Time-Varying General Dynamic Factor Models and the Measurement of Financial Connectedness

Matteo Barigozzi[1] Marc Hallin[2] Stefano Soccorsi[3] Rainer von Sachs[4]

19th April 2020

Abstract

We propose a new time-varying Generalized Dynamic Factor Model for high-dimensional, locally stationary time series. Estimation is based on dynamic principal component analysis jointly with singular VAR estimation, and extends to the locally stationary case the one-sided estimation method proposed by Forni et al. (2017) for stationary data. We prove consistency of our estimators of time-varying impulse response functions as both the sample size T and the dimension n of the time series grow to infinity. This approach is used in an empirical application in order to construct a time-varying measure of financial connectedness for a large panel of adjusted intra-day log ranges of stocks. We show that large increases in long-run connectedness are associated with the main financial turmoils. Moreover, we provide evidence of a significant heterogeneity in the dynamic responses to common shocks in time and over different scales, as well as across industrial sectors.

JEL subject classification: C32, C14.
Key words: locally stationary dynamic factor models, volatility, financial connectedness.

1 Introduction

Together with growing interests in big-data techniques and increased availability of large datasets, high-dimensional statistical methodology has been thriving in the past two decades. Time series analysis and time-series econometrics are no exception and factor models, under their various forms, have emerged as the most successful tools in the analysis of high-dimensional serially dependent observations (see Stock and Watson (2016) and references therein for a recent survey).

The most general approach to factor analysis is, arguably, the so-called General Dynamic Factor Model (henceforth GDFM) initially proposed by Forni et al. (2000) in which common factors or common shocks are loaded via time-invariant filters. The essence of the GDFM

[1]m.barigozzi@lse.ac.uk – Department of Statistics, London School of Economics and Political Science, UK.
[2]mhallin@ulb.ac.be – ECARES, Université Libre de Bruxelles, Belgium.
[3]s.soccorsi@lancaster.ac.uk– Department of Economics, Lancaster University Management School, UK.
[4]rainer.vonsachs@uclouvain.be – ISBA, LIDAM, Université catholique de Louvain, Belgium.

We thank Wei Biao Wu and Danna Zhang for sharing their unpublished manuscript, and two anonymous referees for helpful comments.

is that few unobserved factors drive the main comovements (cross-covariances, whether contemporaneous or lagged) across the panel: a common shock indeed may affect some series at time t but some other one at time $t+1$ only. The GDFM approach encompasses, strictly, the so-called static factor model, where the impact of a factor is assumed to be contemporaneous on all series (see e.g. Stock and Watson, 2002). As shown in Hallin and Lippi (2013), the GDFM actually follows from a very general representation result, in contrast with the static factor model which requires rather stringent assumptions on the dynamics of the data-generating process.[1]

In this paper, we consider the estimation of a time-varying version of the GDFM (henceforth tvGDFM) in which the factors are loaded via time-varying filters. We then develop an application to the measurement through time of connectedness in the financial market, in the spirit of Diebold and Yilmaz (2014).

Our approach builds on the locally stationary framework introduced by Dahlhaus (1997), where it is assumed that the second-order structure is evolving smoothly over time. A slightly different tvGDFM has been previously studied by Eichler et al. (2011). Their method, however, inspired by Forni et al. (2000), is entirely based on dynamic principal component analysis (DPCA). As a result, it suffers from the main drawback of DPCA, which resorts to two-sided filters to recover the space spanned by the factors. Such two-sided filtering makes the Eichler et al. (2011) approach unsuitable for a number of important applications, among which forecasting, impulse response analysis of the dynamic impacts of common shocks, and the analysis of financial connectedness considered in this paper.

To cope with this problem, Forni et al. (2015, 2017), in the stationary case, recur to a combination of spectral estimation and VAR filtering which only involves one-sided filters. In this paper, we extend their approach to the time-varying setting and propose estimators involving one-sided filters only, of time-dependent impulse responses, which we prove to be consistent, uniformly in time, as both the sample size T and the dimension n of the panel grow to infinity. Specifically, we show that the rate of convergence is, up to some multiplicative logarithmic factors in T, of order $\min(T^{\rho_{r^*}}, \sqrt{n})$ where ρ_{r^*} depends on the maximum order r^* of moments that we can assume to exist for the data under study; in particular, $\rho_{r^*} = 1/4$ when $r^* = \infty$ as, e.g., in the Gaussian case. Our asymptotic results build on recent work by Zhang and Wu (2019) on the estimation of large time-varying spectral density matrices.

This tvGDFM approach is particulatly welcome in the study of time-varying connectedness in the financial market. The existing method, as proposed by Diebold and Yilmaz (2014), indeed, is based on variance decompositions in vector moving average models and suffers of two serious limitations. First, it is fully parametric, hence inadequate for the large cross-sections typically affected by systemic events. Second, it is based on a stationary model and, therefore, cannot fully account for time-variation—an essential feature of financial connectedness.[2] Our tvGDFM is ideally designed to overcome both shortcomings. We apply it here to a panel of *adjusted intra-day log ranges*[3] of $n = 329$ constituents of the Standard & Poor's 500 observed from December 31, 1999 to August 31, 2015, for a total of $T = 3939$

[1] Yet another approach has been proposed by Peña and Yohai (2016) where DPCA is used as a data-analytic tool with no reference to the consistent estimation of any underlying factors nor data-generating process.

[2] A rolling estimation is considered but little theoretical justification is provided, and no guidelines are offered for the choice of a window size.

[3] Adjusted intra-day log ranges were defined by Parkinson (1980) and their use as "highly efficient and robust to microstructure noise" log-volatility proxies has been recommended by Alizadeh et al. (2002). It has been shown (Brownlees and Gallo, 2010) that they often outperform more sophisticated alternatives.

observations.

The rest of the paper is organized as follows. In Section 2, we present the time-varying General Dynamic Factor Model. Section 3 proposes an estimation method of the time-varying impulse response functions to common factors, and establishes its consistency. The connectedness measures we derive from the model and the empirical results are discussed in Section 4. Section 5 concludes. Proofs are postponed to an online Appendix.

Notation

We denote by $\boldsymbol{A}^\dagger$ and $\boldsymbol{v}^\dagger$ the transposed complex conjugate of any complex matrix $\boldsymbol{A}$ or column vector $\boldsymbol{v}$, respectively. The imaginary unit is denoted as $\iota := \sqrt{-1}$. Let $\mathbb{N}_0$ stand, as usual, for the set of natural integers. Given two sequences $\{a_n\}$ and $\{b_n\}$, we write $a_n \asymp b_n$ if, for some finite positive constants $\underline{c}$ and $\overline{c}$, there exists an $N_0 \in \mathbb{N}_0$ such that $\underline{c} \leq a_n b_n^{-1} \leq \overline{c}$ for all $n \geq N_0$. Throughout, L stands for the lag operator.

2 A time-varying Generalized Dynamic Factor Model

In this section we present the time-varying Generalized Dynamic Factor Model (tvGDFM) inspired by Eichler et al. (2011). All random variables considered below belong to the space of centered real-valued random variables with finite second-order moments defined over some common probability space.

The factor model approach in the analysis of a (zero-mean) double-indexed process $\mathbf{X} := \{X_{it} : i \in \mathbb{N}_0,\ t \in \mathbb{Z}\}$ (here, the process of intraday log range values; i is a cross-sectional index and t stands for time) is based on a decomposition of X_{it} into the sum

$$X_{it} = \chi_{it} + \xi_{it}, \quad i \in \mathbb{N}_0,\ t \in \mathbb{Z} \tag{1}$$

of two unobserved components: the common component process $\boldsymbol{\chi} := \{\chi_{it}\}$ and the idiosyncratic component process $\boldsymbol{\xi} := \{\xi_{it}\}$. For $\boldsymbol{\chi}$ and $\boldsymbol{\xi}$, we assume the following time-varying MA representations, which account for nonstationarity and the time-varying nature of their second-order structure:

$$\chi_{it} = \sum_{j=1}^{q} \sum_{k=0}^{\infty} c^*_{ijk}(t) u_{j,t-k}, \quad i \in \mathbb{N}_0,\ t \in \mathbb{Z}, \tag{2}$$

$$\xi_{it} = \sum_{j=1}^{\infty} \sum_{k=0}^{\infty} d^*_{ijk}(t) \eta_{j,t-k}, \quad i \in \mathbb{N}_0,\ t \in \mathbb{Z} \tag{3}$$

(see Assumption (A) for identification assumptions). Denoting by $\{\mathbf{X}_{nt} := (X_{1t}, \ldots, X_{nt})'\}$, $\{\boldsymbol{\chi}_{nt} := (\chi_{1t}, \ldots, \chi_{nt})'\}$, and $\{\boldsymbol{\xi}_{nt} := (\xi_{1t}, \ldots, \xi_{nt})'\}$ the n-dimensional subprocesses of $\mathbf{X}$, $\boldsymbol{\chi}$, and $\boldsymbol{\xi}$, we also have

$$\mathbf{X}_{nt} = \boldsymbol{\chi}_{nt} + \boldsymbol{\xi}_{nt}, \quad n \in \mathbb{N}_0,\ t \in \mathbb{Z}$$

with

$$\boldsymbol{\chi}_{nt} := \mathbf{C}^*_n(t, L)\mathbf{u}_t \quad \text{and} \quad \boldsymbol{\xi}_{nt} := \mathbf{D}^*_n(t, L)\boldsymbol{\eta}_t, \quad n \in \mathbb{N}_0,\ t \in \mathbb{Z} \tag{4}$$

where $\mathbf{u}_t := (u_{1t}, \ldots, u_{qt})'$, $\boldsymbol{\eta}_t := (\eta_{1t}, \eta_{2t}, \ldots)'$, and the (i, j) entries of $\mathbf{C}^*_n(t, L)$ and $\mathbf{D}^*_n(t, L)$ are defined by

$$c^*_{ij}(t, L) := \sum_{k=0}^{\infty} c^*_{ijk}(t) L^k, \quad 1 \leq i \leq n,\ 1 \leq j \leq q,$$

and

$$d^*_{ij}(t,L) := \sum_{k=0}^{\infty} d^*_{ijk}(t)L^k, \quad 1 \le i \le n,\ j \in \mathbb{N}_0.$$

The existence of time-independent one-sided filters $\mathbf{C}^*_n(t,L) = \mathbf{C}^*_n(L)$ is justified in the stationary case by the representation results in Hallin and Lippi (2013); here we directly assume (2). The generic element $c^*_{ij}(t,L)$ of $\mathbf{C}^*_n(t,L)$ represents the time-varying impulse response function of variable X_{it} to the jth factor (common shock) u_j; those impulse response functions are the main quantities of interest here.

Throughout, we assume that the shocks are satisfying the following assumption.

Assumption (A). *The common and idiosyncratic shocks are such that*

A1. the q-dimensional process of common shocks $\{\mathbf{u}_t : t \in \mathbb{Z}\}$ is such that $\mathrm{E}[\mathbf{u}_t] = \mathbf{0}$, $\mathrm{E}[\mathbf{u}_t\mathbf{u}'_t] = \mathbf{I}_q$*, and $\mathbf{u}_t$ is independent of $\mathbf{u}_s$ for all $t, s \in \mathbb{Z}$ with $t \neq s$;*

A2. the infinite-dimensional process of idiosyncratic shocks $\boldsymbol{\eta} := \{\boldsymbol{\eta}_t : t \in \mathbb{Z}\}$ is such that, for any n-dimensional subprocess $\{\boldsymbol{\eta}_{nt} = (\eta_{1t}, \ldots, \eta_{nt})' : t \in \mathbb{Z}\}$, $\mathrm{E}[\boldsymbol{\eta}_{nt}] = \mathbf{0}$, $\mathrm{E}[\boldsymbol{\eta}_{nt}\boldsymbol{\eta}'_{nt}] = \mathbf{I}_n$*, and $\boldsymbol{\eta}_{nt}$ is independent of $\boldsymbol{\eta}_{ns}$ for all $t, s \in \mathbb{Z}$ with $t \neq s$;*

A3. the common and idiosyncratic shocks are such that: $\mathrm{E}[\eta_{it}u_{js}] = 0$ *for all $i \in \mathbb{N}_0$, $1 \le j \le q$, and $t, s \in \mathbb{Z}$;*

A4. there exists an $r^* > 4$ and a constant C_0 (independent of i, j, and t) such that $\mathrm{E}[|u_{jt}|^{r^*}] \le C_0$ *and* $\mathrm{E}[|\eta_{it}|^{r^*}] \le C_0$*, for all $i \in \mathbb{N}_0$, $1 \le j \le q$, and $t \in \mathbb{Z}$;*

A5. there exists a $\varphi \in (0, 2]$ and constants K_u and M_u (independent of j and t) and K_η and M_η (independent of i and t) such that, $\mathrm{P}\left(|u_{jt}| > \varepsilon\right) \le K_u \exp\left(-\varepsilon^{\varphi} M_u\right)$ *and* $\mathrm{P}\left(|\eta_{it}| > \varepsilon\right) \le K_\eta \exp\left(-\varepsilon^{\varphi} M_\eta\right)$*, for any $\varepsilon > 0$ and for all $1 \le j \le q$, $i \in \mathbb{N}_0$, and $t \in \mathbb{Z}$.*

Orthonormality and mutual orthogonality of the common and idiosyncratic shocks shocks in (A1)-(A3) are standard in this literature (see e.g. Assumption 1 in Eichler et al., 2011, and Assumption 1 in Forni et al., 2017). The moments conditions in (A4) are instead borrowed from the literature on high-dimensional estimation of spectral densities and we refer here mainly to Wu and Zaffaroni (2018) and Zhang and Wu (2019). Note that moment assumptions in the locally stationary setting are also for example in Dahlhaus (2009), where in fact all moments are assumed to exist. Last, condition (A5) allows for either sub-Gaussian tails ($\varphi = 2$), or heavier tails like sub-exponential ($\varphi = 1$) or even sub-Weibull ($0 < \varphi < 1$) (see e.g. Vershynin, 2018, Chapter 2, and Kuchibhotla and Chakrabortty, 2018).

In practice, observations of $\mathbf{X}$ are available over a finite number T of points. Due to nonstationarity, letting T tend to infinity, that is, extending the process into the future, will not provide further insight into the behavior of the process at the beginning of the time interval. Hence, in this context, we need a different asymptotic scheme in order to assess the quality of inference procedures — typically, in order to study the consistency, as n and T tend to infinity, of estimators of the time-varying impulse response functions $\mathbf{C}^*_n(t,L)$ over the time interval $[1, T]$.

Following Dahlhaus (2009), we consider the *locally stationary* asymptotic scheme, an approach that has been initiated in Dahlhaus (1997). More precisely, for any $\tau \in (0, 1)$, let $\mathbf{X}(\tau) = \{X_{it}(\tau) : \ i \in \mathbb{N}_0, \ t \in \mathbb{Z}\}$ denote the fictitious (i.e., non-observable) stationary process described by the GDFM decomposition

$$X_{it}(\tau) := \chi_{it}(\tau) + \xi_{it}(\tau), \quad i \in \mathbb{N}_0, \ t \in \mathbb{Z}, \tag{5}$$

where

$$\chi_{it}(\tau) = \sum_{j=1}^{q} \sum_{k=0}^{\infty} c_{ijk}(\tau) u_{j,t-k}, \quad i \in \mathbb{N}_0, \ t \in \mathbb{Z}, \tag{6}$$

$$\xi_{it}(\tau) = \sum_{j=1}^{\infty} \sum_{k=0}^{\infty} d_{ijk}(\tau) \eta_{j,t-k}, \quad i \in \mathbb{N}_0, \ t \in \mathbb{Z} \tag{7}$$

where the driving shocks u_{jt} and η_{jt} are the same as in (2) and (3) (hence satisfy Assumption (A)): write $\boldsymbol{\chi}(\tau)$ and $\boldsymbol{\xi}(\tau)$ for $\{\chi_{it}(\tau) : i \in \mathbb{N}_0, \ t \in \mathbb{Z}\}$ and $\{\xi_{it}(\tau) : i \in \mathbb{N}_0, \ t \in \mathbb{Z}\}$, respectively. Letting $\mathbf{X}_{nt}(\tau) := (X_{1t}(\tau), \ldots, X_{nt}(\tau))'$, $\boldsymbol{\chi}_{nt}(\tau) := (\chi_{1t}(\tau), \ldots, \chi_{nt}(\tau))'$, and $\boldsymbol{\xi}_{nt}(\tau) := (\xi_{1t}(\tau), \ldots, \xi_{nt}(\tau))'$, (5)-(7) also can be written, with obvious notation $\mathbf{C}_n(\tau; L)$ and $\mathbf{D}_n(\tau; L)$, as

$$\mathbf{X}_{nt}(\tau) = \boldsymbol{\chi}_{nt}(\tau) + \boldsymbol{\xi}_{nt}(\tau), \quad \tau \in (0,1), \ t \in \mathbb{Z}, \ n \in \mathbb{N}_0 \tag{8}$$

where

$$\boldsymbol{\chi}_{nt}(\tau) := \mathbf{C}_n(\tau; L)\mathbf{u}_t \quad \text{and} \quad \boldsymbol{\xi}_{nt}(\tau) := \mathbf{D}_n(\tau; L)\boldsymbol{\eta}_{nt}, \quad \tau \in (0,1), \ t \in \mathbb{Z}, \ n \in \mathbb{N}_0. \tag{9}$$

As τ ranges over $(0,1)$, the $\mathbf{X}(\tau)$'s thus constitute a collection of stationary processes. Denote by $\mathbf{X}_T := \{X_{it} : \ i \in \mathbb{N}_0, \ 1 \leq t \leq T\}$ the finite-T subprocess of the nonstationary $\mathbf{X}$. The idea consists in approximating the (nonstationary) component X_{it} of $\mathbf{X}_T$ with the value $X_{it}(t/T)$ of the stationary process $\mathbf{X}(\tau) = \{X_{is}(\tau) : \ i \in \mathbb{N}_0, \ s \in \mathbb{Z}\}$, $\tau = t/T$ (the so-called *rescaled time*) at time $s = t$:

$$X_{it} \approx X_{it}(t/T) = \chi_{it}(t/T) + \xi_{it}(t/T), \quad i \in \mathbb{N}_0, \ 1 \leq t \leq T, \tag{10}$$

where $\chi_{it}(t/T)$, defined in (6), depends on the coefficients $c_{ijk}(t/T)$ and $\xi_{it}(t/T)$, defined in (7), similarly depends on the coefficients $d_{ijk}(t/T)$.

If the approximation (10) is to make sense, of course, the coefficients in (6) and (7) need to satisfy some regularity assumptions, and to somehow approximate those in (2) and (3). We require the following regularity conditions for the coefficients of the common component.

Assumption (B part I).

B1. There exists a $\rho_\chi \in [0,1)$ and constants C_1' and C_1 (independent of k) such that

$$\sup_{i \in \mathbb{N}_0} \max_{1 \leq j \leq q} \sup_{t \in \mathbb{N}_0} |c^*_{ijk}(t)| \leq C_1' \rho_\chi^k \quad \textit{and} \quad \sup_{i \in \mathbb{N}_0} \max_{1 \leq j \leq q} \sup_{\tau \in (0,1)} |c_{ijk}(\tau)| \leq C_1 \rho_\chi^k$$

for all $k \in \mathbb{N}$.

B2. The filters $c^*_{ij}(t; L)$ and $c_{ij}(\tau; L)$ are rational for all $i \in \mathbb{N}_0$ and $1 \leq j \leq q$.

B3. There exists a constant C_χ (independent of T and k) such that

$$\sup_{i\in\mathbb{N}_0}\max_{1\le j\le q}\max_{1\le t\le T}|c^*_{ijk}(t)-c_{ijk}(t/T)|\le C_\chi\rho_\chi^k/T$$

for all $T\in\mathbb{N}_0$ and $k\in\mathbb{N}$, where ρ_χ is defined in (B1).

B4. The mapping $\tau\mapsto c_{ij}(\tau;z)$ is twice uniformly continuously differentiable for all $i,j\in\mathbb{N}_0$ and $z\in\mathbb{C}$, and there exists a constant C_2 (independent of k) such that

$$\sup_{i\in\mathbb{N}_0}\max_{1\le j\le q}\sup_{\tau\in(0,1)}\left|\frac{\mathrm{d}^2c_{ijk}(\tau)}{\mathrm{d}\tau^2}\right|\le C_2\rho_\chi^k$$

for all $k\in\mathbb{N}$, where ρ_χ is defined in (B1).

Assumption (B1) is a standard geometric decay requirement for the autocorrelations of the common components. Assumption (B2) allows to apply the results on singular processes used also in Forni et al. (2017). Note that a rational process is a process admitting a VARMA representation of finite (but unspecified) orders and that such processes are dense in the family of stationary processes.

In accordance with Dahlhaus' terminology, a sequence of processes $\{\boldsymbol{\chi}_{nt}(\tau)\}$ satisfying Assumptions (A1) and (B3) will be called *locally stationary*. Although the filters in (6) and (7) are not required to coincide with those in (2) and (3), the approximation in (10) is justified by condition (B3), which is similar to Definition 2.1 in Dahlhaus (1997) and Assumption 4.1 in Dahlhaus (2012). As we show in Section 3, this plays an essential role in the problem of consistent estimation of the impulse response coefficients $c^*_{ijk}(t)$ as n and T tend to infinity.

Assumption (B4) controls the degree of smoothness of the impulse response functions and it is standard in this context (see e.g. Assumption 4.1 in Dahlhaus, 2012). Obviously it implies uniform Lipschitz continuity of the function $c_{ijk}(\cdot)$ and, moreover, implies twofold uniform continuous differentiability of the time-varying spectral density (see Lemma 1 below); both these conditions are also required by Eichler et al. (2011, Assumption 2).

Similarly, we impose the following regularity conditions on the coefficients of the idiosyncratic component.

Assumption (B part II).

B5. There exists a $\rho_\xi\in[0,1)$ and constants B'_{1ij}, B_{1ij} (independent of k) such that

$$\sup_{t\in\mathbb{N}_0}|d^*_{ijk}(t)|\le B'_{1ij}\rho_\xi^k\quad\textit{and}\quad\sup_{\tau\in(0,1)}|d_{ijk}(\tau)|\le B_{1ij}\rho_\xi^k$$

for all $k\in\mathbb{N}$. Moreover, there exist constants B'_1 and B_1 (independent of i and j) such that

$$\sum_{j=1}^{\infty}B'_{1ij}\le B'_1,\quad\sum_{i=1}^{\infty}B'_{1ij}\le B'_1,\quad\sum_{j=1}^{\infty}B_{1ij}\le B_1,\quad\textit{and}\quad\sum_{i=1}^{\infty}B_{1ij}\le B_1$$

for all $i,j\in\mathbb{N}_0$.

B6. There exist constants $B_{\xi ij}$ (independent of T and k) such that

$$\max_{1\leq t\leq T} |d^*_{ijk}(t) - d_{ijk}(t/T)| \leq B_{\xi ij}\rho_\xi^k/T$$

for all $T \in \mathbb{N}_0$ and $k \in \mathbb{N}$, where ρ_ξ is defined in (B5). Moreover, there exists a constant B_ξ (independent of i and j) such that

$$\sum_{j=1}^{\infty} B_{\xi ij} \leq B_\xi \quad \textit{and} \quad \sum_{i=1}^{\infty} B_{\xi ij} \leq B_\xi$$

for all $i, j \in \mathbb{N}_0$.

B7. The mapping $\tau \mapsto d_{ij}(\tau; z)$ is twice uniformly continuously differentiable for all $i, j \in \mathbb{N}_0$ and $z \in \mathbb{C}$, and there exist constants B_{2ij} (independent of k) such that

$$\sup_{\tau\in(0,1)} \left| \frac{\mathrm{d}^2 d_{ijk}(\tau)}{\mathrm{d}\tau^2} \right| \leq B_{2ij}\rho_\xi^k$$

for all $k \in \mathbb{N}$, where ρ_ξ is defined in (B5). Moreover, there exists a constant B_2 (independent of i and j) such that

$$\sum_{j=1}^{\infty} B_{2ij} \leq B_2 \quad \textit{and} \quad \sum_{i=1}^{\infty} B_{2ij} \leq B_2$$

for all $i, j \in \mathbb{N}_0$.

Assumptions (B5), (B6), and (B7) are the analogues of (B1), (B3), and (B4), respectively. In particular, the summability conditions of the constants are a generalization to the tvGDFM of Assumption 4 by Forni et al. (2017), which requires the idiosyncratic components to have "weak" cross-dependence only (see Lemma 2 below).

Unlike the nonstationary $\mathbf{X}$, $\boldsymbol{\chi}$, and $\boldsymbol{\xi}$, the stationary processes $\mathbf{X}(\tau)$, $\boldsymbol{\chi}(\tau)$, and $\boldsymbol{\xi}(\tau)$, for any $\tau \in (0, 1)$, under Assumptions (A) and (B), admit well-defined spectral densities. For any $n \in \mathbb{N}_0$, define the $n \times n$ lag ℓ autocovariance matrix of $\mathbf{X}_n(\tau)$ as

$$\mathbf{\Gamma}^X_n(\tau; \ell) := \mathrm{E}[\mathbf{X}_{nt}(\tau)\mathbf{X}'_{n,t-\ell}(\tau)], \quad \ell \in \mathbb{Z}, \quad \tau \in (0, 1). \tag{11}$$

Then, for any $n \in \mathbb{N}_0$, the $n \times n$ spectral density matrix of $\mathbf{X}_n(\tau)$ is defined as

$$\mathbf{\Sigma}^X_n(\tau; \theta) := \frac{1}{2\pi} \sum_{\ell=-\infty}^{\infty} e^{-\iota\ell\theta}\, \mathbf{\Gamma}^X_n(\tau; \ell), \quad \theta \in [-\pi, \pi], \quad \tau \in (0, 1). \tag{12}$$

Similarly define $\mathbf{\Sigma}^\chi_n(\tau; \theta)$ and $\mathbf{\Sigma}^\xi_n(\tau; \theta)$. For given τ and θ, each of the matrix sequences $\mathbf{\Sigma}^X_n(\tau; \theta)$, $\mathbf{\Sigma}^\chi_n(\tau; \theta)$, and $\mathbf{\Sigma}^\xi_n(\tau; \theta)$ is nested as n increases.

