Asymptotic Theory for Spectral Density Estimates of General Multivariate Time Series [1]

Wei Biao Wu

University of Chicago

Paolo Zaffaroni

Imperial College London and University of Rome La Sapienza

[1]We thank the Co-Editor (Benedikt Potscher) and two anonymous referees for their time and their comments that greatly improved the paper. We also thank Marco Lippi for encouraging us to write this paper and his suggestions on how to improve part of the proofs.

Asymptotic Theory for Spectral Density Estimates of General Multivariate Time Series

corresponding author: Paolo Zaffaroni

January 8, 2017

Abstract

We derive uniform convergence results of lag-window spectral density estimates for a general class of multivariate stationary processes represented by an arbitrary measurable function of iid innovations. Optimal rates of convergence, that hold as both the time series and the cross section dimensions diverge, are obtained under mild and easily verifiable conditions. Our theory complements earlier results, most of which are univariate, which primarily concern in-probability, weak or distributional convergence, yet under a much stronger set of regularity conditions, such as linearity in iid innovations. Based on cross spectral density functions, we then propose a new test for independence between two stationary time series. We also explain the extent to which our results provide the foundation to derive the double asymptotic results for estimation of generalized dynamic factor models.

Keywords: kernel estimator, spectral density matrix, maximum deviations, uniform convergence.

JEL codes: C14,C32

1 Introduction

Consider the n-dimensional stochastic process:

$$\mathbf{Z}_t = (Z_{1t}, \ldots, Z_{it}, \ldots, Z_{nt})^\top = \mathbf{R}(\ldots, \epsilon_{t-1}, \epsilon_t), \tag{1}$$

where the $b \times 1$ vectors $\epsilon_t, t \in \mathbb{Z}$, are independent and identically distributed (iid) and $\mathbf{R}(.)$ is a measurable function such that $\mathbf{Z}_t$ exists (Tong (1990)). Under the above conditions the process $(\mathbf{Z}_t)$ is strictly stationary and ergodic although existence of moments is not warranted. Note that we need not impose $n \geq b$. As a consequence of (1)

$$Z_{it} = R_i(\ldots, \epsilon_{t-1}, \epsilon_t), \quad i = 1, \ldots, n,$$

for a measurable scalar function $R_i(.)$. In the sequel let $\mathcal{F}_t = (\ldots, \epsilon_{t-1}, \epsilon_t)$.

In this paper we are interested in studying uniform convergence, in terms of distribution as well as in terms of moments, of the kernel estimator of the spectral density matrix:

$$\hat{\mathbf{f}}_T(\lambda) = \frac{1}{2\pi} \sum_{u=-T+1}^{T-1} K(\frac{u}{B_T}) e^{-\imath u \lambda} \mathbf{C}(u), \quad -\pi \leq \lambda < \pi, \tag{2}$$

where $\imath = \sqrt{-1}$ denotes the complex unit and

$$\mathbf{C}(u) = \frac{1}{T} \sum^{*} \mathbf{Z}_t \mathbf{Z}_{t+u}^\top, \text{ where the sum } \textstyle\sum^* \text{ is for all } t, t+u \text{ between 1 and } T.$$

Here B_T is the lag-window size and the kernel function satisfies

$$K(0) = 1, \text{ continuous, even and such that } \kappa = \int_{-\infty}^{\infty} K^2(u) du < \infty.$$

The (i,j)-entry of the spectral matrix estimator is denoted by $\hat{f}_{Tij}(\lambda)$ for every $1 \leq i,j \leq n$. Here $\hat{\mathbf{f}}_T(\lambda)$ is an estimator of the true spectral density matrix which, when exists, has the form

$$\mathbf{f}(\lambda) = \frac{1}{2\pi} \sum_{u=-\infty}^{\infty} e^{-\imath u\lambda} \mathbf{\Gamma}(u), \quad -\pi \leq \lambda < \pi,$$

where $\mathbf{\Gamma}(u) = \mathbb{E}(\mathbf{Z}_0 \mathbf{Z}_u^\top)$, $u \in \mathbb{Z}$, is the autocovariance matrix satisfying $\mathbf{\Gamma}(-u) = \mathbf{\Gamma}^\top(u)$. Hereafter we assume $\mathbb{E}\mathbf{Z}_t = 0$ with, at minimum, bounded second moment. As a simple sufficient condition for the existence of the spectral density matrix, we assume that

$$\sum_{u=-\infty}^{\infty} |\mathbf{\Gamma}(u)|_F < \infty, \tag{3}$$

where $|A|_F = (\mathrm{tr}(AA^\top))^{1/2}$ is the Frobenius norm of a matrix A.

Nonparametric estimation of spectral densities appears useful in a large variety of contexts in time series and econometrics; see Hannan (1970), Chapter III, for a review. In general, estimating spectra and cross-spectra would be informative on the degree of persistence and co-movement of the associated variables at the various frequencies. More particularly, it permits efficient estimation of linear regression models through generalized least squares type estimators (see Hannan (1963), Robinson (1972) and Engle (1974) among others), of distributed lags models (Hannan (1965)) and for instrumental variable estimation of general linear systems with disturbances that have serial correlation of unknown form (Robinson (1991)). Accurate estimation of cross-spectral densities is necessary for deriving the optimal out-of-sample forecasts for vector discrete-time stochastic processes by the Wiener-Kolmogorov theory. Similarly, knowledge of the spectral density matrix permits the recovery of missing values for vector observations in discrete time, and interpolation between sampled observations of a continuous time vector process (Rozanov (1967)). Further, spectra and co-spectra are the necessary ingredients for designing the optimal frequency-domain filter when the vector of observations are contaminated by an additive noise (Hannan, Terrell and Tuckwell (1970)).

For all these reasons, statistical properties of spectral density estimates have been explored widely; see Shao and Wu (2007) for further references, such as, for example, Hannan (1970), Anderson (1971), Brillinger (2001), Brockwell and Davis (1991), Grenander and Rosenblatt (1957) and Priestley (1981) among others. Unfortunately, many of the previous results require restrictive conditions on the underlying processes such as linear processes in iid innovations or strong mixing, Gaussianity or existence of all moments. See also contributions in Phillips, Sun and Jin (2006, 2007) and Velasco and Robinson (2001).

In this paper we shall present an asymptotic theory for the kernel matrix estimator $\hat{\mathbf{f}}_T(\lambda)$, and its large deviations, under very mild and natural conditions, thus substantially extending the applicability of spectral analysis to nonlinear, non-Gaussian or non-strong mixing processes. Therefore, in contrast to most of the existing research, our focus is on estimation of multivariate spectral density matrices $\mathbf{f}(\lambda)$ of size $n \times n$, even providing some results for the high-dimensional case of $n \to \infty$. Univariate versions of our results easily follow but will not be presented here.

The paper proceeds as follows. Section 2 recalls the notion of the functional dependence measure of Wu (2005), that is key to the results developed in this paper. In fact, as evinced in the previous section, neither distributional assumptions nor linearity are required for our results to hold, unlike all the results established in the literature. Moreover, a rather mild moment condition on the $\mathbf{Z}_t$ is requested, again in contrast to the existing literature. Our novel multivariate results, for $n \geq 1$, are established in Section 3 under suitable regularity assumptions, which include existence of a finite p-th moment (with $p > 4$). Our most important, novel, results concern the large deviations $\max_{0 \leq \lambda \leq \pi} |\hat{f}_{Tij}(\lambda) - \mathbb{E}[\hat{f}_{Tij}(\lambda)]|^2 / f_{ii}(\lambda) f_{jj}(\lambda), \ \ i, j = 1, \ldots, n$, for which we establish both its asymptotic distribution and (optimal) rate of convergence of its moments of order ν for any $\nu < p/4$. Our results can be used in a variety of contexts in econometrics and mathematical statistics. For this purpose, Section 4 illustrates two possible applications of our results. We first show how to construct a test of serial independence, based on the asymptotic distribution of the maximum deviations of the cross-spectra. Second, in view of the fact that our multivariate results apply uniformly for any n, we illustrate how these results have been used to establish double asymptotic properties (when both T and n diverge to infinity) for estimators of the class of generalized dynamic factor models (GDFM) of Forni, Hallin, Lippi and Zaffaroni (2015a, 2015b). The mathematical appendix contains all the proofs of the lemmas and main theorems.

2 Measuring temporal dependence

To study asymptotic properties of $\hat{\mathbf{f}}_T$, we shall use the concept of functional dependence measure coined by Wu (2005), which measures temporal dependence of a process. Set

$$\mathbf{Z}_{t,\{k\}} = \mathbf{R}(\ldots \epsilon_{k-1}, \epsilon_k^*, \epsilon_{k+1} \ldots, \epsilon_t),\ \ k \leq t,$$

where $\epsilon_i^*, \epsilon_j, i, j \in \mathbb{Z}$, are iid. Let $\mathbf{Z}_{t,\{k\}} = \mathbf{Z}_t$ if $k > t$. Define the component coupled process $Z_{it,\{k\}}$ accordingly. Define the m-dependent approximating sequence

$$\tilde{\mathbf{Z}}_t = \mathbb{E}(\mathbf{Z}_t | \epsilon_{t-m}, \ldots, \epsilon_t) = \mathbb{E}(\mathbf{Z}_t | \mathcal{F}_{t-m,t}),\ m \geq 0,$$

with $\mathcal{F}_{t-m,t} = \sigma(\epsilon_{t-m}, \ldots, \epsilon_t)$ and $\tilde{Z}_{it}$ accordingly. Set the pth norm, for $p \geq 1$, equal to:

$$\|\mathbf{Z}_t\|_p = (\sum_{i=1}^{n} \mathbb{E} \mid Z_{it} \mid^p)^{1/p} \text{ with } \|\mathbf{Z}_t\| = \|\mathbf{Z}_t\|_2.$$

For all $i = 1, \ldots, n$ define the functional dependence measure

$$\delta_{t,p}^{[i]} = \|Z_{it} - Z_{it,\{0\}}\|_p,$$

and its related quantities $\Theta_{m,p}^{[i]} = \sum_{t=m}^{\infty} \delta_{t,p}^{[i]}$, $\Psi_{m,p}^{[i]} = (\sum_{t=m}^{\infty} (\delta_{t,p}^{[i]})^{p'})^{1/p'}$, where $p' = \min(2, p)$, and $d_{m,p}^{[i]} = \sum_{t=0}^{\infty} \min(\Psi_{m,p}^{[i]}, \delta_{t,p}^{[i]})$.

