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

This Version: February 15, 2016

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 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.

☆We would like to thank the editors Serge Darolles, Alain Monfort and Eric Renault and three anonymous referees for their numerous comments on a preliminary version of this paper. We also thank the 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, 4th Long-Memory Symposium at CREATES, 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, see inter alia Haldrup and Vera-Valdés (2015). 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 $\omega \to 0^+$, for some positive c_f. Since Granger and Joyeux (1980), fractional integration of order d, denoted $\mathsf{I}(d)$, has proved the most pervasive example of long memory processes in econometrics. The prototypical example of an $\mathsf{I}(d)$ process is the fractional white noise $z_t = (1-L)^{-d} \epsilon_t$, where L denotes the lag operator and ϵ_t is a white noise sequence. The class of fractionally integrated processes extends to cases where ϵ_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, Gonçalves and Gourieroux (1988), Chambers, 1998, and inter alia Comte and Renault, 1996, 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 individually modelled as fractional white noises as $n \to \infty$. 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, and in particular, from the aggregation mechanism à la Granger (1980). Indeed, the mechanism that leads here to long memory does not rely on heterogeneity assumptions, nor does it involve aggregation.

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.[1] We

[1] The class of Toeplitz matrices is chosen for the technical reason that we use in our proofs the well-established

use the final equation representation of this model (as proposed by Zellner and Palm, 1974, 2004) to prove that, as $n \to \infty$ and for finite t, some of the marginalized processes x_{jt} belonging to $\mathbf{x}_t$ (for $j = 1, ..., n$) may converge in probability to a long memory process of order $\delta \in (0, 1)$. This framework, although is bears some resemblance with that of Onatski and Uhlig (2012), differs from it. We also consider a VAR process whose dimension diverges, but contrary to these authors, our interest does not lie in empirical inference about the vector process. Instead we focus on the population properties of the marginalized series. We introduce three high-level assumptions concerning $(\mathbf{A}_n)$. Under these assumptions, the moving average lag polynomial associated with x_{jt} is asymptotically proportional to the ratio of $\det(\mathbf{I}_{n-1} - \mathbf{A}_{n-1}L)$ over $\det(\mathbf{I}_n - \mathbf{A}_n L)$. We parameterize $\mathbf{A}_n$ by defining a scalar sequence (δ_n) with $\delta_n \in (0, 1)$ such that $\lim_{n\to\infty} \delta_n = \delta \in (0, 1)$, and a circulant matrix $\mathbf{C}_n$ such that the polynomials $\det(\mathbf{I}_n - \mathbf{A}_n z) \sim \det(\mathbf{I}_n - \mathbf{C}_n z)$ as $n \to \infty$. $\mathbf{C}_n$ is assumed to possess about a fraction $\lfloor n\delta_n \rfloor$ of unit eigenvalues ($\lfloor \cdot \rfloor$ denotes the integer part) and $n - \lfloor n\delta_n \rfloor$ zero eigenvalues. We then use the *first theorem* of Szegö (1915) to prove that under our three high-level assumptions, as $n \to \infty$, $\det(\mathbf{I}_{n-1} - \mathbf{A}_{n-1}z) / \det(\mathbf{I}_n - \mathbf{A}_n z) \to (1 - z)^{-\delta}$ and hence $x_{jt} \xrightarrow{p} \kappa \Delta^{-\delta} \epsilon_t$ for some $\kappa > 0$.

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 $n \to \infty$, 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 $n \to \infty$, 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 ϵ_{jt}) dominates the others in terms of magnitude. In this case, we prove that the dominant series j tends to 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 phenomenon 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

First Theorem of Szegö (1915).

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 possibly due to the marginalization of a large system. We illustrate this finding by considering the $\log(MedRV)$ of 49 US stocks, where $MedRV$ 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 and we provide two examples where they hold in the next section. Section 4 presents some Monte Carlo simulations and compares them with some empirical evidences on $\log(MedRV)$. Finally, Section 5 concludes. The appendix contains all the proofs.

In the paper, we use the following notation. $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{C}$ denote the sets of, respectively, natural integers, real and complex scalars. For any $x \in \mathbb{R}$, $\lfloor x \rfloor$ and $\lceil x \rceil$ denote the floor and ceiling of x. 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 $n \to \infty$, $|a_n/b_n|$ is bounded, $b_n/c_n \to 0$, and $a_n/b_n \to 1$. For any complex-valued square matrix $\mathbf{A}$, $\det(\mathbf{A})$ is the determinant of $\mathbf{A}$, $tr(\mathbf{A})$ its trace, $\widetilde{\mathbf{A}}$ its adjugate matrix, $\overline{\mathbf{A}}'$ its conjugate transpose and $|\mathbf{A}| = \sqrt{\text{tr}\left(\overline{\mathbf{A}}'\mathbf{A}\right)}$ its weak (Frobenius) 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 $|\mathbf{A}_n - \mathbf{B}_n| \to 0$ as $n \to \infty$. $1_{\{\cdot\}}$ 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 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., Grenander and Szegö, 1958, 2001 and Gray, 2006).

2.1. Set-up and main results

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

$$\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t, \; t = 1, \ldots, T \tag{1a}$$

$$\epsilon_t = (\epsilon_{1t}, ..., \epsilon_{nt})' \sim \mathsf{IID}(\mathbf{0}, \Omega_n), \tag{1b}$$

where Ω_n is a diagonal matrix with diagonal $\sigma_n^2 = \left(\sigma_{n,1}^2, ..., \sigma_{n,n}^2\right)$ such that $\sigma_{n,k} > 0$ for $k = 1, ..., n$. Hence shocks ϵ_{kt} and $\epsilon_{k't}$ are uncorrelated for $k \neq k'$, $k, k' = 1, ..., n$. We could relax the latter assumption, but at the cost of increased complexity of the exposition, so we maintain it throughout the paper.

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

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

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 $t \leq 0$ and $\mathbf{x}_0 = 0$. In the spirit of Johansen and Nielsen (2016), we therefore consider truncations of lag polynomials and convergence to fractional processes with a fixed start (hence of type II, see Marinucci and Robinson, 1999, and Davidson and Hashimzade, 2009). For $t > 0$, and for any polynomial $P(z)$, we let $[P(z)]^+$ denote the truncation of $P(z)$ for degrees of z^k strictly less than t so that $P(L)\mathbf{x}_t$ only involves $\mathbf{x}_k$ at dates $k > 0$. 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 $n \to \infty$, the univariate process x_{jt} without roots cancellation follows an ARMA(∞,∞). 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 $\mathsf{E}\left[\epsilon_{jt}\epsilon_{kt}\right] = 0$ *for* $j \neq k$. *The FER of* $\mathbf{x}_t$ *is* $\det\left(\mathbf{B}_n(L)\right)\mathbf{x}_t = \widetilde{\mathbf{B}_n(L)}\ \epsilon_t$, *where*

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

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

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

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

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

hence x_{jt} admits the ARMA representation

$$\det\left(\mathbf{B}_n(L)\right)x_{jt} = \widetilde{\mathbf{B}_n(L)}_{j\cdot}\epsilon_t = \sum_{k=1}^{n}\widetilde{\mathbf{B}_n(L)}_{jk}\ \epsilon_{kt}.$$

Therefore, provided $\det(\mathbf{B}_n(L))^{-1}$ can be properly defined, the process x_{jt} admits the representation

$$x_{jt} = \sum_{k=1}^{n} \frac{\widetilde{\mathbf{B}_n(L)}_{jk}}{\det(\mathbf{B}_n(L))} \epsilon_{kt}. \tag{3}$$

Expression (3) constitutes the basis of our theoretical argument. To keep the generality of the exposition, we implicitly consider that (3) is replaced by $x_{jt} = \sum_{k=1}^{n} \left[\frac{\widetilde{\mathbf{B}_n(L)}_{jk}}{\det(\mathbf{B}_n(L))}\right]^{+} \epsilon_{kt}$ when $\det(\mathbf{B}_n(L))^{-1}$ does not exist but $\left[\det(\mathbf{B}_n(L))^{-1}\right]^{+}$ does. In the remainder of the section, we delineate three high-level assumptions which together imply that, as $n \to \infty$, the right-hand side of expression (3) tends in probability to a fractional white noise. The first assumption concerns the summation on the right-hand side of (3).[2]

Assumption B. *There exists $(t, j) \in \mathbb{N}^2$ for which the parameters of the VAR(1) model (1) satisfy as $n \to \infty$,*

(i) $\mathsf{Var}\left(\sum_{k=1,\ k\neq j}^{n} \left[\frac{\widetilde{\mathbf{B}_n(L)}_{jk}}{\det(\mathbf{B}_n(L))}\right] \epsilon_{kt}\right) \to 0;$

(ii) there exists $\kappa > 0$ such that $\widetilde{\mathbf{B}_n(z)}_{jj} \sim \kappa \det(\mathbf{B}_{n-1}(z))$;

(iii) $\sigma_{n,j}^2 \to \sigma_j^2 > 0.$

Assumption B(i) ensures that in the moving average representation for x_{jt}, the only innovation that retains a contribution as $n \to \infty$ is ϵ_{jt}, the others play no role asymptotically. Assumption B(ii) provides a recursive formula that, in the FER representation for x_{jt}, see expression (2), expresses the contribution relative to ϵ_{jt} in the MA lag polynomial of the n-dimensional system in terms of the AR lag polynomial of the $n-1$ dimensional system. Finally, Assumption B(iii) ensures that the distribution of ϵ_{jt} is nondegenerate as $n \to \infty$, so this applies to x_{jt} too.

Under Assumption B, we prove in the Appendix, Subsection 6.1, that the summation on the right-hand side of (3) reduces to its jth element in the sense that for all $\varepsilon > 0$ and as $n \to \infty$,

$$\Pr\left(\left|x_{jt} - \kappa \frac{\det(\mathbf{B}_{n-1}(L))}{\det(\mathbf{B}_n(L))} \epsilon_{jt}\right| > \varepsilon\right) \to 0, \tag{4}$$

where κ is defined in Assumption B(ii).

In (4), x_{jt} tends, as $n \to \infty$, in probability to a moving average process whose lag polynomial 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 a function of a sequence of Toeplitz matrices

[2] 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.

$\mathbf{T}_n$:

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

We make the following assumption regarding the elements $t_k^{(n)}$ of $\mathbf{T}_n$.

Assumption T.

(i) There exists a real-valued 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 such that*

$(i.a)$ $g(\cdot,\cdot)\leq 1$;

$(i.b)$ $\int_0^{2\pi}\left|g\left(\cdot,e^{i\omega}\right)\right|d\omega<\infty$;

(ii) $\forall d\in(0,1)$, $t_{d,k}=\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}g\left(d,e^{i\frac{2\pi j}{n}}\right)e^{-2i\pi jk/n}$ *satisfies* $\sum_{k=-\infty}^{\infty}|t_{d,k}|<\infty$;

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

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

$(iii.b)$ $n\left(t_{\delta,0}^{(n)}-t_0^{(n)}\right)\to 0$ as $n\to\infty$;

(iv) $\forall(d,z)\in(0,1)\times(-\infty,1)$, $\frac{1}{2\pi}\int_0^{2\pi}\log\left(1-g\left(d,e^{i\omega}\right)z\right)d\omega=d\log(1-z)$.