Obviously, because of Assumption (A3),

$$\mathbf{\Sigma}^X_n(\tau; \theta) = \mathbf{\Sigma}^\chi_n(\tau; \theta) + \mathbf{\Sigma}^\xi_n(\tau; \theta), \quad \tau \in (0, 1),\ \theta \in (0, 2\pi].$$

Assuming that the nT-dimensional process $\mathbf{X}_{nT} := \{X_{it} : \ 1 \leq i \leq n,\ 1 \leq t \leq T\}$ (an $n \times T$ panel) is observed, we associate with each $t = 1, \ldots, T$ the spectral density matrices $\mathbf{\Sigma}^X_n(t/T; \theta)$, $\mathbf{\Sigma}^\chi_n(t/T; \theta)$, and $\mathbf{\Sigma}^\xi_n(t/T; \theta)$: those spectral matrices, which depend on rescaled time, will be used as local substitutes for the nonexisting (or meaningless) spectra of $\mathbf{X}_{nT}$.

We denote by $\sigma_{ij}^X(\tau;\theta)$ the (i,j) entry of $\mathbf{\Sigma}_n^X(\tau;\theta)$. The following regularity conditions, proved in the online Appendix, are a consequence of Assumptions (A) and (B).

Lemma 1. *Under Assumptions (A)-(B),*

(i) the mapping $\tau \mapsto \sigma_{ij}^X(\tau;\theta)$ is twice uniformly continuously differentiable for all $i,j \in \mathbb{N}_0$ and $\theta \in [-\pi,\pi]$ and there exists a constant $\mathcal{K}$ (independent of n) such that, for any $n \in \mathbb{N}_0$,

$$\max_{1\le i,j\le n} \sup_{\tau\in(0,1)} \sup_{\theta\in[-\pi,\pi]} \left| \frac{\mathrm{d}^2\sigma_{ij}^X(\tau;\theta)}{\mathrm{d}\tau^2} \right| \le \mathcal{K}.$$

(ii) the mapping $\theta \mapsto \sigma_{ij}^X(\tau;\theta)$ is twice uniformly continuously differentiable for all $i,j \in \mathbb{N}_0$ and $\tau \in (0,1)$ and there exists a constant $\mathcal{K}'$ (independent of n) such that, for any $n \in \mathbb{N}_0$,

$$\max_{1\le i,j\le n} \sup_{\tau\in(0,1)} \sup_{\theta\in[-\pi,\pi]} \left| \frac{\mathrm{d}^2\sigma_{ij}^X(\tau;\theta)}{\mathrm{d}\theta^2} \right| \le \mathcal{K}'.$$

These smoothness requirements are also in Eichler et al. (2011, Assumption 2(i)) and, as a consequence, $\sigma_{ij}^X(\tau;\theta)$ is uniformly Lipschitz continuous both in τ and in θ.

Denote by $\lambda_{j;n}^X(\tau;\theta)$, $\lambda_{j;n}^\chi(\tau;\theta)$, and $\lambda_{j;n}^\xi(\tau;\theta)$ the jth eigenvalues (in decreasing order of magnitude) of the spectral density matrices $\mathbf{\Sigma}_n^X(\tau;\theta)$, $\mathbf{\Sigma}_n^\chi(\tau;\theta)$, and $\mathbf{\Sigma}_n^\xi(\tau;\theta)$, respectively. We make the following assumption.

Assumption (C). *There exist continuous functions $\theta \mapsto \alpha_j^\chi(\tau;\theta)$ and $\theta \mapsto \beta_j^\chi(\tau;\theta)$, $1 \le j \le q$, and an integer N_χ such that, for all $n \ge N_\chi$, any given $\tau \in (0,1)$, and Lebesgue-a.e.*[4] *over $\theta \in [-\pi,\pi]$,*

$$\beta_1^\chi(\tau;\theta) \ge \frac{\lambda_{1;n}^\chi(\tau;\theta)}{n} \ge \alpha_1^\chi(\tau;\theta) > \beta_2^\chi(\tau;\theta) \ge \frac{\lambda_{2;n}^\chi(\tau;\theta)}{n} \ge \dots$$
$$\dots \ge \alpha_{q-1}^\chi(\tau;\theta) > \beta_q^\chi(\tau;\theta) \ge \frac{\lambda_{q;n}^\chi(\tau;\theta)}{n} \ge \alpha_q^\chi(\tau;\theta) > 0.$$

Assumption (C) is a generalization to the time-varying case of the classical assumption of pervasive factors (see also Forni et al., 2000, Assumption 3 in Eichler et al., 2011, and Assumption 3 in Forni et al., 2017).

The eigenvalues $\lambda_{j;n}^\xi(\tau;\theta)$ of $\mathbf{\Sigma}_n^\xi(\tau;\theta)$ and $\lambda_{j;n}^X(\tau;\theta)$ of $\mathbf{\Sigma}_n^X(\tau;\theta)$ are then characterized by the following lemma, which is proved in the online Appendix.

Lemma 2. *Under Assumptions (A)-(C),*

(i) there exists a constant B_ξ such that $\lambda_{1;n}^\xi(\tau;\theta) \le B_\xi$ for all $n \in \mathbb{N}_0$, all $\tau \in (0,1)$ and all $\theta \in [-\pi,\pi]$;

[4] That is, except for a subset of θ values included in a set with Lebesgue measure zero.

(ii) there exist continuous functions $\theta \mapsto \alpha_j(\tau;\theta)$ and $\theta \mapsto \beta_j(\tau;\theta)$, $1 \le j \le q$, and an integer N_X such that, for all $n \ge N_X$, any given $\tau \in (0,1)$, and Lebesgue-a.e. over $\theta \in [-\pi, \pi]$,

$$\beta_1(\tau;\theta) \ge \frac{\lambda^X_{1;n}(\tau;\theta)}{n} \ge \alpha_1(\tau;\theta) > \beta_2(\tau;\theta) \ge \frac{\lambda^X_{2;n}(\tau;\theta)}{n} \ge \dots$$
$$\dots \ge \alpha_{q-1}(\tau;\theta) > \beta_q(\tau;\theta) \ge \frac{\lambda^X_{q;n}(\tau;\theta)}{n} \ge \alpha_q(\tau;\theta) > 0; \qquad (13)$$

(iii) there exists a constant B_X such that $\lambda^X_{q+1;n}(\tau;\theta) \le B_X$ for all $n \in \mathbb{N}_0$, all $\tau \in (0,1)$, and all $\theta \in [-\pi, \pi]$.

Parts *(ii)* and *(iii)* of Lemma 2 imply the presence of an eigen-gap in the spectral density matrix of $\mathbf{X}_n(\tau)$ which is increasing with n, and therefore allows for identification of the tvGDFM as $n \to \infty$.

It is important to stress that Assumption (C) rules out the possibility of a number of factors depending on τ in the stationary processes $\mathbf{X}(\tau)$ and $\boldsymbol{\chi}(\tau)$: irrespective of τ, all spectral density matrices $\mathbf{\Sigma}^X_n(\tau;\theta)$ and $\mathbf{\Sigma}^\chi_n(\tau;\theta)$ have (for $n \ge q+1$) q distinct and exploding (as $n \to \infty$) eigenvalues. The slow variation in time of the tvGDFM parameters implied by Assumptions (B3)-(B4) and (B6)-(B7) is not compatible with a time-varying number of factors and, for this reason, q in Assumption (A1) is fixed over time. Note, however, that a factor model with a time-varying number of factors $q(t)$, say, can always be written as a factor model with a constant number $q := \max_{1 \le t \le T} q(t)$ of factors in which almost all loadings relative to some given common shock are zero over some time period (see Barigozzi et al. (2018)). Such a situation is incompatible with the idea of slowly varying loadings because it corresponds to a cross-sectionally pervasive change in the tvGDFM structure. However, as mentioned in the Introduction, and in agreement with the results by Bates et al. (2013), we expect our locally stationary approach to be robust against the presence of “small” (cross-sectionally non-pervasive) change-points.

We conclude this section by assuming the existence of a singular autoregressive representation for the common components processes $\boldsymbol{\chi}(\tau)$ (for a stationary version, see Assumption 5 in Forni et al., 2017).

Assumption (D). *For any $k \in \mathbb{N}_0$, denote by*

$$\boldsymbol{\chi}^{(k)}(\tau) := \{\boldsymbol{\chi}^{(k)}_t(\tau) := (\chi_{(k-1)(q+1)+1,t}(\tau), \dots, \chi_{k(q+1),t}(\tau))' : t \in \mathbb{Z}\}$$

an arbitrary $(q+1)$-dimensional subprocess of $\boldsymbol{\chi}(\tau)$ (as defined in (6)). Then, for all $\tau \in (0,1)$ and $k \in \mathbb{N}_0$,

D1. there exists a unique autoregressive $(q+1) \times (q+1)$ filter $\mathbf{A}^{(k)}(\tau;L)$ and a $(q+1) \times q$ matrix $\mathbf{H}^{(k)}(\tau)$ of rank q such that

$$\mathbf{A}^{(k)}(\tau;L)\boldsymbol{\chi}^{(k)}_t(\tau) = \mathbf{H}^{(k)}(\tau)\mathbf{u}_t, \quad t \in \mathbb{Z};$$

D2. there exists a constant S (independent of τ and k) such that, denoting by $s_k(\tau)$ the order of $\mathbf{A}^{(k)}(\tau;L)$, $s_k(\tau) \le S$ for all $k \in \mathbb{N}_0$ and $\tau \in (0,1)$;

D3. $\det \mathbf{A}^{(k)}(\tau; z) \neq 0$ *for all* $z \in \mathbb{C}$ *such that* $|z| \leq 1$*;*

D4. letting $\mathbf{\Gamma}^{\chi^{(k)}}(\tau; \ell) := \mathrm{E}[\boldsymbol{\chi}_t^{(k)}(\tau)\boldsymbol{\chi}_{t-\ell}^{(k)\prime}(\tau)]$ *denote the* $(q+1) \times (q+1)$ *lag* ℓ *autocovariance matrix of* $\boldsymbol{\chi}^{(k)}(\tau)$ *and defining, for* $S > 0$*, the* $S(q+1) \times S(q+1)$ *matrix*

$$\mathbf{G}^{(k)}(\tau) := \begin{pmatrix} \mathbf{\Gamma}^{\chi^{(k)}}(\tau; 0) & \mathbf{\Gamma}^{\chi^{(k)}}(\tau; 1) & \dots & \mathbf{\Gamma}^{\chi^{(k)}}(\tau; S-1) \\ \mathbf{\Gamma}^{\chi^{(k)}}(\tau; -1) & \mathbf{\Gamma}^{\chi^{(k)}}(\tau; 0) & \dots & \mathbf{\Gamma}^{\chi^{(k)}}(\tau; S-2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Gamma}^{\chi^{(k)}}(\tau; -S+1) & \mathbf{\Gamma}^{\chi^{(k)}}(\tau; -S+2) & \dots & \mathbf{\Gamma}^{\chi^{(k)}}(\tau; 0) \end{pmatrix},$$

there exists a constant d *(independent of* τ *and* k*) such that* $\det \mathbf{G}^{(k)}(\tau) > d$ *(for* $S = 0$*, let* $\mathbf{G}^{(k)}(\tau) := \mathbf{I}_{q+1}$*).*

Assumption (D) jointly with (B2) is crucial for allowing us to estimate the model by means of one-sided filters. Actually, it has been shown by Anderson and Deistler (2008a,b) that, for rational processes, Assumption (D) holds *generically.*[5] Generically is not enough here, though, and this is why we need to make it an assumption which, however, for the same reason, turns out to be a very mild one (see also Section 4 in Forni et al., 2015).

Now consider the case in which $n = m(q+1)$ for some integer m.[6] The n-dimensional common component $\boldsymbol{\chi}_n(\tau)$ under Assumption (D) admits the autoregressive representation

$$\mathbf{A}_n(\tau; L)\boldsymbol{\chi}_{nt}(\tau) = \mathbf{R}_n(\tau)\mathbf{u}_t, \qquad \tau \in (0,1),\ t \in \mathbb{Z} \tag{14}$$

where, for $\tau \in (0,1)$, $\mathbf{A}_n(\tau; L)$ is block-diagonal with $(q+1)\times(q+1)$ diagonal blocks $\mathbf{A}^{(k)}(\tau; L)$, $1 \leq k \leq m$ satisfying Assumption (D) and $\mathbf{R}_n(\tau)$ (stacking m $(q+1) \times q$ matrices of the type $\mathbf{H}^{(k)}(\tau)$) is of dimension $n \times q$ with full column rank q. Letting $\mathbf{Z}_{nt}(\tau) := \mathbf{A}_n(\tau; L)\mathbf{X}_{nt}(\tau)$, we have

$$\mathbf{Z}_{nt}(\tau) = \mathbf{R}_n(\tau)\mathbf{u}_t + \mathbf{A}_n(\tau; L)\boldsymbol{\xi}_{nt}(\tau) =: \boldsymbol{\psi}_{nt}(\tau) + \boldsymbol{\zeta}_{nt}(\tau), \qquad \tau \in (0,1),\ t \in \mathbb{Z}. \tag{15}$$

This constitutes for $\mathbf{Z}_{nt}(\tau)$ a stationary *static* factor model with the same q common shocks $\{\mathbf{u}_t\}$ as those appearing in the definition (2) of the nonstationary GDFM for $\mathbf{X}$. Let $\mathbf{\Gamma}_n^Z(\tau)$, $\mathbf{\Gamma}_n^\psi(\tau)$, and $\mathbf{\Gamma}_n^\zeta(\tau)$ stand for the $n \times n$ covariance matrices of $\mathbf{Z}_{nt}(\tau)$, $\boldsymbol{\psi}_{nt}(\tau)$, and $\boldsymbol{\zeta}_{nt}(\tau)$, respectively; because of Assumption (A3), we have

$$\mathbf{\Gamma}_n^Z(\tau) = \mathbf{\Gamma}_n^\psi(\tau) + \mathbf{\Gamma}_n^\zeta(\tau).$$

To conclude, let $\mu_{j;n}^\psi(\tau)$ and $\mu_{j;n}^\zeta(\tau)$ denote the jth eigenvalues (in decreasing order of magnitude) of the covariance matrices $\mathbf{\Gamma}_n^\psi(\tau)$ and $\mathbf{\Gamma}_n^\zeta(\tau)$, respectively. The next assumption and the following Lemma, proved in the online Appendix, allow us to identify, for any given $\tau \in (0,1)$, the decomposition (15) as $n \to \infty$ (see, for the stationary case, Assumption 6 and Proposition 4 in Forni et al., 2017).

[5] That is, except for a subset of the parameter space of their VARMA representation contained in a set of Lebesgue measure zero.

[6] This is convenient and does not imply any loss of generality for our asymptotic analysis, see the end Section 3.1 for further details when it does not hold.

Assumption (E). *There exist continuous functions $\alpha_j^\psi(\tau)$ and $\beta_j^\psi(\tau)$, $1 \leq j \leq q$, and an integer N_ψ such that, for all $n \geq N_\psi$, and any given $\tau \in (0,1)$,*

$$\beta_1^\psi(\tau) \geq \frac{\mu_{1;n}^\psi(\tau)}{n} \geq \alpha_1^\psi(\tau) > \beta_2^\psi(\tau) \geq \frac{\mu_{2;n}^\psi(\tau)}{n} \geq \dots$$
$$\dots \geq \alpha_{q-1}^\psi(\tau) > \beta_q^\psi(\tau) \geq \frac{\mu_{q;n}^\psi(\tau)}{n} \geq \alpha_q^\psi(\tau) > 0.$$

Lemma 3. *Under Assumptions (A)-(D), there exists a constant B_ζ such that $\mu_{1;n}^\zeta(\tau) \leq B_\zeta$ for all $n \in \mathbb{N}_0$ and all $\tau \in (0,1)$.*

3 Estimation and consistency

In this section, we show how to adapt the Forni et al. (2017) one-sided estimation method to the time-varying setting described by Assumptions (A)-(E). The substantial advantage over the Eichler et al. (2011) time-varying extension of the simpler dynamic principal component analysis of Forni et al. (2000) is that the approach considered in this paper delivers estimators of the filters $\mathbf{C}_n^*(t, L)$ $(1 \leq t \leq T)$ which are one-sided and therefore can be used directly for time-varying impulse response analysis.

Hereafter, all estimated quantities are denoted with "hats", e.g. $\widehat{c}_{ij;n,T}(t)$ for the estimator, based on the observation of an $n \times T$ realization $\mathbf{X}_{nT}$ of $\mathbf{X}$, of the kth coefficient in the (i,j)th entry of $\mathbf{C}_n^*(t, L)$, etc. Suffixes highlight the dependence on T and (possibly) also on n of those quantities.

3.1 Estimation

Our estimation procedure is based on four main steps; throughout this section, n and T are fixed.

(i) Estimation of the Spectral Density. First, we need an estimator of the spectral density matrices $\mathbf{\Sigma}_n^X(\tau;\theta)$; here, we follow Zhang and Wu (2019). For any given $\tau \in (0,1)$, define, for some bandwidth $b_T \in [0, 1/2]$, $T_1(\tau) := \lfloor T\tau \rfloor - \lfloor Tb_T \rfloor + 1$ and $T_2(\tau) := \lfloor T\tau \rfloor + \lfloor Tb_T \rfloor$. Next, letting $M_T := 2\lfloor Tb_T \rfloor$, define the local estimator of the lag ℓ autocovariance matrix of $\mathbf{X}_n(\tau)$ as

$$\widehat{\mathbf{\Gamma}}_{n,T}^X(\tau;\ell) := \begin{cases} \dfrac{1}{M_T} \displaystyle\sum_{s=T_1(\tau)+\ell}^{T_2(\tau)} \mathrm{J}\left(\dfrac{s - \lfloor \tau T \rfloor}{M_T}\right) \mathbf{X}_{n,s-\ell}\mathbf{X}'_{ns} & \tau \in (0,1),\ 0 \leq \ell \leq (M_T - 1) \\ \widehat{\mathbf{\Gamma}}_{n,T}^{X\prime}(\tau;-\ell) & (-M_T + 1) \leq \ell \leq -1 \end{cases} \tag{16}$$

(see also Rodríguez-Poo and Linton, 2001) where $\mathrm{J}(\cdot)$ is a suitable kernel. Then define, for some $m_T \in \mathbb{N}_0$ such that $m_T < M_T$, the local estimator of the spectral density matrix of $\mathbf{X}_n(\tau)$ as

$$\widehat{\mathbf{\Sigma}}_{n,T}^X(\tau;\theta) := \frac{1}{2\pi} \sum_{\ell=-m_T}^{m_T} \mathrm{K}\left(\frac{|\ell|}{m_T}\right) \widehat{\mathbf{\Gamma}}_{n,T}^X(\tau;\ell) e^{-\iota \ell \theta}, \quad \tau \in (0,1),\ \theta \in [-\pi, \pi], \tag{17}$$

where $\mathrm{K}(\cdot)$ is a suitable kernel. We refer to Assumption (F) in the next section for details on the choice of the kernels, M_T, and m_T.

In practice, the estimator (17) can be computed only at a discrete number of time points and frequencies. Specifically, recalling that M_T by definition is always even, at any given τ, the sum in (16) only involves the M_T time points $t_s := t+s$ such that $(-M_T/2+1) \leq s \leq M_T/2$. As a consequence, we are able to compute the estimator (16) only for the central $(T-M_T+1)$ time points of the sample, i.e., we compute $\widehat{\mathbf{\Gamma}}^X_{n,T}(t/T;\ell)$ for $M_T/2 \leq t \leq (T-M_T/2)$. Similarly, we only are able to consider the m_T frequencies $\theta_j := \pi j/m_T$ with $|j| \leq m_T$; indeed, the maximum achievable resolution for (17) in the frequency domain is $m_T/(2\pi)$.[7]

(ii) Dynamic Principal Component Analysis. Denote by $\widehat{\lambda}^X_{j;n,T}(t/T;\theta_j)$ the jth eigenvalue (in decreasing order of magnitude) of $\widehat{\mathbf{\Sigma}}^X_{n,T}(t/T;\theta_j)$ and by $\widehat{\mathbf{P}}^X_{j;n,T}(t/T;\theta_j)$, the corresponding n-dimensional normalized eigenvector. Then, for a given number q of factors,

$$\widehat{\mathbf{\Sigma}}^\chi_{n,T}(t/T;\theta_j) := \sum_{j=1}^{q} \widehat{\lambda}^X_{j;n,T}(t/T;\theta_j)\widehat{\mathbf{P}}^X_{j;n,T}(t/T;\theta_j)\widehat{\mathbf{P}}^{X\dagger}_{j;n,T}(t/T;\theta_j),$$
$$M_T/2 \leq t \leq (T-M_T/2),\ |j| \leq m_T, \tag{18}$$

is an estimator of the spectral density $\mathbf{\Sigma}^\chi_n(\tau;\theta)$ of the common component.

Lastly, by inverse Fourier transform, compute from (18) the estimators

$$\widehat{\mathbf{\Gamma}}^\chi_{n,T}(t/T;\ell) := \frac{2\pi}{m_T}\sum_{j=-m_T}^{m_T} \widehat{\mathbf{\Sigma}}^\chi_{n,T}(t/T;\theta_j)e^{\iota\ell\theta_j}, \quad M_T/2 \leq t \leq (T-M_T/2),\ \ell \in \mathbb{Z}. \tag{19}$$

of the local autocovariance matrices of the common component.

(iii) VAR filtering. Assuming again, for simplicity, that $n = m(q+1)$ for some integer m, consider the m autoregressive models (each of dimension $(q+1)$: see Assumption (D))

$$\mathbf{A}^{(k)}_n(t/T;L)\boldsymbol{\chi}^{(k)}_{nt}(t/T) = \mathbf{H}^{(k)}_n(t/T)\mathbf{u}_t, \quad M_T/2 \leq t \leq (T-M_T/2),\ 1 \leq k \leq m. \tag{20}$$

Based on the estimated autocovariances (19), compute, using AIC for determining the VAR orders, the Yule-Walker estimates $\widehat{\mathbf{A}}^{(k)}_{n,T}(t/T;L)$ of the autoregressive filters $\mathbf{A}^{(k)}_n(t/T;L)$. Construct the $n \times n$ block-diagonal filter $\widehat{\mathbf{A}}_{n,T}(t/T;L)$ with (see (14)) the m diagonal blocks

$$\widehat{\mathbf{A}}^{(1)}_{n,T}(t/T;L), \ldots, \widehat{\mathbf{A}}^{(m)}_{n,T}(t/T;L)$$

and the filtered process

$$\widehat{\mathbf{Z}}_{nt}(t/T) := \widehat{\mathbf{A}}_{n,T}(t/T;L)\mathbf{X}_{nt}, \quad M_T/2 \leq t \leq (T-M_T/2), \tag{21}$$

to be used as an estimation of $\mathbf{Z}_{nt}(t/T)$ satisfying (15) for $\tau = t/T$.

(iv) Principal Component Analysis. Consider the estimator

$$\widehat{\mathbf{\Gamma}}^{\widehat{Z}}_{n,T}(t/T) := \frac{1}{M_T}\sum_{s=T_1(t/T)}^{T_2(t/T)} \mathrm{J}\left(\frac{s-t}{M_T}\right)\widehat{\mathbf{Z}}_{ns}(t/T)\widehat{\mathbf{Z}}_{ns}(t/T)', \quad M_T/2 \leq t \leq (T-M_T/2), \tag{22}$$

[7] Hence, it makes no sense to compute (17) over the finer grid $\theta_j = \pi j/T$ with $|j| \leq T$.

of the covariance matrix of $\widehat{\mathbf{Z}}_{nt}(t/T)$, where M_T, T_1, T_2, and $\mathrm{J}(\cdot)$ are the same as in (16). Denoting by $\widehat{\mu}^{\widehat{Z}}_{j;n,T}(t/T)$ the jth eigenvalue of $\widehat{\mathbf{\Gamma}}^{\widehat{Z}}_n(t/T)$ (in decreasing order of magnitude), with normalized n-dimensional eigenvector $\widehat{\mathbf{P}}^{\widehat{Z}}_{j;n,T}(t/T)$, define

$$\widehat{\mathbf{R}}_{j;n,T}(t/T) := \widehat{\mathbf{P}}^{\widehat{Z}}_{j;n,T}(t/T)\sqrt{\widehat{\mu}^{\widehat{Z}}_{j;n,T}(t/T)}, \quad M_T/2 \leq t \leq (T - M_T/2),\ 1 \leq j \leq q,$$

and let $\widehat{\mathbf{R}}_{n,T}(t/T) := (\widehat{\mathbf{R}}_{1;n,T}(t/T), \ldots, \widehat{\mathbf{R}}_{q;n,T}(t/T))$. Our estimators of the impulse response functions $\mathbf{C}^*_n(t;L)$ are

$$\widehat{\mathbf{C}}_{n,T}(t;L) := [\widehat{\mathbf{A}}_{n,T}(t/T;L)]^{-1}\widehat{\mathbf{R}}_{n,T}(t/T), \quad M_T/2 \leq t \leq (T - M_T/2), \tag{23}$$

with (i,j) entry $\widehat{c}_{ij;n,T}(t;L) := \sum_{k=0}^{\infty} \widehat{c}_{ijk;n,T}(t)L^k$. As shown in the next section (see Proposition 1), the latter, up to a sign, is a consistent estimator of $c^*_{ij}(t;L) := \sum_{k=0}^{\infty} c^*_{ijk}(t)L^k$, as $n, T \to \infty$.

The estimation procedure just described calls for some comments.

First, the estimator of the time-varying spectral density in step *(i)*, which is the one proposed in Zhang and Wu (2019), consists of a local-in-time estimator of the autocovariances smoothed over time and is then used to compute at each point in time the usual weighted periodogram. Another possible estimator, not considered in this paper, would be the smoothed pre-periodogram proposed by Neumann and von Sachs (1997), where a local-in-time pre-periodogram is computed first, then is smoothed both over time and over frequencies. This latter estimator has also been considered by Dahlhaus (2009) and Eichler et al. (2011).

Second, step *(ii)* is directly taken from Eichler et al. (2011), who propose to estimate the common component by means of time-varying dynamic principal components. Steps *(iii)* and *(iv)*, and the estimator (23) of the time-varying impulse response function to common shocks represent the main novelty of this paper, being the generalization to the time-varying case of the approach proposed by Forni et al. (2017). In particular, step *(iv)* shows how an adequate VAR filtering brings the problem back to a time-varying static factor model in the style of Rodríguez-Poo and Linton (2001) and Motta et al. (2011).

