

Schennach, S. (2013). Long memory via networking. Working paper cwp13/13, CEMMAP.

Shimotsu, K., P. C. Phillips, et al. (2005). Exact local whittle estimation of fractional integration. *The Annals of Statistics 33*(4), 1890–1933.

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.

Szegö, G. (1915). Ein grenzwertsatz über die toeplitzschen determinanten einer reellen positiven funktion. *Mathematische Annalen 76*(4), 490–503.

White, H. (2000). *Asymptotic Theory for Econometricians*. Academic Press Inc. 2nd ed.

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

Zellner, A. and F. Palm (1974). Time series analysis and simultaneous equation econometric models. *Journal of Econometrics 2*(1), 17–54.

Zellner, A. and F. C. Palm (2004). *The structural econometric time series analysis approach*. Cambridge University Press.