Let $(\mathbf{T}_{d,n})$ denote the sequence of Toeplitz matrices whose elements are the $t_{d,k}$ defined in Assumption T(ii) for $k=-(n-1),...,n-1$. For all $d\in(0,1)$, the partial function $g_d(\cdot)=g(d,\cdot)$ is called the *symbol* of the matrices $(\mathbf{T}_{d,n})$: it is such that $g_d(z)=\sum_{k=-\infty}^{\infty}t_{d,k}z^k$. The function $\omega\in\mathbb{R}:\omega\to g_d\left(e^{i\omega}\right)$ is often 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 notation, we refer to $g_d\left(e^{i\omega}\right)$ as the function $\omega\to g\left(d,e^{i\omega}\right)$.

Assumptions T(i) and T(ii) ensure that the First Theorem of Szegö (1915) holds for $\mathbf{I}_n-\mathbf{T}_{\delta,n}z$, for $z\neq 1$. Under Assumption T(ii), 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 all $z<1$ and $d\in(0,1)$, the matrix $\mathbf{I}_n-\mathbf{T}_{d,n}z$ is Toeplitz with symbol $1-g_d(\cdot)z$. Under T(ii), $\mathbf{T}_{d,n}$ belongs to the so-called Wiener class of Toeplitz matrices with absolutely summable entries, so we can use associated results (Gray, 2006, in particular, Section 4.4 and Theorem 4.2). Hence, we show in the Appendix, Subsection 6.2, that as $n\to\infty$,

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

Assumption T(iii) is used in the proof of Theorem 1, where we show that the same limit holds, replacing $\mathbf{T}_{\delta,n}$ with $\mathbf{T}_n$. This assumption also introduces the parameter δ that governs the degree of fractional integration in the limit. We will see what specific role it plays for the matrix $\mathbf{T}_n$ in the examples of Section 3. Finally, Assumption T(iv) implies that the limit in expression (5) equals $(1-z)^{-\delta}$.

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. *The sequence of real matrices* $(\mathbf{A}_n)$ *in the VAR*(1) model (1) *satisfy, for* $z \neq 1$ *and as* $n \to \infty$,

$$\frac{\det(\mathbf{I}_{n-1} - \mathbf{A}_{n-1} z)}{\det(\mathbf{I}_n - \mathbf{A}_n z)} \sim \frac{\det(\mathbf{I}_{n-1} - \mathbf{T}_{n-1} z)}{\det(\mathbf{I}_n - \mathbf{T}_n z)},$$

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

Assumption A is used to link Assumptions B and T so the limit of the ratio of determinants which appears in expression (4) can be expressed using the First Theorem of Szegö.

Together, the three high-level Assumptions B, T and A combine 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. Then, there exist* $(t, j) \in \mathbb{N}^2$ *such that element* x_{jt} *satisfies, for all* $\varepsilon > 0$ *and as* $n \to \infty$,

$$\Pr\left(\left|x_{jt} - \kappa \Delta^{-\delta} \epsilon_{jt}\right| > \varepsilon\right) \to 0,$$

where $\kappa > 0$ *and* $\delta \in (0, 1)$ *are defined in Assumptions B and T.*

Proof: see the Appendix, Subsection 6.3.

In Theorem 1, the marginalized univariate process x_{jt} tends in probability to an $\mathsf{I}(\delta)$ fractional white noise as the dimension n of the system diverges to infinity. Hence individual series may asymptotically (as the cross-sectional dimension increases) exhibit long memory although the infinitely dimensional limiting vector process itself does not. In the theorem, the limit of x_{jt} only depends on one nonzero innovation, the others do not matter (and we show below an example where they disappear, i.e., $\sigma^2_{\epsilon_k} \to 0$ for $k \neq j$). Hence this is in particular distinctly different from the heuristic example of Granger (1980, Section 4) where he generalizes his argument about long memory via aggregation of heterogenous micro-units to a large scale dynamic model similar to our VAR(1) model $\mathbf{B}_n(L) x_t = \varepsilon_t$. He notices that $\mathbf{B}_n^{-1}(L)$ is given by expression (3) above, i.e., the sum of n moving averages of the innovations ϵ_{kt}, $k = 1, ..., n$, and he notices a resemblance with his model of aggregation. It may indeed be possible that some specific assumptions about the parameters may lead to long memory by aggregating heterogenous processes in such a setting. Our framework differs here and, in fact, precludes it: we specify in Assumption B(i) that the contribution of the moving averages of ϵ_{kt}, for $k \neq j$, vanish as $n \to \infty$, so they do not play a role in the long memory of x_{jt}. Theorem 1 shows this distinction clearly since the limiting fractional white noise is $\Delta^{-\delta}\epsilon_{jt}$, so it is entirely driven by the innovation sequence ϵ_{jt} without any contribution from ϵ_{kt}, for $k \neq j$.

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

3. Two examples

The section present two parametric representations where the high-level assumptions B, T and A are satisfied so Theorem 1 and long memory arises in the marginalized representation. Both examples relate to the same underlying Toeplitz sequence and choice of $g(\cdot,\cdot)$. 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 an heterogenous case where the results are not symmetric for all x_{jt} and the distinction with the case of long memory arising from aggregation is clear. Proofs that Assumptions B, T and A hold are provided in the Appendix.

3.1. A symmetric example

In our first example, we specify the VAR as follows. All innovations have finite variance, $\Omega_n = \sigma^2 \mathbf{I}_n$ with $\sigma^2 > 0$.[3] 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; η_n is a real scalar sequence that satisfies $\eta_n = o\left(n^{-1}\right)$, and $\mathbf{D}_n$ is a real antisymmetric Toeplitz matrix with absolutely summable entries.

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 $\omega \geq 0$,

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

and $\omega \to g\left(d, e^{i\omega}\right)$ is even. Let $\delta = 1/2$ so the sequence δ_n is assumed to satisfy:

$$\delta_n = \frac{1}{2} + o\left(n^{-2}\right), \quad \text{with } \delta_n \neq \frac{1}{2}.$$

The coefficients of $\mathbf{T}_n$ are $t_k^{(n)} = \frac{1}{n}\sum_{j=0}^{n-1} g\left(\delta_n, e^{i\frac{2\pi j}{n}}\right) e^{-2i\pi jk/n}$. 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 $c_0^{(n)} = t_0^{(n)}$. As $n \to \infty$, $c_k^{(n)} \sim t_{\delta_n,-k} + t_{\delta_n,n-k}$, for $k \neq 0$, where $t_{\delta_n,k} = t_{d,k}$ given by Assumption T(ii) and evaluated at $d = \delta_n$. Eigenvalues of circulant

[3] We assume that all the diagonal elements of Ω_n are equal for notational ease but this has no incidence on our results. We can relax the assumption to heterogenous and positive diagonal elements.

matrices can be expressed in terms of the associated symbol evaluated at the Fourier ordinates, see Gray (2006, Chapter 3). Hence, $\mathbf{C}_n$ defines a circulant matrix with eigenvalues $g\left(\delta_n, e^{2i\pi j/n}\right)$ for $j = 0, ..., n-1$, 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 $\omega \to g\left(d, e^{i\omega}\right)$ 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 $\mathbf{T}_n^* \equiv \mathrm{Re}\left(\mathbf{T}_n\right)$ with entries $t_k^{*(n)} = \mathrm{Re}\left(t_k^{(n)}\right)$. So are $\mathbf{I}_{n-1} - \mathbf{T}_{n-1}^* z$ and $\mathbf{I}_{n-1} - \mathbf{C}_{n-1} z$. Hence, following the proof of Theorem 1,

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

The entries $t_k^{*(n)}$ of the Toeplitz matrix $\mathbf{T}_n^*$ satisfy:

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

As $n \to \infty$, 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:

$$\sum_{k=1}^{n-1} t_k^*\left(n\right) = 1 - \delta_n + O\left(n^{-1}\right). \tag{9}$$

We show in the Appendix, Subsection 6.4, that Assumptions B, T and A hold so Theorem 1 then implies that all the processes x_{jt}, for all $j \in \mathbb{N}$, tend in probability to uncorrelated fractional white noises of order $1/2$: for all j, t and for all $\varepsilon > 0$,

$$\Pr\left(\left|x_{jt} - \kappa\Delta^{-1/2}\epsilon_{jt}\right| > \varepsilon\right) \underset{n\to 0}{\to} 0. \tag{10}$$

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 circulant matrix which asymptotically presents about $\lfloor n/2 \rfloor$ unit eigenvalues and $\lfloor n/2 \rfloor$ zero eigenvalues.

3.2. 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 $\mathsf{I}\left(\delta\right)$, $\delta \in (0, 1)$. We now give an example of sequence $\mathbf{A}_n$ satisfying Assumptions B, T and A, but were 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 therefore satisfied.

Now, assume that the variance of one innovation ϵ_{jt} dominates the others. For this we assume that there exists $j \in \mathbb{N}$ such that the variances of the innovations satisfy

$$\begin{cases} \sigma_{n,j}^2 = \sigma_j^2 > 0, \\ \sigma_{n,k}^2 = o\left(n^{-1}\right), \quad k \neq j, \end{cases}$$

so Assumption B(iii) holds and Assumption B(i) hence follows. The formal proofs that Assumptions B, T and A hold are collected in the Appendix, Subsection 6.5. Theorem 1 then implies that, for all $\varepsilon > 0$,

$$\Pr\left(\left|x_{jt} - \kappa\Delta^{-\delta}\epsilon_{jt}\right| > \varepsilon\right) \underset{n\to\infty}{\rightarrow} 0.$$

Therefore, when the number of variables n tends to infinity and when one of the innovation processes dominates all the others, then the dominant process entering $\mathbf{x}_t$ tends in probability to an $\mathsf{I}(\delta)$ 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 origins. In particular, despite the multivariate nature of the source of long memory that we present, it is not aggregation that is at play here since only one innovation ϵ_{jt} with nonzero variance remains in the system as $n \to \infty$. The mechanism is closer in a sense to that, which Schennach (2013) delineates, in the case of the impact of a single input that transits through a network (her process is scalar). Following Diebold and Yılmaz (2009, 2014), if we interpret our VAR setting as a network, then there are n nodes, $k = 1, ..., n$, in the system which are in state x_{kt-1} at the end of any period $t-1$. At time t, each node k combines $\mathbf{x}_{t-1}$ with an additional idiosyncratic signal ϵ_{kt} to produce $\mathbf{x}_t$. Contrary to Schennach, all the coefficients of $\mathbf{A}_n$ are strictly less than unity in absolute value, so the signal is transmitted from x_{kt-1} to x_{jt} with attenuation (she requires that some nodes transmit without attenuation, see her Section 3). Also because here, for $k \neq j$, the ratio $\sigma_{\epsilon_j}^2/\sigma_{\epsilon_k}^2 \to \infty$ as $n \to \infty$, we can interpret our framework as providing an example of impact of a relatively "large" innovation sequence within a network. Albeit different, such a setting bears some resemblance with the famous "Noah effect" of Mandelbrot and Wallis (1968) and Mandelbrot (1997) where they show that the existence of outliers arising from heavy tailed distributions can be a source of nonsimilarity (usually in the marginal distribution). Because here nonsimilarity takes the form of long memory (the "Joseph effect"), this interpretation of dynamic networks constitutes a bridge between the Noah and Joseph effects which are considered different forms of non-similarity in the literature.