Third, the matrices $\mathbf{R}_n(\tau)$ and the noise $\mathbf{u}_t$ in (14) are identified up to an arbitrary invertible transformation $\mathbf{P}(\tau)$ only, as $\mathbf{R}_n(\tau)\mathbf{u}_t = \mathbf{R}_n(\tau)\mathbf{P}(\tau)\mathbf{P}^{-1}(\tau)\mathbf{u}_t$. It is shown in the online Appendix that our choice of $\widehat{\mathbf{R}}_{n,T}(t/T)$, together with Assumption (A1) on the orthonormality of $\{\mathbf{u}_t\}$, identifies the impulse responses up to a sign. That sign issue can be solved by imposing identification constraints: see, for instance, Section 4.1 of Forni et al. (2009) in a stationary setting. Since, however, our study of connectedness in Section 4 does not require specifying those signs, we are skipping details.

Fourth, the cross-sectional ordering of the panel has an impact on the selection of the m subvectors $\boldsymbol{\chi}^{(k)}_{nt}(\tau)$ in step *(iii)* and the possible dropping of $n - \lfloor n/(q+1) \rfloor (q+1)$ series at the end of the panel when n is not an exact multiple of $(q+1)$. The $n!$ cross-sectional permutations of the panel, thus, would lead to $n!$ estimators, all sharing the same consistency properties stated in Proposition 1. A Rao-Blackwell argument (see Section 3.5 of Forni et al., 2017 for details) suggests aggregating these estimators into a unique one by simple averaging (after obvious reordering) of the resulting impulse response functions. Although averaging over all $n!$ permutations is clearly unfeasible, as stressed by Forni et al. (2017, Section 4.2) and Forni et al. (2018, Appendix D) in a stationary setting, a few of them are enough, in practice, to deliver stable averages (which therefore are matching the infeasible average over all $n!$ permutations). Such averaging clearly has no impact on consistency.

Fifth, the number q of common shocks throughout has been considered as known, and we assumed it to be constant through time. That number has to be estimated from the observations, though. We suggest using the criterion proposed by Hallin and Liška (2007). However, instead of implementing it on the basis of a classical estimator of the spectral density (invalid in the present context), we suggest running the method on the local estimator $\widehat{\mathbf{\Sigma}}^X_{n,T}(1/2;\theta_j)$ of the spectral density associated with the central part of the observation period ($\tau = 1/2$). The validity of the assumption of a constant number q of common shocks also can be tested heuristically by performing the same analysis for a few values of τ, then comparing the results. This is how we determine q in Section 4.

3.2 Consistency

We now turn to the proof of consistency, as n and T tend to infinity, of the estimated time-varying impulse response functions (23). This, however, requires assumptions on the bandwidths and the kernels used to estimate the autocovariance matrices in (16) and the spectral density matrix in (17). In the previous section we defined $M_T := 2\lfloor Tb_T \rfloor$ for some $b_T \in [0, 1/2]$ so that $0 \leq M_T \leq T$, and, similarly, we now define $m_T := \lfloor 1/h_T \rfloor$ for some $h_T \in (1/M_T, \infty)$, so that $0 \leq m_T < M_T$, as required in Section 3.1.

On kernels and badwidths, we make the following assumptions.

Assumption (F).

F1. The kernels $\mathrm{J}(\cdot)$ *and* $\mathrm{K}(\cdot)$ *are such that*

- *(a)* $\mathrm{J} : [-1/2, 1/2] \to \mathbb{R}^+$ *is continuous, symmetric, and such that* $\int_{-1/2}^{1/2} \mathrm{J}(u)\mathrm{d}u = 1$*;*
- *(b)* $\mathrm{K} : [-1, 1] \to \mathbb{R}^+$ *is continuous, symmetric, and such that* $\int_{-1}^{1} \mathrm{K}(u)\mathrm{d}u = 1$*.*

F2. The bandwidth b_T *is such that* $b_T \asymp T^{-1/3+2/(3r^*)}$ *if* $4 < r^* < 8$*, or* $b_T \asymp T^{-1/4}$ *if* $r^* \geq 8$*, where* r^* *is defined in Assumption (A4).*

F3. The bandwidth h_T *is such that* $h_T \asymp b_T$*.*

Assumption (F1) is standard in the literature (see e.g. Example 4.2 in Dahlhaus, 2009). Assumptions (F2) and (F3) imply the standard asymptotic conditions on the bandwidths, i.e., $b_T \to 0$, $h_T \to 0$, and $Tb_T \to \infty$, $Th_T \to \infty$ as $T \to \infty$. The assumed rates are the optimal ones, in the sense that they deliver the minimum mean-squared-error when estimating the spectral density (see Lemma 4 below and the following discussion). In particular, notice that, because of these assumptions, we must have, neglecting logarithmic in T quantities and powers thereof,

$$\begin{array}{llll} m_T \asymp T^{1/3-2/(3r^*)} & \text{and} \quad M_T \asymp T^{2/3-2/(3r^*)} & if & 4 < r^* < 8 \\ m_T \asymp T^{1/4} & \text{and} \quad M_T \asymp T^{3/4} & if & 8 \leq r^*. \end{array}$$

Although our theory holds for any kernel satisfying (F1), in practice, we adopt

$$\mathrm{J}(u) = \begin{cases} 1 & \text{if } |u| \leq 1/2, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad \mathrm{K}(u) = \begin{cases} 1 - |u| & \text{if } |u| \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

While the choice of a uniform kernel for J(·) is recommended by Zhang and Wu (2019), the choice of a triangular kernel for K(·) is very common (see e.g. Forni et al., 2017 in the stationary case). Notice that, with this choice of kernels, our estimator of the local spectral density is nothing else but the classical weighted periodogram computed using a rolling window of observations of size M_T.

Let $\widehat{\sigma}^X_{ij;T}(\tau;\theta)$ and $\sigma^X_{ij}(\tau;\theta)$ denote the (i,j) entries of $\widehat{\mathbf{\Sigma}}^X_{n,T}(\tau;\theta)$ and $\mathbf{\Sigma}^X_n(\tau;\theta)$, respectively (due to nestedness, $\widehat{\sigma}^X_{ij;T}(\tau;\theta)$ and $\sigma^X_{ij}(\tau;\theta)$ do not depend on n).

We have the following result.

Lemma 4. *Under Assumptions (A), (B), and (F), and given the choice of kernels in* (24)*, there exists constants C_X, C'_X, and C''_X (independent of T and n) such that, for any $n,T \in \mathbb{N}_0$,*

$$\max_{1\le i,j\le n} \mathrm{E}\left[\sup_{\tau\in(0,1)}\sup_{\theta\in[-\pi,\pi]}\left|\widehat{\sigma}^X_{ij;T}(\tau;\theta)-\sigma^X_{ij}(\tau;\theta)\right|^2\right] \le \mathcal{A}_T+\mathcal{B}_{T,r^*}+\Delta_T^2 \tag{25}$$

where

$$\mathcal{A}_T \le C_X\frac{\log T}{Tb_Th_T},\ \mathcal{B}_{T,r^*}\le C'_X\frac{T^{4/r^*}(\log T)^{4+4/r^*}}{T^2b_T^2h_T^2},\ \textit{and}\ \Delta_T^2\le C''_X\left(h_T^2+\frac{b_T^4}{h_T^2}+\frac{1}{T^2b_T^2h_T^2}\right),$$

with r^* defined in Assumption (A4).

This result is proved in the online Appendix; it is based on the Zhang and Wu (2019) approach applied to our tvGDFM. In particular, the proof relies on deriving bounds for the physical dependence of the common and idiosyncratic components, a concept introduced by Wu (2005). The terms $\mathcal{A}_T$ and $\mathcal{B}_{T,r^*}$ are due to the variance of the estimator, while Δ_T is the bias.

Let us show that the choices in Assumptions (F2) and (F3) are optimal. We have a balance between squared bias and variance if, as $T\to\infty$, either

(i) $\mathcal{A}_T \asymp \Delta_T^2$ and $\mathcal{B}_{T,r^*}=o(\mathcal{A}_T)$, or

(ii) $\mathcal{B}_{T,r^*}\asymp\Delta_T^2$ and $\mathcal{A}_T=o(\mathcal{B}_{T,r^*})$.

For simplicity of notation, define $\delta_1(r^*) := 1/3-2/3r^*$ and $\delta_2(r^*) := 1/2-2/r^*$. First, note that $\Delta_T^2 = O(\max(h_T^2, b_T^4/h_T^2))$ since $\min(h_T^2, b_T^4/h_T^2)$ is always dominated by $\mathcal{A}_T$ and $\mathcal{B}_{T,r^*}$. Second, neglecting $\log T$ quantities and powers thereof, we have

$$\begin{array}{rcccc} \mathcal{A}_T\asymp\Delta_T^2 & \text{if} & b_T\asymp T^{-1/4} & \text{and} & h_T\asymp T^{-1/4},\\ \mathcal{B}_{T,r^*}\asymp\Delta_T^2 & \text{if} & b_T\asymp T^{-\delta_1(r^*)} & \text{and} & h_T\asymp T^{-\delta_1(r^*)}. \end{array}$$

Thus, in both cases, we need $h_T\asymp b_T$ in agreement with Assumption (F3). Third, letting $h_T\asymp b_T$, we have, neglecting logarithmic in T quantities and powers thereof, $\mathcal{A}_T\asymp\mathcal{B}_{T,r^*}$ if $b_T\asymp T^{-\delta_2(r^*)}$.

Now, consider first the case $4<r^*<8$ and note that $\delta_2(r^*)<\delta_1(r^*)<1/4$. Then, the optimal bandwidth choice is $b_T\asymp T^{-\delta_1(r^*)}$ since, in this case, $\mathcal{B}_{T,r^*}\asymp\Delta_T^2$ and $\mathcal{A}_T$ is $o(\mathcal{B}_{T,r^*})$, while if $b_T\asymp T^{-1/4}$, we have $\mathcal{A}_T\asymp\Delta_T^2$ but still $\mathcal{A}_T=o(\mathcal{B}_{T,r^*})$, so that the term that dominates is $\mathcal{B}_{T,r^*}$, which is larger than the squared bias. In case $r^*>8$, we have instead $1/4<\delta_1(r^*)<\delta_2(r^*)$ and the optimal bandwidth choice is $b_T\asymp T^{-1/4}$ since now $\mathcal{A}_T\asymp\Delta_T^2$ and $\mathcal{B}_{T,r^*}=o(\mathcal{A}_T)$, while if $b_T\asymp T^{-\delta_1(r^*)}$, we have $\mathcal{B}_{T,r^*}\asymp\Delta_T^2$ but $\mathcal{B}_{T,r^*}$ still is $o(\mathcal{A}_T)$, so that the term that dominates is $\mathcal{A}_T$, which is larger than the squared bias.

Finally, notice that if $r^* = 8$, setting $b_T \asymp T^{-1/4}$ and $h_T \asymp T^{-1/4}$ implies $\mathcal{A}_T \asymp \mathcal{B}_{T,r^*} \asymp \Delta_T^2$. The previous discussion shows that Assumption (F2) is optimal.

Furthermore, for the bandwidth choices in Assumptions (F2) and (F3), the convergence rate in Lemma 4 reduces to

$$\zeta_{T,r^*} := \begin{cases} T^{-2/3+4/(3r^*)}(\log T)^{4+4/r^*} & \text{if} \quad 4 < r^* < 8, \\ T^{-1/2}\log T, & \text{if} \quad r^* \geq 8. \end{cases} \tag{26}$$

In particular, if the data is Gaussian, i.e. $r^* = \infty$, still neglecting $\log T$ quantities and their powers, $\zeta_{T,r^*} = T^{-1/2}$, implying, by Chebychev's inequality, the uniform (in τ and θ) consistency, with rate $T^{1/4}$, of the estimators $\widehat{\sigma}^X_{ij;T}(\tau;\theta)$.

We conclude the discussion of Lemma 4 with three comments.

First, comparing our results with those of Dahlhaus (2009, Example 4.2) on the smoothed pre-periodogram (already suggested by Neumann and von Sachs (1997)) or those in Dahlhaus (1996, Theorem 2.2) and Dahlhaus (2012, Theorem 4.7) on the smoothed segmented periodogram, we conjecture that it is possible to achieve convergence at rate $T^{1/3}$, which would be faster than ours. However, this would require replacing the triangular kernel $K(\cdot)$ by a smoother one, as is well-known already in the spectral estimation of stationary time series. Moreover, the aforementioned results are derived under the assumption that all moments of the innovations of the MA(∞)-representations are finite ($r^* = \infty$). Finally, we preferred to work with the same kernels as in Zhang and Wu (2019), who derive rates (*uniform* in time and frequency) for the mean-squared-errror convergence of the spectral density estimator in the given set-up, whereas it is not clear how such uniform rates can be derived from the results by Dahlhaus (2009).

Second, as in the Dahlhaus approach, our results make use of uniformly bounded second derivatives of the autocovariances with respect to time, which is possible in view of Assumptions (B4) and (B7). We still could achieve consistency by just assuming uniformly bounded first derivatives instead, which implies uniform Lipschitz continuity. However, in this case, we would have $\Delta_T^2 = O(\max(b_T^2/h_T^2, h_T^2))$ and, following a reasoning similar to the one made above, we could show that, in the $r^* = \infty$ case, the optimal bandwidths should be (neglecting again logarithmic quantities in T and their powers) of order $b_T \asymp T^{-2/5}$ and $h_T \asymp T^{-1/5}$, implying $\zeta_{T,r^*} = T^{-2/5}$ and, therefore, the consistency rate $T^{1/5}$.

Third, in the stationary case, we can just set $b_T = 1/2$, so that $M_T = T$ and, from Lemma 4, up to $\log T$ quantities and their powers, obtain $\mathcal{A}_T = O((Th_T)^{-1}) = O(m_T/T)$. Moreover, we would have $\Delta_T^2 = O(h_T^2)$ and, in the $r^* = \infty$ case, the optimal bandwidth choice would be $h_T \asymp T^{-1/3}$, yielding, as expected, a faster convergence of the mean-squared-errorwith rate $T^{2/3}$, implying the consistency rate $T^{1/3}$. This result coincides, still up to $\log T$ quantities, with the corresponding result in Wu and Zaffaroni (2018) (see also Proposition 6 in Forni et al., 2017).

In the following proposition, we consider consistency in terms of the estimation of the coefficients $c^*_{ijk}(t)$.

Proposition 1. *Let Assumptions (A)-(F) hold with, in Assumptions (A4) and (A5),* $r^* > 4$ *and* $\varphi \in (0, 2]$ *. Define*

$$H(n, T, r^*, \varphi) := \max\left(\zeta_{T,r^*}^{1/2}, n^{-1/2}\right) \log^{1/\varphi} T$$

and assume that $n = O(T^\omega)$ *for some* $\omega > 0$*. Then, for any* $\varepsilon > 0$ *there exist* $\eta(\varepsilon)$, $T^* = T^*(\varepsilon)$ *and* $N^* = N^*(\varepsilon)$*, all independent of* i *and* j*, such that, for all* $n \geq N^*$ *and* $T \geq T^*$*, and for* *any given $k \geq 0$, there exists a sequence $\{s_j(t)\}_{j=1}^q$, with $s_j(t) = \pm 1$, for which*

$$\mathrm{P}\left(\max_{M_T/2 \leq t \leq (T-M_T/2)} \frac{\left|\widehat{c}_{ijk;n,T}(t) - s_j(t)c^*_{ijk}(t)\right|}{H(n,T,r^*,\varphi)} \geq \eta(\varepsilon)\right) < \varepsilon$$

for all $1 \leq i \leq n$ and $1 \leq j \leq q$.

This consistency result is proved in the online Appendix and justifies, for large n and T, the analysis of connectedness conducted in Section 4 on the basis of $\widehat{\mathbf{C}}^*_{n,T}(t;z)$. Three final comments are in order.

First, due to the identification issue mentioned is Section 3.1, consistency holds up to post-multiplication by a sign. While that issue can be fixed by means of identification constraints, we do not resolve it here, since the sign indetermination does not affect our empirical results in Section 4.

Second, by Assumptions (F2) and (F3), $H(n,T,r^*,\varphi) \to 0$, as $n,T \to \infty$. In particular, in the Gaussian case, (up to powers of $\log T$ quantities) we have convergence with a rate $\max(T^{-1/4}, n^{-1/2})$, to be compared with the rate $\max(T^{-1/3}, n^{-1/2})$ in the stationary case, which directly follows from Proposition 10 in Forni et al. (2017). Let us stress here that the adjusted intra-day log range observations we are considering in Section 4 are well approximated by Gaussian variables (see e.g. Alizadeh et al., 2002).

Third, due to the two-sided kernel used for smoothing in time, the above result only holds for the central $(T - M_T + 1)$ observations, not for the beginning nor the end of the sample. A consequence is that we only do recover the impulse response functions for the same central values of t.

4 An analysis of time-varying financial connectedness

4.1 Financial connectedness and the tvGDFM

The use of a tvGDFM in a study of time-varying financial connectedness is motivated by two stylized facts we observe for the adjusted intra-day log ranges under study.

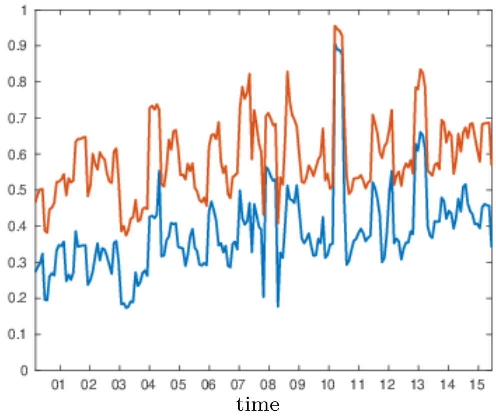
(1) *Co-movements.* Figure 1 reports, as a function of time, the proportions of variance accounted for by the first factor (blue line) and the first three factors (red line), respectively, computed from the tvGDFM as described in Sections 2 and 3. Both proportions exhibit a visible evolution in time.

(2) *Time-variation.* Figure 2 reports rolling estimates of the 329×329 sample covariance matrix of these log ranges computed at selected dates. The time-variation in the magnitude of covariances very clearly appears and indicates increased interdependencies during the financial crisis in 2008.

Common factors in the tvGDFM can be considered as "market-wide" and generate most of the dynamic interdependencies across the observed log ranges which are the focus in this section. The dynamic specification of the loadings in our GDFM is particularly useful in this context since filters naturally induce measures of connectedness at different time scales obtained from the impulse response functions of the observed data to the factors, and the implied variance decomposition.

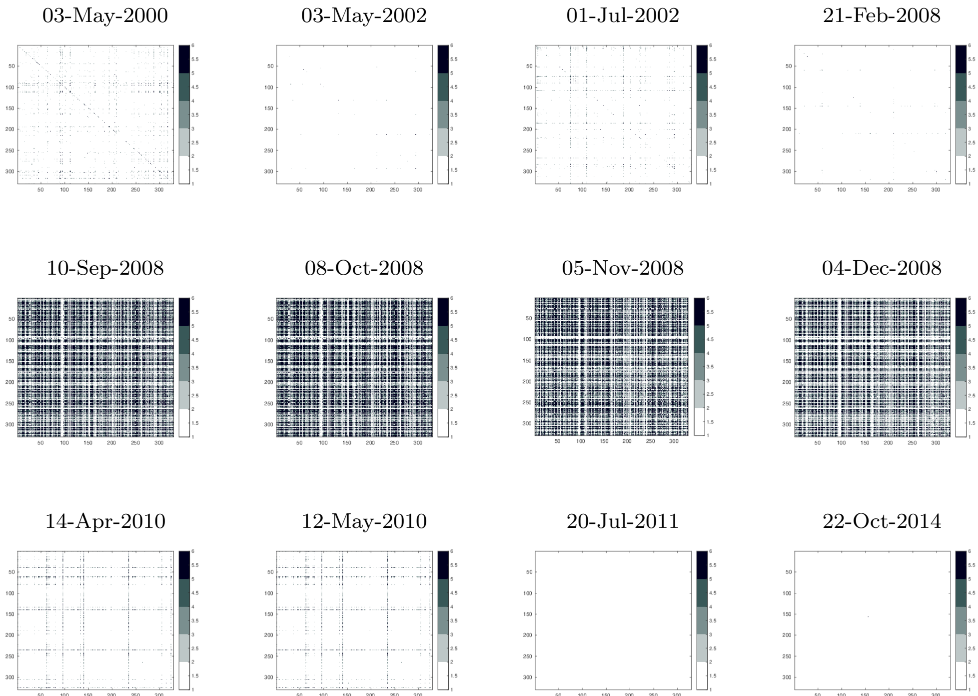
In our approach, motivated by the fact that cross-dependencies in the observed data are predominantly driven by common factors, the connectedness of each variable lies in its

Figure 1: Share of variance explained by common factors.



Standard & Poor's 500 from December 31, 1999 to August 31, 2015. Evolution in time of the estimated shares of variance accounted for by the first factor (blue line) and the first three factors (red line). The share of variance explained at time t by the first k factors is defined as $n^{-1}\sum_{\ell=1}^{k}\sum_{j=-5}^{5}\widehat{\lambda}_{\ell;n,T}^{X}(t/T;\theta_j)$, $\theta_j = j\pi/5$ where $\widehat{\lambda}_{\ell;n,T}^{X}(t/T;\theta_j)$ is the ℓth largest (in decreasing order of magnitude) eigenvalue of the estimated time-varying spectral density matrix of intra-day adjusted log ranges (see (17) for details).

Figure 2: Covariance matrix at selected dates.



Heatmaps of sample covariance matrices of the intra daily adjusted log ranges estimated at selected dates using a window of 22 days (see (16) for details). Same scale in each plot.

degree of commonality and is undirectional. In line with previous works on systemic risk, the connectedness measure associated with each variable in the dataset indeed consists of its own contribution to the total connectedness of the whole system. Note that, even though *systemic risk* is not uniquely defined (as reviewed by Benoit et al. (2017)), the standard approach to its quantitative analysis is essentially a measurement of comovements. Acharya et al. (2017), extending previous works of Acharya et al. (2012) and Brownlees and Engle (2017), consider individual capitalization with respect to that of the market. Similarly, Adrian and Brunnermeier (2016) measure the conditional effect of deviations from median value-at-risk on the system value-at-risk.

This application is mainly related to three strands of the financial econometrics literature. First, earlier works have also considered a multi-scale approach in systemic risk analysis (Bandi and Tamoni, 2017) in close relation to fields like asset pricing (Balke and Wohar, 2002; Ortu et al., 2013; Dew-Becker and Giglio, 2016), risk management (Engle, 2010), investment, employment, and R&D (Barrero et al., 2017). Second, two extensions of the Diebold and Yilmaz (2014) have been proposed quite recently: Baruník and Křehlík (2018) consider a frequency-domain decomposition of connectedness matrices in a low-dimensional stationary setting; Korobilis and Yilmaz (2018) consider time-varying estimation of high-dimensional connectedness matrices by means of Bayesian shrinkage, based, however, on a restrictive parametric VAR model in which connectedness is directional between any two given variables. Third, following Diebold and Yilmaz (2014) who strongly recommend the adoption of network measures in econometric models for financial connectedness, our approach also applies to network analysis of time series: see Acemoglu et al. (2010); Billio et al. (2012); Allen et al. (2012); Barigozzi and Hallin (2017); Barigozzi and Brownlees (2019), to quote only a few.

The main findings of our empirical analysis of the Standard & Poor's 500 adjusted intra-day log ranges in the next sections are the following:

(a) connectedness is much stronger at mid to low frequencies;
(b) large increases in long-run connectedness are associated with, and often anticipate, the main financial downturns;
(c) the largest spike in long-run connectedness associated with the great crisis of 2007-2009 is much amplified in banks, firms in related financial sectors, and real estate;
(d) during periods of crisis, the factors tend to affect all stocks contemporaneously, while during calm periods we find evidence of lead-lag relations between the stocks and the market.

4.2 Data and model specification

We apply the tvGSDM methodology developed in the previous sections to an analysis of the daily volatility of stocks which have been constituents of the Standard & Poor's 500 from December 31, 1999 to August 31, 2015. In order to do so, we retain the daily maximum and minimum prices of $n = 329$ stocks observed over a sample of $T = 3939$ daily observations. Specifically, as log-volatility proxy, we consider the intra-day adjusted log ranges for each of those 329 stocks, defined as (Parkinson, 1980)

$$X_{it} := \frac{(p_{it,\text{high}} - p_{it,\text{low}})^2}{4 \log 2}, \tag{27}$$

where $p_{it,\text{high}}$ and $p_{it,\text{low}}$ are the maximum and minimum log prices, respectively, of the i-th stock on day t. On the resulting panel $\{\mathbf{X}_{nt} = (X_{1t}, \dots, X_{nt})' | \, 1 \leq t \leq T\}$, we fit the tvGDFM studied in Sections 2 and 3, then compute the connectedness measures as described in Section 4.2.

The spectral density matrix $\mathbf{\Sigma}^X$ is estimated as in (17) with the kernels $\mathrm{J}(\cdot)$ and $\mathrm{K}(\cdot)$ given in (24). This yields

$$\widehat{\mathbf{\Sigma}}^X_{n,T}(t/T;\theta_j) = \frac{1}{2\pi M_T}\sum_{\ell=-m_T}^{m_T}\left(1-\frac{|\ell|}{m_T}\right)\sum_{s=t-M_T/2+1+\ell}^{t+M_T/2}\mathbf{X}_{n,s-\ell}\mathbf{X}'_{ns}e^{-\iota\ell\theta_j}, \tag{28}$$

where $M_T/2 \le t \le (T-M_T/2)$, $\theta_j = \pi j/m_T$, and $|j| \le m_T$. In particular, we set $M_T = 22$, corresponding to a month of trading days and $m_T = 5$ corresponding to one week of trading.

The estimated impulse response functions (23) then are computed by truncating the infinite sum at lag $k_{\max}$; all connectedness measures defined in this section thus are computed from the matrix $\widehat{\mathbf{C}}_{n,T}(t;L)$ with entries

$$\widehat{c}_{ij;n,T}(t;L) = \sum_{k=0}^{k_{\max}}\widehat{c}_{ijk;n,T}(t)L^k, \quad 1\le i\le n, \quad 1\le j\le q. \tag{29}$$

In the sequel, we set $k_{\max} = 20$.

