Biometrika (2023), **110**, 4, *pp.* 1077–1098 https://doi.org/10.1093/biomet/asad018
Advance Access publication 15 March 2023

√2-estimation for smooth eigenvectors of matrix-valued functions

By GIOVANNI MOTTA

Department of Statistics, Texas A&M University,
155 Ireland Street, College Station, Texas 77843-3143, U.S.A.
g.motta@stat.tamu.edu

WEI BIAO WU

Department of Statistics, University of Chicago,
5734 S. University Avenue, Chicago, Illinois 60637, U.S.A.
wbwu@galton.uchicago.edu

and MOHSEN POURAHMADI

Department of Statistics, Texas A&M University,
3143 TAMU, College Station, Texas 77843-3143, U.S.A.
pourahm@stat.tamu.edu

Summary

Modern statistical methods for multivariate time series rely on the eigendecomposition of matrix-valued functions such as time-varying covariance and spectral density matrices. The curse of indeterminacy or misidentification of smooth eigenvector functions has not received much attention. We resolve this important problem and recover smooth trajectories by examining the distance between the eigenvectors of the same matrix-valued function evaluated at two consecutive points. We change the sign of the next eigenvector if its distance with the current one is larger than the square root of 2. In the case of distinct eigenvalues, this simple method delivers smooth eigenvectors. For coalescing eigenvalues, we match the corresponding eigenvectors and apply an additional signing around the coalescing points. We establish consistency and rates of convergence for the proposed smooth eigenvector estimators. Simulation results and applications to real data confirm that our approach is needed to obtain smooth eigenvectors.

Some key words: Local stationarity; Matrix-valued function; Spectral decomposition; Spectral density matrix.

1. Introduction

It is well known that unit-norm eigenvectors of a time-invariant matrix are unique up to a ± sign; see Wicklin (2017). Albeit different software packages, such as SAS, MATLAB®, R and Mathematica®, might produce different eigenvectors for the same input matrix, this indeterminacy can be resolved by fixing the first row of the eigenvector matrix to be positive; see Lawley & Maxwell (1971, p. 18). Unfortunately, when dealing with matrix-valued functions, following this simple convention leads to unsmooth eigenvector functions.

A matrix-valued function is a path of matrices $A(x)$ whose entries depend on a real variable $x \in \mathbb{R}$; see, e.g., Bunse-Gerstner et al. (1991). The second source of misidentification is the phenomenon of eigenvalue functions, eigencurves from here on, intersecting or coalescing at a point x_c say, in which case the corresponding eigenvector functions, eigenfunctions from here on, cannot be identified uniquely due to the lack of unique ordering of the eigencurves at the coalescing point x_c.

For a symmetric matrix-valued function $A(x)$, we denote by $\lambda_1(x), \ldots, \lambda_N(x)$ its possibly unordered eigencurves and by $\ell_1(x) \geqslant \cdots \geqslant \ell_N(x)$ the ordered eigenvalues. The eigenfunctions corresponding to the eigencurves are denoted by $\Pi(x)$, whereas eigenvectors corresponding to the ordered eigenvalues are denoted by $P(x)$. Then $A(x)$ admits two possibly different spectral decompositions:

$$A(x) = \Pi(x)\Lambda(x)\Pi(x)^{\mathrm{T}} = \sum_{j=1}^{N} \lambda_j(x)\Pi_j(x)\Pi_j(x)^{\mathrm{T}}$$

$$= P(x)L(x)P(x)^{\mathrm{T}} = \sum_{j=1}^{N} \ell_j(x)P_j(x)P_j(x)^{\mathrm{T}}.$$