Throughout the paper we assume that the above quantities are finite for some $p \geq 2$. Then by Lemma 8 in Xiao and Wu (2012), letting $\mathbf{\Gamma}(u)_{ij}$ be the (i,j)th element of $\mathbf{\Gamma}(u)$,

$$\sum_{u=0}^{\infty} |\mathbf{\Gamma}(u)_{ii}| \leq \sum_{u=0}^{\infty} \sum_{t=0}^{\infty} \delta_{t,2}^{[i]} \delta_{t+u,2}^{[i]} \leq (\Theta_{0,2}^{[i]})^2 < \infty.$$

Hence (3) holds and thus the spectral density matrix $\mathbf{f}(\cdot)$ exists. Finally, set

$$\begin{aligned}
\delta_{t,p} &= \max_{1\leq i\leq n} \delta_{t,p}^{[i]}, \Theta_{m,p} = \max_{1\leq i\leq n} \Theta_{m,p}^{[i]}, \\
\Psi_{m,p} &= \max_{1\leq i\leq n} \Psi_{m,p}^{[i]}, d_{m,p} = \max_{1\leq i\leq n} d_{m,p}^{[i]}.
\end{aligned}$$

The $\delta_{t,p}$ quantify the dependence of $\mathbf{Z}_t$ on ϵ_0. Our main results in the paper need conditions on the decay of $\delta_{t,p}$.

3 Main results

Throughout this section we assume that there exists $c_0 > 0$ such that $\mathbf{f}(\lambda) - c_0 \mathbf{I}_n$ is positive definite for all λ. The following are the needed conditions on the kernel K and the lag B_T.

Assumption 1. K is an even and bounded function with bounded support in $(-1, 1)$, continuous in $(-1, 1)$, $K(0) = 1$, $\kappa = \int_{-1}^{1} K^2(u)du < 1$ and $\sum_{l\in\mathbb{Z}} \sup_{|s-l|<1} |K(lw) - K(sw)| = O(1)$ as $w \to 0$.

Assumption 2 (Condition 4 of Liu and Wu (2010)) There exist constants $0 < \underline{b} < \bar{b} < 1$ *and* $\underline{c}, \bar{c} > 0$ *such that* $\underline{c}T^{\underline{b}} \leq B_T \leq \bar{c}T^{\bar{b}}$ *holds for all large* T*.*

Assumption 1 is actually quite mild. It holds for the commonly used kernels including uniform or rectangle, triangle, Epanechnikov and quartic kernels. It implies Conditions 2 and 3 of Liu and Wu (2010).

We now derive the asymptotic distribution of the maximum deviation of the spectral density matrix estimator $\hat{\mathbf{f}}(\lambda)$.

Theorem 1. *Let Assumptions 1 and 2 hold. Assume* $\mathbb{E}\mathbf{Z}_0 = 0, \|\mathbf{Z}_0\|_p < \infty, p > 4$ *and*

$$\delta_{m,p} = O(\rho^m) \text{ for some } 0 < \rho < 1. \tag{4}$$

Let $\lambda_l^* = \pi|l|/B_T$*. Then for all* $x \in \mathbb{R}$

$$\mathbb{P}\left(\max_{0<l<B_T} \frac{T}{B_T} \frac{|\hat{f}_{Tij}(\lambda_l^*) - \mathbb{E}[\hat{f}_{Tij}(\lambda_l^*)]|^2}{\kappa f_{ii}(\lambda_l^*) f_{jj}(\lambda_l^*)} - 2\log B_T + \log(\pi \log B_T) \leq x\right) \to e^{-e^{-x/2}},$$

for every $i, j = 1, \ldots, n$*.*

Proof. We generalize the proof of Theorem 5 of Liu and Wu (2010). This requires the extension of a number of preliminary lemmas that are established in the Appendix. The proof then follows. ■

Remark.

(i) *The form of the limiting distributions of* $\hat{\mathbf{f}}_T(\lambda)$*, and functionals of this quantity, will be the same regardless of whether the underlying stochastic process* $\mathbf{Z}_t$ *in (1) is linear or nonlinear in iid innovations. The difference between our approach and the contributions listed in Section 1 relies on the ease with which one can verify the required regularity conditions. A simple example of nonlinear process is the transformed linear* *process, for example $Z_t = |X_t| - \mathbb{E}|X_t|$, where X_t is a linear process. We can easily compute its functional dependence measures, and thus establish the desired statistical properties, as a special case of our results. In contrast, the aforementioned techniques for spectral density estimation of linear processes do not extend easily now.*

(ii) *In Theorem 5 of Liu and Wu (2010) Gumbel convergence with $n = 1$ is derived. Theorem 1 implies the Gumbel convergence for cross spectral density function estimates for $n > 1$. In Section 4.1 we shall apply this result for testing dependence between two stationary time series. In practice researchers use coherence to study functional connectivity of networks; see for example Bastos and Schoffelen (2016) and references therein.*

(iii) *Theorem 1 holds also under the weaker condition (12).*

(iv) *Theorem 1 permits to evaluate simultaneous confidence intervals for any subset of elements of $\mathbf{f}(\lambda_l^*)$ via the Bonferroni method.*

(v) *Theorem 1 implies*

$$\max_{0 \le l \le B_T} |\hat{f}_{Tij}(\lambda_l^*) - \mathbb{E}[\hat{f}_{Tij}(\lambda_l^*)]|^2 = O_{\mathbb{P}}\left(\frac{B_T \, \log B_T}{T} \max_{0 \le l \le B_T} f_{ii}(\lambda_l^*) f_{jj}(\lambda_l^*) \right).$$

Without additional difficulties, Theorems 1 and 2 of Liu and Wu (2010) can be generalized as follows:

Theorem 2. *(Theorem 1 of Liu and Wu (2010)) Let Condition 1 of Liu and Wu (2010) hold. Assume $\mathbb{E}\mathbf{Z}_0 = 0, \|\mathbf{Z}_0\|_p < \infty, p \ge 2$ and $\Theta_{0,p} < \infty$. Let $1/B_T + B_T/T \to 0$. Then for every $i, j = 1, \ldots, n$*

$$\sup_{0 \le \lambda \le \pi} \| \hat{f}_{Tij}(\lambda) - f_{ij}(\lambda) \|_{p/2} \to 0.$$

Theorem 3. *Let Condition 2 of Liu and Wu (2010) hold. Assume* $\mathbf{Z}_0 = 0, \|\mathbf{Z}_0\|_4 < \infty$ *and* $\Theta_{0,4} < \infty$. *Let* $1/B_T + B_T/T \to 0$. *Then for every* $i, j = 1, \ldots, n$, *if* $0 < \lambda < \pi$,

$$\sqrt{\frac{T}{B_T}}\left(\hat{f}_{Tij}(\lambda) - \mathbb{E}[\hat{f}_{Tij}(\lambda)]\right) \to_d \kappa^{1/2}(\zeta_1 + \imath\zeta_2),$$

where (ζ_1, ζ_2) *is bivariate normal with mean* 0, $\mathbb{E}(\zeta_1^2) = (f_{ii}(\lambda)f_{jj}(\lambda) + \mathrm{Re} f_{ij}^2(\lambda))/2$, $\mathbb{E}(\zeta_2^2) = (f_{ii}(\lambda)f_{jj}(\lambda) - \mathrm{Re} f_{ij}^2(\lambda))/2$, *and* $\mathbb{E}(\zeta_1\zeta_2) = \mathrm{Im} f_{ij}^2(\lambda)/2$. *If* $\lambda = 0$ *or* π, *then*

$$\sqrt{\frac{T}{B_T}}\left(\hat{f}_{Tij}(\lambda) - \mathbb{E}[\hat{f}_{Tij}(\lambda)]\right) \to_d \kappa^{1/2} N(0, f_{ii}(\lambda)f_{jj}(\lambda) + f_{ij}^2(\lambda)).$$

Remark.

(i) (Remark 5 of Liu and Wu (2010)) If $K(x) - 1 = O(x)$ *as* $x \to 0$ *and* $\sum_{k\geq 1} k\delta_{k,2} < \infty$ *then* $\mathbb{E}\hat{f}_{Tij}(\lambda) - f_{ij}(\lambda) = O(B_T^{-1})$ *and we can replace* $\mathbb{E}\hat{f}_{Tij}(\lambda)$ *by* $f_{ij}(\lambda)$ *for a sufficiently smooth model spectra, in particular whenever* $T \log T = o(B_T^3)$. *Moreover, if* $\sum_{k\geq 1} k^q\delta_{k,2} < \infty$, *implying that the model spectra is q-differentiable, then* $\mathbb{E}\hat{f}_{Tij}(\lambda) - f_{ij}(\lambda) = O(B_T^{-q})$. *Note that under (4), it trivially holds that* $\sum_{k\geq 1} k^q\delta_{k,2} < \infty$ *for every* $q > 1$. *In this case we can replace* $\mathbb{E}\hat{f}_{Tij}(\lambda)$ *by* $f_{ij}(\lambda)$ *whenever* $T \log T = o(B_T^{q+1})$. *Note, however, that q will also depend on the choice of the kernel* $K(.)$, *see Hannan (1970), Chapter V, Theorem 10; namely q must satisfy*

$$\lim_{x\to 0} \frac{1 - K(x)}{|x|^q} = K_q < \infty.$$

For the truncated (resp. Bartlett) estimator, $q = \infty$ *(resp.* $q = 2$*).*

(ii) We wish to have B_T *as small as possible in order to achieve a quasi parametric rate but q (smoothness of the spectra) as large as possible, such that* $T \log T = o(T^{b(q+1)})$,

which is satisfied if $\underline{b}(q+1) > 1$.

We now consider the asymptotic behaviour of the moments of the maximum deviation of the kernel estimator $\hat{\mathbf{f}}_T(\lambda)$, which holds for any, arbitrarily large, n. This feature makes this theorem of independent interest. We have relegated the proof of Theorem 4 to the Appendix.

Theorem 4. *Let Assumptions 1 and 2 hold. Assume $\mathbb{E}\mathbf{Z}_0 = \mathbf{0}, \|\mathbf{Z}_0\|_p < \infty, p > 4$ and*

$$\delta_{m,p} \leq A\rho^m \text{ for some } 0 < \rho < 1, \;\; A > 0. \tag{5}$$

Let $0 < \nu < p/2$. Then there exists a constant C, only depending on ν, p, b, ρ, such that

$$\max_{1\leq i,j\leq n} \| \max_{0\leq\lambda\leq\pi} T|\hat{f}_{Tij}(\lambda) - \mathbb{E}[\hat{f}_{Tij}(\lambda)]|\|_\nu \leq C\theta_T, \text{ where } \theta_T = (TB_T \log B_T)^{1/2}.$$

Remark.

Theorem 4 can be extended to the case when $\delta^{[i]}_{m,p} = O(m^{-\alpha_i})$, for some $\alpha_i > 0$, by suitable modification of (A.3), (A.4), (A.10) and (A.11).