4. 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.

4.1. Monte Carlo

We provide here simulations that examine the validity of our theoretical asymptotic results when the dimensions of the cross-section and 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 $k = 0, \ldots, n-1$), 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).[4] We deliberately choose a large bandwidth, as 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 3.1 (i.e., Equation (6) with $\eta_n = 0$), where $\mathbf{T}_n^*$ has symbol g_d. We denote by d the value taken by δ_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 $\mathsf{I}(d)$ 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 (9) 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

[4]All estimations are performed in OxMetrics 7.0 (see Doornik, 2013).

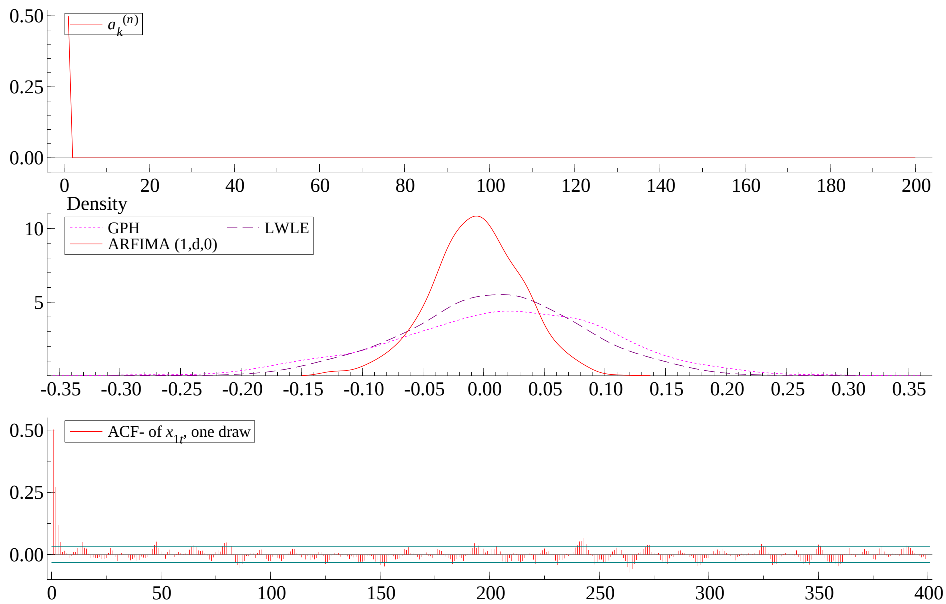

Figure 1: Simulation results for a n-dimensional diagonal VAR(1) $\mathbf{x}_t = \mathbf{A}_n\mathbf{x}_{t-1} + \epsilon_t$, with $\mathbf{A}_n = d\mathbf{I}_n$, where $d = 0.499$, $\epsilon_t \overset{iid}{\sim} N(\mathbf{0}, \mathbf{I}_n)$, $n = 200$ and $t = 1, \ldots, 4000$. The panels report respectively, (a) the value of the elements of the first row of $\mathbf{A}_n$, denoted $a_k^{(n)}$ (for $k = 0, \ldots, n-1$); (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.

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 as 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 $n \to \infty$ and $d \to 1/2$. 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 $x_{1t}, \ldots, x_{200t}$, 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$(0, d, 0)$.

4.2. Empirical illustration

The presence of long memory in the volatility is now considered as a stylized fact of the log-returns of financial assets (see Baillie, Bollerslev, and Mikkelsen, 1996, Breidt, Crato, and de Lima, 1998, and Comte and Renault, 1998, among others). As reported in Lieberman and Phillips (2008)

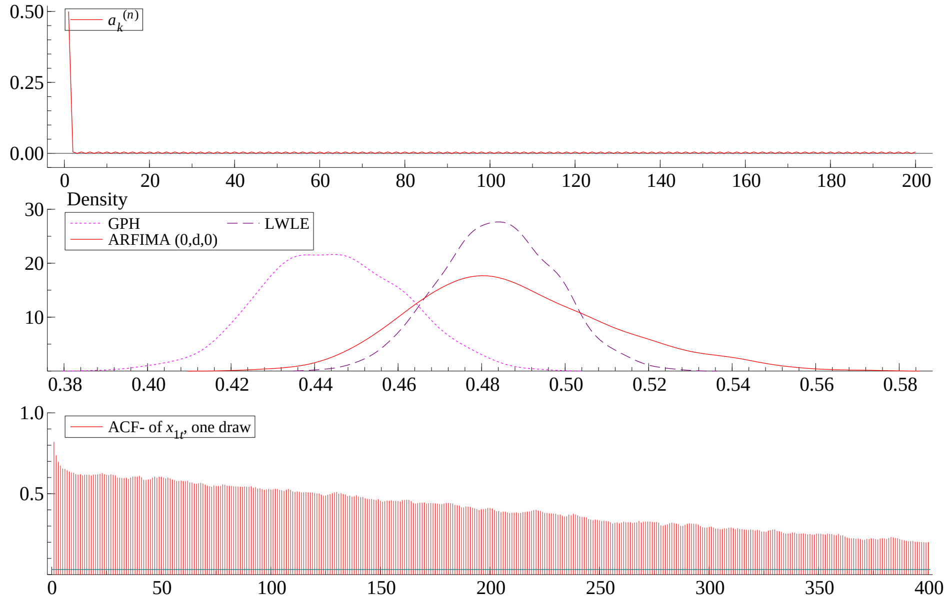

Figure 2: Simulation results for a n-dimensional diagonal VAR(1) $\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t$, with $\mathbf{A}_n = \mathbf{T}_n^*$, where $\mathbf{T}_n^* \equiv \mathrm{Re}(\mathbf{T}_n)$, $\mathbf{T}_n$ has symbol defined by (7), $d = 0.499$, $\epsilon_t \overset{iid}{\sim} N(\mathbf{0}, \mathbf{I}_n)$, $n = 200$ and $t = 1, \ldots, 4000$. The panels report respectively, (a) the value of the elements of the first row of $\mathbf{A}_n$, denoted $a_k^{(n)}$ (for $k = 0, \ldots, n-1$); (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.

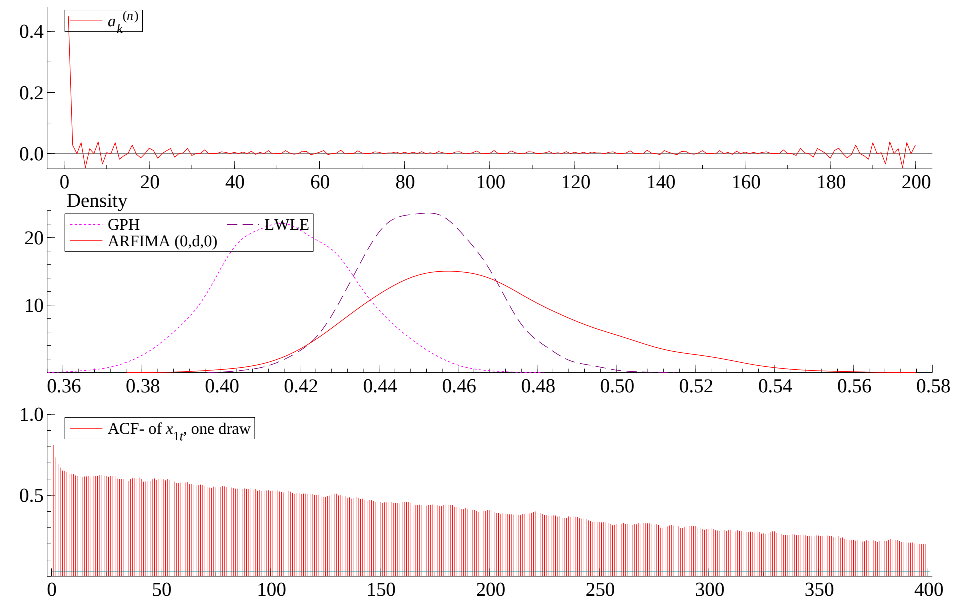

Figure 3: Simulation results for a n-dimensional diagonal VAR(1) $\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t$, with $\mathbf{A}_n = \mathbf{T}_n^*$, where $\mathbf{T}_n^* \equiv \mathrm{Re}(\mathbf{T}_n)$, $\mathbf{T}_n$ has symbol defined by (7), $d = 0.45$, $\epsilon_t \overset{iid}{\sim} N(\mathbf{0}, \mathbf{I}_n)$, $n = 200$ and $t = 1, \ldots, 4000$. The panels report respectively, (a) the value of the elements of the first row of $\mathbf{A}_n$, denoted $a_k^{(n)}$ (for $k = 0, \ldots, n-1$); (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.

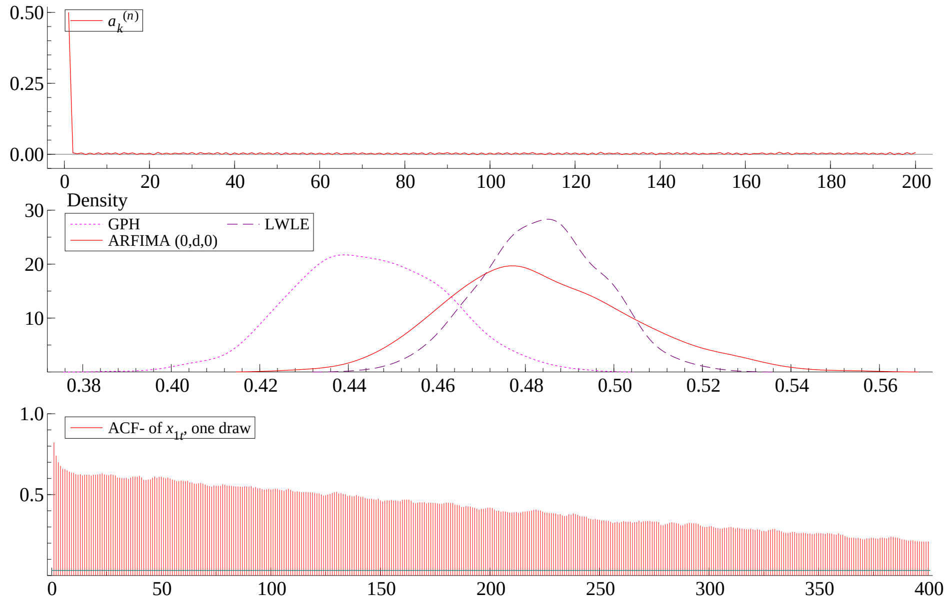

Figure 4: Simulation results for a n-dimensional diagonal VAR(1) $\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t$, with $\mathbf{A}_n = \mathbf{T}_n^* + \eta_n \mathbf{D}_n$, where $\mathbf{T}_n^* \equiv \mathrm{Re}\,(\mathbf{T}_n)$, $\mathbf{T}_n$ has symbol defined by (7), $\eta_n = 1/(n \log(n))$, $\mathbf{D}_n$ is an antisymmetric Toeplitz matrix with above diagonal elements drawn from a standard normal distribution, $d = 0.499$, $\epsilon_t \overset{iid}{\sim} N(\mathbf{0}, \mathbf{I}_n)$, $n = 200$ 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.