Once the spectrum is estimated, we need to determine the number of factors. The number q is estimated by applying the criterion of Hallin and Liška (2007) to the local estimate of the spectral density matrix defined in (28). Estimation at various points in time (various values of t) supports the evidence that $q = 3$ throughout the observation period, hence is compatible with the assumption made of a "constant q".[8]

Finally, as explained at the end of Section 3.2, in order to avoid the finite-sample dependence of the results on the cross-sectional ordering, we average the estimated impulse response computed from 100 random permutations of the observed cross-sectional units. Hereafter, for simplicity of notation, we do not always indicate explicitly the dependence of the estimators on T and/or n.

4.3 Connectedness

Our connectedness measurements, in analogy with Diebold and Yilmaz (2014), are based on the $n \times n$ estimated matrices

$$\widehat{\mathbf{Q}}_n(t;z) := \widehat{\mathbf{C}}_n(t;z)\widehat{\mathbf{C}}'_n(t;z), \qquad M_T/2 \le t \le (T-M_T/2), \quad z\in\mathbb{C}. \tag{30}$$

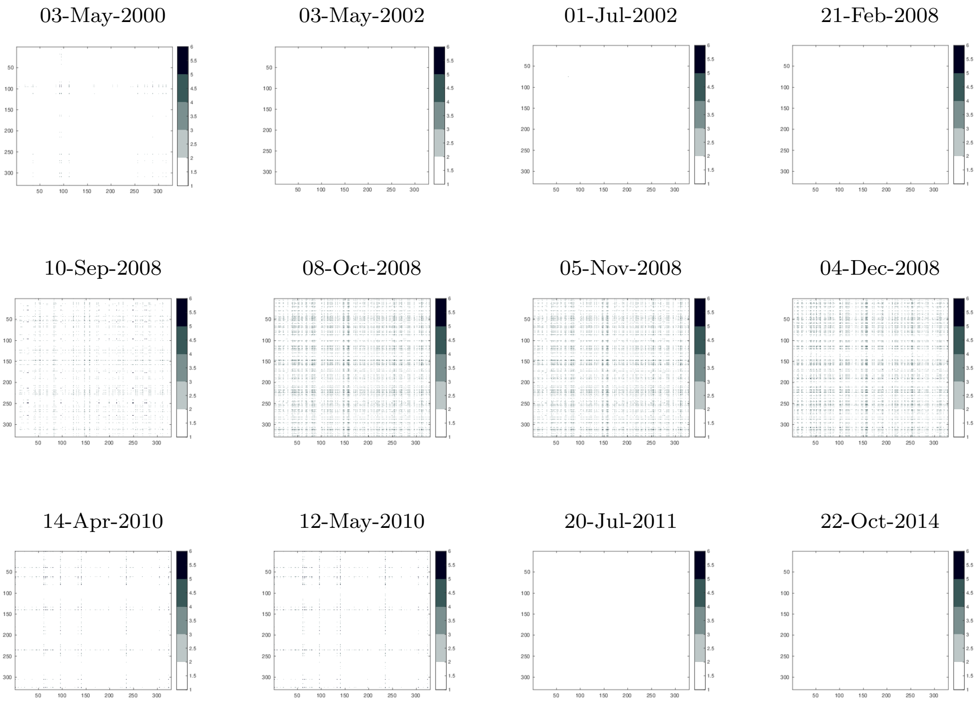
Note that, due to its quadratic nature, $\widehat{\mathbf{Q}}_n(t;z)$ is not impacted by the sign indeterminacy in the estimation of $\widehat{\mathbf{C}}_n(t;z)$. Since $\widehat{\mathbf{C}}_n(t;z)$ represents the dynamic impact of the common "market-wide" shocks, considering (30) at different horizons yields connectedness measurements at different horizons; namely,

(a) *a long-run connectedness matrix* at time t measured as $\widehat{\mathbf{Q}}_n(t;1)$ which is generated by the long-run effects of the market shocks;

(b) *an instantaneous connectedness matrix* at time t measured as $\widehat{\mathbf{Q}}_n(t;0)$ which is generated by the instantaneous effects of the market shocks;

[8] Additional results under alternative penalty functions and related settings proposed by Hallin and Liška (2007) lead to the same conclusion.

Figure 3: Instantaneous connectedness at selected dates.



Heatmaps of the estimated instantaneous connectedness matrices $\widehat{\mathbf{Q}}_n\,(t;0)$ defined in (30), for selected points in time. Same scale in each plot.

(c) *spectral connectedness matrices within specific frequency bands* $\Theta \subset [0, \pi]$

$$\widehat{\mathbf{Q}}_n\,(t;\Theta) := \frac{1}{|\Theta|} \sum_{j:\theta_j \in \Theta} \widehat{\mathbf{Q}}_n\left(t; e^{-\iota\theta_j}\right), \qquad M_T/2 \le t \le (T - M_T/2) \quad \theta_j = \pi j/m_T, \tag{31}$$

where $|\Theta|$ stands for the size of the frequency band Θ, measuring connectedness in the components with period $2\pi/\theta_j$, for $\theta_j \in \Theta$, of the spectral representation of $\mathbf{X}_{nt}$.[9]

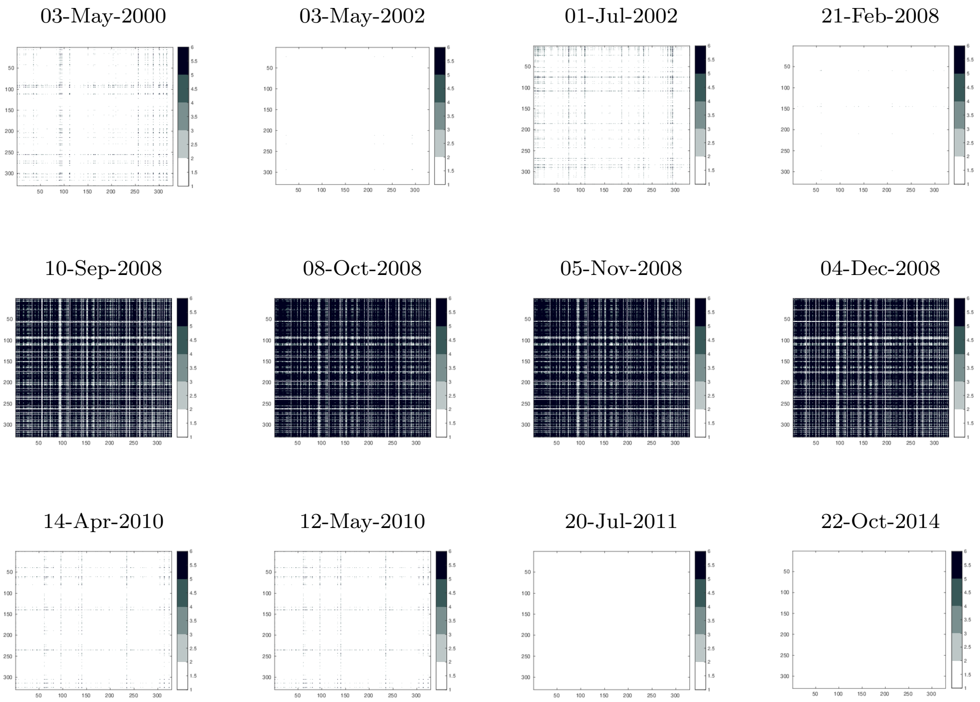
In Figures 3 and 4, we present heatmaps of the estimated long-run and instantaneous connectedness matrices $\widehat{\mathbf{Q}}_n(t;0)$ and $\widehat{\mathbf{Q}}_n(t;1)$, as defined in (30). We clearly see evidence of time-variation, with higher connectedness during the crisis periods as the great financial crisis of 2007-2008. Compared with the results based on the sample covariance matrices in Figure 2, these figures reveal that connectedness in the long-run absorbs, or arguably even amplifies, the time-variation in the data, while the short-run is relatively smoother over time.

To better appreciate the dynamics of connectedness, we can consider cross-sectional aggregation of the connectedness matrices. Figure 5 provides plots of the Frobenius norms[10] of the instantaneous and long-run connectedness matrices $\widehat{\mathbf{Q}}_n\,(t;0)$ and $\widehat{\mathbf{Q}}_n\,(t;1)$, respectively, along with the daily values of the S&P500 index. The plots reveal that both long-run and instantaneous connectedness have spikes in conjunction with important financial crashes. The turbulence at the beginning of our sample is related to a series of events starting with the burst

[9] See Theorem 11.8.2 in Brockwell and Davis (1991).

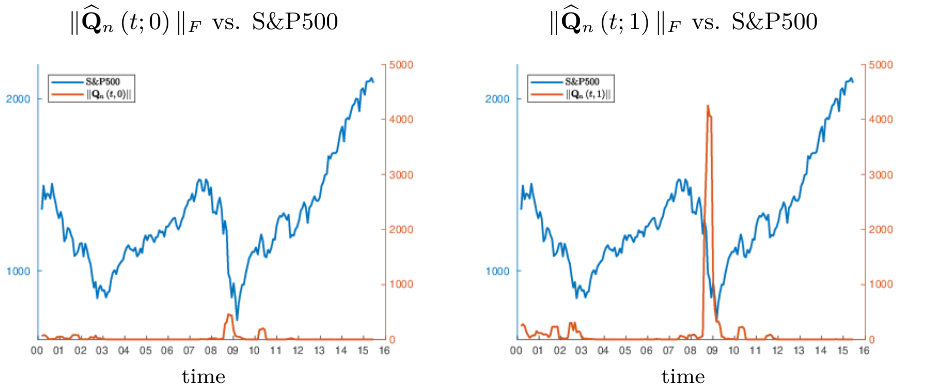
[10] Other norms would be equally suitable, and actually yield very similar results.

Figure 4: LONG-RUN CONNECTEDNESS, $\widehat{\mathbf{Q}}_n\,(t;1)$, AT SELECTED DATES.



Heatmaps of the estimated long-run connectedness matrices $\widehat{\mathbf{Q}}_n\,(t;1)$ defined in (30), for selected points in time. Same scale in each plot.

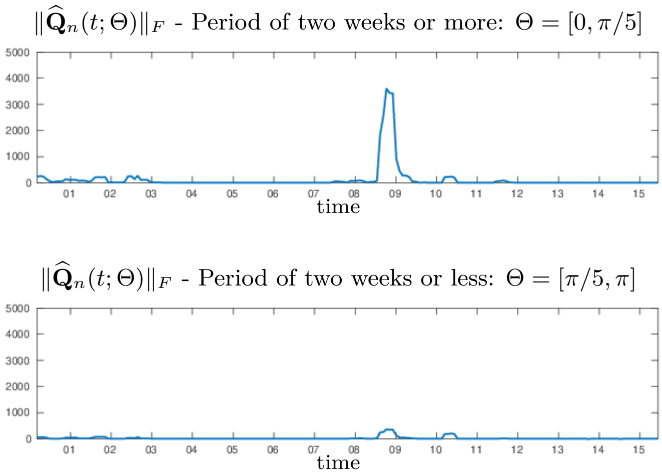
Figure 5: INSTANTANEOUS AND LONG-RUN AGGREGATE CONNECTEDNESS.



Frobenius norms (red) of the estimated instantaneous (left) and long-run (right) connectedness matrices $\widehat{\mathbf{Q}}_n(t;0)$ and $\widehat{\mathbf{Q}}_n(t;1)$ as defined in (30), plotted against time, along with the S&P 500 index (blue).

of the dot-com bubble early in 2000 and then followed by the 2002-2003 US recession and the US stock market downturn of 2002. Connectedness then stays low and stable until 2007 and the onset of the great financial crisis, yielding the maximal connectedness values recorded in the observation period. Consistently with the view that financial risk is a forward-looking concept affecting future investment strategies, Figure 5 suggests that long-run connectedness

Figure 6: Spectral connectedness.



Frobenius norm of the estimated spectral connectedness matrices $\widehat{\mathbf{Q}}_n(t;\Theta)$ as defined in (31).

is, quantitatively, the most relevant concept. Nevertheless, it should be noticed that short-run dynamics also may reveal different patterns: see, for instance, the 2010 connectedness spike which, at instantaneous level, is almost as pronounced as during the great financial crisis.

This finding is confirmed when looking at Figure 6, where we report the Frobenius norms of the spectral connectedness matrices $\widehat{\mathbf{Q}}_n\,(t;\Theta)$ for two selected frequency bands Θ, as defined in (31). In particular, we consider frequencies corresponding to cycles of period at least two weeks ($\Theta = [0, \pi/5]$, top panel) and to cycles of period less than two weeks ($\Theta = [\pi/5, 0]$, bottom panel). Since spectral connectedness are normalized by the size of the frequency band considered, their scales allow for meaningful comparisons: we observe that connectedness gets stronger and stronger as we filter out high-frequency components. When focussing on cycles of at least two weeks, we see that the norm of spectral connectedness is very similar to that of long-run connectedness in Figure 5, while it is indeed very small at high frequencies.

4.4 Sectoral connectedness

The connectedness of a specific cross-sectional item i attributable to shock j and the *mean connectedness*[11] of series i at time t are measured by

$$\widehat{Q}_{ij}\,(t;z) := \{\widehat{c}_{ij}\,(t;z)\}^2 \quad \text{and} \quad \widehat{Q}_i\,(t;z) := \frac{1}{q}\sum_{j=1}^{q}\widehat{Q}_{ij}\,(t;z)\,, \tag{32}$$

respectively; long-run, instantaneous, and spectral versions of the same concepts follow in an obvious way. By means of (32), we can evaluate connectedness within a group of cross-sectional items. Let $\mathcal{S}(\kappa)$, with cardinality n_κ, denote the set of cross-sectional indexes of the series belonging to some given sector κ. We can measure the corresponding sector-specific mean connectedness at time t as

$$\widehat{Q}_{\mathcal{S}(\kappa)}\,(t;z) := \frac{1}{n_\kappa}\sum_{i\in\mathcal{S}(\kappa)}\widehat{Q}_i\,(t;z)\,, \tag{33}$$

[11] Mean connectedness actually is what Diebold and Yilmaz (2014) call *total connectedness*.

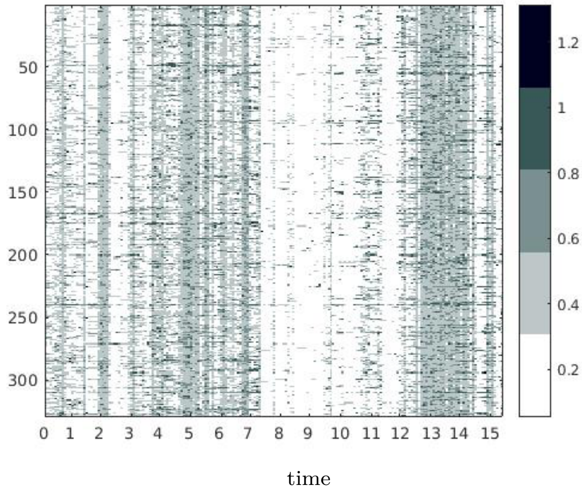
Figure 7: Difference between sectoral and overall long-run connectedness.

Evolution through time of the differences between long-run sectoral mean connectedness $\widehat{Q}_{\mathcal{S}(\mathcal{K})}(t;1)$ and the overall long-run mean connectedness $\widehat{Q}_{(n)}(t;1)$ (thick line); the reference level zero is the light grey horizontal line.

from which we can compute sector-specific long-run, instantaneous, and spectral mean connectedness.

In Figure 7, we consider the evolution over time of connectedness within the main industry-specific sectors, as defined in equation (33). Sector-specific connectedness essentially is an

Figure 8: Average absolute phase.



Evolution in time (horizontal axis) of the average absolute phases $\bar{\phi}_i^{\chi}(t/T;\theta_j)$, $1 \leq i \leq n$ (on the vertical axis), as defined in (35).

average mean connectedness within the sector. Plotting the differences between industry-specific connectedness and the panel-wide average

$$\widehat{Q}_{(n)}\left(t;z\right) := \frac{1}{n}\sum_{i=1}^{n}\widehat{Q}_i\left(t;z\right), \tag{34}$$

mean connectedness provides interesting insights into the heterogeneity of dynamics across the various sectors. It tells us, for instance, whether the long-run connectedness in any given sector comoves with the overall market long-run connectedness, exceeding, subceeding, or remaining constant with it.

During the great financial crisis of 2007-2009, the connectedness of Financial Services and Real Estate and Investment Trusts are among the largest. Similar dynamics are observed for the sectors of Industrial Metals and Mining, Industrial Engineering, Oil Equipment and Services, and Chemicals. While a number of sectors display roughly the same amount of connectedness as the market average (e.g. Food and Drug Retailers, Food Producers, General Retailers), some others display more specific dynamics, with lower than the market average connectedness during the crises: see Pharmaceuticals and Biotechnology, Software and Computer Services, Technology, Hardware and Equipment. The turmoils of the early 2000's are associated with high connectedness in some sectors which are clearly related to the dot-com bubble (Software and Computer Services, Technology Hardware and Equipment). Finally, it should be stressed that we find less heterogeneity across sectors in calm times than during financial turmoils.

4.5 Dynamic effects of common factors

As argued in the Introduction, the GDFM, unlike static factor models, does not require any particular restriction on the dynamic impacts of the factors. In this section, we analyze such dynamics over time for the panel of S&P500 intra-day adjusted log ranges. Specifically, given the estimated time-varying common spectral density matrix $\widehat{\mathbf{\Sigma}}^{\chi}_{n,T}(t/T;\theta_j)$ defined in (18), with entries $\widehat{\sigma}^{\chi}_{ij}(t/T;\theta_j)$, we consider the time-varying *phase spectrum* which is an $n \times n$ matrix with generic elements

$$\widehat{\phi}^{\chi}_{ij}(t/T;\theta_j) := \tan^{-1}\left(\frac{\mathrm{Im}(\widehat{\sigma}^{\chi}_{ij}(t/T;\theta_j))}{\mathrm{Re}(\widehat{\sigma}^{\chi}_{ij}(t/T;\theta_j))}\right), \quad M_T/2 \leq t \leq (T - M_T/2)),\ |j| \leq m_T.$$

Evaluating this quantity is most useful at frequencies where the coherence between $\widehat{\chi}_i(t/T)$ and $\widehat{\chi}_j(t/T)$—i.e. the pairwise correlation among the spectral components of the two processes at the same frequency—is high (see e.g. Granger and Hatanaka, 1964, Section 5.6). For this reason, given the results in Figure 6, we here focus only on the frequency range $\Theta = [0, \pi/5]$.

Specifically, in Figure 8, for $1 \leq i \leq n$, we report

$$\bar{\phi}^{\chi}_{i}(t/T;\theta_j) := \frac{1}{n}\sum_{j:\theta_j\in\Theta}\sum_{\ell=1}^{n}\left|\widehat{\phi}^{\chi}_{i\ell}(t/T;\theta_j)\right|, \quad M_T/2 \leq t \leq (T - M_T/2), \quad \theta_j = \pi j/m_T, \quad (35)$$

a quantity which is zero when at time t there is no lead-lag relationship between the common component $\widehat{\chi}_i(t/T)$ and all other common components of $\widehat{\boldsymbol{\chi}}_\ell(t/T)$. It is intuitively clear that, during the turmoil periods as the great financial crisis 2007-2009, all stocks tend to be in phase (white areas in the figure), thus are comoving instantaneously, while during quieter periods the effects of the common factors are more heterogeneous in time. Inspection of Figure 8 confirms this intuition.

5 Conclusions

We introduce a new time-varying version of the General Dynamic Factor Model (tvGDFM) for high-dimensional locally stationary processes in the sense of Dahlhaus (1997, 2009), thereby extending previous work on dynamic factor models (especially recent results in Forni et al., 2015, 2017) by allowing for time-varying one-sided loading filters. We propose an estimation method and, based on recent work of Zhang and Wu (2019), establish its consistency (with rates). Unlike the related approach by Eichler et al. (2011), our time-varying GDFM does allow for impulse response estimation and analysis.

This local stationarity approach, in a sense, is the opposite of the change-point setting where abrupt changes in parameters are assumed (for the case of factor models, see Barigozzi et al., 2018, and references therein). In this respect, the main advantage of a smoothly varying model is that it does not require to determine the number and exact location in time of possible change-points, while, admittedly, a disadvantage is that, in our framework, the number of factors cannot be time-varying (such variation is tantamount to an abrupt change in the factor loadings). It should be noted, however, that "small change-points" (violating the assumption of stationarity) do not necessarily preclude consistent recovery of the common-idiosyncratic decomposition; see Bates et al. (2013) for more precise results on this (in a stationary static factor model context). Intuitively, "small change-points" here is to be interpreted as "cross-sectionally non-pervasive" ones. Providing a formal statement for this is beyond the scope of this paper but we reasonably can expect that our locally stationary approach similarly is robust to the presence of such "small change-points".

Our tvGDFM is then employed in an analysis of financial connectedness along the lines of Diebold and Yilmaz (2014). The main difference between their approach and ours is that we can handle large datasets where episodes of systemic risk are typically pervasive, thus implying a factor structure, and accomodate locally stationary time series. We find that the comovements in a high-dimensional dataset of intra-day adjusted log ranges of the constituents of the Standard & Poor's 500 index are remarkably strong; this suggests basing the analysis of connectedness on (estimators of) the impulse response functions with respect to common factors or shocks, a straightforward source of systemic risk also considered (in a fixed low-dimensional context) by Billio et al. (2012). We show that large increases in connectedness, especially in their mid to low frequencies, are associated with the most important turmoils in the stock market, as the great financial crisis of 2007-2009. Moreover, the dynamic effect of factors is heterogeneous across industrial sectors and time. During crisis periods the financial and real estate sectors are the most affected, and all stocks react in a synchronous way to market shocks.

Our empirical analysis opens the way for two important empirical questions. First, since connectedness is a measure of systemic risk, its ultimate use is to predict rare financial events. Exploring such predictability requires an appropriate definition of "rare event" and the specification of the forecasting equation. Second, a structural analysis along the lines of Barigozzi et al. (2019), but based on our novel time-varying framework, would allow us to attach an economic meaning to each individual common shock and shed more light on the sources of connectedness and their propagation mechanisms. Both extensions are non-trivial, especially in a nonstationary setting, and are left for further research.

References

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2010). Cascades in Networks and Aggregate Volatility. Working Paper 16516, National Bureau of Economic Research.

Acharya, V., R. Engle, and M. Richardson (2012). Capital Shortfall: A New Approach to Ranking and Regulating Systemic Risks. *American Economic Review 102*(3), 59–64.

Acharya, V. V., L. H. Pedersen, T. Philippon, and M. Richardson (2017). Measuring Systemic Risk. *The Review of Financial Studies 30*(1), 2–47.

Adrian, T. and M. K. Brunnermeier (2016). CoVaR. *American Economic Review 106*(7), 1705–1741.

Alizadeh, S., M. W. Brandt, and F. X. Diebold (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance 57*(3), 1047–1091.

Allen, F., A. Babus, and E. Carletti (2012). Asset commonality, debt maturity and systemic risk. *Journal of Financial Economics 104*(3), 519–534.

Anderson, B. D. and M. Deistler (2008a). Generalized linear dynamic factor models-A structure theory. In *Proceedings of the 47th IEEE Conference on Decision and Control*, pp. 1980–1985.

Anderson, B. D. and M. Deistler (2008b). Properties of zero-free transfer function matrices. *SICE Journal of Control, Measurement and System Integration 1*, 284–292.

Balke, N. S. and M. E. Wohar (2002). Low-Frequency Movements in Stock Prices: A State-Space Decomposition. *The Review of Economics and Statistics 84*(4), 649–667.

Bandi, F. M. and A. Tamoni (2017). The Horizon of Systematic Risk: A New Beta Representation. SSRN Scholarly Paper ID 2337973, Rochester, NY.

Barigozzi, M. and C. Brownlees (2019). NETS: Network estimation for time series. *Journal of Applied Econometrics 34*(3), 347–364.

Barigozzi, M., H. Cho, and P. Fryzlewicz (2018). Simultaneous multiple change-point and factor analysis for high-dimensional time series. *Journal of Econometrics 206*, 187–225.

Barigozzi, M. and M. Hallin (2017). A network analysis of the volatility of high dimensional financial series. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 66*(3), 581–605.

Barigozzi, M., M. Hallin, and S. Soccorsi (2019). Identification of Global and Local Shocks in International Financial Markets via General Dynamic Factor Models. *Journal of Financial Econometrics 17*(3), 462–494.

Barrero, J. M., N. Bloom, and I. Wright (2017). Short and Long Run Uncertainty. Working Paper 23676, National Bureau of Economic Research.

Baruník, J. and T. Křehlík (2018). Measuring the Frequency Dynamics of Financial Connectedness and Systemic Risk. *Journal of Financial Econometrics 16*(2), 271–296.

Bates, B. J., M. Plagborg-Møller, J. H. Stock, and M. W. Watson (2013). Consistent factor estimation in dynamic factor models with structural instability. *Journal of Econometrics 177*(2), 289–304.

Benoit, S., J.-E. Colliard, C. Hurlin, and C. Pérignon (2017). Where the Risks Lie: A Survey on Systemic Risk. *Review of Finance 21*(1), 109–152.

Billio, M., M. Getmansky, A. W. Lo, and L. Pelizzon (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics 104*(3), 535–559.

Brockwell, P. J. and R. A. Davis (1991). *Time Series: Theory and Methods* (2 ed.). Springer Series in Statistics. New York: Springer-Verlag.

Brownlees, C. and R. F. Engle (2017). SRISK: A Conditional Capital Shortfall Measure of Systemic Risk. *The Review of Financial Studies 30*(1), 48–79.

Brownlees, C. T. and G. M. Gallo (2010). Comparison of Volatility Measures: a Risk Management Perspective. *Journal of Financial Econometrics 8*(1), 29–56.

Dahlhaus, R. (1996). Asymptotic statistical inference for nonstationary processes with evolutionary spectra. In P. Robinson and M. Rosenblatt (Eds.), *Athens Conference on Applied Probability and Time Series Analysis; Lecture Notes in Statistics*, Volume 115, pp. 145–159. Springer.

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *The Annals of Statistics 25*(1), 1–37.

Dahlhaus, R. (2009). Local inference for locally stationary time series based on the empirical spectral measure. *Journal of Econometrics 151*(2), 101–112.