Theorem 5. *Let the assumptions of Theorem 1 be satisfied and $p > 4$. Then, for all ν^* such that $1 \leq \nu^* < p/4$, and every $i, j = 1, \ldots, n$:*

$$\left\| \max_{0<l<B_T} \frac{T}{B_T \log B_T} \frac{|\hat{f}_{Tij}(\lambda_l^*) - \mathbb{E}[\hat{f}_{Tij}(\lambda_l^*)]|^2}{\kappa f_{ii}(\lambda_l^*) f_{jj}(\lambda_l^*)} - 2 \right\|_{\nu^*} \to 0. \tag{6}$$

Proof. We first perform the Fourier expansion $1/f^{1/2}_{jj}(\lambda) = \sum_{k=-\infty}^{\infty} a_{jk} e^{\imath k\lambda}$. Let $g_{jB}(\lambda) = \sum_{k=-B_T}^{B_T} a_{jk} e^{\imath k\lambda}$. Then $\max_{0\leq\lambda\leq\pi} |1/f^{1/2}_{jj}(\lambda) - g_{jB}(\lambda)| \to 0$ as $B_T \to \infty$. Note that $h_{ij}(\lambda) := (\hat{f}_{Tij}(\lambda) - \mathbb{E}[\hat{f}_{Tij}(\lambda)])(g_{jB}(\lambda) g_{iB}(\lambda))^{\frac{1}{2}}$ is a trigonometric polynomial of order $3B_T$.

By Zygmund (2002), Chapter X, Theorem 7.28, we have

$$\max_{0\le\lambda\le\pi} |h_{ij}(\lambda)| \le (1+G^{-1}) \max_{0\le l\le J} |h_{ij}(\pi l/J)|, \tag{7}$$

where $J = 6(1+G)B_T$ and $G \in \mathbb{N}$ is a constant. By Theorem 1

$$\mathbb{P}\left(\max_{0\le l\le J} \frac{T}{J\kappa'}|h_{ij}(\pi l/J)|^2 - 2\log J + \log(\pi \log J) \le x\right) \to e^{-e^{-x/2}}, \tag{8}$$

where $\kappa' = \int_{-1}^{1} K^2(Ju/B_T)du = B_T\kappa/J$. Thus we have

$$\max_{0\le l\le J} \frac{T}{B_T\kappa\log B_T}|h_{ij}(\pi l/J)|^2 - 2 \to 0 \text{ in probability.} \tag{9}$$

By Theorem 4, we have the uniform integrability: for all ν with $1 \le \nu < p/2$,

$$\left\|\max_{0\le\lambda\le\pi} |\hat{f}_{Tij}(\lambda) - \mathbb{E}[\hat{f}_{Tij}(\lambda)]|\right\|_{\nu} = O((B_T \log B_T/T)^{1/2}). \tag{10}$$

Hence by the in-probability convergence (9) we obtain

$$\left\|\max_{0\le l\le J} \frac{T|h_{ij}(\pi l/J)|^2}{B_T\kappa\log B_T} - 2\right\|_{\nu^*} \to 0, \tag{11}$$

which implies (6) in view of (7) by choosing a sufficiently large G. ∎

Remark.

Condition (4) can be weakened to

$$\begin{aligned} d_{m,p} &= O(m^{-\alpha_1}), \;\; \alpha_1 > \max\left[1/2 - (p-4)/(2\delta p), 2\delta/p\right], \\ \Theta_{m,p} &= O(m^{-\alpha_2}), \;\; \alpha_2 > \max\left[1 - (p-4)/(2\delta p), 0\right], \end{aligned} \tag{12}$$

where $B_T = O(T^b)$ for some $b < 1$ by Assumption 2.

The following theorem concerns the uniform convergence of the spectral density function matrix estimate $\hat{\mathbf{f}}_T(\cdot)$, in the maximum deviations form, for the high dimensional case in which the dimension n is allowed to grow. Let

$$\Delta_{nT} = \max_{1 \le i,j \le n} \max_{0 \le \lambda \le \pi} |\hat{f}_{Tij}(\lambda) - \mathbb{E}[\hat{f}_{Tij}(\lambda)]|. \tag{13}$$

Theorem 6. *Let the assumptions in Theorem 1 be satisfied. Then*

$$\Delta_{nT} = O_P(n^{4/p}(T^{-1} B_T \log B_T)^{1/2}). \tag{14}$$

Consequently the uniform consistency $\max_{1 \le i,j \le n} \max_{0 \le \lambda \le \pi} |\hat{f}_{Tij}(\lambda) - f_{ij}(\lambda)| \to 0$ *in probability holds if* $n^{8/p} B_T \log B_T = o(T)$.

Proof. By Zygmund (2002), Chapter X, Theorem 7.28, it suffices to consider the frequencies $\lambda_l = \pi l/(2B_T)$. Let $Q_{ij}(\lambda) = T[\hat{f}_{Tij}(\lambda) - \mathbb{E}\hat{f}_{Tij}(\lambda)]$. Note that

$$\mathbb{P}(\Delta_{nT} \ge 2x) \le \sum_{i,j=1}^{n} \sum_{l=1}^{2B} \mathbb{P}(|Q_{ij}(\lambda_l)| \ge x). \tag{15}$$

We shall apply (A.8) to deal with the probability on the right hand side of (15). Recall Theorem 4 for θ_T. For any $\delta > 0$, elementary calculations show that there exists a constant $C_\delta > 0$ such that for $x = C_\delta \Theta_{0,p}^2 n^{4/p} \theta_T$, the right hand side of (15) is less than δ. Hence (14) follows. Under the decay condition (4), it follows that, as $B_T \to \infty$, the bias $\mathbb{E}\hat{f}_{Tij}(\lambda) - f_{ij}(\lambda) \to 0$ uniformly over $1 \le i, j \le n$ and $0 \le \lambda \le \pi$. Thus the proof is complete. ■

4 Theoretical applications

Our results can be used in a variety of problems arising in econometrics and multivariate statistics. In this section we develop two different theoretical applications. First, we show how to construct a test for independence based on our large deviation results for cross-spectra estimators. Second, we illustrate the use of our uniform convergence results for cross-spectra estimators for developing the asymptotic theory for estimators of generalized dynamic factor models of Forni et al. (2015a, 2015b), as both the number of time series T and cross-section n diverge to infinity.

4.1 Testing for independence

In this section we shall apply Theorem 1 to test independence between stationary time series. To fix the idea let $n = 2$. Our goal is to test the null hypothesis H_0: the processes $(Z_{1t})_{t\in\mathbb{Z}}$ and $(Z_{2t})_{t\in\mathbb{Z}}$ are independent. To this end, we shall construct the test statistic

$$Q = \max_{0\le l\le B_T} \frac{T}{B_T} \frac{|\hat{f}_{T12}(\lambda_l^*)|^2}{\kappa \hat{f}_{T11}(\lambda_l^*)\hat{f}_{T22}(\lambda_l^*)},$$

where we recall $\lambda_l^* = \pi|l|/B_T$. Note that $\hat{f}_{T12}(\lambda)/(\hat{f}_{T11}(\lambda)\hat{f}_{T22}(\lambda))^{1/2}$ is an estimate of the cross correlation spectral density $f_{12}(\lambda)/(f_{11}(\lambda)f_{22}(\lambda))^{1/2}$. By Theorem 1, we can reject the null hypothesis H_0 at level α, $0 < \alpha < 1$, if

$$Q > 2\log B_T - \log(\pi \log B_T) - 2\log\log(1-\alpha)^{-1}. \tag{16}$$

As a classical approach, Haugh (1976) proposed the following test statistic

$$H = T\sum_{u=-B}^{B} (\hat{\rho}(u)_{12})^2, \text{ where } \hat{\rho}(u)_{12} = \frac{\hat{\Gamma}(u)_{12}}{\hat{\Gamma}(0)_{11}^{1/2}\hat{\Gamma}(0)_{22}^{1/2}}.$$

In Haugh's test, one rejects H_0 if $H \geq \chi^2_{1+2B,1-\alpha}$. Our test (16) is based on the L^∞ norm, a different criterion from Haugh's quadratic test statistic. In particular, one should expect our test to have certainly more power in cases when the cross-correlation spectral density diverges sharply over a narrow set of frequencies.

4.2 Estimation of generalized dynamic factor models

GDFM, as introduced in Forni, Hallin, Lippi and Reichlin (2000), consist in modeling a panel $\{x_{it}, 1 \leq i \leq n, 1 \leq t \leq T\}$, namely a n-tuple of time series observed over a time period of length T as a finite realization of a stochastic process of the form

$$x_{it} = \chi_{it} + \xi_{it} = \sum_{f=1}^{q} b_{if}(L) u_{ft} + \xi_{it}, i \in \mathbb{N},\ t \in \mathbb{Z}, \tag{17}$$

where unobserved common component χ_{it} is made by the $u_t = (u_{1t}, \ldots, u_{qt})'$, an unobservable q-dimensional orthonormal white noise called the dynamic factors, and by the square-summable filters $b_{if}(L), i \in \mathbb{N}, f = 1, \ldots, q$, ($L$, as usual, stands for the lag operator). Here ξ_{it} is called the idiosyncratic components and is assumed weakly cross-correlated (cross-sectional orthogonality being an extreme case). Moreover the idiosyncratic components ξ_{it} and the common shocks u_{fs} are mutually orthogonal at any lead and lag.

Much of the literature on GDFM is based under the assumption that the space spanned by the common component χ_{it} is finite-dimensional, implying that the common component can be estimated consistently using the first r standard principal components; see Stock and Watson (2002a,b), Bai and Ng (2002). The assumption of a finite-dimensional factor space, however, lacks generality. Forni et al. (2000) relax the finite-dimensional assumption and propose to use q principal components in the frequency domain, the so-called Brillinger's dynamic principal components. However, their estimator involves the application of twosided filters acting on the observations x_{it} and, as a consequence, it leads to poor out of sample forecasting performance. Forni et al. (2015a) show how to retain the flexibility of the GDFM without the finite-dimensional assumption and, at the same time, without the need to use bilateral filters providing one-sided representations.