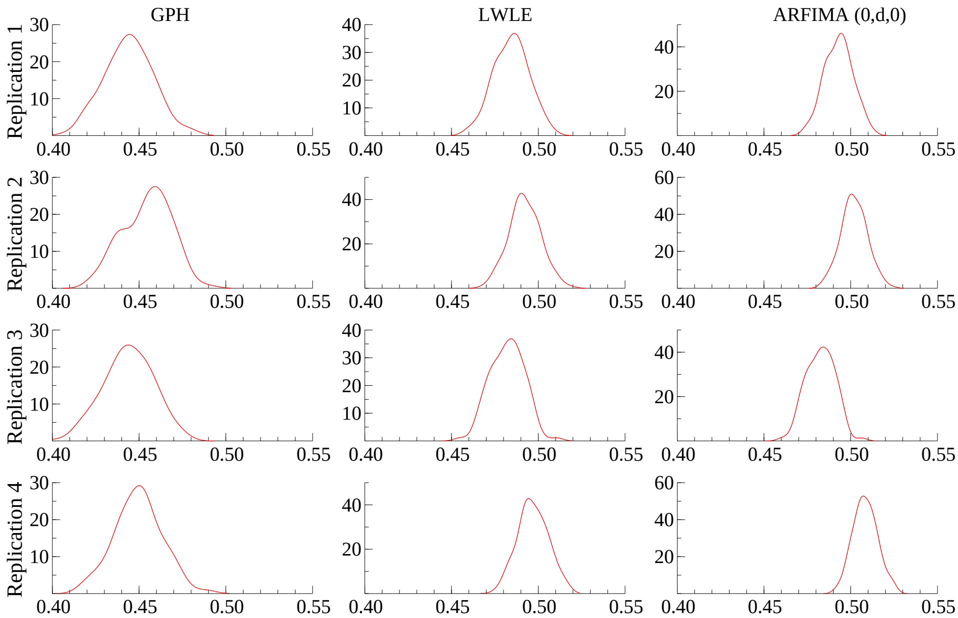

Figure 5: Simulation results for a n-dimensional diagonal VAR(1) $\mathbf{x}_t = \mathbf{A}_n \mathbf{x}_{t-1} + \epsilon_t$, with $\mathbf{A}_n = \mathbf{T}_n^*$, where $\mathbf{T}_n^* \equiv \mathrm{Re}\,(\mathbf{T}_n)$, $\mathbf{T}_n$ has symbol defined by (7), $d = 0.499$, $\epsilon_t \overset{iid}{\sim} N(\mathbf{0}, \mathbf{I}_n)$, $n = 200$ 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.

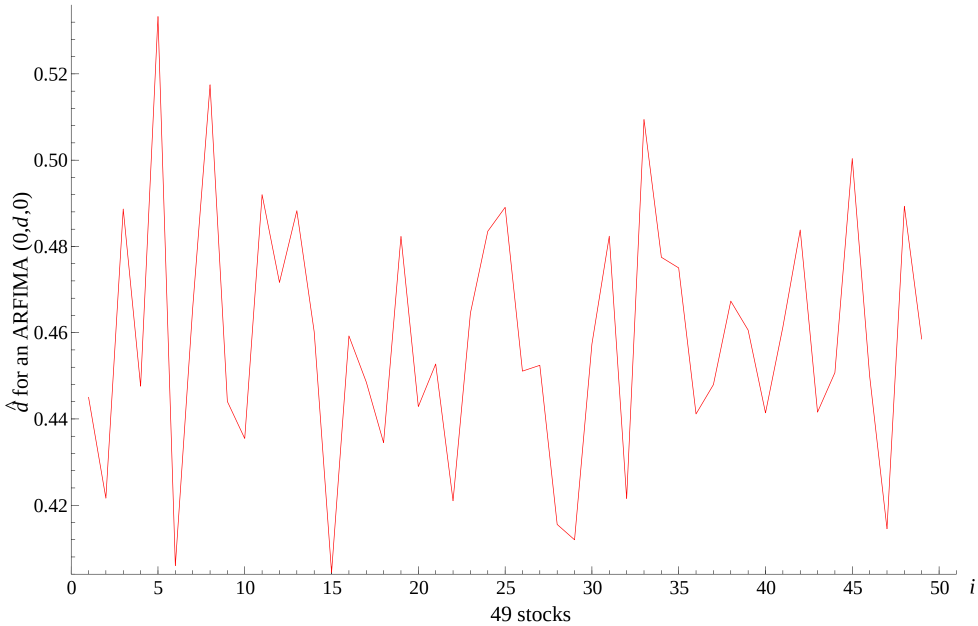

Figure 6: Long memory parameter of an ARFIMA$(0, d, 0)$ model estimated by maximum likelihood on $\log(MedRV_{it})$ for the 49 stocks $i = 1, \ldots, 49$.

"*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]).*"

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).[5] 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.[6]

Figure 6 plots the long memory parameter estimated using an ARFIMA model on $\log(MedRV_{it})$ for $i = 1, \ldots, 49$.[7] 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

[5]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.

[6]If $r_{t,i}$ is the ith 5-minutes return of a day t containing M of such returns, the MedRV of day t is computed as $MedRV_t = \frac{\pi}{6-4\sqrt{3}+\pi} \frac{M}{M-2} \sum_{i=3}^{M} med(|r_{t,i}|, |r_{t,i-1}|, |r_{t,i-2}|)^2$, where $med(\cdot)$ denotes the median.

[7]Similar to the previous section, the ARFIMA model is estimated on $(1-L)^{1/2} \log(MedRV_{it})$ and $1/2$ is added ex-post to the estimated value to ensure the estimated d to lie in the covariance stationarity region.

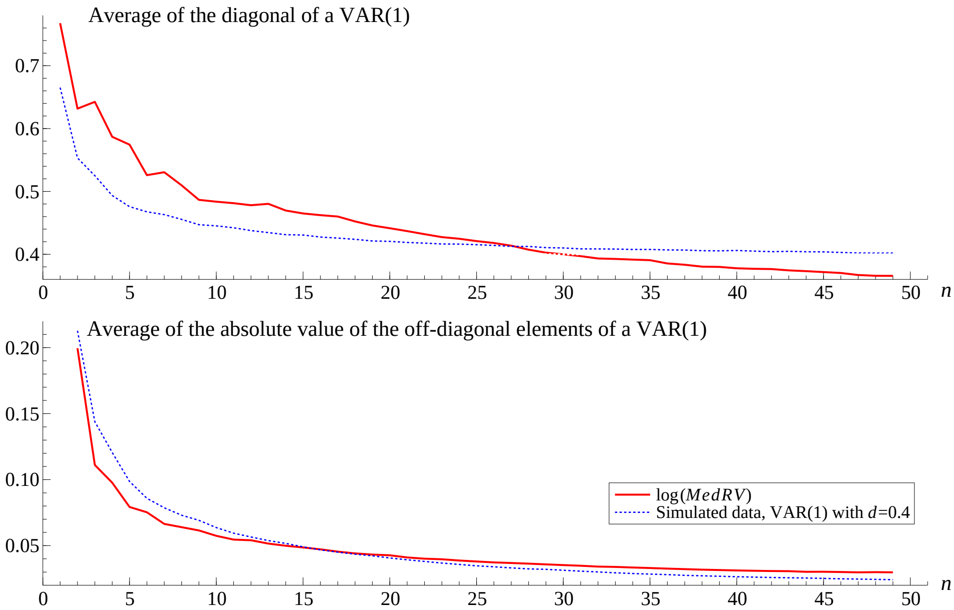

Figure 7: Average of the diagonal elements (upper panel) and average of the absolute value of the off-diagonal elements (lower panels) of a VAR(1) estimated on $\log(MedRV_{it})$ while progressively increasing the dimension of the VAR)

Vahid (2007). Figure 7 plots some summary statistics on the estimated parameters of a VAR(1) model estimated on $\log(MedRV_{it})$, 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_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.

5. 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 whose MA lag polynomial is expressed as a sum that is derived from the determinant and the adjugate of the matrix lag polynomial of the VAR. We then develop three high-level assumptions ensuring that at least one of the elements of the vector process converges in probability to a fractional white noise of degree $\delta \in (0,1)$. We show that these assumptions are satisfied for two specific examples of an n-dimensional VAR(1) model where either (i) all univariate processes tend in probability to an $\mathsf{I}\left(\frac{1}{2}\right)$ fractional white noise as $n \to \infty$, or (ii) one univariate process tends to an $\mathsf{I}(\delta)$ 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. There exist several possible routes for extending our results. For instance, one could relax some of the assumptions on the correlation structure of the VAR innovations, or we could replace $\mathbf{A}_n$, the matrix parameter of the VAR(1), with $\mathbf{V}_n\mathbf{A}_n\mathbf{V}_n^{-1}$ where $\mathbf{V}_n$ denotes a sequence of orthonormal matrices. This would modify the adjugate matrices but not the determinant of $\mathbf{B}_n(L)$. Alternatively, it is possible to consider convergence to long memory processes with richer short-memory dynamics.