Dahlhaus, R. (2012). Locally stationary processes. In T. Subba Rao, S. Subba Rao, and C. R. Rao (Eds.), *Handbook of Statistics: Time Series Analysis, Methods and Applications*, Volume 30, pp. 351–413. Elsevier.

Dew-Becker, I. and S. Giglio (2016). Asset Pricing in the Frequency Domain: Theory and Empirics. *The Review of Financial Studies 29*(8), 2029–2068.

Diebold, F. X. and K. Yilmaz (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics 182*(1), 119–134.

Eichler, M., G. Motta, and R. von Sachs (2011). Fitting dynamic factor models to non-stationary time series. *Journal of Econometrics 163*(1), 51–70.

Engle, R. F. (2010). The Risk that Risk Will Change. SSRN Scholarly Paper ID 1539166, Rochester, NY.

Forni, M., D. Giannone, M. Lippi, and L. Reichlin (2009). Opening the black box: Structural factor models with large cross sections. *Econometric Theory 25*(5), 1319–1347.

Forni, M., A. Giovannelli, M. Lippi, and S. Soccorsi (2018). Dynamic factor model with infinite-dimensional factor space: Forecasting. *Journal of Applied Econometrics 33*(5), 625–642.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). The generalized dynamic-factor model: Identification and estimation. *Review of Economics and Statistics 82*(4), 540–554.

Forni, M., M. Hallin, M. Lippi, and P. Zaffaroni (2015). Dynamic factor models with infinite-dimensional factor spaces: One-sided representations. *Journal of Econometrics 185*(2), 359–371.

Forni, M., M. Hallin, M. Lippi, and P. Zaffaroni (2017). Dynamic factor models with infinite-dimensional factor spaces: Asymptotic analysis. *Journal of Econometrics 199*(1), 74–92.

Granger, C. W. J. and M. Hatanaka (1964). *Spectral Analysis of Economic Time Series. (PSME-1)*. Princeton University Press.

Hallin, M. and M. Lippi (2013). Factor models in high-dimensional time series. A time-domain approach. *Stochastic Processes and their Applications 123*(7), 2678–2695.

Hallin, M. and R. Liška (2007). Determining the Number of Factors in the General Dynamic Factor Model. *Journal of the American Statistical Association 102*(478), 603–617.

Korobilis, D. and K. Yilmaz (2018). Measuring Dynamic Connectedness with Large Bayesian VAR Models. SSRN Scholarly Paper ID 3099725, Rochester, NY.

Kuchibhotla, A. K. and A. Chakrabortty (2018). Moving beyond sub-Gaussianity in high dimensional statistics: Applications in covariance matrix estimation and linear regressions. arXiv preprint arXiv:1804.02605.

Motta, G., C. M. Hafner, and R. von Sachs (2011). Locally stationary factor models: Identification and nonparametric estimation. *Econometric Theory 27*(6), 1279–1319.

Neumann, M. H. and R. von Sachs (1997). Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. *The Annals of Statistics 25*(1), 38–76.

Ortu, F., A. Tamoni, and C. Tebaldi (2013). Long-Run Risk and the Persistence of Consumption Shocks. *The Review of Financial Studies 26*(11), 2876–2915.

Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *The Journal of Business 53*(1), 61–65.

Peña, D. and V. J. Yohai (2016). Generalized dynamic principal components. *Journal of the American Statistical Association 111*(515), 1121–1131.

Rodríguez-Poo, J. M. and O. Linton (2001). Nonparametric factor analysis of residual time series. *Test 10*(1), 161–182.

Stock, J. H. and M. W. Watson (2002). Forecasting Using Principal Components From a Large Number of Predictors. *Journal of the American Statistical Association 97*(460), 1167–1179.

Stock, J. H. and M. W. Watson (2016). Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics. In J. B. Taylor and H. Uhlig (Eds.), *Handbook of Macroeconomics*, Volume 2, pp. 415–525. Elsevier.

Vershynin, R. (2018). *High dimensional probability. An introduction with applications in Data Science.* Cambridge University Press.

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

Wu, W. B. and P. Zaffaroni (2018). Asymptotic theory for spectral density estimates of general multivariate time series. *Econometric Theory 34*(1), 1–22.

Zhang, D. and W. B. Wu (2019). Convergence of covariance and spectral density estimates for high dimensional locally stationary processes. Working Paper, UCSD.

Time-Varying General Dynamic Factor Models and the Measurement of Financial Connectedness Technical Appendix

Matteo Barigozzi[1] Marc Hallin[2] Stefano Soccorsi[3] Rainer von Sachs[4]

19th April 2020

A1 Notation

Denote by $\|\boldsymbol{A}\| = \sqrt{\mu^{(1)}(\boldsymbol{A}^\dagger\boldsymbol{A})}$, where $\mu^{(1)}(\boldsymbol{A}^\dagger\boldsymbol{A})$ is the largest eigenvalue (which is always real) of $\boldsymbol{A}^\dagger\boldsymbol{A}$, the spectral norm of a given complex $p\times p$ matrix $\boldsymbol{A}$ and by $\|\boldsymbol{A}\|_F = \sqrt{\mathrm{tr}(\boldsymbol{A}^\dagger\boldsymbol{A})}$ its Frobenius norm. Similarly, write $\|\boldsymbol{v}\| = \sqrt{\sum_{i=1}^p v_i^2}$ for the Euclidean norm of a p-dimensional vector $\boldsymbol{v} = (v_1,\dots,v_p)'$

A2 Proof of Lemma 1

Let $\gamma_{ij}^\chi(\tau;\ell) = \mathrm{E}[\chi_{it}(\tau)\chi_{jt-\ell}(\tau)]$ and $\gamma_{ij}^\xi(\tau;\ell) = \mathrm{E}[\xi_{it}(\tau)\xi_{jt-\ell}(\tau)]$. By Assumptions (A1), (B1), and (B4), for any $\ell\in\mathbb{Z}$,

$$\begin{aligned}\left|\frac{\mathrm{d}^2\gamma_{ij}^\chi(\tau;\ell)}{\mathrm{d}\tau^2}\right| &= \left|\sum_{k=0}^\infty\sum_{s=1}^q \frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\left\{c_{isk}(\tau)c_{js,k+|\ell|}(\tau)\right\}\right|\\ &= \left|\sum_{k=0}^\infty\sum_{s=1}^q \Big[\frac{\mathrm{d}^2c_{isk}(\tau)}{\mathrm{d}\tau^2}c_{js,k+|\ell|}(\tau) + \frac{\mathrm{d}^2c_{js,k+|\ell|}(\tau)}{\mathrm{d}\tau^2}c_{isk}(\tau)\Big]\right|\\ &\le 2qC_2\rho_\chi^{|\ell|}\sum_{k=0}^\infty \rho_\chi^{2k}C_1 = \frac{2qC_2C_1\rho_\chi^{|\ell|}}{1-\rho_\chi^2} =: \mathcal{K}_1\rho_\chi^{|\ell|}, \text{ say.}\end{aligned}\tag{A.1}$$

Similarly, by Assumptions (A2), (B5), and (B7), for any $\ell\in\mathbb{Z}$,

$$\begin{aligned}\left|\frac{\mathrm{d}^2\gamma_{ij}^\xi(\tau;\ell)}{\mathrm{d}\tau^2}\right| &= \left|\sum_{k=0}^\infty\sum_{s=1}^\infty \frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\left\{d_{isk}(\tau)d_{js,k+|\ell|}(\tau)\right\}\right|\\ &= \left|\sum_{k=0}^\infty\sum_{s=1}^\infty \Big[\frac{\mathrm{d}^2d_{isk}(\tau)}{\mathrm{d}\tau^2}d_{js,k+|\ell|}(\tau) + \frac{\mathrm{d}^2d_{js,k+|\ell|}(\tau)}{\mathrm{d}\tau^2}d_{isk}(\tau)\Big]\right|\\ &\le 2\rho_\xi^{|\ell|}\sum_{k=0}^\infty\rho_\xi^{2k}\sum_{s=1}^\infty B_{2is}B_{1js} = \frac{2B_2B_1\rho_\xi^{|\ell|}}{1-\rho_\xi^2} =: \mathcal{K}_2\rho_\xi^{|\ell|}, \text{ say.}\end{aligned}\tag{A.2}$$

Moreover, by Assumption (A3),

$$\sigma_{ij}^X(\tau;\theta) = \sigma_{ij}^\chi(\tau;\theta) + \sigma_{ij}^\xi(\tau;\theta) = \frac{1}{2\pi}\sum_{\ell=-\infty}^\infty (\gamma_{ij}^\chi(\tau;\ell) + \gamma_{ij}^\xi(\tau;\ell))e^{-\iota\ell\theta}.\tag{A.3}$$

Therefore, from (A.1) and (A.2) and noticing that $\mathcal{K}_1$ and $\mathcal{K}_2$ are independent of i, j, ℓ, and τ, we get

$$\begin{aligned}\max_{1\le i,j\le n}\sup_{\tau\in(0,1)}\sup_{\theta\in[-\pi,\pi]}\left|\frac{\mathrm{d}^2\sigma^X_{ij}(\tau;\theta)}{\mathrm{d}\tau^2}\right| &\le \frac{1}{2\pi}\left\{\mathcal{K}_1\sum_{\ell=-\infty}^{\infty}\rho_\chi^{|\ell|}+\mathcal{K}_2\sum_{\ell=-\infty}^{\infty}\rho_\xi^{|\ell|}\right\}\\ &\le \frac{1}{2\pi}\left\{\frac{2\mathcal{K}_1}{1-\rho_\chi}+\frac{2\mathcal{K}_2}{1-\rho_\xi}\right\}=:\mathcal{K},\text{ say,}\end{aligned}\tag{A.4}$$

since $|e^{-\iota\theta\ell}|=1$ for all $\theta\in[-\pi,\pi]$ and all $\ell\in\mathbb{N}_0$. This proves part *(i)* of the lemma.

Then, because of Assumptions (A1) and (B1), for any $\ell\in\mathbb{Z}$,

$$|\gamma^\chi_{ij}(\tau;\ell)|=\left|\sum_{k=0}^{\infty}\sum_{s=1}^{q}c_{isk}(\tau)c_{js,k+|\ell|}(\tau)\right|\le\sum_{k=0}^{\infty}\sum_{s=1}^{q}C_1^2\rho_\chi^k\rho_\chi^{k+|\ell|}\le\frac{C_1^2\rho_\chi^{|\ell|}}{1-\rho_\chi^2}=:\mathcal{K}_3\rho_\chi^{|\ell|},\ \text{ say.}\tag{A.5}$$

Similarly, because of Assumptions (A2) and (B5), for any $h\in\mathbb{Z}$,

$$|\gamma^\xi_{ij}(\tau;\ell)|=\left|\sum_{k=0}^{\infty}\sum_{s=1}^{\infty}d_{isk}(\tau)d_{js,k+|\ell|}(\tau)\right|\le\sum_{k=0}^{\infty}\sum_{s=1}^{\infty}B_{1is}B_{1js}\rho_\xi^k\rho_\xi^{k+|\ell|}\le\frac{B_1^2\rho_\xi^{|\ell|}}{1-\rho_\xi^2}=:\mathcal{K}_4\rho_\xi^{|\ell|},\text{ say.}\tag{A.6}$$

Therefore, using again (A.3), from (A.5) and (A.6) and noticing that $\mathcal{K}_3$ and $\mathcal{K}_4$ are independent of i, j, ℓ, and τ, we get

$$\begin{aligned}\max_{1\le i,j\le n}\sup_{\tau\in(0,1)}\sup_{\theta\in[-\pi,\pi]}\left|\frac{\mathrm{d}^2\sigma^X_{ij}(\tau;\theta)}{\mathrm{d}\theta^2}\right| &\le \frac{1}{2\pi}\left\{\mathcal{K}_3\sum_{\ell=-\infty}^{\infty}|\ell|^2\rho_\chi^{|\ell|}+\mathcal{K}_4\sum_{\ell=-\infty}^{\infty}|\ell|^2\rho_\xi^{|\ell|}\right\}\\ &= \frac{1}{2\pi}\left\{\frac{4\mathcal{K}_3}{(1-\rho_\chi)^3}+\frac{4\mathcal{K}_4}{(1-\rho_\xi)^3}\right\}=:\mathcal{K}',\ \text{ say.}\end{aligned}\tag{A.7}$$

This proves part *(ii)* of the lemma. □

A3 Proof of Lemma 2

Denote as $\sigma^\xi_{ij}(\tau;\theta)$ the generic (i,j) entry of $\mathbf{\Sigma}^\xi_n(\tau;\theta)$. For any $n\in\mathbb{N}_0$, $\tau\in(0,1)$, and $\theta\in[-\pi,\pi]$, we have

$$\begin{aligned}\lambda^\xi_{1;n}(\tau;\theta)=\|\mathbf{\Sigma}^\xi_n(\tau;\theta)\|&\le\max_{1\le j\le n}\sum_{i=1}^{n}|\sigma^\xi_{ij}(\tau;\theta)|\le\frac{1}{2\pi}\max_{1\le j\le n}\sum_{i,\ell=1}^{n}|d_{i\ell}(\tau;e^{-\iota\theta})d^\dagger_{\ell j}(\tau;e^{-\iota\theta})|\\ &\le\frac{1}{2\pi}\max_{1\le j\le n}\sum_{i,\ell=1}^{n}\sum_{k=0}^{\infty}|d_{i\ell k}(\tau)e^{-\iota\theta k}|\sum_{h=0}^{\infty}|d^\dagger_{\ell jh}(\tau)e^{\iota\theta h}|\le\frac{1}{2\pi}\max_{1\le j\le n}\sum_{i,\ell=1}^{n}\sum_{k=0}^{\infty}|d_{i\ell k}(\tau)|\sum_{h=0}^{\infty}|d^\dagger_{\ell jh}(\tau)|\\ &\le\frac{1}{2\pi}\max_{1\le j\le n}\sum_{i,\ell=1}^{n}\sum_{k,h=0}^{\infty}B_{1i\ell}B_{1\ell j}\rho_\xi^k\rho_\xi^h\le\frac{B_1^2}{2\pi(1-\rho_\xi)^2},\end{aligned}$$

because of Assumptions (A2) and (B5). Part *(i)* is proved by defining $B_\xi:=B_1^2/(2\pi(1-\rho_\xi)^2)$ and noting that it is independent of n, τ, and θ.

Parts *(ii)* and *(iii)* readily follow from Assumption (C) and part *(i)*, and an application of Weyl's inequality. □

A4 Proof of Lemma 3

Denote as $\lambda^\zeta_{1;n}(\tau;\theta)$ the largest eigenvalue of the spectral density of $\boldsymbol{\zeta}_{nt}(\tau)$. Then, for any $n\in\mathbb{N}_0$, $\tau\in(0,1)$, and $\theta\in[-\pi,\pi]$ (see also (15))

$$\lambda^\zeta_{1;n}(\tau;\theta)=\max_{\boldsymbol{a}\,:\,\boldsymbol{a}^\dagger\boldsymbol{a}=1}\boldsymbol{a}^\dagger\mathbf{A}_n(\tau;e^{-\iota\theta})\mathbf{\Sigma}^\xi_n(\tau;\theta)\mathbf{A}'_n(\tau;e^{\iota\theta})\boldsymbol{a}\le\lambda^\xi_{1;n}(\tau;\theta)\lambda^A_{1;n}(\tau;\theta)\tag{A.8}$$

where $\lambda^A_{1;n}(\tau;\theta)$ is the largest eigenvalue of $\mathbf{A}_n(\tau;e^{-\iota\theta})\mathbf{A}'_n(\tau;e^{\iota\theta})$. Moreover, denoting by $\lambda_1^{A^{(k)}}(\tau;\theta)$ the largest eigenvalue of $\mathbf{A}^{(k)}(\tau;e^{-\iota\theta})\mathbf{A}^{(k)\prime}(\tau;e^{\iota\theta})$ and recalling that $\mathbf{A}_n(\tau;L)$ is block-diagonal with diagonal blocks $\mathbf{A}^{(1)}(\tau;L),\dots,\mathbf{A}^{(m)}(\tau;L)$, we have

$$\lambda^A_{1;n}(\tau;\theta) \le \max_{1\le k\le n} \lambda_1^{A^{(k)}}(\tau;\theta) \le D_\zeta \tag{A.9}$$

where D_ζ is a constant independent of n, τ, and θ, because of Assumptions (D3) and (D4). By using Lemma 2 and (A.9) in (A.8), we have $\lambda^\zeta_{1;n}(\tau;\theta) \le B_\xi D_\zeta$. Therefore,

$$\mu^\zeta_{1;n}(\tau) = \max_{\boldsymbol{w}:\boldsymbol{w}'\boldsymbol{w}=1} \boldsymbol{w}'\boldsymbol{\Gamma}^\zeta_n(\tau)\boldsymbol{w} \le \int_{-\pi}^{\pi} \lambda^\zeta_{1;n}(\tau;\theta)\mathrm{d}\theta \le 2\pi B_\xi D_\zeta.$$

The proof is completed by defining $B_\zeta := 2\pi B_\xi D_\zeta$ and noting that it is independent of n, τ, and θ. □

A5 Proof of Lemma 4

The proof requires two intermediate results.

Lemma A1. *Under Assumptions (A) and (B) there exists a constant A_1 (independent of i and t) such that* $\sup_{\tau\in(0,1)} \mathrm{E}[|X_{it}(\tau)|^{r^*}] \le A_1$ *for all $i \in \mathbb{N}_0$ and $t \in \mathbb{Z}$, with r^* as defined in Assumption (A4).*

Proof of Lemma A1. By Minkowski inequality,

$$\sup_{\tau\in(0,1)} \left\{\mathrm{E}[|X_{it}(\tau)|^{r^*}]\right\}^{1/r^*} \le \sup_{\tau\in(0,1)} \left\{\mathrm{E}[|\chi_{it}(\tau)|^{r^*}]\right\}^{1/r^*} + \sup_{\tau\in(0,1)} \left\{\mathrm{E}[|\xi_{it}(\tau)|^{r^*}]\right\}^{1/r^*}. \tag{A.10}$$

Then, from (6)

$$\begin{aligned}
\sup_{\tau\in(0,1)} \left\{\mathrm{E}[|\chi_{it}(\tau)|^{r^*}]\right\}^{1/r^*} &= \sup_{\tau\in(0,1)} \left\{\mathrm{E}\left[\left|\sum_{j=1}^{q}\sum_{k=0}^{\infty} c_{ijk}(\tau)u_{j,t-k}\right|^{r^*}\right]\right\}^{1/r^*} \\
&\le \sup_{\tau\in(0,1)} \sum_{j=1}^{q}\sum_{k=0}^{\infty} |c_{ijk}(\tau)| \left\{\mathrm{E}\left[|u_{j,t-k}|^{r^*}\right]\right\}^{1/r^*} \\
&\le C_0^{1/r^*} C_1 \sum_{k=0}^{\infty} \rho_\chi^k \le \frac{C_0^{1/r^*}C_1}{1-\rho_\chi} =: A_{11}, \text{ say,}
\end{aligned} \tag{A.11}$$

because of Assumptions (A4) and (B1). Similarly, from (7),

$$\begin{aligned}
\sup_{\tau\in(0,1)} \left\{\mathrm{E}[|\xi_{it}(\tau)|^{r^*}]\right\}^{1/r^*} &= \sup_{\tau\in(0,1)} \left\{\mathrm{E}\left[\left|\sum_{j=1}^{\infty}\sum_{k=0}^{\infty} d_{ijk}(\tau)\eta_{j,t-k}\right|^{r^*}\right]\right\}^{1/r^*} \\
&\le \sup_{\tau\in(0,1)} \sum_{j=1}^{\infty}\sum_{k=0}^{\infty} |d_{ijk}(\tau)| \left\{\mathrm{E}\left[|\eta_{j,t-k}|^{r^*}\right]\right\}^{1/r^*} \\
&\le C_0^{1/r^*} \sum_{j=1}^{\infty}\sum_{k=0}^{\infty} \rho_\xi^k B_{1ij} \le \frac{C_0^{1/r^*}B_1}{1-\rho_\xi} =: A_{12}, \text{ say,}
\end{aligned} \tag{A.12}$$

because of Assumptions (A4) and (B5). Substituting (A.11) and (A.12) into (A.10) and defining $A_1 := A_{11} + A_{12}$ completes the proof. □

For all $t \in \mathbb{Z}$, let $\boldsymbol{\varepsilon} := \{\boldsymbol{\varepsilon}_t = (\mathbf{u}'_t\ \boldsymbol{\eta}'_t)'\}$, and define $\mathcal{F}_t := (\dots,\boldsymbol{\varepsilon}_{t-1},\boldsymbol{\varepsilon}_t)$. Moreover, denoting by $\boldsymbol{\varepsilon}^* = \{\boldsymbol{\varepsilon}^*_t = (\mathbf{u}^{*\prime}_t\ \boldsymbol{\eta}^{*\prime}_t)'\}$ an independent copy of $\boldsymbol{\varepsilon}$, define $\mathcal{F}^*_t := (\dots,\boldsymbol{\varepsilon}_{-1},\boldsymbol{\varepsilon}^*_0,\boldsymbol{\varepsilon}_1,\dots,\boldsymbol{\varepsilon}_{t-1},\boldsymbol{\varepsilon}_t)$, which is a version of $\mathcal{F}_t$ where $\boldsymbol{\varepsilon}_0$ is replaced with $\boldsymbol{\varepsilon}^*_0$. Note that $\mathcal{F}^*_t = \mathcal{F}_t$ if $t < 0$.

Then, from (6) and (7), it is clear that, for any $\tau \in (0,1)$, $i \in \mathbb{N}_0$, and $t \in \mathbb{Z}$, we have that $X_{it}(\tau) =: g_i(\tau; \mathcal{F}_t)$, where $g_i : (0,1) \times \mathbb{R}^\infty \to \mathbb{R}$ is a measurable function. Put $X^*_{it}(\tau) := g_i(\tau; \mathcal{F}^*_t)$. Then, for any $r > 0$, we define the physical dependence measure (see also Wu, 2005 and Zhang and Wu, 2019) as

$$\begin{aligned}\delta_{t,r,i} &= \sup_{\tau\in(0,1)} \left\{\mathrm{E}\left[(g_i(\tau;\mathcal{F}_t) - g_i(\tau;\mathcal{F}^*_t))^r\right]\right\}^{1/r} \\ &= \sup_{\tau\in(0,1)} \left\{\mathrm{E}\left[(X_{it}(\tau) - X^*_{it}(\tau))^r\right]\right\}^{1/r}, \ i \in \mathbb{N}_0, \ t \in \mathbb{Z}. \end{aligned} \tag{A.13}$$

Lemma A2. *Under Assumptions (A) and (B) there exists a $\rho \in [0,1)$ and a constant A_2 (independent of r, i, and t) such that $\delta_{t,r,i} \leq A_2\rho^t$, $i \in \mathbb{N}_0$, and $t \in \mathbb{Z}$, for all $r \leq r^*$ where r^* is defined in Assumption (A4).*

Proof of Lemma A2. First, notice that

$$\begin{aligned}X_{it}(\tau) - X^*_{it}(\tau) &= \sum_{j=1}^{q} c_{ijt}(\tau)(u_{j0} - u^*_{j0}) + \sum_{j=1}^{\infty} d_{ijt}(\tau)(\eta_{j0} - \eta^*_{j0}) \\ &=: (\chi_{it}(\tau) - \chi^*_{it}(\tau)) + (\xi_{it}(\tau) - \xi^*_{it}(\tau)), \text{ say.}\end{aligned} \tag{A.14}$$

By Minkowski inequality, it follows from (A.13) that

$$\delta_{t,r,i} \leq \sup_{\tau\in(0,1)} \left\{\mathrm{E}\left[(\chi_{it}(\tau) - \chi^*_{it}(\tau))^r\right]\right\}^{1/r} + \sup_{\tau\in(0,1)} \left\{\mathrm{E}\left[(\xi_{it}(\tau) - \xi^*_{it}(\tau))^r\right]\right\}^{1/r} =: \delta^\chi_{t,r,i} + \delta^\xi_{t,r,i}, \text{ say.} \tag{A.15}$$

Then, from (A.14), for any $r \leq r^*$,

$$\begin{aligned}\delta^\chi_{t,r,i} &= \sup_{\tau\in(0,1)} \left\{\mathrm{E}\left[\left|\sum_{j=1}^{q} c_{ijt}(\tau)(u_{j0} - u^*_{j0})\right|^r\right]\right\}^{1/r} \\ &\leq \sup_{\tau\in(0,1)} \sum_{j=1}^{q} |c_{ijt}(\tau)| \left\{\mathrm{E}\left[\left|(u_{j0} - u^*_{j0})\right|^r\right]\right\}^{1/r} \leq 2C_0^{1/r} C_1 \rho^t_\chi,\end{aligned} \tag{A.16}$$

because of Assumptions (A4) and (B1), and Minkowski inequality. Similarly, still from (A.14), for any $r \leq r^*$,

$$\begin{aligned}\delta^\xi_{t,r,i} &= \sup_{\tau\in(0,1)} \left\{\mathrm{E}\left[\left|\sum_{j=1}^{\infty} d_{ijt}(\tau)(\eta_{j0} - \eta^*_{j0})\right|^r\right]\right\}^{1/r} \\ &\leq \sup_{\tau\in(0,1)} \sum_{j=1}^{\infty} |d_{ijt}(\tau)| \left\{\mathrm{E}\left[\left|(\eta_{j0} - \eta^*_{j0})\right|^r\right]\right\}^{1/r} \leq 2C_0^{1/r} \sum_{j=1}^{\infty} B_{1ij}\rho^t_\xi \leq 2C_0^{1/r} B_1 \rho^t_\xi,\end{aligned} \tag{A.17}$$

because of Assumptions (A4) and (B5), and Minkowski inequality. Substituting (A.16) and (A.17) into (A.15) and defining $\rho := \max(\rho_\chi, \rho_\xi)$ and $A_2 := 2C_0^{1/r}(C_1 + B_1)$ completes the proof. □

For any $n \in \mathbb{N}_0$, the mean-squared-error satisfies

$$\max_{1\leq i,j\leq n} \mathrm{E}\left[\sup_{\tau\in(0,1)} \sup_{\theta\in[-\pi,\pi]} \left|\widehat{\sigma}^X_{ij;T}(\tau;\theta) - \sigma^X_{ij}(\tau;\theta)\right|^2\right] \leq 2\left(\mathcal{V}_T + \Delta^2_T\right), \tag{A.18}$$

where

$$\mathcal{V}_T := \max_{1\leq i,j\leq n} \mathrm{E}\left[\sup_{\tau\in(0,1)} \sup_{\theta\in[-\pi,\pi]} \left|\widehat{\sigma}^X_{ij;T}(\tau;\theta) - \mathrm{E}[\widehat{\sigma}^X_{ij;T}(\tau;\theta)]\right|^2\right] \quad \text{and} \tag{A.19}$$

$$\Delta^2_T := \max_{1\leq i,j\leq n} \sup_{\tau\in(0,1)} \sup_{\theta\in[-\pi,\pi]} \left\{\mathrm{E}[\widehat{\sigma}^X_{ij;T}(\tau;\theta)] - \sigma^X_{ij}(\tau;\theta)\right\}^2. \tag{A.20}$$

First, let us consider (A.19). For any $\alpha > 0$ and $r > 0$, define the dependent adjusted norm

$$\Phi_{r,\alpha} := \max_{1\le i\le n} \sup_{k\in\mathbb{N}} (k+1)^{\alpha} \sum_{t=k}^{\infty} \delta_{t,r,i}$$

(see Zhang and Wu, 2019). For any $r \le r^*$,

$$\Phi_{r,\alpha} \le A_2 \sup_{k\in\mathbb{N}} (k+1)^{\alpha} \sum_{t=k}^{\infty} \rho^t = \frac{A_2}{1-\rho} \sup_{k\in\mathbb{N}} (k+1)^{\alpha} \rho^k = \frac{A_2}{1-\rho}, \tag{A.21}$$

because of Lemma A2. Moreover, because of Lemma A1, we can apply Corollary 4.4 in Zhang and Wu (2019) in the case $\alpha > 1/2 - 2/r^*$, which, along with (A.21), implies

$$\begin{aligned}\mathcal{V}_T &\le \Phi^4_{r,\alpha}\left(\mathcal{K}_1 \frac{m_T \log T}{M_T} + \mathcal{K}_2 \frac{m_T^2 T^{4/r} (\log M_T)^{4+4/r}}{M_T^2}\right)\\ &\le \left(\frac{A_2}{1-\rho}\right)^4 \left(\mathcal{K}_1 \frac{m_T \log T}{M_T} + \mathcal{K}_2 \frac{m_T^2 T^{4/r} (\log M_T)^{4+4/r}}{M_T^2}\right),\end{aligned} \tag{A.22}$$

for some constants $\mathcal{K}_1$ and $\mathcal{K}_2$ (independent of T) and any $r \le r^*$. Recalling that $M_T = \lfloor T b_T \rfloor$ and $m_T = \lfloor 1/h_T \rfloor$, we thus obtain, for any $r \le r^*$,

$$\mathcal{V}_T \le C_X \frac{\log T}{T b_T h_T} + C'_X \frac{T^{4/r} (\log T)^{4+4/r}}{T^2 b_T^2 h_T^2} = \mathcal{A}_T + \mathcal{B}_{T,r} \tag{A.23}$$

for some constants C_X and C'_X (independent of T and n).