Although computationally simple, the unobserved nature of χ_{it} and ξ_{it} makes the theoretical analysis of the estimation procedure for the GDFM non trivial. For instance, Forni et al. (2015b) need to make use of the kernel estimator $\hat{\mathbf{f}}_T(\lambda)$ as both n and T diverge. This is non-standard since typically in multivariate analysis one considers the limiting behaviour of estimators when T diverges for a given n. The present paper provides theoretical support to Forni et al. (2015b) since it establishes primitive regularity assumptions for the bound

$$\max_{1\leq i,j\leq n} \mathbb{E}\left(\max_{|h|\leq B_T} \left| \hat{f}_{Tij}(\theta_h^*) - f_{ij}(\theta_h^*) \right|^2 \right) \leq C \left(T^{-1} B_T \log B_T \right), \tag{18}$$

where $\theta_h^* = \pi h / B_T$ and $C > 0$ is a constant. This result is a special case of Theorem 4 above. The strength of this result is that it holds *uniformly* for every T and every i, j in n and, therefore, it applies also when both $n, T \to \infty$. Neither distributional assumptions nor linearity of the x_{it} are required whereas a bounded $(4+\epsilon)$-moment of the x_{it} is required, among other regularity conditions. Using this bound, one can then establish a sharp bound, as both n and T diverge, for quantities like

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \mathbb{E}\left(\max_{|h|\leq B_T} \left| \hat{f}_{Tij}(\theta_h^*) - f_{ij}(\theta_h^*) \right|^2 \right).$$

In turn, results of this type are necessary to derive consistency, with rate, for the large deviation of the common component cross-spectral densities $\max_{|h|\leq B_T} |\hat{f}^{\chi}_{Tij}(\theta_h^*) - f^{\chi}_{ij}(\theta_h^*)| = O_{\mathbb{P}}(\max(1/\sqrt{n}, 1/\sqrt{T B_T \log T}))$ of the unobserved χs. Noticeably, the same rate of convergence has been preserved for all the other quantities of interest of the GFDM model (see Forni et al. (2015b)).

Bounds similar to (18) have been established in the literature under conditions that are either not verified by GDFM or that are unduly strong. Bentkus (1985), Theorem 2.2, obtains uniform rates of convergence for the mean square of the spectral density kernel estimator under Gaussianity. Assuming that all moments are finite and adopting Rosenblatt's type conditions on the summability of high-order cumulants, Brillinger (2001), Theorem 7.7.3, derives uniform strong consistency for the multivariate case. Woodroofe and Van Ness (1967), Theorems 3.1 and 3.3, establish uniform consistency, with rates, as well as the asymptotic distribution of the maximum deviation $\max_\lambda |\hat{f}_{Tii}(\lambda) - f_{ii}(\lambda)|$ when the scalar observable process is a linear process in iid innovations with bounded eighth moment. Under similar conditions, in particular linearity in iid innovations, Hannan (1970) extends their result to the multivariate case. However, unless Gaussianity of χ_{it} and ξ_{it} is assumed, it is not guaranteed that the observable x_{it} in (17) can be represented as a linear process in iid innovations even if both the common and idiosyncratic components χ_{it} and ξ_{it} do satisfy such representation.

A Appendix

We establish here the lemmas required to prove the theorems. We then report the proof of Theorem 4.

Lemma 1. *(Lemma 1 of Liu and Wu (2010)) Assume $\|\mathbf{Z}_t\|_p < \infty$ for $p > 1$ and $\mathbb{E}\mathbf{Z}_t = 0$. Let $C_p = 18p^{3/2}(p-1)^{-1/2}$ and $p' = \min(2, p)$. Let $\alpha_1, \alpha_2, \ldots$ be an arbitrary sequence of complex numbers. Then for every $j = 1, \ldots, n$,*

$$\| \sum_{k=1}^{T} \alpha_k (Z_{jk} - \tilde{Z}_{jk}) \|_p \leq C_p A_T \Theta_{m+1,p}, \text{ where } A_T \equiv \Big(\sum_{k=1}^{T} |\alpha_k|^{p'} \Big)^{1/p'}.$$

Also, we have (i) $\| \sum_{k=1}^{T} \alpha_k Z_{jk} \|_p \leq C_p A_T \Theta_{0,p}$ *and (ii)* $\| \sum_{k=1}^{T} \alpha_k \tilde{Z}_{jk} \|_p \leq C_p A_T \Theta_{0,p}$.

Proof. Each component of $\mathbf{Z}_t$ satisfies the assumptions of Lemma 1 of Liu and Wu (2010).

■

Lemma 2. *(Proposition 1 of Liu and Wu (2010)) Assume* $\|\mathbf{Z}_t\|_{2p} < \infty$ *for* $p \geq 2$, $\mathbb{E}\mathbf{Z}_t = 0$ *and* $\Theta_{0,2p} < \infty$. *Let*

$$A_T^{[ij]} = \sum_{1 \leq l < l' \leq T} \alpha_{l-l'} Z_{il} Z_{jl'}, \tilde{A}_T^{[ij]} = \sum_{1 \leq l < l' \leq T} \alpha_{l-l'} \tilde{Z}_{il} \tilde{Z}_{jl'},$$

where the α_t *are complex numbers, and* $D_T = (\sum_{s=1}^{T-1} |\alpha_s|^2)^{\frac{1}{2}}$. *Then*

$$\frac{\| A_T^{[ij]} - \mathbb{E}A_T^{[ij]} - (\tilde{A}_T^{[ij]} - \mathbb{E}\tilde{A}_T^{[ij]}) \|}{T^{\frac{1}{2}} D_T \Theta_{0,2p}} \leq C_{2p} d_{m,2p} \text{ for every } i, j = 1, .., n.$$

Proof. Let $E_{it-1} = \sum_{l=1}^{t-1} \alpha_{t-l} Z_{il}, \tilde{E}_{it-1} = \sum_{l=1}^{t-1} \alpha_{t-l} \tilde{Z}_{il}$ and

$$A_T^{[ij]*} = \sum_{1 \leq l < l' \leq T} \alpha_{l-l'} \tilde{Z}_{il} Z_{jl'} = \sum_{t=2}^{T} Z_{jt} \tilde{E}_{it-1}.$$

Let $\mathcal{P}_l(\cdot) \equiv \mathbb{E}(\cdot|\mathcal{F}_l) - \mathbb{E}(\cdot|\mathcal{F}_{l-1})$. Then

$$\| \mathcal{P}_l(A_T^{[ij]} - A_T^{[ij]*}) \|_p \leq I_l + II_l,$$

setting

$$
\begin{aligned}
I_l &= \| \sum_{t=2}^{T} Z_{jt,\{l\}} \left[(E_{it-1} - \tilde{E}_{it-1}) - (E_{it-1,\{l\}} - \tilde{E}_{it-1,\{l\}}) \right] \|_p, \\
II_l &= \sum_{t=2}^{T} \| (Z_{jt} - Z_{jt,\{l\}})(E_{it-1} - \tilde{E}_{it-1}) \|_p .
\end{aligned}
$$

Since $\| E_{it} - \tilde{E}_{it} \|_{2p} \leq C_{2p} D_T \Theta^{[i]}_{m+1,2p}$ by Lemma 1, and $\| \tilde{Z}_{it} - \tilde{Z}_{it,\{l\}} \|_{2p} \leq \delta^{[i]}_{t-l,2p}$ with $\sum_{t=2}^{T} \delta^{[i]}_{t-l,2p} \leq \Theta^{[i]}_{0,2p}$

$$
\sum_{l=-\infty}^{T} II_l^2 \leq C_{2p}^2 D_T^2 (\Theta^{[i]}_{m+1,2p})^2 \sum_{l=-\infty}^{T} \Theta^{[j]}_{0,2p} (\sum_{l'=1}^{T-1} \delta^{[j]}_{l'-l,2p}) \leq C_{2p}^2 D_T^2 T (\Theta^{[i]}_{m+1,2p})^2 (\Theta^{[j]}_{0,2p})^2.
$$

$$
\| \sum_{t=1}^{T-1} \left[Z_{it} - \tilde{Z}_{it} - Z_{it,\{l\}} + \tilde{Z}_{it,\{l\}}) \sum_{s=1+t}^{T} \alpha_{s-t} Z_{js,\{l\}} \right] \|_p \leq 2 \sum_{t=1}^{T-1} \min(\delta^{[i]}_{t-l,2p}, \Psi^{[i]}_{m+1,2p}) C_{2p} D_T \Theta^{[j]}_{0,2p},
$$

then

$$
\sum_{l=-\infty}^{T} I_l^2 \leq C_{2p}^2 D_T^2 (\Theta^{[j]}_{0,2p})^2 \sum_{t=-\infty}^{T} \Theta^{[i]}_{0,2p} \sum_{s=1}^{T-1} \min(\delta^{[i]}_{s-t,2p}, \Psi^{[i]}_{m+1,2p}) \leq C_{2p}^2 D_T^2 T (\Theta^{[j]}_{0,2p})^2 \Theta^{[i]}_{m+1,2p} d^{[i]}_{m,2p}.
$$

Since $\Theta^{[i]}_{m+1,p} \leq d^{[i]}_{m,p}$

$$
\begin{aligned}
&\| A_T^{[ij]} - \mathbb{E} A_T^{[ij]} - (A_T^{[ij]*} - \mathbb{E} A_T^{[ij]*}) \|_p^2 \leq \sum_{l=-\infty}^{T} \| \mathcal{P}_l (A_T^{[ij]} - A_T^{[ij]*}) \|_p^2 \\
&\leq 2 C_{2p}^2 D_T^2 T (\Theta^{[j]}_{0,2p})^2 (d^{[i]}_{m,2p})^2 \leq 2 C_{2p}^2 D_T^2 T \Theta^2_{0,2p} d^2_{m,2p}.
\end{aligned}
$$

The same bound applies to $\| A_T^{[ij]*} - \mathbb{E} A_T^{[ij]*} - (\tilde{A}_T^{[ij]} - \mathbb{E} \tilde{A}_T^{[ij]}) \|_p^2$. ■

Lemma 3. *(Proposition 2 of Liu and Wu (2010)) Assume* $\mathbb{E}\mathbf{Z}_0 = 0, \| \mathbf{Z}_0 \|_4 < \infty, \Theta_{0,4} < \infty$. *Let* $\alpha_l = \beta_l e^{\imath l\lambda}$ *for* $\lambda \in \mathbb{R}, \beta_l \in \mathbb{R}, 1-T \le l \le T-1, m \in \mathbb{N}$. *Define for every* $i = 1, \ldots, n$

$$D_l^{[i]} = A_l^{[i]} - \mathbb{E}(A_l^{[i]}|\mathcal{F}_{l-1}), \;\; A_l^{[i]} = \sum_{t=0}^{\infty} \mathbb{E}(\tilde{Z}_{it+l}|\mathcal{F}_l)e^{\imath t\lambda}$$

and

$$M_T^{[ij]} = \sum_{t=1}^{T} \bar{D}_t^{[i]} \sum_{l=1}^{t-1} \alpha_{l-t} D_l^{[j]}, \;\; i,j = 1, \ldots, n,$$

where $^-$ *denotes complex conjugate. Then*

$$\frac{\| \tilde{A}_T^{[ij]} - \mathbb{E}\tilde{A}_T^{[ij]} - M_T^{[ij]} \|}{m^{\frac{3}{2}} T^{\frac{1}{2}} \| Z_{i\,0} \|_4 \| Z_{j\,0} \|_4} \le C V_m^{\frac{1}{2}}(\beta) \text{ for every } i,j = 1, \ldots, n.$$

setting

$$V_m(\beta) = \max_{1-T \le l \le T-1} \beta_l^2 + m \sum_{l'=-1}^{-T-1} |\beta_{l'} - \beta_{l'-1}|^2.$$