6. Appendix

6.1. *Proof of Expression (4)*

Develop expression (4) as

$$\begin{aligned}\left|x_{jt}-\kappa\frac{\det\left(\mathbf{B}_{n-1}\left(L\right)\right)}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\epsilon_{jt}\right| &= \left|\sum_{k\neq j}\frac{\widetilde{\mathbf{B}_{n}\left(L\right)}_{jk}}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\epsilon_{kt}+\frac{\widetilde{\mathbf{B}_{n}\left(L\right)}_{jj}-\kappa\det\left(\mathbf{B}_{n-1}\left(L\right)\right)}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\epsilon_{jt}\right| \\ &\leq \left|\sum_{k\neq j}\frac{\widetilde{\mathbf{B}_{n}\left(L\right)}_{jk}}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\epsilon_{kt}\right|+\left|\frac{\widetilde{\mathbf{B}_{n}\left(L\right)}_{jj}-\kappa\det\left(\mathbf{B}_{n-1}\left(L\right)\right)}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\epsilon_{jt}\right|.\end{aligned}$$

The errors are uncorrelated so

$$\begin{aligned}&\Pr\left(\left|x_{jt}-\kappa\frac{\det\left(\mathbf{B}_{n-1}\left(L\right)\right)}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\epsilon_{jt}\right|>\varepsilon\right) \\ \leq\ &\Pr\left(\left|\sum_{k\neq j}\frac{\widetilde{\mathbf{B}_{n}\left(L\right)}_{jk}}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\epsilon_{kt}\right|>\varepsilon\right)+\Pr\left(\left|\frac{\widetilde{\mathbf{B}_{n}\left(L\right)}_{jj}-\kappa\det\left(\mathbf{B}_{n-1}\left(L\right)\right)}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\epsilon_{jt}\right|>\varepsilon\right).\end{aligned}$$

Assume first that $\det\left(\mathbf{B}_{n}\left(1\right)\right)\neq 0$ and $\det\left(\mathbf{B}_{n}\left(1\right)\right)\not\to 0$ as $n\to\infty$. Then for n large enough, Chebyshev's inequality implies that

$$\begin{aligned}&\Pr\left(\left|x_{jt}-\kappa\frac{\det\left(\mathbf{B}_{n-1}\left(L\right)\right)}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\epsilon_{jt}\right|>\varepsilon\right) \\ \leq\ &\frac{1}{\varepsilon^{2}}\mathsf{Var}\left[\sum_{k\neq j}\left(\frac{\widetilde{\mathbf{B}_{n}}\left(L\right)_{jk}}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\right)\epsilon_{jt}\right]+\frac{1}{\varepsilon^{2}}\mathsf{Var}\left[\frac{\widetilde{\mathbf{B}_{n}\left(L\right)}_{jj}-\kappa\det\left(\mathbf{B}_{n-1}\left(L\right)\right)}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\epsilon_{jt}\right].\end{aligned}$$

Assumption B(i) states that $\mathsf{Var}\left[\sum_{k\neq j}\left(\frac{\widetilde{\mathbf{B}_{n}}\left(L\right)_{jk}}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\right)\epsilon_{jt}\right]\to 0$. Assumption B($ii$) implies that $\frac{\widetilde{\mathbf{B}_{n}\left(z\right)}_{jj}-\kappa\det\left(\mathbf{B}_{n-1}\left(z\right)\right)}{\det\left(\mathbf{B}_{n}\left(z\right)\right)}\to 0$ for all z such that the denominator remains bounded away from zero as $n\to\infty$ hence, since $\det\left(\mathbf{B}_{n}\left(1\right)\right)\not\to 0$, $\mathsf{Var}\left[\frac{\widetilde{\mathbf{B}_{n}\left(L\right)}_{jj}-\kappa\det\left(\mathbf{B}_{n-1}\left(L\right)\right)}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\epsilon_{jt}\right]\to 0$ in particular . Hence, for all $\varepsilon>0$,

$$\Pr\left(\left|x_{jt}-\kappa\frac{\det\left(\mathbf{B}_{n-1}\left(L\right)\right)}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\epsilon_{jt}\right|>\varepsilon\right)\to 0. \tag{11}$$

Now, if the polynomial $\det\left(\mathbf{B}_{n}\left(1\right)\right)=0$ or $\det\left(\mathbf{B}_{n}\left(1\right)\right)\to 0$, then $\left[\det\left(\mathbf{B}_{n}\left(1\right)\right)^{-1}\right]^{+}=O\left(t\right)=O\left(1\right)$ for fixed t and Assumption B(ii) still implies that

$$\mathsf{Var}\left[\frac{\widetilde{\mathbf{B}_{n}\left(L\right)}_{jj}-\kappa\det\left(\mathbf{B}_{n-1}\left(L\right)\right)}{\det\left(\mathbf{B}_{n}\left(L\right)\right)}\epsilon_{jt}\right]\to 0$$

so expression (11) holds here too.

6.2. *Proof of Expression (5)*

To prove the expression, we start by recalling the first Theorem of Szegö, as stated in Grenander and Szegö (1958, reprinted in 2001, Section 5.2 p. 64). For this, we need to define Grenander and

Szegö's class L (2001, Section 1.2, p. 4). Class L denotes the set of all complex-valued functions F which are measurable in the Lebesgue sense and for which[8]

$$\int_0^{2\pi} |F(u)|\, du$$

exists. We denote by $\lambda_k^{(n)}$, $1 \le k \le n+1$, the eigenvalues associated with the $(n+1)$-dimensional Toeplitz Ψ_{n+1} matrix with (ℓ, k) entries $\psi_{\ell-k}$, where $\psi_k = \frac{1}{2\pi}\int_0^{2\pi} e^{-ikx} f\left(e^{iu}\right) du$ and $f(\cdot)$ denotes the symbol of (Ψ_n).[9] We are ready to state the required theorem.

Theorem 2 (Grenander and Szegö, 2001, Sec. 5.2)**.** *Let $x \to f\left(e^{ix}\right)$ be a real-valued function of the class L. We denote by m and M the 'essential' lower and upper bounds of $x \to f\left(e^{ix}\right)$. respectively, and assume that m and M are finite. if $F(\lambda)$ is any continuous function defined in the finite interval $m \le \lambda \le M$, we have*

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n+1} F\left(\lambda_k^{(n)}\right)}{n+1} = \frac{1}{2\pi}\int_0^{2\pi} F\left[f\left(e^{ix}\right)\right] dx.$$

As a corollary, using $F(\cdot) = \log(\cdot)$, the determinant of Ψ_{n+1} is $\det(\Psi_{n+1}) = \prod_{j=1}^{n+1} \lambda_j^{(n)}$ so, provided all eigenvalues are strictly positive,

$$\lim_{n\to\infty} \log\det(\Psi_{n+1})^{\frac{1}{n+1}} = \frac{1}{2\pi}\int_0^{2\pi} \log f\left(e^{iu}\right) du$$

and

$$\lim_{n\to\infty} \det(\Psi_{n+1})^{\frac{1}{n+1}} = \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log f\left(e^{iu}\right) du\right\}.$$

Hence $\det(\Psi_{n+1})^{\frac{1}{n+1}} \sim \det(\Psi_n)^{\frac{1}{n}}$ so, noticing that

$$\frac{\det(\Psi_{n+1})}{\det(\Psi_n)} = \det(\Psi_n)^{\frac{1}{n}} \left[\frac{\det(\Psi_{n+1})^{\frac{1}{n+1}}}{\det(\Psi_n)^{\frac{1}{n}}}\right]^{n+1},$$

it follows that $\frac{\det(\Psi_{n+1})}{\det(\Psi_n)} \sim \det(\Psi_n)^{\frac{1}{n}}$ and

$$\lim_{n\to\infty} \frac{\det(\Psi_{n+1})}{\det(\Psi_n)} = \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log f\left(e^{iu}\right) du\right\}. \tag{12}$$

Now we need to show that the symbol associated with $\Psi_n = \mathbf{I}_n - \mathbf{T}_{\delta,n} z$ is of class L with $m > 0$ for all $z \in (-\infty, 1)$. For this, we notice that the required symbol is $1 - g_\delta(\cdot)\, z$ which is of class L if g_δ is of class L too and its minimum is $1 - x > 0$. This concludes the proof.

6.3. Proof of Theorem 1

Together, Assumptions B, T and A, and hence expressions (4) and (5), imply that there exist $(t, j) \in \mathbb{N}^2$ such that for all $\varepsilon > 0$,

$$\Pr\left(\left|x_{jt} - \frac{\det(\mathbf{I}_{n-1} - \mathbf{T}_{n-1}L)}{\det(\mathbf{I}_n - \mathbf{T}_n L)}\epsilon_{jt}\right| > \varepsilon\right) \underset{n\to\infty}{\to} 0,$$

[8] The original definition is for the integral over $[-\pi, \pi]$. To match the setting of this paper, we shift it to $[0, 2\pi]$.

[9] Grenander and Szegö let f denote the spectral density, here it denotes the symbol of (Ψ_n) for notational consistency. The spectral density evaluated at ω coincides with the symbol evaluated at $e^{i\omega}$.

and

$$\Pr\left(\left|\left[\frac{\det\left(\mathbf{I}_{n-1}-\mathbf{T}_{\delta,n-1}L\right)}{\det\left(\mathbf{I}_{n}-\mathbf{T}_{\delta,n}L\right)}-\Delta^{-\delta}\right]\epsilon_{jt}\right|>\varepsilon\right)\underset{n\to\infty}{\to}0.$$

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

$$\frac{\det\left(\mathbf{I}_{n-1}-\mathbf{T}_{n-1}z\right)}{\det\left(\mathbf{I}_{n}-\mathbf{T}_{n}z\right)}-\frac{\det\left(\mathbf{I}_{n-1}-\mathbf{T}_{\delta,n-1}z\right)}{\det\left(\mathbf{I}_{n}-\mathbf{T}_{\delta,n}z\right)}\to 0.$$

Assumption T(i) states that $g(\cdot,\cdot)$ is real-valued, which implies that $t_{d,k}=\int_0^{2\pi}g\left(d,e^{i\omega}\right)e^{-ik\omega}d\omega=\overline{t_{d,-k}}$, i.e., $\mathbf{T}_{d,n}$ is Hermitian. This entails in particular that $t_k^{(n)}+t_{n-k}^{(n)}=t_k^{(n)}+t_{-k}^{(n)}\in\mathbb{R}$ with the notations of Assumption T. Also $g_d(\cdot)=g(d,\cdot)$ 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 $g_d\left(e^{i\omega}\right)$ which characterizes $\mathbf{T}_{d,n}$ as belonging to the Wiener Class (Gray, 2006, p. 40). Under Assumption T(iii), $\mathbf{T}_n$ and $\mathbf{T}_{\delta_n,n}$ — the matrix with entries $t_{\delta_n,k}$ defined as $t_{d,k}$ evaluated at $d=\delta_n$ — are asymptotically equivalent under the weak norm, 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 $\mathbf{T}_{\delta,n}$ within the Wiener class, we can associate a Circulant matrix $\mathbf{C}_{\delta,n}$ such that $\mathbf{C}_{\delta,n}\sim\mathbf{T}_{\delta,n}$ defined as