The following Lemma which is similar to Theorem 2.1 in Dahlhaus (2012) (see also Dahlhaus, 1996, Theorem 2.1), is needed to bound the bias.

Lemma A3. *Under Assumptions (A), (B), and (F), there exists constants $\mathcal{M}$, $\mathcal{L}$, and $\mathcal{H}$ (independent of i, j, ℓ and τ), such that*

$$\max_{1\le i,j\le n} \sup_{\tau\in(0,1)} \max_{|\ell|\le(M_T-1)} \left| \mathrm{E}[\widehat{\gamma}^X_{ij;T}(\tau;\ell)] - \gamma^X_{ij}(\tau;\ell) \right| \le \mathcal{M} b_T^2 + o(b_T^2) + \frac{\mathcal{L}}{T b_T} + \frac{\mathcal{H}}{T}. \tag{A.24}$$

Proof of Lemma A3. Consider (A.20). Let $\gamma^X_{ij}(\tau;\ell)$ and $\widehat{\gamma}^X_{ij;T}(\tau;\ell)$ be the (i,j) entries of the lag ℓ autocovariance matrix $\mathbf{\Gamma}^X_n(\tau;\ell)$ and its estimator $\widehat{\mathbf{\Gamma}}^X_n(\tau;\ell)$ as defined in (16), respectively. By Assumptions (A1), (B1), and (B3):

$$\begin{aligned}\left|\gamma^{\chi}_{ij}(\tau;\ell) - \mathrm{E}[\chi_{i,\lfloor\tau T\rfloor}\chi'_{j,\lfloor\tau T\rfloor-\ell}]\right| &\le \sum_{k=0}^{\infty}\sum_{s=1}^{q} \left| c_{isk}(\tau) c_{js,k+|\ell|}(\tau) - c^*_{isk}(\lfloor\tau T\rfloor) c^*_{js,k+|\ell|}(\lfloor\tau T\rfloor)\right|\\ &\le \frac{2qC_\chi C'_1 \rho_\chi^{|\ell|}}{T} \sum_{k=0}^{\infty} \rho_\chi^{2k} + o(T^{-1})\\ &\le \frac{2qC_\chi C'_1}{T(1-\rho_\chi^2)} + o(T^{-1}) =: \frac{\mathcal{H}_1}{T} + o(T^{-1}), \text{ say.}\end{aligned} \tag{A.25}$$

Similarly, by Assumptions (A2), (B5), and (B6),

$$\begin{aligned}\left|\gamma^{\xi}_{ij}(\tau;\ell) - \mathrm{E}[\xi_{i,\lfloor\tau T\rfloor}\xi'_{j,\lfloor\tau T\rfloor-\ell}]\right| &\le \sum_{k=0}^{\infty}\sum_{s=1}^{\infty} \left| d_{isk}(\tau) d_{js,k+|\ell|}(\tau) - d^*_{isk}(\lfloor\tau T\rfloor) d^*_{js,k+|\ell|}(\lfloor\tau T\rfloor)\right|\\ &\le \frac{2\rho_\xi^{|\ell|}}{T} \sum_{s=1}^{\infty} B_{\xi is} B'_{1js} \sum_{k=0}^{\infty} \rho_\chi^{2k} + o(T^{-1})\\ &\le \frac{2B_\xi B'_1}{T(1-\rho_\xi^2)} + o(T^{-1}) =: \frac{\mathcal{H}_2}{T} + o(T^{-1}), \text{ say.}\end{aligned} \tag{A.26}$$

Notice that, in (A.25) and (A.26), $\mathcal{H}_1$ and $\mathcal{H}_2$, as well as the remainders, are independent of i, j, ℓ, and τ. Therefore, from (A.25) and (A.26), because of Assumption (A3) (on the mutual uncorrelatedness of the common and idiosyncratic shocks) we have

$$\max_{1\le i,j\le n}\sup_{\tau\in(0,1)}\sup_{\ell\in\mathbb{Z}}\left|\gamma_{ij}^X(\tau;\ell)-\mathrm{E}[X_{i,\lfloor\tau T\rfloor}X'_{j,\lfloor\tau T\rfloor-\ell}]\right|\le\frac{\mathcal{H}_1}{T}+\frac{\mathcal{H}_2}{T}+o(T^{-1})=:\frac{\mathcal{H}}{T},\text{ say.}\tag{A.27}$$

This is the same result as in, for example, Dahlhaus (2012, equation (73)).

Then, for any $\tau\in(0,1)$ and $|\ell|\le(M_T-1)$, using (16) and (A.27),

$$\begin{aligned}\left|\mathrm{E}[\widehat{\gamma}_{ij;T}^X(\tau;\ell)]-\gamma_{ij}^X(\tau;\ell)\right|&=\left|\left\{\frac{1}{M_T}\sum_{s=T_1(\tau)+\ell}^{T_2(\tau)}\mathrm{J}\left(\frac{s-\lfloor\tau T\rfloor}{M_T}\right)\mathrm{E}[X_{i,s-\ell}X'_{js}]\right\}-\gamma_{ij}^X(\tau;\ell)\right|\\&\le\left|\left\{\frac{1}{M_T}\sum_{s=T_1(\tau)+\ell}^{T_2(\tau)}\mathrm{J}\left(\frac{s-\lfloor\tau T\rfloor}{M_T}\right)\gamma_{ij}^X(s/T;\ell)\right\}-\gamma_{ij}^X(\tau;\ell)\right|+\frac{\mathcal{H}}{T}.\end{aligned}\tag{A.28}$$

Moreover, by a Taylor expansion of order two of $\sigma_{ij}^X(s/T;\theta)$ in its first argument in a neighborhood of τ, we have, in view of Assumption (F1), for any $T_1(\tau)+\ell\le s\le T_2(\tau)$

$$\begin{aligned}\frac{1}{M_T}\sum_{s=T_1(\tau)+\ell}^{T_2(\tau)}\mathrm{J}\left(\frac{s-\lfloor\tau T\rfloor}{M_T}\right)\gamma_{ij}^X(s/T;\ell)&=\frac{1}{2\lfloor Tb_T\rfloor}\sum_{s=T_1(\tau)+\ell}^{T_2(\tau)}\mathrm{J}\left(\frac{s-\lfloor\tau T\rfloor}{2\lfloor Tb_T\rfloor}\right)\int_{-\pi}^{\pi}e^{\iota\theta\ell}\sigma_{ij}^X(s/T;\theta)\mathrm{d}\theta\\&\le\int_{-1/2}^{1/2}\mathrm{J}(u)\mathrm{d}u\int_{-\pi}^{\pi}e^{\iota\theta\ell}\sigma_{ij}^X(\tau;\theta)\mathrm{d}\theta+\frac{b_T^2}{2}\int_{-1/2}^{1/2}u^2\mathrm{J}(u)\mathrm{d}u\int_{-\pi}^{\pi}e^{\iota\theta\ell}\frac{\mathrm{d}^2\sigma_{ij}^X(\tau;\theta)}{\mathrm{d}\tau^2}\mathrm{d}\theta+o(b_T^2)+\frac{\mathcal{L}}{Tb_T}\\&=\gamma_{ij}^X(\tau;\ell)+\frac{b_T^2}{2}\int_{-1/2}^{1/2}u^2\mathrm{J}(u)\mathrm{d}u\int_{-\pi}^{\pi}e^{\iota\theta\ell}\frac{\mathrm{d}^2\sigma_{ij}^X(\tau;\theta)}{\mathrm{d}\tau^2}\mathrm{d}\theta+o(b_T^2)+\frac{\mathcal{L}}{Tb_T},\end{aligned}\tag{A.29}$$

where $\mathcal{L}$ and the remainders are independent of i,j,ℓ, and τ due to Assumptions (B4) and (B7). Notice also that the first-order term of the Taylor expansion of $\sigma_{ij}^X(s/T;\theta)$ drops out due to the symmetry of the kernel J about the origin.

Substituting (A.29) into (A.28), we get, in view of Lemma 1*(i)* (where $\mathcal{K}$ is defined),

$$\max_{1\le i,j\le n}\sup_{\tau\in(0,1)}\max_{|\ell|\le(M_T-1)}\left|\mathrm{E}[\widehat{\gamma}_{ij;T}^X(\tau;\ell)]-\gamma_{ij}^X(\tau;\ell)\right|\le b_T^2\frac{\mathcal{K}}{2}\int_{-1/2}^{1/2}u^2\mathrm{J}(u)\mathrm{d}u+o(b_T^2)+\frac{\mathcal{H}}{T}+\frac{\mathcal{L}}{Tb_T},$$

since $|e^{\iota\theta\ell}|=1$ for all $\theta\in[-\pi,\pi]$ and $\ell\in\mathbb{Z}$. Defining $\mathcal{M}:=\mathcal{K}/2\int_{-1/2}^{1/2}u^2\mathrm{J}(u)\mathrm{d}u$ (which is finite by Assumption (F1)) completes the proof. □

Because of Lemma A3, using in (17) the triangular kernel given in (24), for any $\theta\in[-\pi,\pi]$ and $\tau\in(0,1)$, the bias of our spectral estimator satisfies

$$\begin{aligned}2\pi\left\{\mathrm{E}[\widehat{\sigma}_{ij;T}^X(\tau;\theta)]-\sigma_{ij}^X(\tau;\theta)\right\}&=\sum_{\ell=-m_T}^{m_T}\left(1-\frac{|\ell|}{m_T}\right)\mathrm{E}[\widehat{\gamma}_{ij}^X(\tau;\ell)]e^{-\iota\ell\theta}-\sum_{\ell=-\infty}^{\infty}\gamma_{ij}^X(\tau;\ell)e^{-\iota\ell\theta}\\&=\sum_{\ell=-m_T}^{m_T}\left(1-\frac{|\ell|}{m_T}\right)\gamma_{ij}^X(\tau;\ell)e^{-\iota\ell\theta}-\sum_{\ell=-m_T}^{m_T}\gamma_{ij}^X(\tau;\ell)e^{-\iota\ell\theta}\\&\quad+\sum_{\ell=-m_T}^{m_T}\left(1-\frac{|\ell|}{m_T}\right)\left\{\mathrm{E}[\widehat{\gamma}_{ij}^X(\tau;\ell)]-\gamma_{ij}^X(\tau;\ell)\right\}e^{-\iota\ell\theta}-\sum_{|\ell|>m_T}\gamma_{ij}^X(\tau;\ell)e^{-\iota\ell\theta}\\&\le\sum_{\ell=-m_T}^{m_T}\frac{|\ell|}{m_T}|\gamma_{ij}^X(\tau;\ell)|+\sum_{\ell=-m_T}^{m_T}\left(1-\frac{|\ell|}{m_T}\right)\left(\mathcal{M}b_T^2+\frac{\mathcal{L}}{Tb_T}\right)+\sum_{|\ell|>m_T}|\gamma_{ij}^X(\tau;\ell)|+o(m_Tb_T^2)+\frac{m_T\mathcal{H}}{T}\\&=:\mathcal{S}_{1ijT}(\tau;\theta)+\mathcal{S}_{2ijT}(\tau;\theta)+\mathcal{S}_{3ijT}(\tau;\theta)+o(m_Tb_T^2)+o(m_TT^{-1}b_T^{-1}),\text{ say.}\end{aligned}\tag{A.30}$$

Then, from (A.5) and (A.6) in the proof of Lemma 1, because of Assumption (A3), we have

$$\max_{1\le i,j\le n}\sup_{\tau\in(0,1)}|\gamma_{ij}^X(\tau;\ell)|\le\mathcal{K}_3\rho_\chi^{|\ell|}+\mathcal{K}_4\rho_\xi^{|\ell|},\tag{A.31}$$

hence

$$\max_{1\le i,j\le n}\sup_{\tau\in(0,1)}\sup_{\theta\in[-\pi,\pi]}\mathcal{S}_{1ijT}(\tau;\theta)\le\sum_{\ell=-\infty}^{\infty}\left\{\mathcal{K}_3\rho_\chi^{|\ell|}+\mathcal{K}_4\rho_\xi^{|\ell|}\right\}\frac{|\ell|}{m_T}\le\frac{2\mathcal{K}_3\rho_\chi}{m_T(1-\rho_\chi)^2}+\frac{2\mathcal{K}_4\rho_\xi}{m_T(1-\rho_\xi)^2}.\tag{A.32}$$

Similarly, from Lemma A3,

$$\max_{1\le i,j\le n}\sup_{\tau\in(0,1)}\sup_{\theta\in[-\pi,\pi]}\mathcal{S}_{2ijT}(\tau;\theta)\le\mathcal{M}m_Tb_T^2+\frac{\mathcal{L}}{Tb_T}m_T.\tag{A.33}$$

Finally,

$$\max_{1\le i,j\le n}\sup_{\tau\in(0,1)}\sup_{\theta\in[-\pi,\pi]}\mathcal{S}_{3ijT}(\tau;\theta)\le\sum_{|\ell|>m_T}\left\{\mathcal{K}_3\rho_\chi^{|\ell|}+\mathcal{K}_4\rho_\xi^{|\ell|}\right\}\frac{|\ell|}{m_T}\le\frac{F}{m_T},\tag{A.34}$$

for some constant F (independent of i, j, ℓ, and τ). Substituting (A.32), (A.33), and (A.34) into (A.30), squaring the bias, and noticing that all those results are independent of i, j, τ, and θ, we obtain, from (A.20) (recall that $M_T = 2\lfloor Tb_T\rfloor$ and $m_T = \lfloor 1/h_T\rfloor$),

$$\Delta_T^2\le C_X''\left(h_T^2+\frac{b_T^4}{h_T^2}+\frac{1}{T^2b_T^2h_T^2}\right)\;+\;o\left(\frac{b_T^4}{h_T^2}\right)+o\left(\frac{1}{T^2b_T^2h_T^2}\right)\tag{A.35}$$

for some constant C_X'' (independent of T and n).

Substituting (A.23) and (A.35) into (A.18) completes the proof. □

A6 Proof of Proposition 1

We divide the proof into four steps corresponding to the four steps of the estimation procedure.

(i) - Estimation of Spectral Density. Recall that, as discussed in Section 3.1, the estimator of the spectral density can be computed only for t/T with $M_T/2\le t\le(T-M_T/2)$ and for $\theta_j=\pi jh_T$ with $|j|\le m_T$. For simplicity of notation, we let $\mathcal{T}_T:=\{M_T/2,\dots,(T-M_T/2)\}$. Then, a straightforward implication of Lemma 4 is that there exists a constant C^* (independent of T and n) such that, for any $n, T\in\mathbb{N}_0$,

$$\max_{1\le i,k\le n}\mathrm{E}\left[\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\left|\widehat{\sigma}_{ik;T}^X(t/T;\theta_j)-\sigma_{ik}^X(t/T;\theta_j)\right|^2\right]\le C^*\zeta_{T,r^*}\tag{A.36}$$

with ζ_{T,r^*} defined in (26). From (A.36), for any $n\in\mathbb{N}_0$,

$$\begin{aligned}
&\mathrm{E}\left[\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\frac{1}{n^2}\left\|\widehat{\mathbf{\Sigma}}_{n,T}^X(t/T;\theta_j)-\mathbf{\Sigma}_n^X(t/T;\theta_j)\right\|^2\right]\\
&\qquad\le\mathrm{E}\left[\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\frac{1}{n^2}\left\|\widehat{\mathbf{\Sigma}}_{n,T}^X(t/T;\theta_j)-\mathbf{\Sigma}_n^X(t/T;\theta_j)\right\|_F^2\right]\\
&\qquad=\frac{1}{n^2}\mathrm{E}\left[\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\sum_{i=1}^n\sum_{k=1}^n\left|\widehat{\sigma}_{ik;T}^X(t/T;\theta_j)-\sigma_{ik}^X(t/T;\theta_j)\right|^2\right]\\
&\qquad\le\frac{1}{n^2}\sum_{i=1}^n\sum_{k=1}^n\mathrm{E}\left[\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\left|\widehat{\sigma}_{ik;T}^X(t/T;\theta_j)-\sigma_{ik}^X(t/T;\theta_j)\right|^2\right]\\
&\qquad\le C^*\zeta_{T,r^*}.
\end{aligned}$$

Therefore, by Chebychev's inequality, for any $n \in \mathbb{N}_0$, as $T \to \infty$,

$$\max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \frac{1}{n} \left\| \widehat{\mathbf{\Sigma}}^X_{n,T}(t/T; \theta_j) - \mathbf{\Sigma}^X_n(t/T; \theta_j) \right\| = O_{\mathrm{P}}\left(\zeta_{T,r^*}^{1/2} \right). \tag{A.37}$$

Let $\boldsymbol{\ell}_i$ denote the ith vector in the n-dimensional canonical basis, i.e. the vector with 1 in entry i and 0 elsewhere. Then, again from (A.36), for any $n \in \mathbb{N}_0$ and $1 \le i \le n$,

$$\begin{aligned}
&\mathrm{E}\left[\max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \frac{1}{n} \left\| \boldsymbol{\ell}_i' \Big(\widehat{\mathbf{\Sigma}}^X_{n,T}(t/T; \theta_j) - \mathbf{\Sigma}^X_n(t/T; \theta_j) \Big) \right\|^2 \right] \\
&= \mathrm{E}\left[\max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \frac{1}{n} \boldsymbol{\ell}_i' \Big(\widehat{\mathbf{\Sigma}}^X_{n,T}(t/T; \theta_j) - \mathbf{\Sigma}^X_n(t/T; \theta_j) \Big) \Big(\widehat{\mathbf{\Sigma}}^X_{n,T}(t/T; \theta_j) - \mathbf{\Sigma}^X_n(t/T; \theta_j) \Big)' \boldsymbol{\ell}_i \right] \\
&= \mathrm{E}\left[\max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \frac{1}{n} \boldsymbol{\ell}_i' \Big(\widehat{\mathbf{\Sigma}}^X_{n,T}(t/T; \theta_j) - \mathbf{\Sigma}^X_n(t/T; \theta_j) \Big) \Big(\widehat{\mathbf{\Sigma}}^X_{n,T}(t/T; \theta_j) - \mathbf{\Sigma}^X_n(t/T; \theta_j) \Big) \boldsymbol{\ell}_i \right] \\
&= \frac{1}{n} \mathrm{E}\left[\max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \sum_{k=1}^n \left| \widehat{\sigma}^X_{ik;n,T}(t/T; \theta_j) - \sigma^X_{ik}(t/T; \theta_j) \right|^2 \right] \\
&\le \frac{1}{n} \sum_{k=1}^n \mathrm{E}\left[\max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \left| \widehat{\sigma}^X_{ik;n,T}(t/T; \theta_j) - \sigma^X_{ik}(t/T; \theta_j) \right|^2 \right] \le C^* \zeta_{T,r^*},
\end{aligned} \tag{A.38}$$

where $\| \cdot \|$ in this case denotes the Euclidean norm of a vector. Therefore, by Chebychev's inequality, and since C^* in (A.38) does not depend on i, for any $\varepsilon > 0$, there exists $\eta(\varepsilon)$ and an integer $T^* = T(\varepsilon)$, both independent of i, such that

$$\mathrm{P}\left(\max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \frac{1}{\sqrt{n \zeta_{T,r^*}}} \left\| \boldsymbol{\ell}_i' \Big(\widehat{\mathbf{\Sigma}}^X_{n,T}(t/T; \theta_j) - \mathbf{\Sigma}^X_n(t/T; \theta_j) \Big) \right\| \ge \eta(\varepsilon) \right) < \varepsilon, \tag{A.39}$$

for all $n \in \mathbb{N}_0$, $1 \le i \le n$, and $T \ge T^*$. Equivalently, hereafter, we say that

$$\max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \frac{1}{\sqrt{n}} \left\| \boldsymbol{\ell}_i' \Big(\widehat{\mathbf{\Sigma}}^X_{n,T}(t/T; \theta_j) - \mathbf{\Sigma}^X_n(t/T; \theta_j) \Big) \right\| = O_{\mathrm{P}}\left(\zeta_{T,r^*}^{1/2} \right) \tag{A.40}$$

as $T \to \infty$, uniformly in i. This proves the analogue of Lemma 1*(i)* and 1*(ii)* in Forni et al. (2017).

(ii) - Dynamic Principal Components. By using (A.37) and Lemma 2*(i)*, for any $n \in \mathbb{N}_0$, we have

$$\begin{aligned}
&\max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \frac{1}{n} \left\| \widehat{\mathbf{\Sigma}}^X_{n,T}(t/T; \theta_j) - \mathbf{\Sigma}^\chi_n(t/T; \theta_j) \right\| \\
&\qquad \le \max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \frac{1}{n} \left\| \widehat{\mathbf{\Sigma}}^X_{n,T}(t/T; \theta_j) - \mathbf{\Sigma}^X_n(t/T; \theta_j) \right\| + \max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \frac{1}{n} \left\| \mathbf{\Sigma}^\xi_n(t/T; \theta_j) \right\| \\
&\qquad = O_{\mathrm{P}}(\zeta_{T,r^*}^{1/2}) + O(n^{-1}) = O_{\mathrm{P}}\left(\max\left(\zeta_{T,r^*}^{1/2}, n^{-1} \right) \right),
\end{aligned} \tag{A.41}$$

as $T \to \infty$. Similarly, we can show that

$$\begin{aligned}
&\max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \frac{1}{\sqrt{n}} \left\| \boldsymbol{\ell}_i' \Big(\widehat{\mathbf{\Sigma}}^X_{n,T}(t/T; \theta_j) - \mathbf{\Sigma}^\chi_n(t/T; \theta_j) \Big) \right\| \\
&\le \max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \frac{1}{\sqrt{n}} \left\| \boldsymbol{\ell}_i' \Big(\widehat{\mathbf{\Sigma}}^X_{n,T}(t/T; \theta_j) - \mathbf{\Sigma}^X_n(t/T; \theta_j) \Big) \right\| + \max_{t \in \mathcal{T}_T} \max_{|j| \le m_T} \frac{1}{\sqrt{n}} \left\| \boldsymbol{\ell}_i' \mathbf{\Sigma}^\xi_n(t/T; \theta_j) \right\| \\
&= O_{\mathrm{P}}\left(\max\left(\zeta_{T,r^*}^{1/2}, n^{-1/2} \right) \right)
\end{aligned} \tag{A.42}$$

as $T \to \infty$, uniformly in i, because of (A.40), while for the second term we have (recall that $\boldsymbol{\ell}_i' \boldsymbol{\ell}_i = 1$)

$$\max_{1\le i\le n} \max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} \frac{1}{n} \left\| \boldsymbol{\ell}_i' \boldsymbol{\Sigma}_n^{\xi}(t/T;\theta_j) \right\|^2 = \max_{1\le i\le n} \max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} \frac{1}{n} \boldsymbol{\ell}_i' \boldsymbol{\Sigma}_n^{\xi}(t/T;\theta_j) \boldsymbol{\Sigma}_n^{\xi}(t/T;\theta_j) \boldsymbol{\ell}_i$$
$$\le \max_{\boldsymbol{w}:\boldsymbol{w}'\boldsymbol{w}=1} \max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} \frac{1}{n} \boldsymbol{w}' \boldsymbol{\Sigma}_n^{\xi}(t/T;\theta_j) \boldsymbol{\Sigma}_n^{\xi}(t/T;\theta_j) \boldsymbol{w}$$
$$= \max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} \frac{1}{n} \|\boldsymbol{\Sigma}_n^{\xi}(t/T;\theta_j)\|^2 = O(n^{-1}) \tag{A.43}$$

by definition of the largest eigenvalue of a matrix and Lemma 2*(i)*. This proves the analogue of Lemma 1*(iii)* and *(iv)* in Forni et al. (2017).