Proof. Note that $A_l^{[i]} = \sum_{t=0}^{m} \mathbb{E}(\tilde{Z}_{it+l}|\mathcal{F}_l)e^{\imath t\lambda}$ and that $D_l^{[i]}$ is a m-dependent martingale difference sequence. Then, setting $U_l^{[i]} = e^{\imath(l-t)\lambda}\mathbb{E}(A_l^{[i]}|\mathcal{F}_{l-1})$, by summation by parts:

$$\begin{aligned} &\| \sum_{l=1}^{t-8m} \alpha_{l-t}(\tilde{Z}_{il} - D_l^{[i]}) \| \le Cm \| Z_{i\,0} \|_2^{\frac{1}{2}} \max_l | \beta_l | + \| \sum_{l=1}^{t-8m} (\beta_{l-t} - \beta_{l-t-1}) U_l^{[i]} \| \\ &\le C V_m^{\frac{1}{2}}(\beta) m \| Z_{i0} \|_2 . \end{aligned}$$

Likewise

$$\| \sum_{l=1}^{t-8m} \alpha_{l-t}(\tilde{Z}_{il} - \bar{D}_l^{[i]}) \| \le C V_m^{\frac{1}{2}}(\beta) m \| Z_{i0} \|_2 .$$

For $W_{1t}^{[ij]} = \tilde{Z}_{it} \sum_{l=1}^{t-8m} \beta_{l-t} e^{\imath(l-t)\lambda}(\tilde{Z}_{jl} - D_l^{[j]})$ then

$$\| W_{1t}^{[ij]} \| \le C V_m^{\frac{1}{2}}(\beta) m \| Z_{i0} \|_2 \| Z_{j0} \|_2$$

yielding

$$\| \sum_{t=1}^{T} W_{1t}^{[ij]} \| \leq \sum_{s=1}^{4m-1} \| \sum_{l=0}^{(T-s)/4m} W_{1s+4ml}^{[ij]} \| \leq C\Delta,$$

setting $\Delta = \max_{1\leq i,j\leq n} \Delta^{[ij]}$, $\Delta^{[ij]} = V_m^{\frac{1}{2}}(\beta) m^{\frac{3}{2}} T^{\frac{1}{2}} \parallel Z_{i0} \parallel_2 \parallel Z_{j0} \parallel_2$. Except for replacing $\parallel Z_{i0} \parallel_2 \parallel Z_{j0} \parallel_2$ with $\parallel Z_{i0} \parallel_4 \parallel Z_{j0} \parallel_4$, the same bound applies to $\| \sum_{t=1}^{T} (W_{2t}^{[ij]} - \mathbb{E}W_{2t}^{[ij]}) \|$ and $\| \sum_{t=1}^{T} (W_{3t}^{[ij]} - \mathbb{E}W_{3t}^{[ij]}) \|$ setting $W_{2t}^{[ij]} = \tilde{Z}_{it} \sum_{l=t-8m+1}^{t-1} \beta_{l-t} e^{\imath(l-t)\lambda} (\tilde{Z}_{jl} - D_l^{[j]})$ and $W_{3t}^{[ij]} = (\tilde{Z}_{it} - \bar{D}_t^{[i]}) \sum_{l=1}^{t-1} \beta_{l-t} e^{\imath(l-t)\lambda} D_l^{[j]}$. ■

Lemma 4. *(Lemma 2 of Liu and Wu (2010)) Assume* $\|\mathbf{Z}_t\|_p < \infty$ *for* $p \geq 2$ *and* $\mathbb{E}\mathbf{Z}_t = 0$. *Then Lemma 2 holds for every* Z_{it}, $i = 1, \ldots, n$.

Proof. Trivial since each component of Z_t satisfies the assumptions of Lemma 2 of Liu and Wu (2010). ■

Lemma 5. *(Proposition 3 of Liu and Wu (2010)) Let* $\mathbf{Z}_t$ *be* m*-dependent with* $\mathbb{E}\mathbf{Z}_t = 0, |Z_{it}| \leq M$ *a.s.,* $m \leq T$ *and* $M \geq 1$. *Let* $S_{r,l}^{[ij]} = \sum_{t=l+1}^{l+r} Z_{it} \sum_{s=1}^{t-1} a_{T,t-s} Z_{js}$, *where* $l \geq 0, l+r \leq T$ *and assume* $\max_{1\leq t\leq T} |a_{T,t}| \leq K_0$, $\max_{1\leq t\leq T} \max_{1\leq i\leq n} \mathbb{E}Z_{it}^4 \leq K_0$ *for some* $K_0 > 0$. *Then for any* $x, y \geq 1$ *and* $Q > 0$,

$$P(|S_{r,l}^{[ij]} - \mathbb{E}S_{r,l}^{[ij]}| \geq x) \leq 2e^{-y/4} + C_1 T^3 M^2 \left(x^{-2} y^2 m^3 (M^2 + r) \sum_{s=1}^{T} a_{T,s}^2 \right)^Q$$
$$+ C_1 T^4 M^2 \max_{1\leq i\leq n} P\left(|Z_{it}| \geq \frac{C_2 x}{y m^2 (M + r^{\frac{1}{2}})} \right) \text{ for every } i, j = 1, \ldots, n.$$

Proof. Trivial since each component of $\mathbf{Z}_t$ satisfies the assumptions of Proposition 3 of Liu and Wu (2010). ■

Lemma 6. *(Theorem 6 of Liu and Wu (2010)) Let* $a_{T,l} = b_{T,l} e^{\imath l\lambda}$, *where* $\lambda \in \mathbb{R}, b_{T,l} \in \mathbb{R}$

with $b_{T,l} = b_{T,-l}$ *and*

$$L_T^{[ij]} = \sum_{1 \le l, l' \le T} a_{T,l-l'} Z_{il} Z_{jl'} \text{ and } \sigma_T^2 = \omega(\lambda) \sum_{r=1}^{T} \sum_{t=1}^{T} b_{T,t-r}^2.$$

where $\omega(u) = 2$ *if* $u/\pi \in Z$ *and* $\omega(u) = 1$ *otherwise. Assume* $\mathbb{E}\mathbf{Z}_t = 0, \|\mathbf{Z}_0\|_4 < \infty, \Theta_{0,4} < \infty$ *and*

$$\begin{aligned}
&\max_{0 \le t \le T} b_{T,t}^2 = o(\zeta_T^2),\ \zeta_T^2 = \sum_{t=1}^{T} b_{T,t}^2, \\
&T\zeta_T^2 = O(\sigma_T^2), \\
&\sum_{r=1}^{T} \sum_{t=1}^{r-1} \left| \sum_{l=1+r}^{T} a_{T,r-l} a_{T,t-l} \right|^2 = o(\sigma_T^4), \\
&\sum_{r=1}^{T} |b_{T,r} - b_{T,r-1}|^2 = o(\zeta_T^2).
\end{aligned}$$

Recall Theorem 3 for the Gaussian vector $(\zeta_1, \zeta_2)^\top$*. Then for* $0 < \lambda < \pi$

$$\frac{L_T^{[ij]} - \mathbb{E} L_T^{[ij]}}{\sigma_T} \to_d 2\pi(\zeta_1 + \imath \zeta_2).$$

Proof. Note that

$$L_T^{[ij]} = A_T^{[ij]} + \bar{A}_T^{[ji]} + a_{T,0} \sum_{t=1}^{T} Z_{it} Z_{jt},$$

where by Lemma 1

$$\| \sum_{t=1}^{T} Z_{it} Z_{jt} - T\gamma_{ij}(0) \| \le CT^{\frac{1}{2}} (\| Z_{i0} \|_4 \Theta_{0,4}^{[j]} + \| Z_{j0} \|_4 \Theta_{0,4}^{[i]}),$$

$\gamma_{ij}(0)$ denoting the (i,j)-entry of $\mathbf{\Gamma}(0)$. It suffices to show that for any m

$$\frac{M_T^{[ij]} + \bar{M}_T^{[ji]}}{\sigma_T} \to_d 2\pi(\tilde{\zeta}_1 + \imath\tilde{\zeta}_2),$$

and then use Bernstein's lemma, where $(\tilde{\zeta}_1, \tilde{\zeta}_2)^\top$ is a mean 0 Gaussian vector similarly defined as $(\zeta_1, \zeta_2)^\top$ with $f_{ij}(\lambda)$ in the latter replaced by

$$\tilde{f}_{ij}(\lambda) = \frac{1}{2\pi}\sum_{l=-m}^{m} e^{\imath l\lambda}\mathbb{E}(\tilde{Z}_{i0}\tilde{Z}_{jl}).$$

Since $\| \sum_{t=1}^T \bar{D}_t^{[i]} U_t^{[j]*} \| \leq CT^{\frac{1}{2}} \max_{1\leq t\leq T} |b_{T,t}|$, setting $U_t^{[i]*} = \sum_{l=(t-4m+1)\vee 1}^{t-1} a_{T,l-t} D_l^{[i]}$, we need to show that

$$\frac{1}{\sigma_T}\sum_{t=1+4m}^{T} (\bar{D}_t^{[i]} U_t^{[j]\diamond} + D_t^{[j]}\bar{U}_t^{[i]\diamond}) \to_d 2\pi(\tilde{\zeta}_1 + \imath\tilde{\zeta}_2),$$

setting $U_t^{[i]\diamond} = \sum_{l=1}^{t-4m} a_{T,l-t} D_l^{[i]}$. We now apply the martingale central limit theorem (cf. Corollary 3.1 in Hall and Heyde (1980)). Since $\sum_{t=1+4m}^T \| \bar{D}_t^{[i]} U_t^{[j]\diamond} \|_4^4 \leq CT\zeta_T^4 = o(\sigma_T^4)$, the Lindeberg condition is satisfied. By the covariance matrix of $(\tilde{\zeta}_1, \tilde{\zeta}_2)^\top$, we have $\mathbb{E}(\tilde{\zeta}_1^2 + \tilde{\zeta}_2^2) = \tilde{f}_{ii}(\lambda)\tilde{f}_{jj}(\lambda)$ and $\mathbb{E}(\tilde{\zeta}_1 + \imath\tilde{\zeta}_2)^2 = \tilde{f}_{ij}^2(\lambda)$. Note that the latter two identities also imply the stated covariance structure of $(\tilde{\zeta}_1, \tilde{\zeta}_2)^\top$. By the Cramét-wold device, it remains to verify

$$\frac{1}{\sigma_T^2}\sum_{t=1+4m}^{T} \mathbb{E}\left(|\bar{D}_t^{[i]} U_t^{[j]\diamond} + D_t^{[j]}\bar{U}_t^{[i]\diamond}|^2 | \mathcal{F}_{t-1}\right) \to_p 4\pi^2 \tilde{f}_{ii}(\lambda)\tilde{f}_{jj}(\lambda). \tag{A.1}$$

and

$$\frac{1}{\sigma_T^2}\sum_{t=1+4m}^{T} \mathbb{E}\left([\bar{D}_t^{[i]} U_t^{[j]\diamond} + D_t^{[j]}\bar{U}_t^{[i]\diamond}]^2 | \mathcal{F}_{t-1}\right) \to_p 4\pi^2 \tilde{f}_{ij}^2(\lambda). \tag{A.2}$$