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

The sequence $(\mathbf{C}_{\delta,n})$ is not uniquely defined, see Grenander and Szegö (1958, Section 7.6). Gray (2006, Lemma 4.6) shows for instance that choosing $t_{\delta,-k}+t_{\delta,n-k}$ for the entries $c_{\delta,k}^{(n)}$ yields a matrix which is asymptotically equivalent to $\mathbf{T}_{\delta,n}$ but this is not how we define $c_{\delta,k}^{(n)}$ here. Instead, we define circulant matrices $\mathbf{C}_{\delta,n}$ and $\mathbf{C}_n$ with entries $c_{\delta,k}^{(n)}$ and $c_k^{(n)}$ such that, respectively,

$$c_{\delta,k}^{(n)}=t_{\delta,-k}^{(n)}+t_{\delta,n-k}^{(n)} \tag{14a}$$

$$c_{k}^{(n)}=t_{-k}^{(n)}+t_{n-k}^{(n)} \tag{14b}$$

with $t_{\delta,k}^{(n)}=\frac{1}{n}\sum_{j=0}^{n-1}g\left(\delta,e^{i\frac{2\pi j}{n}}\right)e^{-2i\pi jk/n}$. Since $t_k^{(n)}$, defined in Assumption T$(iii.a)$, converges to $t_{\delta,k}$, the matrix $\mathbf{C}_n$ is asymptotically equivalent to that with entries $t_{\delta_n,-k}+t_{\delta_n,n-k}$. 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, it also holds that $\mathbf{I}_n-\mathbf{C}_nz\sim\mathbf{I}_n-\mathbf{T}_nz$ and from Theorem 2.1 of Gray (2006, p. 17), $\left(\mathbf{I}_n-\mathbf{C}_nz\right)^{-1}\sim\left(\mathbf{I}_n-\mathbf{T}_nz\right)^{-1}$ for all z such that the inverse of the determinants remain bounded in strong norm. Hence, also from the same theorem,

$$\left(\mathbf{I}_n-\mathbf{C}_nz\right)^{-1}\left(\mathbf{I}_{n-1}-\mathbf{C}_{n-1}z\right)\sim\left(\mathbf{I}_n-\mathbf{T}_nz\right)^{-1}\left(\mathbf{I}_{n-1}-\mathbf{T}_{n-1}z\right).$$

Now for the set $z \in \mathcal{Z}$ such that $\left|\det\left[(\mathbf{I}_n - \mathbf{C}_n z)^{-1}(\mathbf{I}_{n-1} - \mathbf{C}_{n-1} z)\right]\right| < \infty$ (we show later in the proof that $\mathcal{Z} = (-\infty, 1)$), we may use Corollary 2.4 of Gray (2006, p. 24) that[10]

$$\det\left((\mathbf{I}_{n-1} - z\mathbf{C}_{n-1})^{-1}(\mathbf{I}_n - z\mathbf{C}_n)\right) \sim \det\left((\mathbf{I}_{n-1} - z\mathbf{T}_{n-1})^{-1}(\mathbf{I}_{n-1} - z\mathbf{T}_{n-1})\right)$$

and as the limit of $\det\left((\mathbf{I}_{n-1} - z\mathbf{C}_{n-1})^{-1}(\mathbf{I}_n - z\mathbf{C}_n)\right)$ is finite for $z \in \mathcal{Z}$ (Szegö's first theorem), we conclude that

$$\det\left((\mathbf{I}_n - \mathbf{C}_n z)^{-1}(\mathbf{I}_{n-1} - \mathbf{C}_{n-1} z)\right) \sim \det\left((\mathbf{I}_n - \mathbf{T}_n z)^{-1}(\mathbf{I}_{n-1} - \mathbf{T}_{n-1} z)\right),$$

i.e.

$$\frac{\det(\mathbf{I}_{n-1} - z\mathbf{C}_{n-1})}{\det(\mathbf{I}_n - z\mathbf{C}_n)} \sim \frac{\det(\mathbf{I}_{n-1} - z\mathbf{T}_{n-1})}{\det(\mathbf{I}_n - z\mathbf{T}_n)}. \tag{15}$$

The symbol of $\mathbf{I}_n - z\mathbf{T}_{\delta,n}$ is $1 - g_\delta(\cdot)\, z$. Hence Gray (2006, Lemma 4.6) implies that:

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

Note that

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

and

$$\begin{aligned}\frac{\det(\mathbf{I}_{n-1} - z\mathbf{C}_{n-1})}{\det(\mathbf{I}_n - z\mathbf{C}_n)} &= \frac{\det(\mathbf{I}_{n-1} - z\mathbf{C}_{\delta,n-1})}{\det(\mathbf{I}_n - z\mathbf{C}_{\delta,n})} \\ &\times \frac{\det\left(\mathbf{I}_{n-1} + z(\mathbf{I}_{n-1} - z\mathbf{C}_{\delta,n-1})^{-1}(\mathbf{C}_{\delta,n-1} - \mathbf{C}_{n-1})\right)}{\det\left(\mathbf{I}_n + z(\mathbf{I}_n - z\mathbf{C}_{\delta,n})^{-1}(\mathbf{C}_{\delta,n} - \mathbf{C}_n)\right)}.\end{aligned} \tag{16}$$

We now consider the limit of

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

Let $\mathbf{H}_n(z) = z(\mathbf{I}_{n-1} - z\mathbf{C}_{\delta,n-1})^{-1}(\mathbf{C}_{\delta,n-1} - \mathbf{C}_{n-1})$. For all $z \in \mathcal{Z}$, by Assumption T(*iii.b*) $t_{\delta,0}^{(n)} - t_0^{(n)} = o\left(n^{-1}\right)$ so $c_{\delta,0}^{(n)} - c_0^{(n)} = o\left(n^{-1}\right)$ and the diagonal elements of $\mathbf{H}_n(z)$ are $o\left(n^{-1}\right)$ so $\mathrm{tr}\mathbf{H}_n(z) = o(1)$ .Let $\mu_k^{(n)}$, $k = 1, ..., n$, denote the eigenvalues of $\mathbf{H}_n(z)$. Then

$$\det(\mathbf{I}_n + \mathbf{H}_n(z)) = \prod_{k=1}^n \left(1 + \mu_k^{(n)}\right).$$

[10] The corollary is that

$$\lim_{n\to\infty} \sqrt[n]{\det\left((\mathbf{I}_{n-1} - z\mathbf{C}_{n-1})^{-1}(\mathbf{I}_n - z\mathbf{C}_n)\right)} = \lim_{n\to\infty} \sqrt[n]{\det\left(\left(\mathbf{I}_{n-1} - z\mathbf{T}_{n-1}^*\right)^{-1}\left(\mathbf{I}_{n-1} - z\mathbf{T}_{n-1}^*\right)\right)}$$

when the matrices do not have nonpositive eigenvalues. Here we take the nth power of the expression and consider asymptotic equivalence only, so negative eigenvalues are not precluded.

We know that as $n \to \infty$, the eigenvalues $\mu_k^{(n)} \to 0$ since $\mathbf{C}_{\delta,n-1} - \mathbf{C}_{n-1} \to 0$, hence for n large enough, $\log\det(\mathbf{I}_n + \mathbf{H}_n(z))$ exists and we may use Jacobi's formula (Abadir and Magnus, 2005, result 13.36 p 372) that implies that since $\mathrm{tr}\mathbf{H}_n(z) = o(1)$,

$$\det(\mathbf{I}_n + \mathbf{H}_n(z)) = O(1).$$

Therefore, there exists $\kappa_0 \in (0,\infty)$ such that $\det(\mathbf{I}_{n-1} + \mathbf{H}_n(z)) \to \kappa_0$ and 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 $\lim_{n\to\infty} \frac{\det(\mathbf{I}_{n-1} - z\mathbf{C}_{\delta,n-1})}{\det(\mathbf{I}_n - z\mathbf{C}_{\delta,n})} = (1-z)^{-\delta}$, it follows that $\forall\varepsilon > 0$ and $z \in (-\infty, 1)$, there exist $\eta = \varepsilon(1-z)^{\delta}$ and N such that for $n > N$

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

It follows from Equation (15) that

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

There remains to consider the set $\mathcal{Z}$ defined above as the set of z such that $\mathbf{I}_n - z\mathbf{C}_n$ does not asymptotically possess zero eigenvalues. As we showed above, for $z \in \mathcal{Z}$, $\frac{\det(\mathbf{I}_{n+1} - z\mathbf{T}_{\delta,n+1})}{\det(\mathbf{I}_n - z\mathbf{T}_{\delta,n})} \sim \frac{\det(\mathbf{I}_{n-1} - z\mathbf{C}_{n-1})}{\det(\mathbf{I}_n - z\mathbf{C}_n)}$ and $\frac{\det(\mathbf{I}_{n+1} - z\mathbf{T}_{\delta,n+1})}{\det(\mathbf{I}_n - z\mathbf{T}_{\delta,n})} \sim \det(\mathbf{I}_n - z\mathbf{T}_{\delta,n})^{\frac{1}{n}}$ and similarly for $\det(\mathbf{I}_n - z\mathbf{C}_n)^{\frac{1}{n}}$. Hence the geometric averages of the eigenvalues of $\mathbf{I}_n - z\mathbf{T}_{\delta,n}$ and $\mathbf{I}_n - z\mathbf{C}_n$ coincide. The eigenvalues of $\mathbf{I}_n - z\mathbf{C}_n$ cannot therefore be negative or zero if those of $\mathbf{I}_n - z\mathbf{T}_{\delta,n}$ are not, and the latter are positive if $1 - g_\delta(\cdot)\, z > 0$, i.e. for $z < 1$. Hence $\mathcal{Z} = (-\infty, 1)$, which is the domain of definition of $(1-z)^{-\delta}$ for $\delta \in (0,1)$.

It follows that

$$\Pr\left(\left| x_{jt} - (1-L)^{-\delta} \epsilon_{jt} \right| > \varepsilon\right) \underset{n\to\infty}{\to} 0,$$

and the theorem holds.

6.4. Proofs relative to Subsection 3.1

We collect here the proofs related to Section 3.1 that show that Assumptions B, T and A are satisfied for $\mathbf{A}_n = \mathbf{T}_n^* + \eta_n \mathbf{D}_n$, where $\mathbf{T}_n^*$, η_n and $\mathbf{D}_n$ are specified as in Section 3.1. We need the following lemmas whose proofs are provided in Subsection 6.6.

Lemma L(i): Under the assumptions of Subsection 3.1 and as $n \to \infty$, the coefficients of $\mathbf{T}_n^*$ satisfy,

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

Lemma L(ii): Under the assumptions of both examples of Section 3 and as $n \to \infty$, the coefficients of $\mathbf{T}_n^*$ satisfy for all k, $-n < k < n$,

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

Proofs of the lemmas: see Subsection 6.6.

6.4.1. Proof of the validity of Assumption T for the matrix $\mathbf{T}_n$

Assumptions T(i), T(iii) and T(iv) follow from the definitions of g and δ_n and of Lemma L(i). 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

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

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

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

Hence Assumption T holds.

6.4.2. Proof of the validity of Assumption B for the matrix $\mathbf{A}_n = \mathbf{T}_n^* + \eta_n \mathbf{D}_n$

We let $j = 1$ without loss of generality and start by showing Assumption B(ii) holds.