It follows from (A.41) and Weyl's inequality that, for all $1 \le \ell \le q$, as $n, T \to \infty$,

$$\max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} \frac{1}{n} \left| \widehat{\lambda}_{\ell;n,T}^X(t/T;\theta_j) - \lambda_{\ell;n}^{\chi}(t/T;\theta_j) \right| \le \max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} \frac{1}{n} \left\| \widehat{\boldsymbol{\Sigma}}_{n,T}^X(t/T;\theta_j) - \boldsymbol{\Sigma}_n^{\chi}(t/T;\theta_j) \right\|$$
$$= O_{\mathrm{P}}\left(\max\left(\zeta_{T,r^*}^{1/2}, n^{-1} \right) \right). \tag{A.44}$$

Let $\widehat{\boldsymbol{\Lambda}}_{n,T}^X(t/T;\theta_j)$ and $\boldsymbol{\Lambda}_n^{\chi}(t/T;\theta_j)$ be the $q \times q$ diagonal matrices with the q largest eigenvalues of $\widehat{\boldsymbol{\Sigma}}_{n,T}^X(t/T;\theta_j)$ and $\boldsymbol{\Sigma}_n^{\chi}(t/T;\theta_j)$, respectively. Then, since q is finite, (A.44) holds uniformly in ℓ and, as $n, T \to \infty$, we have

$$\max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} \left\| \left(n^{-1} \boldsymbol{\Lambda}_n^{\chi}(t/T;\theta_j) - n^{-1} \widehat{\boldsymbol{\Lambda}}_{n,T}^X(t/T;\theta_j) \right) \right\|^2$$
$$\le \sum_{\ell=1}^{q} \max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} \frac{1}{n^2} \left(\widehat{\lambda}_{\ell;n,T}^X(t/T;\theta_j) - \lambda_{\ell;n}^{\chi}(t/T;\theta_j) \right)^2 = O_{\mathrm{P}}\left(\max\left(\zeta_{T,r^*}, n^{-2} \right) \right). \tag{A.45}$$

From Assumption (C), as $n \to \infty$

$$\max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} \left\| n (\boldsymbol{\Lambda}_n^{\chi}(t/T;\theta_j))^{-1} \right\| = \max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} n (\lambda_{q;n}^{\chi}(t/T;\theta_j))^{-1} \le D, \tag{A.46}$$

with $D > 0$ independent of n. And from (A.45), (A.46), and Lemma 2*(ii)*, as $n, T \to \infty$,

$$\max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} \left\| n (\widehat{\boldsymbol{\Lambda}}_{n,T}^X(t/T;\theta_j))^{-1} \right\|$$
$$\le \max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} \left\| n \big((\widehat{\boldsymbol{\Lambda}}_{n,T}^X(t/T;\theta_j))^{-1} - (\boldsymbol{\Lambda}_n^{\chi}(t/T;\theta_j))^{-1} \big) \right\| + \max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} \left\| n (\boldsymbol{\Lambda}_n^{\chi}(t/T;\theta_j))^{-1} \right\|$$
$$= O_{\mathrm{P}}\left(\max\left(\zeta_{T,r^*}^{1/2}, n^{-1} \right) \right) + O(1) = O_{\mathrm{P}}(1). \tag{A.47}$$

This proves the analogue of Lemma 2 in Forni et al. (2017).

Let $\widehat{\mathbf{P}}_{n,T}^X(t/T;\theta_j)$ be the $n \times q$ matrix having as columns the normalized eigenvectors of $\widehat{\boldsymbol{\Sigma}}_{n,T}^X(t/T;\theta_j)$ corresponding to its q largest eigenvalues. Let $\mathbf{P}_n^{\chi}(t/T;\theta_j)$ be the $n \times q$ matrix having as columns the normalized eigenvectors of $\boldsymbol{\Sigma}_{n,T}^{\chi}(t/T;\theta_j)$ corresponding to its q largest eigenvalues. By "normalized" we mean that the q columns $\mathbf{p}_{j;n}^{\chi}(t/T;\theta_j)$ of $\mathbf{P}_n^{\chi}(t/T;\theta_j)$ are such that $\mathbf{p}_{j;n}^{\chi\dagger}(t/T;\theta_j) \mathbf{p}_{j;n}^{\chi}(t/T;\theta_j) = 1$.

Now, by Assumption (C),

$$\max_{1\le \ell\le q} \max_{t\in\mathcal{T}_T} \max_{|j|\le m_T} \left(\lambda_{\ell;n}^{\chi}(t/T;\theta_j) n^{-1} \right) \ge C, \tag{A.48}$$

for some constant C independent of n. Then, by Theorem 2 and Corollary 2 in Yu et al. (2015), there exists a $q\times q$ complex diagonal matrix $\boldsymbol{\mathcal{J}}(t/T;\theta_j)$ with unit modulus entries, such that, as $n,T\to\infty$,

$$
\begin{aligned}
&\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\left\|\widehat{\mathbf{P}}^X_{n,T}(t/T;\theta_j)-\mathbf{P}^\chi_n(t/T;\theta_j)\boldsymbol{\mathcal{J}}(t/T;\theta_j)\right\|\\
&\qquad\le \max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\frac{2^{3/2}\sqrt{q}\left\|\widehat{\boldsymbol{\Sigma}}^X_{n,T}(t/T;\theta_j)-\boldsymbol{\Sigma}^\chi_n(t/T;\theta_j)\right\|}{\lambda^\chi_{q;n}(t/T;\theta_j)}\\
&\qquad\le 2^{3/2}\sqrt{q}\,Cn^{-1}\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\left\|\widehat{\boldsymbol{\Sigma}}^X_{n,T}(t/T;\theta_j)-\boldsymbol{\Sigma}^\chi_n(t/T;\theta_j)\right\|=O_{\mathrm{P}}\left(\max\left(\zeta^{1/2}_{T,r^*},n^{-1}\right)\right),
\end{aligned}
\tag{A.49}
$$

because of (A.41) and (A.48). Noting that $\|\mathbf{P}^\chi_n(t/T;\theta_j)\boldsymbol{\mathcal{J}}(t/T;\theta_j)\|=1$ and $\|\boldsymbol{\ell}_i'\boldsymbol{\Sigma}^\chi_n(t/T;\theta_j)\|=O(\sqrt{n})$, moreover, for all $t\in\mathcal{T}_T$ and $|j|\le m_T$, as $n,T\to\infty$,

$$
\begin{aligned}
&\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\sqrt{n}\left\|\boldsymbol{\ell}_i'\Big(\widehat{\mathbf{P}}^X_{n,T}(t/T;\theta_j)-\mathbf{P}^\chi_n(t/T;\theta_j)\boldsymbol{\mathcal{J}}(t/T;\theta_j)\Big)\right\|\\
&=\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\frac{1}{\sqrt{n}}\Big\|\boldsymbol{\ell}_i'\Big\{\widehat{\boldsymbol{\Sigma}}^X_{n,T}(t/T;\theta_j)\widehat{\mathbf{P}}^X_{n,T}(t/T;\theta_j)\,n(\widehat{\boldsymbol{\Lambda}}^X_{n,T}(t/T;\theta_j))^{-1}\\
&\quad-\boldsymbol{\Sigma}^\chi_n(t/T;\theta_j)\mathbf{P}^\chi_n(t/T;\theta_j)\,n(\boldsymbol{\Lambda}^\chi_n(t/T;\theta_j))^{-1}\Big\}\Big\|\\
&\le\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\frac{1}{\sqrt{n}}\left\|\boldsymbol{\ell}_i'\Big(\widehat{\boldsymbol{\Sigma}}^X_{n,T}(t/T;\theta_j)-\boldsymbol{\Sigma}^\chi_n(t/T;\theta_j)\Big)\right\|\,\|n(\boldsymbol{\Lambda}^\chi_n(t/T;\theta_j))^{-1}\|\\
&\quad+\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\frac{1}{\sqrt{n}}\left\|\boldsymbol{\ell}_i'\boldsymbol{\Sigma}^\chi_n(t/T;\theta_j)\right\|\,\left\|n\Big((\boldsymbol{\Lambda}^\chi_n(t/T;\theta_j))^{-1}-(\widehat{\boldsymbol{\Lambda}}^X_{n,T}(t/T;\theta_j))^{-1}\Big)\right\|\\
&\quad+\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\frac{1}{\sqrt{n}}\left\|\boldsymbol{\ell}_i'\boldsymbol{\Sigma}^\chi_n(t/T;\theta_j)\right\|\,\|n(\boldsymbol{\Lambda}^\chi_n(t/T;\theta_j))^{-1}\|\,\left\|\widehat{\mathbf{P}}^X_{n,T}(t/T;\theta_j)-\mathbf{P}^\chi_n(t/T;\theta_j)\boldsymbol{\mathcal{J}}(t/T;\theta_j)\right\|\\
&\quad+o_{\mathrm{P}}\left(\max\left(\zeta^{1/2}_{T,r^*},n^{-1/2}\right)\right)=O_{\mathrm{P}}\left(\max\left(\zeta^{1/2}_{T,r^*},n^{-1/2}\right)\right),
\end{aligned}
\tag{A.50}
$$

uniformly in i, because of (A.42), (A.45), (A.46), and (A.49).

Furthermore, because of Assumptions (A1) and (B1),

$$
\max_{1\le i\le n}\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\sigma^\chi_{ii}(t/T;\theta_j)=\max_{1\le i\le n}\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\sum_{\ell=1}^q c_{i\ell}(t/T;e^{-\iota\theta_j})c^\dagger_{\ell i}(t/T;e^{-\iota\theta_j})\le\frac{C_1^2}{(1-\rho_\chi)^2}
\tag{A.51}
$$

and, since (B1) holds for all $i\in\mathbb{N}_0$, (A.51) is independent of n. Therefore, denoting by $p^\chi_{i\ell}(t/T;\theta_j)$ the (i,ℓ) entry of $\mathbf{P}^\chi_n(t/T;\theta_j)$, we also have

$$
\max_{1\le i\le n}\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\sigma^\chi_{ii}(t/T;\theta_j)=\max_{1\le i\le n}\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\sum_{\ell=1}^q\Big(\lambda^\chi_{\ell;n}(t/T;\theta_j)\Big)\big|p^\chi_{i\ell}(t/T;\theta_j)\big|^2\le\frac{C_1^2}{(1-\rho_\chi)^2}.
\tag{A.52}
$$

Therefore, replacing (A.48) into (A.52), we see that for all $1\le\ell\le q$ we must have

$$
\max_{1\le i\le n}\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\left(n\big|p^\chi_{i\ell}(t/T;\theta_j)\big|^2\right)\le A
$$

where the constant A is also independent of n since the constants in (A.48) and (A.52) are. It follows that

$$
\max_{1\le i\le n}\max_{t\in\mathcal{T}_T}\max_{|j|\le m_T}\sqrt{n}\left\|\boldsymbol{\ell}_i'\mathbf{P}^\chi_n(t/T;\theta_j)\right\|\le M
\tag{A.53}
$$

for some constant M that does not depend on n.

Finally, since $\boldsymbol{\mathcal{J}}^\dagger(t/T;\theta_j)\boldsymbol{\mathcal{J}}(t/T;\theta_j)=\mathbf{I}_q$ and $\boldsymbol{\mathcal{J}}^\dagger(t/T;\theta_j)\boldsymbol{\Lambda}^X_n(t/T;\theta_j)\boldsymbol{\mathcal{J}}(t/T;\theta_j)=\boldsymbol{\Lambda}^X_n(t/T;\theta_j)$, then, for any $n \in \mathbb{N}_0$,

$$\begin{aligned}
&\max_{t\in\mathcal{T}_T}\max_{|j|\leq m_T}\left\|\boldsymbol{\ell}_i'\Big(\mathbf{P}_n^\chi(t/T;\theta_j)(\boldsymbol{\Lambda}_n^\chi(t/T;\theta_j))^{1/2}\boldsymbol{\mathcal{J}}(t/T;\theta_j)-\widehat{\mathbf{P}}_{n,T}^X(t/T;\theta_j)(\widehat{\boldsymbol{\Lambda}}_{n,T}^X(t/T;\theta_j))^{1/2}\Big)\right\| \\
&\quad\leq \max_{t\in\mathcal{T}_T}\max_{|j|\leq m_T}\left\|\boldsymbol{\ell}_i'\sqrt{n}\Big(\mathbf{P}_n^\chi(t/T;\theta_j)\boldsymbol{\mathcal{J}}(t/T;\theta_j)-\widehat{\mathbf{P}}_{n,T}^X(t/T;\theta_j)\Big)\left(n^{-1}\boldsymbol{\Lambda}_n^\chi(t/T;\theta_j)\right)^{1/2}\right\| \\
&\quad+\max_{t\in\mathcal{T}_T}\max_{|j|\leq m_T}\left\|\sqrt{n}\boldsymbol{\ell}_i'\widehat{\mathbf{P}}_{n,T}^X(t/T;\theta_j)\Big(n^{-1/2}\big(\boldsymbol{\Lambda}_n^\chi(t/T;\theta_j)\big)^{1/2}-n^{-1/2}\big(\widehat{\boldsymbol{\Lambda}}_{n,T}^X(t/T;\theta_j)\big)^{1/2}\Big)\right\| \\
&\quad=: I+II,\quad \text{say.}
\end{aligned}\tag{A.54}$$

It follows from (A.46) and (A.50) that I is $O_{\mathrm{P}}(\max(\zeta_{T,r^*}^{1/2},n^{-1/2}))$ uniformly in i. For II, we have

$$\begin{aligned}
II &\leq \max_{t\in\mathcal{T}_T}\max_{|j|\leq m_T}\left\{\left\|\sqrt{n}\boldsymbol{\ell}_i'\widehat{\mathbf{P}}_{n,T}^X(t/T;\theta_j)\right\|\left\|\Big(n^{-1/2}\big(\boldsymbol{\Lambda}_n^\chi(t/T;\theta_j)\big)^{1/2}-n^{-1/2}\big(\widehat{\boldsymbol{\Lambda}}_{n,T}^X(t/T;\theta_j)\big)^{1/2}\Big)\right\|\right\} \\
&=: \max_{t\in\mathcal{T}_T}\max_{|j|\leq m_T}\left(II_a\times II_b\right),\quad \text{say.}
\end{aligned}\tag{A.55}$$

It follows from (A.45) that II_b in (A.55) is $O_{\mathrm{P}}(\max(\zeta_{T,r^*}^{1/2},n^{-1})$ uniformly in t and j, while for II_a we have (recall that $\|\boldsymbol{\mathcal{J}}(t/T;\theta_j)\|=1$)

$$\begin{aligned}
\max_{t\in\mathcal{T}_T}\max_{|j|\leq m_T} II_b &\leq \max_{t\in\mathcal{T}_T}\max_{|j|\leq m_T}\left\|\sqrt{n}\boldsymbol{\ell}_i'\Big(\widehat{\mathbf{P}}_{n,T}^X(t/T;\theta_j)-\mathbf{P}_n^\chi(t/T;\theta_j)\boldsymbol{\mathcal{J}}(t/T;\theta_j)\Big)\right\| \\
&+\max_{t\in\mathcal{T}_T}\max_{|j|\leq m_T}\left\|\sqrt{n}\boldsymbol{\ell}_i'\mathbf{P}_n^\chi(t/T;\theta_j)\boldsymbol{\mathcal{J}}(t/T;\theta_j)\right\| \\
&=O_{\mathrm{P}}(\max(\zeta_{T,r^*}^{1/2},n^{-1/2})+O(1),
\end{aligned}\tag{A.56}$$

uniformly in i, because of (A.53) and (A.50). Hence, $II=O_{\mathrm{P}}(\max(\zeta_{T,r^*}^{1/2},n^{-1})$ uniformly in i, and therefore, from (A.54), as $n,T\to\infty$,

$$\begin{aligned}
&\max_{t\in\mathcal{T}_T}\max_{|j|\leq m_T}\left\|\boldsymbol{\ell}_i'\Big(\mathbf{P}_n^\chi(t/T;\theta_j)(\boldsymbol{\Lambda}_n^\chi(t/T;\theta_j))^{1/2}\boldsymbol{\mathcal{J}}(t/T;\theta_j)-\widehat{\mathbf{P}}_{n,T}^X(t/T;\theta_j)(\widehat{\boldsymbol{\Lambda}}_{n,T}^X(t/T;\theta_j))^{1/2}\Big)\right\| \\
&\quad=O_{\mathrm{P}}\left(\max\left(\zeta_{T,r^*}^{1/2},n^{-1/2}\right)\right),
\end{aligned}\tag{A.57}$$

uniformly in i. This proves the analogue of Lemma 4 in Forni et al. (2017).

Hereafter, let $\vartheta_{n,T,r^*}:=\max\left(\zeta_{T,r^*},n^{-1}\right)$. The spectral density matrix of the common component has rank q for all $\theta\in[-\pi,\pi]$ and $\tau\in(0,1)$ because of Assumption (C); for any $n\in\mathbb{N}_0$, ot can be expressed as

$$\begin{aligned}
\boldsymbol{\Sigma}_n^\chi(t/T;\theta_j) &= \left\{\mathbf{P}_n^\chi(t/T;\theta_j)[\boldsymbol{\Lambda}_n^\chi(t/T;\theta_j)]^{1/2}\boldsymbol{\mathcal{J}}(t/T;\theta_j)\right\}\left\{\boldsymbol{\mathcal{J}}^\dagger(t/T;\theta_j)[\boldsymbol{\Lambda}_n^\chi(t/T;\theta_j)]^{1/2}\mathbf{P}_n^{\chi\dagger}(t/T;\theta_j)\right\} \\
&=\mathbf{P}_n^\chi(t/T;\theta_j)\boldsymbol{\Lambda}_n^\chi(t/T;\theta_j)\mathbf{P}_n^{\chi\dagger}(t/T;\theta_j),
\end{aligned}\tag{A.58}$$

with entries $\sigma_{ik}^\chi(t/T;\theta_j)=\boldsymbol{\ell}_i'\boldsymbol{\Sigma}_n^\chi(t/T;\theta_j)\boldsymbol{\ell}_k$. The estimator of the spectral density matrix of the common component is obtained by principal component analysis as

$$\begin{aligned}
\widehat{\boldsymbol{\Sigma}}_{n,T}^\chi(t/T;\theta_j) &:= \widehat{\mathbf{P}}_{n,T}^X(t/T;\theta_j)[\widehat{\boldsymbol{\Lambda}}_{n,T}^X(t/T;\theta_j)]^{1/2}[\widehat{\boldsymbol{\Lambda}}_{n,T}^X(t/T;\theta_j)]^{1/2}\widehat{\mathbf{P}}_{n,T}^{X\dagger}(t/T;\theta_j) \\
&=\widehat{\mathbf{P}}_{n,T}^X(t/T;\theta_j)\widehat{\boldsymbol{\Lambda}}_{n,T}^X(t/T;\theta_j)\widehat{\mathbf{P}}_{n,T}^{X\dagger}(t/T;\theta_j)
\end{aligned}\tag{A.59}$$

with entries $\widehat{\sigma}_{ik;n,T}^\chi(t/T;\theta_j)=\boldsymbol{\ell}_i'\widehat{\boldsymbol{\Sigma}}_{n,T}^\chi(t/T;\theta_j)\boldsymbol{\ell}_k$. Then, by comparing (A.58) with (A.59) and because of (A.57), for any $\varepsilon>0$, there exists $\eta(\varepsilon)$, $T^*=T^*(\varepsilon)$, and $N^*=N^*(\varepsilon)$, all independent of i and k, such that

$$\mathrm{P}\left(\max_{t\in\mathcal{T}_T}\max_{|j|\leq m_T}\frac{\left|\widehat{\sigma}_{ik;n,T}^\chi(t/T;\theta_j)-\sigma_{ik}^\chi(t/T;\theta_j)\right|}{\vartheta_{n,T,r^*}^{1/2}}\geq\eta(\varepsilon)\right)<\varepsilon$$

for all $n \geq N^*$ and $T \geq T^*$. Equivalently as $n, T \to \infty$,

$$\max_{t\in\mathcal{T}_T}\max_{|j|\leq m_T}\left|\widehat{\sigma}^{\chi}_{ik;n,T}(t/T;\theta_j)-\sigma^{\chi}_{ik}(t/T;\theta_j)\right| = \max_{t\in\mathcal{T}_T}\max_{|j|\leq m_T}\left\|\boldsymbol{\ell}_i'\left(\widehat{\boldsymbol{\Sigma}}^{\chi}_{n,T}(t/T;\theta_j)-\boldsymbol{\Sigma}^{\chi}_{n}(t/T;\theta_j)\right)\boldsymbol{\ell}_k\right\|$$
$$= O_{\mathrm{P}}\left(\max\left(\zeta^{1/2}_{T,r^*}, n^{-1/2}\right)\right), \tag{A.60}$$

uniformly in i and k. This proves the analogue of Proposition 7 in Forni et al. (2017).

The (i,k) entry of the estimated lag ℓ autocovariance matrix $\widehat{\boldsymbol{\Gamma}}^{\chi}_{n,T}(\tau;k)$ defined in (19) is

$$\widehat{\gamma}^{\chi}_{ik;n,T}(t/T;\ell) = \frac{2\pi}{2m_T+1}\sum_{j=-m_T}^{m_T}\widehat{\sigma}^{\chi}_{ik;n,T}(t/T;\theta_k)e^{\iota\ell\theta_j} \tag{A.61}$$

and, by definition of a lag ℓ autocovariance, its population counterpart satisfies

$$\gamma^{\chi}_{ik}(t/T;\ell) = \int_{-\pi}^{\pi}\sigma^{\chi}_{ik}(t/T;\theta)e^{\iota\ell\theta}\mathrm{d}\theta. \tag{A.62}$$

Therefore, for any given lag ℓ (putting $\theta_{-m_T-1} := -\pi$), we have

$$\begin{aligned}
&\max_{t\in\mathcal{T}_T}|\widehat{\gamma}^{\chi}_{ik;n,T}(t/T;\ell)-\gamma^{\chi}_{ik}(t/T;\ell)|\\
&\leq \max_{t\in\mathcal{T}_T}\frac{2\pi}{2m_T+1}\sum_{j=-m_T}^{m_T}\left|e^{\iota\ell\theta_j}\widehat{\sigma}^{\chi}_{ik;n,T}(t/T;\theta_j)-e^{\iota\ell\theta_j}\sigma^{\chi}_{ik}(t/T;\theta_j)\right|\\
&\quad+\max_{t\in\mathcal{T}_T}\left|\frac{2\pi}{2m_T+1}\sum_{j=-m_T}^{m_T}e^{\iota\ell\theta_j}\sigma^{\chi}_{ik}(t/T;\theta_j)-\int_{-\pi}^{\pi}e^{\iota\ell\theta}\sigma^{\chi}_{ik}(t/T;\theta)\mathrm{d}\theta\right|\\
&\leq \max_{t\in\mathcal{T}_T}\frac{2\pi}{2m_T+1}\sum_{j=-m_T}^{m_T}\left|\widehat{\sigma}^{\chi}_{ik;n,T}(t/T;\theta_j)-\sigma^{\chi}_{ik}(t/T;\theta_j)\right|\\
&\quad+\max_{t\in\mathcal{T}_T}\frac{2\pi}{2m_T+1}\sum_{j=-m_T}^{m_T}\max_{\theta_{j-1}\leq\theta\leq\theta_j}\left|e^{\iota\ell\theta_j}\sigma^{\chi}_{ik}(t/T;\theta_j)-e^{\iota\ell\theta}\sigma^{\chi}_{ik}(t/T;\theta)\right|\\
&\leq \max_{t\in\mathcal{T}_T}2\pi\max_{|j|\leq m_T}\left|\widehat{\sigma}^{\chi}_{ik;n,T}(t/T;\theta_j)-\sigma^{\chi}_{ik}(t/T;\theta_j)\right|+\frac{2\pi}{2m_T+1}\frac{C_1^2}{(1-\rho_\chi)^2}\sum_{j=-m_T}^{m_T}\max_{\theta_{j-1}\leq\theta\leq\theta_j}\left|e^{\iota\ell\theta_j}-e^{\iota\ell\theta}\right|\\
&\quad+\max_{t\in\mathcal{T}_T}\frac{2\pi}{2m_T+1}\sum_{j=-m_T}^{m_T}\max_{\theta_{j-1}\leq\theta\leq\theta_j}|\sigma^{\chi}_{ik}(t/T;\theta_j)-\sigma^{\chi}_{ik}(t/T;\theta)| = O_{\mathrm{P}}(\vartheta^{1/2}_{n,T,r^*})+O(m_T^{-1}),
\end{aligned} \tag{A.63}$$

as $n, T \to \infty$, uniformly in i and k. For proving (A.63) we used (A.60) for the first term on the right-hand side, Assumption (B1) and the fact that the exponential function has bounded variation for the second, and Lemma 1*(ii)* for the third, which implies that the spectral density is Lipschitz continuous in θ uniformly in t. Moreover, the last term on the right-hand side of (A.63) is dominated by the first one because of Assumptions (F2) and (F3). Summing up, for any $\ell \in \mathbb{Z}$,

$$\max_{t\in\mathcal{T}_T}\left|\widehat{\gamma}^{\chi}_{ik;n,T}(t/T;\ell)-\gamma^{\chi}_{ik}(t/T;\ell)\right| = O_{\mathrm{P}}(\vartheta^{1/2}_{n,T,r^*}), \tag{A.64}$$

as $n, T \to \infty$, uniformly in i and k. This extends Proposition 8 in Forni et al. (2017) to the time-varying case.

(iii) - VAR filtering. Assuming that n factorizes, for some integer m, into $n = m(q+1)$, we estimate via Yule-Walker m distinct $(q+1)$-dimensional VAR models of order at most S (in view of Assumption

(D2)). For the sake of simplicity, let us assume $S = 1$: the Yule-Walker estimators of the VAR(1) coefficients (see also (20)) then are

$$\widehat{\mathbf{A}}^{(k)}_{n,T}(t/T) = \widehat{\mathbf{\Gamma}}^{\chi^{(k)}}_{n,T}(t/T;1)\left[\widehat{\mathbf{\Gamma}}^{\chi^{(k)}}_{n,T}(t/T;0)\right]^{-1}, \qquad 1 \le k \le m,$$

where $\widehat{\mathbf{\Gamma}}^{\chi^{(k)}}_{n,T}(t/T;\ell)$ is the $(q+1)\times(q+1)$ sub-matrix of $\widehat{\mathbf{\Gamma}}^{\chi}_{n,T}(t/T;\ell)$ corresponding to the lag ℓ autocovariance matrix of the sub-vector $\boldsymbol{\chi}^{(k)}_{n,T;t/T}$.