In the sequel we only prove (A.1) since the proof of (A.2) is similar. Rewriting $\mathbb{E}(\cdot|\mathcal{F}_{t-1}) =$

$\sum_{r=1}^{m}(\mathbb{E}(\cdot|\mathcal{F}_{t-r})-\mathbb{E}(\cdot|\mathcal{F}_{t-r-1}))+\mathbb{E}(\cdot|\mathcal{F}_{t-m-1})$, note that for $-m\leq r\leq m-1$,

$$\begin{aligned}&\| \sum_{t=1+4m}^{T}\Big(\mathbb{E}[|\bar{D}_t^{[i]}U_t^{[j]\diamond}+D_t^{[j]}\bar{U}_t^{[i]\diamond}|^2|\mathcal{F}_{t-r}]-\mathbb{E}[|\bar{D}_t^{[i]}U_t^{[j]\diamond}+D_t^{[j]}\bar{U}_t^{[i]\diamond}|^2|\mathcal{F}_{t-r-1}]\Big)\|^2\\ &\leq 4\sum_{t=1+4m}^{T}\| D_t^{[i]}\|_4^4\| U_t^{[j]\diamond}\|_4^4\leq CT\zeta_T^4=o(\sigma_T^4).\end{aligned}$$

Since the $D_t^{[i]}$ are $\mathcal{F}_{t-m,t}$-measurable whilst the $U_t^{[i]\diamond}$ are $\mathcal{F}_{t-4m}$-measurable, we obtain $\mathbb{E}((D_t^{[i]})^2(U_t^{[j]\diamond})^2|\mathcal{F}_{t-m,t})=(U_t^{[j]\diamond})^2\mathbb{E}((D_t^{[i]})^2)$ and (A.1) is equivalent to

$$\begin{aligned}&\frac{1}{\sigma_T^2}\sum_{t=1+4m}^{T}\Big(U_t^{[i]\diamond}U_t^{[j]\diamond}\mathbb{E}(\bar{D}_t^{[i]}\bar{D}_t^{[j]})+\bar{U}_t^{[i]\diamond}\bar{U}_t^{[j]\diamond}\mathbb{E}(D_t^{[i]}D_t^{[j]})\\ &+|U_t^{[j]\diamond}|^2\mathbb{E}(|D_t^{[i]}|^2)+|U_t^{[i]\diamond}|^2\mathbb{E}(|D_t^{[j]}|^2)\Big)\to_p ||D_t^{[i]}||^2||D_t^{[j]}||^2,\end{aligned}$$

since $||D_t^{[i]}||^2=2\pi\tilde{f}_{ii}(\lambda)$. Since $\| \sum_{t=1+4m}^{T}(U_t^{[i]\diamond}U_t^{[j]\diamond})-\mathbb{E}(U_t^{[i]\diamond}U_t^{[j]\diamond}))\ \|=o_p(\sigma_T^2)$ and $\sum_{t=1+4m}^{T}|\mathbb{E}(U_t^{[i]\diamond}U_t^{[j]\diamond})|=o(\sigma_T^2)$, the result follows noticing that

$$\mathbb{E}(|U_t^{[i]\diamond}|^2)=\sum_{l=1}^{t-4m}b_{T,l-t}^2\| D_t^{[i]}\|^2 .$$

■

Set

$$g_T^{[ij]}(\lambda)=L_T^{[ij]}-\mathbb{E}L_T^{[ij]}-\sum_{t=1}^{T}(Z_{it}Z_{jt}-\mathbb{E}Z_{it}Z_{jt}),$$

and

$$g_{T,m}^{[ij]}(\lambda)=\tilde{L}_T^{[ij]}-\mathbb{E}\tilde{L}_T^{[ij]}-\sum_{t=1}^{T}(\tilde{Z}_{it}\tilde{Z}_{jt}-\mathbb{E}\tilde{Z}_{it}\tilde{Z}_{jt}),$$

noticing that unless $i=j$ then $g_T^{[ij]}(\lambda)\neq\bar{g}_T^{[ij]}(\lambda)=g_T^{[ij]}(-\lambda)$. Set

$$\tau_T=\sqrt{TB_T}/\log B_T.$$

Lemma 7. *(Lemma A.1 and Remark A.2 of Liu and Wu (2010)) Let Assumptions 1 and 2 hold and* $\mathbb{E}\mathbf{Z}_0 = 0, \|\mathbf{Z}_0\|_p < \infty, p > 4$ *hold. Further assume* $\delta_{s,p} = O(\rho^s)$ *for some* $0 < \rho < 1$. *Then for any* $0 < \mathcal{C} < 1$, *there exists* $\gamma \in (0, \mathcal{C})$ *such that, for* $m = [T^\gamma]$, *for every* $i, j = 1, \dots, n$

$$\max_{1 \le l \le B_T} |g_T^{[ij]}(\lambda_l^*) - g_{T,m}^{[ij]}(\lambda_l^*)| = o_p(\sqrt{TB_T / \log B_T}),$$

Proof. This follows precisely Liu and Wu (2010), by setting

$$Y_{t,m}^{[ij]}(\lambda) = \tilde{Z}_{it} \sum_{s=1}^{t-1} a_{T,t-s} \tilde{Z}_{js}.$$

and $Y_{t,s_l}^{[ij]}(\lambda)$ accordingly, where $s_l = [T^{\rho^l}]$, $1 \le l \le r$, $r \in \mathbb{N}$ such that $0 < \rho^r < \mathcal{C}$. Also, we replace their definition of $\breve{u}_r(\lambda)$ with

$$\breve{u}_r(\lambda) = \sum_{t \in H_r} \left((Y_{t,s_l}^{[ij]}(\lambda) - Y_{t,s_{l+1}}^{[ij]}(\lambda)) + (Y_{t,s_l}^{[ij]}(-\lambda) - Y_{t,s_{l+1}}^{[ij]}(-\lambda)) \right).$$

■

Remark 1. Lemmas 4,5,6 of Liu and Wu (2010) extend without any additional difficulty.

Lemma 8. *(Lemma A.5 of Liu and Wu (2010)) Suppose* $\mathbb{E}\mathbf{Z}_0 = 0, \|\mathbf{Z}_0\|_4 < \infty$ *and* $d_{T,4} = O((\log T)^{-2})$. *Let* $\iota = B_T^{-1}(\log B_T)^2$. *For every* $i, j = 1, \dots, n$ *we have*

(i) uniformly on $\{(\lambda_1, \lambda_2) : 0 \le \lambda_l \le \pi - \iota, l = 1, 2$ *and* $|\lambda_1 - \lambda_2| \ge \iota\}$,

$$|\mathbb{E}[g_T^{[ij]}(\lambda_1) - \mathbb{E}g_T^{[ij]}(\lambda_1)][\bar{g}_T^{[ij]}(\lambda_2) - \mathbb{E}\bar{g}_T^{[ij]}(\lambda_2)]| = O(TB_T/(\log B_T)^2)$$

(ii) uniformly on $\{(\lambda_1, \lambda_2) \in [\iota, \pi - \iota]^2, |\lambda_1 - \lambda_2| \ge B_T^{-1}\}$ *for* α_T *satisfying* $\limsup_{T \to \infty} \alpha_T <$

1,

$$|\mathbb{E}[g_T^{[ij]}(\lambda_1) - \mathbb{E}g_T^{[ij]}(\lambda_1)][\bar{g}_T^{[ij]}(\lambda_2) - \mathbb{E}\bar{g}_T^{[ij]}(\lambda_2)]| = O(\alpha_T T B_T \kappa f_{ii}(\lambda_1) f_{jj}(\lambda_2)),$$

(iii) uniformly on $\{\iota \leq \lambda \leq \pi - \iota\}$,

$$\left|\mathbb{E}|g_T^{[ij]}(\lambda) - \mathbb{E}g_T^{[ij]}(\lambda)|^2 - 4\pi^2 T B_T f_{ii}(\lambda) f_{jj}(\lambda)\right| = O(TB_T(\log B_T)^{-2}).$$

Proof. (i) and (ii). Since $\| M_T^{[ij]}(\lambda) - N_T^{[ij]}(\lambda) \| = O(\sqrt{nm})$, where

$$N_T^{[ij]}(\lambda) = \sum_{t=1}^{T} \bar{D}_{t,\lambda}^{[i]} \sum_{l=1}^{t-1-m} \alpha_{T,l-t} D_{t,\lambda}^{[j]},$$

and $M_T^{[ij]}(\lambda) = M_T^{[ij]}, D_{t,\lambda}^{[j]} = \bar{D}_t^{[j]}$ as defined in Lemma 3, we need to show that

$$r^*_{T,\lambda_1,\lambda_2} = \left|\mathbb{E}(N_T^{[ij]}(\lambda_1) + \bar{N}_T^{[ij]}(\lambda_1))(\bar{N}_T^{[ij]}(\lambda_2) + N_T^{[ij]}(\lambda_2))\right| = O(TB_T(\log B_T)^{-2}),$$

since

$$\left|\mathbb{E}(M_T^{[ij]}(\lambda_1) + \bar{M}_T^{[ij]}(\lambda_1))(\bar{M}_T^{[ij]}(\lambda_2) + M_T^{[ij]}(\lambda_2))\right| \leq r^*_{T,\lambda_1,\lambda_2} + O(T\sqrt{mB_T} + \sqrt{Tm}B_T).$$

Elementary calculations yield

$$\begin{aligned}
&r^*_{T,\lambda_1,\lambda_2} = \\
&\mathbb{E}(\bar{D}_{t,\lambda_1}^{[i]} D_{t,\lambda_2}^{[i]})\mathbb{E}(D_{t,\lambda_1}^{[j]} \bar{D}_{t,\lambda_2}^{[j]}) \sum_{t=1}^{T} \sum_{l=1}^{t-m-1} K^2((t-l)/B_T)\cos((t-l)(\lambda_1 - \lambda_2)) \\
&+\mathbb{E}(\bar{D}_{t,\lambda_1}^{[i]} \bar{D}_{t,\lambda_2}^{[j]})\mathbb{E}(D_{t,\lambda_1}^{[j]} D_{t,\lambda_2}^{[i]}) \sum_{t=1}^{T} \sum_{l=1}^{t-m-1} K^2((t-l)/B_T)\cos((t-l)(\lambda_1 + \lambda_2)).
\end{aligned}$$

Then follows the proof of Liu and Wu (2010).