Elements $\widetilde{\mathbf{B}_n(z)}_{1k}$, for $k = 1, ..., n$, of the first row of $\widetilde{\mathbf{B}_n(z)}$, satisfy $\widetilde{\mathbf{B}_n(z)}_{1k} = (-1)^{k+1} \det\left(\mathbf{CoB}_n(z)_{1k}\right)$, where $\mathbf{CoB}_n(z)_{\ell k}$ is the (ℓ, k) 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 $\gamma_{\ell k}$ denotes the (ℓ, k) 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

$$\mathbf{CoB}_n(z)_{11} = \mathbf{B}_{n-1}(z) + \left(\mathbf{T}_{n-1}^* - \mathbf{T}_n^{*(1)} + \eta_{n-1} \mathbf{D}_{n-1} - \eta_n \mathbf{D}_n^{(1)}\right) z. \tag{17}$$

Consider first the case where $\det(\mathbf{B}_{n-1}(z)) \nrightarrow 0$ and $\det(\mathbf{B}_{n-1}(z)) \neq 0$ as $n \to \infty$, then (17) implies

$$\widetilde{\mathbf{B}_n(z)}_{11} = \det(\mathbf{B}_{n-1}(z)) \det\left(\mathbf{I}_{n-1} + [\mathbf{B}_{n-1}(z)]^{-1}\left(\mathbf{T}^*_{n-1} - \mathbf{T}^{*(1)}_n + \eta_{n-1}\mathbf{D}_{n-1} - \eta_n \mathbf{D}^{(1)}_n\right) z\right). \tag{18}$$

Now the elements of $\mathbf{T}^*_{n-1} - \mathbf{T}^{*(1)}_n$ tend to zero as $n \to \infty$, hence the eigenvalues of $\mathbf{T}^*_{n-1} - \mathbf{T}^{*(1)}_n$ also tend to zero and there exists N such that for $n \geq N$, the eigenvalues of $\mathbf{I}_{n-1} + [\mathbf{B}_{n-1}(z)]^{-1}\left(\mathbf{T}^*_{n-1} - \mathbf{T}^{*(1)}_n\right) z$ are strictly positive. We now use result 12.30 of Abadir and Magnus (2005) that states that if matrix $\mathbf{I}+\mathbf{M}$ is positive definite then

$$\det(\mathbf{I}+\mathbf{M}) \leq \exp \operatorname{tr}(\mathbf{M}). \tag{19}$$

Hence for $n \geq N$, Lemma L(ii) implies that the second element on the right-hand side of expression (18) is bounded. Hence

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

where $\kappa = \lim_{n\to\infty} \det\left(\mathbf{I}_{n-1} + [\mathbf{B}_{n-1}(z)]^{-1}\left(\mathbf{T}^*_{n-1} - \mathbf{T}^{*(1)}_n + \eta_{n-1}\mathbf{D}_{n-1} - \eta_n \mathbf{D}^{(1)}_n\right) z\right) > 0$. Now let $\det(\mathbf{B}_{n-1}(z)) = 0$ or $\det(\mathbf{B}_{n-1}(z)) \to 0$, then also $\widetilde{\mathbf{B}_n(z)}_{11} \to 0$ and $\widetilde{\mathbf{B}_n(z)}_{11} \sim \kappa \det(\mathbf{B}_{n-1}(z))$ as $n \to \infty$, and expression B(ii) holds. This constitutes the first part of the proof.

We now turn to B(i). We first consider $\widetilde{\mathbf{B}_n(z)}_{1k}$, $\forall k \neq 1$, for $z < 1$. 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 $n \to \infty$:

$$\widetilde{\mathbf{B}_n(z)}_{12} \sim -\det\left(\begin{bmatrix} -t^{*(n)}_1 z & -t^{*(n)}_1 z & \cdots & -t^{*(n)}_{n-2} z \\ -t^{*(n)}_2 z & 1 - t^{*(n)}_0 z & \ddots & \vdots \\ \vdots & \ddots & \ddots & -t^{*(n)}_1 z \\ -t^{*(n)}_{n-1} z & -t^{*(n)}_{n-3} z & \cdots & 1 - t^{*(n)}_0 z \end{bmatrix}\right).$$

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

$$\widetilde{\mathbf{B}_n(z)}_{12} \sim -\left(\max_{0<k<n}\left|t^{*(n)}_k\right| z\right) \det\left(\begin{bmatrix} \frac{-t^{*(n)}_1}{\max_{0<k<n}\left|t^{*(n)}_k\right|} & -t^{*(n)}_2 z & \cdots & -t^{*(n)}_{n-2} z \\ \frac{-t^{*(n)}_1}{\max_{0<k<n}\left|t^{*(n)}_k\right|} & 1 - t^{*(n)}_0 z & \ddots & \vdots \\ \vdots & \vdots & \ddots & -t^{*(n)}_1 z \\ \frac{-t^{*(n)}_{n-1}}{\max_{0<k<n}\left|t^{*(n)}_k\right|} & -t^{*(n)}_{n-3} z & \cdots & 1 - t^{*(n)}_0 z \end{bmatrix}\right).$$

Without loss of generality, we assume for instance that $\max_{0<k<n}\left|t^{*(n)}_k\right| = \left|t^{*(n)}_1\right|$. Lemma L($i$) shows that for $k \neq 0$, $t^{*(n)}_k = O\left(n^{-1}\right)$, hence,

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

The formula for the determinant of partitioned matrices is

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

where in our case the second determinant on the right, i.e., $\det\left(\mathbf{B}_{n-2}(z) + O(1)\,O\left(n^{-1}\right)\right) \sim \det(\mathbf{B}_{n-2}(z))$ as $n\to\infty$. Therefore

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

Now, we introduce the Hadamard polynomial product, which is defined for $P(z) = \sum_{k\geq 0} p_k z^k$ and $Q(z) = \sum_{k\geq 0} q_k z^k$ as $P\circ Q(z) = \sum_{k=0}^{\min(\deg P, \deg Q)} p_k q_k z^k$ and $P(z)^{\circ 2} = P(z)\circ P(z)$. Below $[P(z)]_{z=1}$ refers to $P(z)$ evaluated at $z=1$. Hence

$$\mathsf{Var}\left[\sum_{k\neq j}\frac{\widetilde{\mathbf{B}_n}(z)_{jk}}{\det(\mathbf{B}_n(z))}\epsilon_{kt}\right] = \sum_{k\neq j}\left|\frac{\widetilde{\mathbf{B}_n}(z)_{jk}}{\det(\mathbf{B}_n(z))}\right|_{z=1}^{\circ 2}\sigma_{\epsilon_k}^2 \tag{20}$$

where

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

Hence as $n\to\infty$,

$$\sum_{k=2}^{n}\left|\frac{\widetilde{\mathbf{B}_n(z)}_{1\ell}}{\det(\mathbf{B}_n(z))}\right|^{\circ 2} = \frac{\sigma_\epsilon^2}{2\pi}\left|\frac{\det(\mathbf{B}_{n-2}(z))}{\det(\mathbf{B}_n(z))}\right|^{\circ 2} O\left(n^{-1}\right) + o\left(\frac{1}{n}\left|\frac{\det(\mathbf{B}_{n-2}(z))}{\det(\mathbf{B}_n(z))}\right|^{\circ 2}\right). \tag{21}$$

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 $n\to\infty$, under assumption T (using the same argument as used in proving Theorem 1),

$$\frac{\det(\mathbf{B}_{n-1}(z))}{\det(\mathbf{B}_n(z))} \sim \frac{\det(\mathbf{I}_{n-1} - z\mathbf{C}_{n-1})}{\det(\mathbf{I}_n - z\mathbf{C}_n)}. \tag{22}$$

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

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

hence the second term on the right-hand side in expression (21) is $o\left(n^{-1}\right)$ when $z\neq 1$.

Now for $z\to 1$, the truncated polynomial $\left[(1-z)^{-1}\right]^{+}$ evaluated at $z=1$ takes the value t so

$$\sum_{k=2}^{n}\left|\frac{\widetilde{\mathbf{B}_n(z)}_{1k}}{\det(\mathbf{B}_n(z))}\right|_{z=1}^{\circ 2} = O\left(n^{-1}\right).$$

Together with expression (20), this shows that Assumption B(i) holds.

6.4.3. Proof of the validity of Assumption A

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

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

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

6.5. Proofs relative to Subsection 3.2

In this example, many of the proofs follow from the results above where we notice that Lemma L(ii) has been shown to hold also under the assumptions of Subsection 3.2. By construction Assumption A holds since $\mathbf{A}_n=\mathbf{T}_n^*$. Also, Assumptions T$(i)$ and T(ii) , T$(iii.a)$ and $T\,(iv)$ follow as shown in Subsection 3.1. Now $\delta_n=\delta$ for all n so $t_{\delta,0}^{(n)}-t_0^{(n)}=0$ and Assumption T$(iii.b)$ holds also.

We now consider Assumption B, starting with B(ii). We follow for this the lines of the proof provided in the previous subsection. First, if $\det\left(\mathbf{B}_{n-1}\left(z\right)\right)\nrightarrow 0$ and $\det\left(\mathbf{B}_{n-1}\left(z\right)\right)\neq 0$,

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

where the elements and eigenvalues of $\mathbf{T}_{n-1}^{*}-\mathbf{T}_{n}^{*(1)}$ tend to zero so there exists N such that for $n\geq N$, $\mathbf{I}_{n-1}+\left[\mathbf{B}_{n-1}\left(z\right)\right]^{-1}\left(\mathbf{T}_{n-1}^{*}-\mathbf{T}_{n}^{*(1)}\right)z$ is positive definite. We showed previously using Lemma L(ii) which holds here too, that

$$\det\left(\mathbf{I}_{n-1}+\left[\mathbf{B}_{n-1}\left(z\right)\right]^{-1}\left(\mathbf{T}_{n-1}^{*}-\mathbf{T}_{n}^{*(1)}\right)z\right)=O\left(1\right).$$

Hence there exists $\kappa_0>0$ such that $\det\left(\mathbf{I}_{n-1}+\left[\mathbf{B}_{n-1}\left(z\right)\right]^{-1}\left(\mathbf{T}_{n-1}^{*}-\mathbf{T}_{n}^{*(1)}\right)z\right)\rightarrow\kappa_0$, as $n\rightarrow\infty$. Therefore

$$\begin{aligned}\widetilde{\mathbf{B}_n\left(z\right)}_{11}&=\det\left(\mathbf{B}_{n-1}\left(z\right)\right)\det\left(\mathbf{I}_{n-1}+\left[\mathbf{B}_{n-1}\left(z\right)\right]^{-1}\left(\mathbf{T}_{n-1}^{*}-\mathbf{T}_{n}^{*(1)}\right)z\right)\\&\underset{n\rightarrow\infty}{\sim}\kappa_0\det\left(\mathbf{B}_{n-1}\left(z\right)\right).\end{aligned}$$

The result also holds if $\det\left(\mathbf{B}_{n-1}\left(z\right)\right)\rightarrow 0$ or $\det\left(\mathbf{B}_{n-1}\left(z\right)\right)=0$ since it then also holds that $\widetilde{\mathbf{B}_n\left(z\right)}_{11}\rightarrow 0$.