Since $\widehat{\mathbf{\Gamma}}^{\chi^{(k)}}_{n,T}(t/T;0)$ is finite-dimensional, (A.64) implies that it is consistent uniformly in t; together with Assumption (D4), it also implies that $\det[\widehat{\mathbf{\Gamma}}^{\chi^{(k)}}_{n,T}(t/T;0)] > d/2$, uniformly in t, with probability arbitrarily close to one for T large enough. The same arguments as in Appendix C of Forni et al. (2017) and (A.64) then entail, as $n, T \to \infty$,

$$\max_{t\in\mathcal{T}_T}\max_{1\le k\le m}\left\|\widehat{\mathbf{A}}^{(k)}_{n,T}(t/T) - \mathbf{A}^{(k)}_{n}(t/T)\right\| = O_{\mathrm{P}}(\vartheta^{1/2}_{n,T,r^*}). \tag{A.65}$$

Moreover, denoting by $\widehat{\mathbf{A}}_{n,T}(t/T)$ the $n\times n$ block-diagonal matrix having diagonal blocks $\widehat{\mathbf{A}}^{(k)}_{n,T}(t/T)$ for $1 \le k \le m$, we have, from (A.65),

$$\max_{t\in\mathcal{T}_T}\frac{1}{\sqrt{n}}\left\|\widehat{\mathbf{A}}_{n,T}(t/T) - \mathbf{A}_{n}(t/T)\right\| = O_{\mathrm{P}}(\vartheta^{1/2}_{n,T,r^*}) \tag{A.66}$$

and

$$\max_{t\in\mathcal{T}_T}\left\|\boldsymbol{\ell}_i'\left(\widehat{\mathbf{A}}_{n,T}(t/T) - \mathbf{A}_{n}(t/T)\right)\right\| = O_{\mathrm{P}}(\vartheta^{1/2}_{n,T,r^*}) \tag{A.67}$$

as $n, T \to \infty$, uniformly in i, since by construction $\mathbf{A}_n(t/T)$ has only $m(q+1)^2$ non-zero entries. This extends Proposition 9 in Forni et al. (2017) to the time-varying setting.

We now establish the following two lemmas.

Lemma A4. *Under Assumptions (A) and (B), for any $n \in \mathbb{N}_0$,*
(i) $\max_{t\in\mathcal{T}_T}\max_{1\le i\le n}|X_{it}| = O_{\mathrm{P}}(\log^{1/\varphi} T)$ *and*
(ii) $\max_{t\in\mathcal{T}_T}\max_{1\le i\le n}|X_{it}(t/T)| = O_{\mathrm{P}}(\log^{1/\varphi} T)$
where φ is defined in Assumption (A5).

Proof. First notice that, because of Assumption (A5), for any $\varepsilon > 0$,

$$\mathrm{P}\left(\max_{t\in\mathcal{T}_T}\max_{1\le j\le q}|u_{jt}| > \varepsilon\right) \le Tq\mathrm{P}\left(|u_{jt}| > \varepsilon\right) \le K_u Tq\exp\left(-\varepsilon^{\varphi}K_u\right).$$

Hence, $\max_{t\in\mathcal{T}_T}\max_{1\le j\le q}|u_{jt}| = O_{\mathrm{P}}(\log^{1/\varphi} T)$. Likewise, since we assume $n = O(T^{\omega})$ for some $\omega > 0$, then $\max_{t\in\mathcal{T}_T}\max_{1\le j\le n}|\eta_{jt}| = O_{\mathrm{P}}(\log^{1/\varphi} T)$. The proof then follows from the absolute summability of the coefficients in (2), (3), (6), and (7) due to Assumptions (B1) and (B5). □

Lemma A5. *Under Assumptions (A) and (B), for any $n \in \mathbb{N}_0$,*

$$\max_{t\in\mathcal{T}_T}\max_{1\le i\le n}|X_{it} - X_{it}(t/T)| = O_{\mathrm{P}}(T^{-1}\log^{1/\varphi} T).$$

Proof. First let us show that

$$\max_{t\in\mathcal{T}_T}\max_{1\le i\le n}|\chi_{it} - \chi_{it}(t/T)| = O_{\mathrm{P}}(T^{-1}\log^{1/\varphi} T). \tag{A.68}$$

Without loss of generality, let us assume $q = 1$. From (2) and (6), for any $1 \le i \le n$ and $K \ge 0$,

$$|\chi_{it} - \chi_{it}(t/T)| \le \sum_{k=0}^{K}|c^*_{i1k}(t) - c_{i1k}(t/T)|\,|u_{t-k}| + \left|\sum_{k=K+1}^{\infty}(c^*_{i1k}(t) - c_{i1k}(t/T))u_{t-k}\right|.$$

Assumption (B1) implies that, for any $\varepsilon > 0$ and $\eta > 0$, there exists a constant $K^* = K(\varepsilon, \eta)$ independent of i, t, and T, such that

$$\mathrm{P}\left[\left|\sum_{k=K^*+1}^{\infty}(c^*_{i1k}(t) - c_{i1k}(t/T))u_{t-k}\right| > \eta/2\right] \leq \varepsilon/2.$$

Hence,

$$\begin{aligned}\mathrm{P}\left[|\chi_{it} - \chi_{it}(t/T)| > \eta\right] \leq & \,\mathrm{P}\left[\sum_{k=0}^{K^*}|c^*_{i1k}(t) - c_{i1k}(t/T)|\,|u_{t-k}| > \eta/2\right] \\ & + \mathrm{P}\left[\left|\sum_{k=K^*+1}^{\infty}(c^*_{i1k}(t) - c_{i1k}(t/T))u_{t-k}\right| > \eta/2\right] \\ \leq & \,\mathrm{P}\left[\sum_{k=0}^{K^*}|c^*_{i1k}(t) - c_{i1k}(t/T)|\,|u_{t-k}| > \eta/2\right] + \varepsilon/2. \end{aligned} \tag{A.69}$$

Now, from Assumption (B3), since $\rho_\chi < 1$,

$$\mathrm{P}\left[\sum_{k=0}^{K^*}|c^*_{i1k}(t) - c_{i1k}(t/T)|\,|u_{t-k}| > \eta/2\right] \leq \mathrm{P}\left[\frac{K^* C_\chi}{T}\max_{t\in\mathcal{T}_T}|u_t| > \eta/2\right]$$

where (see the proof of Lemma A4) $\max_{t\in\mathcal{T}_T}|u_t| = O_{\mathrm{P}}(\log^{1/\varphi} T)$. It follows that there exists $T^* = T(\varepsilon, \eta)$ independent of i and t such that

$$\mathrm{P}\left[\sum_{k=0}^{K^*}|c^*_{i1k}(t) - c_{i1k}(t/T)|\,|u_{t-k}| > \eta/2\right] \leq \varepsilon/2 \tag{A.70}$$

for all $T \geq T^*$; (A.68) follows from putting together (A.69) and (A.70). The proof of

$$\max_{t\in\mathcal{T}_T}\max_{1\leq i\leq n}|\xi_{it} - \xi_{it}(t/T)| = O_{\mathrm{P}}(T^{-1}\log^{1/\varphi} T)$$

follows along the same steps, using Assumption (B6). The claim follows. □

(iv) - Principal Component Analysis. Since, for simplicity, we assumed $S = 1$ in (21),

$$\widehat{\mathbf{Z}}_{nt}(t/T) = \left[\mathbf{I}_n - \widehat{\mathbf{A}}_{n,T}(t/T)L\right]\mathbf{X}_{nt}, \qquad t \in \mathcal{T}_T. \tag{A.71}$$

Defining

$$\widetilde{\mathbf{Z}}_{nt}(t/T) := \left[\mathbf{I}_n - \mathbf{A}_n(t/T)L\right]\mathbf{X}_{nt}, \qquad t \in \mathcal{T}_T, \tag{A.72}$$

it follows from (A.66) and Lemma A4 that, as $n, T \to \infty$ (note that the filters in (A.71) and (A.72) just load $(q+1)$ series at a time)

$$\max_{t\in\mathcal{T}_T}\frac{1}{\sqrt{n}}\left\|\widehat{\mathbf{Z}}_{nt}(t/T) - \widetilde{\mathbf{Z}}_{nt}(t/T)\right\| = O_{\mathrm{P}}(\vartheta^{1/2}_{n,T,r^*}\log^{1/\varphi} T). \tag{A.73}$$

Lemma A5 moreover implies that, as $n, T \to \infty$,

$$\max_{t\in\mathcal{T}_T}\frac{1}{\sqrt{n}}\left\|\widetilde{\mathbf{Z}}_{nt}(t/T) - \mathbf{Z}_{nt}(t/T)\right\| = O_{\mathrm{P}}(T^{-1}\log^{1/\varphi} T). \tag{A.74}$$

By combining (A.73) and (A.74), as $n, T \to \infty$, we get

$$\max_{t\in\mathcal{T}_T}\frac{1}{\sqrt{n}}\left\|\widehat{\mathbf{Z}}_{nt}(t/T) - \mathbf{Z}_{nt}(t/T)\right\| = O_{\mathrm{P}}(\vartheta^{1/2}_{n,T,r^*}\log^{1/\varphi} T). \tag{A.75}$$

Similarly, from (A.67) and Lemma A5,

$$\max_{t\in\mathcal{T}_T}\left\|\boldsymbol{\ell}_i'\left(\widehat{\mathbf{Z}}_{nt}(t/T)-\mathbf{Z}_{nt}(t/T)\right)\right\|=O_{\mathrm{P}}(\vartheta_{n,T,r^*}^{1/2}\log^{1/\varphi}T), \tag{A.76}$$

uniformly in i.

Next, consider the rolling estimator

$$\widehat{\boldsymbol{\Gamma}}_{n,T}^{Z}(t/T):=\frac{1}{M_T}\sum_{s=T_1(t/T)}^{T_2(t/T)}\mathbf{Z}_{ns}(t/T)\mathbf{Z}_{ns}'(t/T),\quad t\in\mathcal{T}_T, \tag{A.77}$$

based on the uniform kernel as in (24) (see also (22)) of the covariance matrix of the unobservable $\mathbf{Z}_{nt}(t/T)$ and similarly define the estimator

$$\widehat{\boldsymbol{\Gamma}}_{n,T}^{\widehat{Z}}(t/T):=\frac{1}{M_T}\sum_{s=T_1(t/T)}^{T_2(t/T)}\widehat{\mathbf{Z}}_{ns}(t/T)\widehat{\mathbf{Z}}_{ns}'(t/T),\quad t\in\mathcal{T}_T \tag{A.78}$$

of the covariance matrix of the estimated $\widehat{\mathbf{Z}}_{nt}(t/T)$. Comparing (A.77) with (A.78), we obtain

$$\frac{1}{n}\left\|\widehat{\boldsymbol{\Gamma}}_{n,T}^{\widehat{Z}}(t/T)-\widehat{\boldsymbol{\Gamma}}_{n,T}^{Z}(t/T)\right\|=\frac{1}{nM_T}\left\|\sum_{s=T_1(t/T)}^{T_2(t/T)}\left[\widehat{\mathbf{Z}}_{ns}(t/T)\widehat{\mathbf{Z}}_{ns}'(t/T)-\mathbf{Z}_{ns}(t/T)\mathbf{Z}_{ns}'(t/T)\right]\right\|. \tag{A.79}$$

By (A.75), the right-hand side of (A.79) can be bounded uniformly in $t\in\mathcal{T}_T$, so that, as $n,T\to\infty$,

$$\max_{t\in\mathcal{T}_T}\frac{1}{n}\left\|\widehat{\boldsymbol{\Gamma}}_{n,T}^{\widehat{Z}}(t/T)-\widehat{\boldsymbol{\Gamma}}_{n,T}^{Z}(t/T)\right\|=O_{\mathrm{P}}(\vartheta_{n,T,r^*}^{1/2}\log^{1/\varphi}T). \tag{A.80}$$

and similarly, by (A.67) and (A.76),

$$\max_{t\in\mathcal{T}_T}\frac{1}{\sqrt{n}}\left\|\boldsymbol{\ell}_i'\left(\widehat{\boldsymbol{\Gamma}}_{n,T}^{\widehat{Z}}(t/T)-\widehat{\boldsymbol{\Gamma}}_{n,T}^{Z}(t/T)\right)\right\|=O_{\mathrm{P}}(\vartheta_{n,T,r^*}^{1/2}\log^{1/\varphi}T), \tag{A.81}$$

uniformly in i. Now, let $\boldsymbol{\Gamma}_n^Z(t/T)$ be the time-varying covariance matrix of the filtered process $\mathbf{Z}_n(t/T)$ obtained from $\mathbf{X}_n(t/T)$ as defined in (5) and (6), with (i,j) entry $\gamma_{ij}^Z(t/T)$ and denote as $\widehat{\gamma}_{ij;T}^Z(t/T)$ the (i,j) entry of $\widehat{\boldsymbol{\Gamma}}_{n,T}^Z(t/T)$. Starting from Corollary 3.4 in Zhang and Wu (2019), because of Lemmas A1 and A2, we can follow the same steps as those used by those authors for the spectral density estimation and leading from Corollary 4.4 to Lemma 4 in that paper. As a result, it is possible to show that there exists a constant C^{**} (independent of T and n) such that, for any $n,T\in\mathbb{N}_0$ (recalling that $M_T=2\lfloor Tb_T\rfloor$),

$$\max_{1\le i,j\le n}\mathrm{E}\left[\sup_{\tau\in(0,1)}\left|\widehat{\gamma}_{ij;T}^Z(\tau)-\gamma_{ij}^Z(\tau)\right|^2\right]\le C^{**}\psi_{T,r^*}, \tag{A.82}$$

where

$$\psi_{T,r^*}:=\max\left(\frac{\log T}{Tb_T},\frac{T^{4/r^*}(\log T)^{4+4/r^*}}{T^2b_T^2},b_T^4\right).$$

Note that ψ_{T,r^*} is the maximum of three quantities. The first of them also appears in Rodríguez-Poo and Linton (2001, Proposition 3.2) and Motta et al. (2011, Theorem 1), while the third one is the square of the first quantity in (A.24).

Therefore, from (A.82), for any $n \in \mathbb{N}_0$,

$$\begin{aligned}
\mathrm{E}\left[\max_{t\in\mathcal{T}_T}\frac{1}{n^2}\left\|\widehat{\boldsymbol{\Gamma}}^Z_{n,T}(t/T)-\boldsymbol{\Gamma}^Z_n(t/T)\right\|^2\right] &\le \mathrm{E}\left[\max_{t\in\mathcal{T}_T}\frac{1}{n^2}\left\|\widehat{\boldsymbol{\Gamma}}^Z_{n,T}(t/T)-\boldsymbol{\Gamma}^Z_n(t/T)\right\|^2_F\right]\\
&= \frac{1}{n^2}\mathrm{E}\left[\max_{t\in\mathcal{T}_T}\sum_{i=1}^n\sum_{j=1}^n\left|\widehat{\gamma}^Z_{ij;T}(t/T)-\gamma^Z_{ij}(t/T)\right|^2\right]\\
&\le \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\mathrm{E}\left[\max_{t\in\mathcal{T}_T}\left|\widehat{\gamma}^Z_{ij;T}(t/T)-\gamma^Z_{ij}(t/T)\right|^2\right]\\
&\le C^{**}\psi_{T,r^*}.
\end{aligned}$$

Thus, by Chebychev's inequality and since $\psi_{T,r^*}=o(\zeta_{T,r^*})$ as $T\to\infty$,

$$\max_{t\in\mathcal{T}_T}\frac{1}{n}\left\|\widehat{\boldsymbol{\Gamma}}^Z_{n,T}(t/T)-\boldsymbol{\Gamma}^Z_n(t/T)\right\| = O_{\mathrm{P}}(\zeta^{1/2}_{T,r^*}). \tag{A.83}$$

Following similar steps as for (A.38), we also have

$$\max_{t\in\mathcal{T}_T}\frac{1}{\sqrt{n}}\left\|\boldsymbol{\ell}'_i\left(\widehat{\boldsymbol{\Gamma}}^Z_{n,T}(t/T)-\boldsymbol{\Gamma}^Z_n(t/T)\right)\right\| = O_{\mathrm{P}}(\zeta^{1/2}_{T,r^*}), \tag{A.84}$$

uniformly in i. Therefore, from (A.80), (A.81), (A.83), and (A.84),

$$\max_{t\in\mathcal{T}_T}\frac{1}{n}\left\|\widehat{\boldsymbol{\Gamma}}^{\widehat{Z}}_{n,T}(t/T)-\boldsymbol{\Gamma}^Z_{n,T}(t/T)\right\| = O_{\mathrm{P}}(\vartheta^{1/2}_{n,T,r^*}\log^{1/\varphi}T) \tag{A.85}$$

and

$$\max_{t\in\mathcal{T}_T}\frac{1}{\sqrt{n}}\left\|\boldsymbol{\ell}'_i\left(\widehat{\boldsymbol{\Gamma}}^{\widehat{Z}}_{n,T}(t/T)-\boldsymbol{\Gamma}^Z_{n,T}(t/T)\right)\right\| = O_{\mathrm{P}}(\vartheta^{1/2}_{n,T,r^*}\log^{1/\varphi}T), \tag{A.86}$$

as $n,T\to\infty$, uniformly in i. This proves the analogue of Lemma 11 in Forni et al. (2017).

Now, while we used (A.37) and (A.40) to estimate the dynamic model (8)-(9) via dynamic principal components, estimating the static model (15) via principal components can be achieved using (A.85) and (A.86). In particular, following similar steps as those leading to (A.41) and (A.42) and in view of Assumption (E) and Lemma 3, we can prove that

$$\max_{t\in\mathcal{T}_T}\frac{1}{n}\left\|\widehat{\boldsymbol{\Gamma}}^{\widehat{Z}}_{n,T}(t/T)-\boldsymbol{\Gamma}^{\psi}(t/T)\right\| = O_{\mathrm{P}}(\vartheta^{1/2}_{n,T,r^*}\log^{1/\varphi}T)$$

and

$$\max_{t\in\mathcal{T}_T}\frac{1}{\sqrt{n}}\left\|\boldsymbol{\ell}'_i\left(\widehat{\boldsymbol{\Gamma}}^{\widehat{Z}}_{n,T}(t/T)-\boldsymbol{\Gamma}^{\psi}(t/T)\right)\right\| = O_{\mathrm{P}}(\vartheta^{1/2}_{n,T,r^*}\log^{1/\varphi}T)$$

as $n,T\to\infty$, uniformly in i.

Then, let $\mathbf{M}^{\psi}(t/T)$ be the $q\times q$ diagonal matrix with the q largest eigenvalues of $\boldsymbol{\Gamma}^{\psi}_n(t/T)$ and $\mathbf{V}^{\psi}_n(t/T)$ the $n\times q$ matrix of the corresponding normalized eigenvectors. Similarly let $\widehat{\mathbf{M}}^{\widehat{Z}}_{n,T}(t/T)$ be the $q\times q$ diagonal matrix with entries the q largest eigenvalues of $\widehat{\boldsymbol{\Gamma}}^Z_{n,T}(t/T)$ and $\widehat{\mathbf{V}}^{\widehat{Z}}_{n,T}(t/T)$ the $n\times q$ matrix of the corresponding normalized eigenvectors. Following similar arguments as those leading to (A.45) and (A.50), we have

$$\max_{t\in\mathcal{T}_T}\frac{1}{n}\left\|\widehat{\mathbf{M}}^{\widehat{Z}}_{n,T}(t/T)-\mathbf{M}^{\psi}(t/T)\right\| = O_{\mathrm{P}}(\vartheta^{1/2}_{n,T,r^*}\log^{1/\varphi}T) \tag{A.87}$$

and

$$\max_{t\in\mathcal{T}_T}\sqrt{n}\left\|\boldsymbol{\ell}'_i\left(\widehat{\mathbf{V}}^{\widehat{Z}}_{n,T}(t/T)-\mathbf{V}^{\psi}(t/T)\mathbf{S}(t/T)\right)\right\| = O_{\mathrm{P}}(\vartheta^{1/2}_{n,T,r^*}\log^{1/\varphi}T) \tag{A.88}$$

as $n, T \to \infty$, uniformly in i, where $\mathbf{S}(t/T)$ is a $q \times q$ diagonal matrix with entries $s_j(t) = \pm 1$.

Now, by Assumption (A1), we have $\mathrm{E}[\mathbf{u}_t \mathbf{u}_t'] = \mathbf{I}_q$. Therefore, for all $\tau \in (0, 1)$, the common component of the static factor model (15) has covariance $\boldsymbol{\Gamma}_n^{\psi}(\tau) = \mathbf{R}_n(\tau)\mathbf{R}_n'(\tau)$ and, by construction, we have $\mathbf{R}_n(\tau) := \mathbf{V}_n^{\psi}(\tau)[\mathbf{M}^{\psi}(\tau)]^{1/2}$, where $\mathbf{M}^{\psi}(\tau)$ is the $q \times q$ diagonal matrix with the q largest eigenvalues of $\boldsymbol{\Gamma}_n^{\psi}(\tau)$ and $\mathbf{V}_n^{\psi}(\tau)$ the $n \times q$ matrix of the corresponding normalized eigenvectors. Since, by definition, $\widehat{\mathbf{R}}_n(t/T) := \widehat{\mathbf{V}}_{n,T}^{\widehat{Z}}(t/T)[\widehat{\mathbf{M}}_{n,T}^{\widehat{Z}}(t/T)]^{1/2}$, from (A.87) and (A.88), it follows that

$$\max_{t \in \mathcal{T}_T} \left\| \boldsymbol{\ell}_i' \left(\widehat{\mathbf{R}}_{n,T}(t/T) - \mathbf{R}_n(t/T)\mathbf{S}(t/T) \right) \right\| = O_{\mathrm{P}}(\vartheta_{n,T,r^*}^{1/2} \log^{1/\varphi} T) \tag{A.89}$$

as $n, T \to \infty$, uniformly in i.

Finally, recall the definitions

$$\mathbf{C}_n(t/T; L) := [\mathbf{A}_n(t/T; L)]^{-1}\mathbf{R}_n(t/T) \quad \text{and} \quad \widehat{\mathbf{C}}_{n,T}(t; L) := [\widehat{\mathbf{A}}_{n,T}(t/T; L)]^{-1}\widehat{\mathbf{R}}_{n,T}(t/T)$$

of the impulse response functions and their estimators. Since $[\widehat{\mathbf{A}}_{n,T}(t/T; L)]$ is block-diagonal, so is $[\widehat{\mathbf{A}}_{n,T}(t/T; L)]^{-1}$. For any given $1 \leq i \leq n$, label as k_i the block containing the ith component x_{it} of $\mathbf{x}_{nt}$, with $1 \leq k_i \leq m$ and let $\mathcal{I}_i := \{h \in \mathbb{N}_0 : (k_i - 1)(q + 1) + 1 \leq h \leq k_i(q + 1)\}$ denote the set of indexes corresponding to those components of $\mathbf{x}_{nt}$ belonging to block k_i. Letting $\boldsymbol{m}_j$ stand for the jth vector in the q-dimensional canonical basis (the vector with 1 in entry j and 0 elsewhere), the (i, j) entry of $\widehat{\mathbf{C}}_{n,T}(t; L)$ is

$$\widehat{c}_{ij;n,T}(t; L) = \sum_{h \in \mathcal{I}_i} \boldsymbol{\ell}_i' [\widehat{\mathbf{A}}_{n,T}(t/T; L)]^{-1} \boldsymbol{\ell}_h \boldsymbol{\ell}_h' \widehat{\mathbf{R}}_{n,T}(t/T)\boldsymbol{m}_j;$$

note that the sum is only over $(q + 1)$ elements, which is finite for any $n \in \mathbb{N}_0$. Therefore, we can use (A.67) and (A.89) to show that, for any given lag $k \geq 0$,

$$\max_{t \in \mathcal{T}_T} |\widehat{c}_{ijk;n,T}(t) - s_j(t)c_{ijk}(t/T)| = O_{\mathrm{P}}(\vartheta_{n,T,r^*}^{1/2} \log^{1/\varphi} T), \tag{A.90}$$

as $n, T \to \infty$, uniformly in i and j. Moreover, from Assumption (B3), for any given $k \geq 0$, as $T \to \infty$,

$$\sup_{i \in \mathbb{N}_0} \max_{1 \leq j \leq q} \max_{t \in \mathcal{T}_T} \left| c_{ijk}^*(t) - c_{ijk}(t/T) \right| \leq C_\chi \rho_\chi^k / T = O(T^{-1}). \tag{A.91}$$

Combining (A.90) and (A.91) completes the proof. □

References

Dahlhaus, R. (1996). Asymptotic statistical inference for nonstationary processes with evolutionary spectra. In P. Robinson and M. Rosenblatt (Eds.), *Athens Conference on Applied Probability and Time Series Analysis; Lecture Notes in Statistics*, Volume 115, pp. 145–159. Springer.

Dahlhaus, R. (2012). Locally stationary processes. In T. Subba Rao, S. Subba Rao, and C. R. Rao (Eds.), *Handbook of Statistics: Time Series Analysis, Methods and Applications*, Volume 30, pp. 351–413. Elsevier.

Forni, M., M. Hallin, M. Lippi, and P. Zaffaroni (2017). Dynamic factor models with infinite-dimensional factor spaces: Asymptotic analysis. *Journal of Econometrics 199*(1), 74–92.

Motta, G., C. M. Hafner, and R. von Sachs (2011). Locally stationary factor models: Identification and nonparametric estimation. *Econometric Theory 27*(6), 1279–1319.

Rodríguez-Poo, J. M. and O. Linton (2001). Nonparametric factor analysis of residual time series. *Test 10*(1), 161–182.

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

Yu, Y., T. Wang, and R. J. Samworth (2015). A useful variant of the Davis–Kahan theorem for statisticians. *Biometrika 102*, 315–323.

Zhang, D. and W. B. Wu (2019). Convergence of covariance and spectral density estimates for high dimensional locally stationary processes. Working Paper, UCSD.