(iii) From (i)

$$
\begin{aligned}
& r^*_{T,\lambda,\lambda} = \\
& \mathbb{E}(|D^{[i]}_{t,\lambda}|^2)\mathbb{E}(|D^{[j]}_{t,\lambda_1}|^2)\sum_{t=1}^{T}\sum_{l=1}^{t-m-1} K^2((t-l)/B_T) \\
& +|\mathbb{E}(D^{[i]}_{t,\lambda}D^{[j]}_{t,\lambda})|^2\sum_{t=1}^{T}\sum_{l=1}^{t-m-1} K^2((t-l)/B_T)cos((t-l)(2\lambda)) \\
& = O(TB_T(\log B_T)^{-2}) + \parallel D^{[i]}_{0,\lambda} \parallel^2 \parallel D^{[j]}_{0,\lambda} \parallel^2 T \sum_{s=-B_T}^{B_T} K^2(s/B_T) \\
& = O(TB_T(\log B_T)^{-2}) + 4\pi^2\kappa \tilde{f}_{ii}(\lambda)\tilde{f}_{jj}(\lambda),
\end{aligned}
$$

where recall that $\mathbb{E}|D^{[i]}_{0,\lambda}|^2 = 2\pi\tilde{f}_{ii}(\lambda)$. ■

Lemma 9. *(Lemma 8 of Liu and Wu (2010)) Set* $E_T = B_T - (\log B_T)^2$ *and* $J_T = \sqrt{TB_T}/(\log B_T)^4$. *Under the conditions of Theorem 1 for every* $i, j = 1, \ldots, n$

$$
\mathbb{P}\left(\max_{(\log B_T)^2 \le r \le E_T} \frac{|\sum_{l=1}^{k_T} \hat{u}^{[ij]}_l(\lambda^*_r)|^2}{4\pi^2\kappa T B_T f_{ii}(\lambda_r) f_{jj}(\lambda_r)} - 2\log(B_T) + \log(\pi \log B_T) \le x\right) \to e^{-e^{-x/2}},
$$

with

$$
\hat{u}^{[ij]}_l(\lambda) = u^{[ij]}_l(\lambda) I\{|u^{[ij]}_l(\lambda)| \le J_T\} - \mathbb{E}\left(u^{[ij]}_l(\lambda) I\{|u^{[ij]}_l(\lambda)| \le J_T\}\right),
$$

and

$$
u^{[ij]}_l(\lambda) = \sum_{t \in H_l} (\bar{Y}^{[ij]}_{t,m}(\lambda) - \mathbb{E}\bar{Y}^{[ij]}_{t,m}(\lambda)) + \sum_{t \in H_l} (\bar{Y}^{[ij]}_{t,m}(-\lambda) - \mathbb{E}\bar{Y}^{[ij]}_{t,m}(-\lambda)),
$$

with $H_l = [(l-1)(p_T + q_T) + 1, p_T + (l-1)q_T]$, $p_T = [B_T^{1+\beta}]$ *some small* $\beta > 0, q_T = B_T + m, k_T = T/(p_T + q_T)$, *and*

$$
\bar{Y}^{[ij]}_{t,m}(\lambda) = \bar{Z}^{[i]}_{t,m} \sum_{s=1}^{t-1} a_{T,t-s} \bar{Z}^{[j]}_{s,m}
$$

$$\bar{Z}_{s,m}^{[k]} = Z_{s,m}^{[i]\ \prime} - \mathbb{E}Z_{s,m}^{[i]\ \prime},\ Z_{s,m}^{[i]\ \prime} = Z_{s,m}^{[i]} I\{|Z_{s,m}^{[i]}| \leq (TB_T)^{\alpha}\},\ \alpha < 1/4.$$

Proof. This follows the proof of Lemma 8 in Liu and Wu (2010). ■

Proof of Theorem 4. Let $Q_{ij}(\lambda) = T|\hat{f}_{Tij}(\lambda) - E[\hat{f}_{Tij}(\lambda)]|$. Note that (5) implies $\Theta_{m,p}^{[i]} \leq Am^{-\alpha}$ for any $\alpha > 0$. So we can assume without loss of generality that α satisfies

$$b < \alpha p/2 \text{ and } (1-2\alpha)b < 1 - 4/p. \tag{A.3}$$

In fact, set $\alpha = \max(B_1, B_2) + 1$ where $B_1 = 2b/p, B_2 = 1 - (1-4/p)/(2b)$. In turn, (A.3) implies that there exists a $\beta \in (0,1)$ such that

$$b < \alpha\beta p/2 \text{ and } (p/4 - \alpha\beta p/2)b < p/4 - 1. \tag{A.4}$$

In fact, β can be obtained as $\beta = \max(B_1, B_2)/\alpha + 1/2$ where $B_1/\alpha = 2b/(p\alpha), B_2/\alpha = 1/\alpha - (1-4/p)/(2b\alpha)$. Therefore α and β only depend on p, b.

We then follow the arguments of Theorem 10 in Xiao and Wu (2012) where, in particular, their Lemma 9 is replaced by our Lemma 2 (see Remark S.2 in Xiao and Wu (2012b)) and their Lemmas 11 and 12 are generalized using our Lemmas 2, 5 and Corollary 1.6 and 1.7 of Nagaev (1979). It remains to show that their result (41) is replaced by

$$\begin{aligned}
&\| \sum_{t,s=1}^{T} c_{s,t}(Z_{it}Z_{js} - \gamma_{ij}(t-s)) \|_{p/2} \leq \\
&C_{p/2}\mathcal{D}_T\Big(\sqrt{20}C_p\sqrt{T}\Theta_{0,p}^{[i]}\Theta_{0,p}^{[j]} + 2^{1-2/p}(\Theta_{0,p}^{[i]}\|Z_{j0}\|_p + \Theta_{0,p}^{[j]}\|Z_{i0}\|_p)\Big) \\
&\leq C_{p/2}\mathcal{D}_T\Big(\sqrt{20}C_p\sqrt{T}A^2/(1-\rho)^2 + 2^{2-2/p}C_pA/(1-\rho)\Big),
\end{aligned} \tag{A.5}$$

where $\gamma_{ij}(u)$ denotes the (ij)th entry of $\mathbf{\Gamma}(u)$ and

$$\mathcal{D}_T^2 \equiv \max\big(\max_{1\le s\le T}\sum_{t=1}^{T} c_{s,t}^2, \max_{1\le t\le T}\sum_{s=1}^{T} c_{s,t}^2\big).$$

Inequality (A.5) is a consequence of Lemma 1 and Lemma 2, as follows. First, notice that one can rewrite

$$\sum_{t=1}^{T}\sum_{s=1}^{T} c_{s,t}(Z_{it}Z_{js}-\gamma_{ij}(t-s)) = \sum_{t=2}^{T}\sum_{s=1}^{t-1} c_{s,t}(Z_{it}Z_{js}-\gamma_{ij}(t-s)) \tag{A.6}$$

$$+\sum_{s=2}^{T}\sum_{t=1}^{s-1} c_{s,t}(Z_{it}Z_{js}-\gamma_{ij}(t-s))$$

$$+\sum_{t=1}^{T} c_{t,t}(Z_{it}Z_{jt}-\gamma_{ij}(0)) \tag{A.7}$$

$$= A_{1T}^{[ij]} + A_{2T}^{[ij]} + \sum_{t=1}^{T} c_{t,t}(Z_{it}Z_{jt}-\gamma_{ij}(0)).$$

We deal with the right hand side of (A.6), namely $A_{1T}^{[ij]}$, the other two terms following along the same lines. We can apply the argument in the proof of Lemma 2. For simplicity set $E_{jt-1} = \sum_{s=1}^{t-1} c_{s,t}Z_{js}$ and $D_T = (\max_{1\le s\le T}\sum_{t=1}^{T} c_{s,t}^2)^{\frac{1}{2}}$. Then, for $\mathcal{P}_l(\cdot) \equiv E(\cdot|\mathcal{F}_l) - E(\cdot|\mathcal{F}_{l-1})$,

$$\| \mathcal{P}_l A_{1T}^{[ij]} \|_p \le I_l + II_l,$$

setting

$$I_l = \| \sum_{t=2}^{T} Z_{it,\{l\}} \big[(E_{jt-1} - E_{jt-1,\{l\}})\big] \|_p,$$

$$II_l = \sum_{t=2}^{T} \| (Z_{it} - Z_{it,\{l\}})E_{jt-1} \|_p \,.$$

Since $\| E_{jt} \|_{2p} \leq C_{2p} D_T \Theta_{0,2p}^{[j]}$ by Lemma 1 noticing that $2p > 2$, and $\| \tilde{Z}_{it} - \tilde{Z}_{it,\{l\}} \|_{2p} \leq \delta_{t-l,2p}^{[i]}$ with $\sum_{t=2}^T \delta_{t-l,2p}^{[i]} \leq \Theta_{0,2p}^{[i]}$,

$$\sum_{l=-\infty}^{T} II_l^2 \leq C_{2p}^2 D_T^2 (\Theta_{0,2p}^{[j]})^2 \sum_{l=-\infty}^{T} \Theta_{0,2p}^{[i]} (\sum_{l'=1}^{T-1} \delta_{l'-l,2p}^{[i]}) \leq C_{2p}^2 D_T^2 T (\Theta_{0,2p}^{[i]})^2 (\Theta_{0,2p}^{[j]})^2.$$

Similarly, since

$$\| \sum_{t=1}^{T-1} [Z_{it} - Z_{it,\{l\}}] \sum_{s=1+t}^{T} c_{s,t} Z_{js,\{l\}} \|_p \leq 2 \sum_{t=1}^{T-1} \delta_{t-l,2p}^{[i]} C_{2p} D_T \Theta_{0,2p}^{[j]},$$

then

$$\sum_{l=-\infty}^{T} I_l^2 \leq 4 C_{2p}^2 D_T^2 (\Theta_{0,2p}^{[j]})^2 \sum_{t=-\infty}^{T} \Theta_{0,2p}^{[i]} \sum_{s=1}^{T-1} \delta_{s-t,2p}^{[i]} \leq 4 C_{2p}^2 D_T^2 T (\Theta_{0,2p}^{[j]})^2 (\Theta_{0,2p}^{[i]})^2.$$

Finally, the result follows by using $\| A_{1T}^{[ij]} \|_p^2 \leq C_p^2 \sum_{l=-\infty}^{T} \| \mathcal{P}_l A_{1T}^{[ij]} \|_p^2$. The same bound applies to $\| A_{2T}^{[ij]} \|_p^2$ where now D_T must be replaced by $\max_{1 \leq t \leq T} (\sum_{s=1}^{T} c_{s,t}^2)^{\frac{1}{2}}$. The third term follows by a straight application of Lemma 1. Hence (A.5) is now established.