We now consider Assumption B(i) , if $\sigma_{n,k}^2=o\left(n^{-1}\right)$ then $\sum_{\substack{k=1\\k\neq j}}^{n}\sigma_{n,k}^2=o\left(1\right)$ so whether or not $\det\left(\mathbf{B}_n\left(1\right)\right)\nrightarrow 0$,

$$\sum_{\substack{k=1\\k\neq j}}^{n}\left[\frac{\widetilde{\mathbf{B}_n\left(z\right)}_{jk}}{\det\left(\mathbf{B}_n\left(z\right)\right)}\right]_{z=1}^{+\circ 2}\sigma_{n,k}^2=O\left(\sum_{\substack{k=1\\k\neq j}}^{n}\sigma_{n,k}^2\right)=o\left(1\right).$$

And finally B(iii) holds by construction.

6.6. Proofs the Lemmas

Proof of L(i)**.** Since $\mathbf{T}_n^* = \mathrm{Re}\left(\mathbf{T}_n\right)$, we start expressing the coefficients of the latter matrix. Since $g\left(d,x\right) = 1_{\{0\le x<\pi d\}} + 1_{\left\{\pi\left(\frac{3}{2}-d\right)<x\le\frac{3\pi}{2}\right\}}$ for $d\in(0,1)$ and $x\in[0,2\pi]$, the coefficients of $\mathbf{T}_n$ satisfy:

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

Hence for $k=0$,

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

and therefore when $n^2\left(\delta-\delta_n\right)\to 0$, $t_0^{*(n)} = \delta + O\left(n^{-1}\right)$.

Now, when $k\neq 0$,

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

so

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

i.e.

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

we keep simplifying

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

so finally

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

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

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

where

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

Therefore

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

where

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

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

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

and therefore when $n^2(1/2-\delta_n) \to 0$, $\xi_k(\delta_n, n) = o\left(n^{-1}\right)$ while

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

We notice that

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

Hence

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

Now,

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

Notice that $k\left(\frac{1}{2}-\delta_n\right)=o\left(n^{-1}\right)\ \forall k<n$, hence for k odd, $t_{-k}^{*(n)}=O\left(n^{-1}\right)$. 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 $\sin\frac{\pi k\delta_n}{2}=\sin 2\pi m\delta_n=O\left(m\left(\frac{1}{2}-\delta_n\right)\right)=o\left(n^{-1}\right)$ and if $k=4m+2$, then $\cos\frac{\pi k}{4}=\cos\left(m\pi+\frac{\pi}{2}\right)=0$. Hence for all k such that $0<|k|<n$,

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

which concludes the proof of L(i).

Proof of L(ii).

Recall that $t_{-k}^{*(n)}=\mathrm{Re}\frac{1}{n}\sum_{\ell=0}^{n-1}g\left(\delta_n,e^{i\frac{2\pi\ell}{n}}\right)e^{-2i\pi\ell k/n}$ and notice that $n^{-1}=(n-1)^{-1}\left(1-n^{-1}\right)$ so, whether or not δ_n is constant,

$$\begin{aligned}
t_{-k}^{*(n-1)}-t_{-k}^{*(n)} &= \frac{1}{n(n-1)}\mathrm{Re}\left[n\sum_{\ell=0}^{n-2}g\left(\delta_{n-1},e^{i\frac{2\pi\ell}{n-1}}\right)e^{-2i\pi\ell k/(n-1)}-(n-1)\sum_{\ell=0}^{n-1}g\left(\delta_n,e^{i\frac{2\pi\ell}{n}}\right)e^{-2i\pi\ell k/n}\right] \\
&= \frac{1}{n(n-1)}\mathrm{Re}\left[n\sum_{\ell=0}^{n-2}\left[g\left(\delta_{n-1},e^{i\frac{2\pi\ell}{n-1}}\right)e^{-2i\pi\ell k/(n-1)}-g\left(\delta_n,e^{i\frac{2\pi\ell}{n}}\right)e^{-2i\pi\ell k/n}\right]\right. \\
&\quad\left.-(n-1)\,g\left(\delta_n,e^{i\frac{2\pi(n-1)}{n}}\right)e^{-2i\pi(n-1)k/n}+\sum_{\ell=0}^{n-1}g\left(\delta_n,e^{i\frac{2\pi\ell}{n}}\right)e^{-2i\pi\ell k/n}\right].
\end{aligned}$$

Hence,

$$\begin{aligned}
t_{-k}^{*(n-1)}-t_{-k}^{*(n)} &= \frac{1}{n(n-1)}\mathrm{Re}\left[n\sum_{\ell=0}^{n-2}\left[g\left(\delta_{n-1},e^{i\frac{2\pi\ell}{n-1}}\right)-g\left(\delta_n,e^{i\frac{2\pi\ell}{n}}\right)e^{2i\pi\ell k\frac{1}{n(n-1)}}\right]e^{-2i\pi\ell k/(n-1)}\right.\\
&\quad\left.+\sum_{\ell=0}^{n-1}g\left(\delta_n,e^{i\frac{2\pi\ell}{n}}\right)e^{-2i\pi\ell k/n}-(n-1)\,g\left(\delta_n,e^{i\frac{2\pi(n-1)}{n}}\right)e^{-2i\pi(n-1)k/n}\right]\\
&= \frac{1}{n-1}\mathrm{Re}\left[\sum_{\ell=0}^{n-2}\left[g\left(\delta_{n-1},e^{i\frac{2\pi\ell}{n-1}}\right)-g\left(\delta_n,e^{i\frac{2\pi\ell}{n}}\right)e^{2i\pi\ell k\frac{1}{n(n-1)}}\right]e^{-2i\pi\ell k/(n-1)}\right]\\
&\quad+\frac{1}{n-1}t_{-k}^{*(n)}-\frac{1}{n}g\left(\delta_n,e^{i\frac{2\pi(n-1)}{n}}\right)e^{-2i\pi(n-1)k/n}
\end{aligned}$$

i.e., using the definition of $g(\cdot,\cdot)$

$$\begin{aligned}
t_{-k}^{*(n-1)}-t_{-k}^{*(n)} &= \frac{1}{n-1}\mathrm{Re}\left[\sum_{\ell=0}^{n-2}\left[1_{\left\{\ell<\frac{(n-1)\delta_{n-1}}{2}\right\}}+1_{\left\{\left(\frac{3(n-1)}{4}-\frac{(n-1)\delta_{n-1}}{2}\right)<\ell\leq\frac{3(n-1)}{4}\right\}}\right.\right.\\
&\quad\left.\left.-\left(1_{\left\{\ell<\frac{n\delta_n}{2}\right\}}+1_{\left\{\left(\frac{3n}{4}-\frac{n\delta_n}{2}\right)<\ell\leq\frac{3n}{4}\right\}}\right)e^{2i\pi\ell k\frac{1}{n(n-1)}}\right]e^{-2i\pi\ell k/(n-1)}\right]\\
&\quad+\frac{1}{n-1}t_{-k}^{*(n)}-\frac{1}{n}g\left(\delta_n,e^{i\frac{2\pi(n-1)}{n}}\right)e^{-2i\pi(n-1)k/n},
\end{aligned}$$

and,

$$\begin{aligned}
t_{-k}^{*(n-1)}-t_{-k}^{*(n)} &\sim \frac{1}{n-1}\left(-\cos\pi\delta k+\cos\frac{3-2\delta}{2}\pi k-\cos\frac{3\pi}{2}k\right)\\
&\quad+\frac{1}{n-1}\mathrm{Re}\left[\sum_{\ell=0}^{n-2}\left[1_{\left\{\ell<\frac{n\delta}{2}\right\}}+1_{\left\{\left(\frac{3n}{4}-\frac{n\delta}{2}\right)<\ell\leq\frac{3n}{4}\right\}}\right]\left(1-e^{+2i\pi\ell k\frac{1}{n(n-1)}}\right)e^{-2i\pi\ell k/(n-1)}\right]\\
&\quad+\frac{1}{n-1}t_{-k}^{*(n)}\\
&= O\left(n^{-1}\right),
\end{aligned}$$

where we have used that for n large and $\ell<n$, $1-e^{+2i\pi\ell k\frac{1}{n(n-1)}}=O(1)$. This concludes the proof of L(ii).

7. 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., Bollerslev, T., Mikkelsen, H., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 74, 3–30.

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.

Breidt, F. J., Crato, N., de Lima, P., 1998. The detection and estimation of long memory in stochastic volatility. Journal of Econometrics 83 (1-2), 325–348.

Chambers, M. J., 1998. Long memory and aggregation in macroeconomic time series. International Economic Review 39 (4), pp. 1053–1072.

Chevillon, G., Mavroeidis, S., 2013. Learning can generate long memory. Working paper, ESSEC Business School.

Comte, F., Renault, E., 1996. Long memory continuous time models. Journal of Econometrics 73, 101–149.

Comte, F., Renault, E., 1998. Long memory in continuous-time stochastic volatility models. Mathematical Finance 8, 291–323.

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.

Diebold, F. X., Yılmaz, K., 2009. Measuring financial asset return and volatility spillovers, with application to global equity markets. The Economic Journal 119, 158–171.

Diebold, F. X., Yılmaz, K., 2014. On the network topology of variance decompositions: Measuring the connectedness of financial firms. Journal of Econometrics 182, 119–134.

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.

Gonçalves, E., Gourieroux, C., 1988. Agrégation de processus autorégressifs d'ordre 1. Annales d'Economie et Statistiques 12, 127–149.

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., Szegö, G., 1958. Toeplitz forms and their applications. Vol. 321. University of California Press.

Grenander, U., Szegö, G., 2001. Toeplitz forms and their applications. AMS Chelsea Publishing, Providence, RI, 2nd ed.

Hecq, A., Laurent, S., Palm, F. C., 2012. On the univariate representation of BEKK models with common factors.

Tech. rep., Forthcoming in *Journal of Time Series Econometrics*.

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.

Johansen, S., Nielsen, M. N., 2016. The role of initial values in conditional sum-of-squares estimation of nonstationary fractional time series models. Econometric Theory forthcoming.

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.

Mandelbrot, B., 1997. Fractals and Scaling in Finance: Discontinuity, Concentration, Risk. Springer.

Mandelbrot, B., J.R., W., 1968. Noah, Joseph and operational hydrology. Water Resources Research 4, 909–918.

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.

Niels, H., Vera-Valds, J. E., 2015. Long memory, fractional integration, and cross-sectional aggregation. Tech. Rep. 59, CREATES Research Paper.

Onatski, A., Uhlig, H., 2012. Unit roots in white noice. Econometric Theory 28, 485–508.

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.

Sowell, F., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. Journal of Econometrics 53, 165–188.

Szegö, 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.