For any $K > 1$, there exists constants $C_{p,K,\beta}$, $C_{K,\beta}$ and C_p, such that, for all $x \geq \theta_T$, we have

$$\begin{aligned} \mathbb{P}(|Q_{ij}(\lambda)| \geq x) \quad \leq \quad & C_{p,K,\beta} x^{-p/2} (\Theta_{0,p}^{[i]} \Theta_{0,p}^{[j]})^{p/2} (L_T \log T) \\ & + C_{K,\beta} (x^{-p/2} (\Theta_{0,p}^{[i]} \Theta_{0,p}^{[j]})^{p/2} H_T)^K + e^{-C_p x^2 / (T B_T (\Theta_{0,4}^{[i]} \Theta_{0,4}^{[j]})^2)}, \quad \text{(A.8)} \end{aligned}$$

where $L_T = (T B_T)^{p/4} T^{-\alpha\beta p/2} + T B_T^{p/2-1-\alpha\beta p/2} + T$ and $H_T = T^{1+\sqrt{\beta}(p/4-1)} B_T^{p/4}$. Specifically, the second and the third terms in the right hand side of (A.8) correspond to the last two terms in inequality (44) in Xiao and Wu (2012) whereas the first term refers to the combination of theirs (50) and (51). Hence (A.8) follows from the generalization of inequalities (43), (44), (45) in Xiao and Wu (2012).

We shall now use the large deviation inequality (A.8) and conclude the proof by using $EX^a = a^{-1}\int_0^\infty x^{a-1}\mathbb{P}(X > x)dx$ which holds for any positive random variable X with finite ath moment. By Theorem 7.28 in Zygmund (2002), let $Q_{ij}^* = \max_{0\le\lambda\le\pi}|Q_{ij}(\lambda)|$ and $\lambda_l = \pi l/(2B)$, then $Q_{ij}^* \le 2\max_{0\le l\le 2B}|Q_{ij}(\lambda_l)|$ since $Q_{ij}(\lambda)$ is a trigonometric polynomial with order B. Hence by (A.8), for a sufficiently large constant $K > 0$,

$$\begin{aligned}\int_{K\theta_T}^\infty x^{\nu-1}\mathbb{P}(Q_{ij}^* \ge 2x)dx &\le (1+2B_T)\int_{K\theta_T}^\infty x^{\nu-1}\max_\lambda \mathbb{P}(|Q_{ij}(\lambda)| \ge x)dx \\ &\le C_{p,K,\beta,\nu}(1+2B_T)(\theta_T^{\nu-p/2}(\Theta_{0,p}^{[i]}\Theta_{0,p}^{[j]})^{p/2}L_T\log T \\ &\quad +((\Theta_{0,p}^{[i]}\Theta_{0,p}^{[j]})^{p/2}K^{-p/2}H_T)^K\theta_T^{\nu-pK/2} \\ &\quad +\theta_T^\nu B_T^{-C_{p,\nu}K^2/(\Theta_{0,4}^{[i]}\Theta_{0,4}^{[j]})^2}). \qquad \text{(A.9)}\end{aligned}$$

Elementary calculations show that, under (A.4), the right hand side of (A.9) is $O(\theta_T^\nu)$ if we choose a large enough K. Hence we have $\|Q_{ij}^*\|_\nu = O(\theta_T)$ since $\int_0^{K\theta_T} x^{\nu-1}\mathbb{P}(Q_{ij}^* \ge 2x)dx \le (K\theta_T)^\nu/\nu$. In particular the two inequalities in (A.4) allow to bound the terms associated with the first and the second component of L_T. The last term of L_T does not require any restrictions since $p/4 > 1$. The term involving H_T requires K large enough such that

$$A_1 = \frac{b}{(p/4-1)(1-\sqrt{\beta})} < K, \qquad \text{(A.10)}$$

and the third, last, term on the right hand side of (A.9) requires K large enough such that

$$\Big(\frac{(\Theta_{0,4}^{[i]}\Theta_{0,4}^{[j]})^2}{C_{p,\nu}}\Big)^{\frac{1}{2}} < K. \qquad \text{(A.11)}$$

Since $\Theta_{0,p}^{[i]} = \sum_{t=0}^{\infty} \delta_{t,p}^{[i]} \leq A/(1-\rho)$ for every $i \leq n$, it follows that (A.11) is implied by

$$A_2 = \frac{A^2}{C_{p,\nu}^{1/2}(1-\rho)^2} < K.$$

Then set $K = \max(A_1, A_2) + 1$. This implies that K only depends on ν, p, b, ρ. Since the same applies to α and β, it follows that we can construct a constant $C_{\nu,p,b,\rho}$ that satisfies our statement. ■

B References

Anderson, T.W. (1971) *The Statistical Analysis of Time Series.* New York: Wiley.

Bai, J. and Ng, S. (2002) Determining the number of factors in approximate factor models. *Econometrica* 70, 191–221.

Bastos, A.M. and Schoffelen, J.M. (2016) A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls. *Front. Syst. Neurosci.* http://dx.doi.org/10.3389/fnsys.2015.00175

Bentkus, R. (1985) Rate of uniform convergence of statistical estimators of spectral density in spaces of differentiable functions.*Lithuanian Mathematical Journal* 25, 209–219.

Bentkus, R.Y. and Rudzkis, R.A. (1982) On the distribution of some statistical estimates of spectral density. *Theory of Probability and Its Applications* 27, 795–814

Brillinger, D.R. (2001) *Time series: data analysis and theory.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Brockwell, P.J. and Davis, R.A. (1991) *Time series: theory and methods.* Second edition. Springer Series in Statistics. Springer-Verlag, New York.

Engle, R.F. (1974) Band Spectrum Regression. *International Economic Review* 15, 1–11.

Forni, M., Hallin M., Lippi M. and Reichlin, L. (2000) The Generalized Dynamic-Factor

Model: Identification and Estimation. *The Review of Economics and Statistics* 82, 540–554.

Forni, M., Hallin M., Lippi M. and Zaffaroni, P. (2015a) Dynamic factor models with infinite-dimensional factor space: one-sided representations. *Journal of Econometrics* 185, 359–371.

Forni, M., Hallin M., Lippi M. and Zaffaroni, P. (2015b) Dynamic factor models with infinite-dimensional factor space: asymptotic analysis. Preprint.

Grenander, U. and Rosenblatt, M. (1957) *Statistical Analysis of Stationary Time Series.* New York: Wiley.

Hall, P. and Heyde, C.C. (1980) *Martingale Limit Theory and Its Application.* New York: Academic Press.

Hannan, E.J. (1963), Regression for time series. In *Proc. Sympos. Time Series Analysis*, pp. 17–37. New York: Wiley.

Hannan, E.J. (1965) The Estimation of Relationship Involving Distributed Lags. *Econometrica* 33, 206–224.

Hannan , E. J. (1970) *Multiple Time Series.* New York: Wiley.

Hannan, E.J., Terrell, R.D. and Tuckwell, N.E. (1970) The seasonal adjustment of economic time series. *International Economic Review* 11, 24–52.

Haugh, L.D. (1976) Checking the Independence of Two Covariance Stationary Time Series: A Univariate Residual Cross-Correlation Approach. *Journal of the American Statistical Association* 71, 378–385.

Liu, W. and Wu, W. B. (2010) Asymptotics of Spectral Density Estimates. *Econometric Theory* 26, 1218–1245.

Nagaev, S. V. (1979) Large Deviations of Sums of Independent Random Variables. *Annals of Probabability* 7, 745–789.

Phillips, P.C.B., Sun, Y. and Jin, S. (2006) Spectral density estimation and robust hypothesis testing using steep origin kernels without truncation. *International Economic Review* 47, 837–894.

Phillips, P.C.B., Sun, Y. and Jin, S. (2007) Long run variance estimation and robust regression testing using sharp origin kernels with no truncation. *Journal of Statistical Planning and Inference* 137, 985–1023.

Priestley, M. B. (1981) *Spectral analysis and time series.* London-New York: Academic Press.

Robinson, P. M. (1972) Non-linear regression for multiple time-series. *Journal of Applied Probability* 9, 758–768.

Robinson, P M. (1991) Automatic Frequency Domain Inference on Semiparametric and Nonparametric Models. *Econometrica* 59, 1329–63.

Rosenblatt, M. (1984) Asymptotic normality, strong mixing, and spectral density estimates. *Annals of Probability* 12, 1167–1180

Rozanov, A. (1967) *Stationary Random Processes.* San Francisco: Holden-Day.

Shao, X. and Wu, W.B. (2007) Asymptotic spectral theory for nonlinear time series. *Annals of Statistics* 35, 1773–1801.

Stock, J.H. and Watson, M.W. (2002a) Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* 20, 147–162.

Stock, J.H. and Watson, M.W. (2002b) Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97, 1167–1179.

Tong, H. (1990) *Non-linear Time Series: A Dynamical System Approach.* Oxford University Press.

Velasco, C. and Robinson, P.M. (2001) Edgeworth expansions for spectral density estimates and Studentized sample mean. *Econometric Theory* 17, 497–539.

Xiao, H. and Wu, W.B. (2012) Covariance Matrix Estimation for Stationary Time Series. *Annals of Statistics* 40, 466–493.

Xiao, H. and Wu, W.B. (2012b) Supplement to: Covariance Matrix Estimation for Stationary Time Series.

Woodroofe, Michael B. and Van Ness, John W. (1967) The maximum deviation of sample spectral densities. *Ann. Math. Statist* 38, 1558–1569.

Wu, W.B. (2005) Nonlinear system theory: Another look at dependence. *Proceedings of the National Academy of Sciences USA* 102, pp. 14150–14154.

Zygmund, A. (2002) *Trigonometric series. Vol. I, II.* Third edition. Cambridge: Cambridge University Press.