Here $\Lambda(x) = \mathrm{diag}\{\lambda_1(x), \ldots, \lambda_N(x)\}$, $L(x) = \mathrm{diag}\{\ell_1(x), \ldots, \ell_N(x)\}$, $P(x) = [P_1(x)| \cdots |P_N(x)]$ and $\Pi(x) = [\Pi_1(x)| \cdots |\Pi_N(x)]$ are orthonormal matrices for each x, $\lambda_j(x)$ and $\Pi_j(x)$ are smooth functions of x and $\ell_1(x) \geqslant \cdots \geqslant \ell_N(x)$ are ordered values of the eigencurves $\lambda_1(x), \ldots, \lambda_N(x)$. Note that $\ell_j(x)$ and $P_j(x)$ can be obtained from conventional eigendecomposition of $A(x)$ and they are generally not differentiable in x. Moreover, unless $\ell_j(x) \equiv \lambda_j(x)$, the unordered eigencurves $\lambda_j(x)$ and the corresponding eigenfunctions $\Pi_j(x)$ cannot be obtained from conventional eigendecomposition of $A(x)$.

The two sources of misidentification, discontinuity and mismatch, that we deal with in this paper, are the choices of signing and swapping the eigenfunctions of $\{A(x),\ x \in \mathbb{R}\}$. Let $\Pi_{1j}(x)$ be the first entry of $\Pi_j(x)$ and assume that the sign of $\Pi_{1j}(x)$ changes around x_0. Forcing $\Pi_{1j}(x)$ to be positive, as in the case of a single matrix A, would reduce its degree of smoothness around x_0. The second source of misidentification is due to the lack of a unique ordering/ranking of the eigenvalues in the presence of coalescing/crossing points, that is, in the case of local multiplicity. The eigencurves $\lambda_1(x), \ldots, \lambda_N$ of $A(x)$ can be ranked at each x in decreasing order. However, due to changing order at the coalescing points, the ensuing trajectories of $\Lambda(x)$ and $\Pi(x)$ may not inherit the degree of smoothness of $A(x)$. Consequently, one needs to consider all possible combinations, or swappings, of the eigenvectors corresponding to different rankings of the eigenvalues.

In fact, the interactive effects of both sources, signing and swapping, translates into a high-dimensional combinatorial problem that renders estimation infeasible even for small-dimensional time series. If c denotes the number of coalescing points ($c \geqslant 0$), the total number of possible combinations of signed eigenvectors is $2^{N(c+1)}$, whereas the total number of possible combinations of swapped eigenvectors is $N!^{(c+1)}$. In the Supplementary Material we provide more details and numerical examples about the combinatorial nature of signing and swapping. To properly introduce the problem we solve in this paper, we also need to clarify its mathematical root. To this end, we distinguish between distinct and coalescing eigencurves.

Let $A(x) \in \mathcal{C}^k(\mathbb{R}, \mathbb{C}_H^{N\times N})$, where $\mathcal{C}^k(\mathbb{R}, \mathbb{C}_H^{N\times N})$ denotes the set of Hermitian $N \times N$ matrices that are k times differentiable functions of the real variable x. Dieci & Eirola (1999,

Proposition 2.4) proved that if the eigencurves are distinct, there exist $P, L \in \mathcal{C}^k$ unitary and diagonal, respectively, such that $P(x)^* A(x) P(x) = L(x)$ for all x, where M^* denotes the conjugate transpose of matrix M. However, although mathematically such a smooth $P(x)$ does exist, the identification up to sign renders the eigenvectors computationally unsmooth, even in the case of distinct eigenvalues. In this paper, our key idea for solving this problem is to change the sign of the jth eigenvector $P_j(x_{t+1})$ if $\|P_j(x_{t+1}) - P_j(x_t)\| > \sqrt{2}$, where the x_t are grid points of a partitioning of the unit interval. In the case of distinct eigenvalues, this simple method delivers smooth eigenvectors; see § 2.2. In the case of coalescing eigenvalues, elaborating on the condition in (1) below, we use a similar distance inequality that delivers smooth eigenvectors.

Suppose now that the eigencurves $\lambda_1(x), \ldots, \lambda_N(x)$ of $A(x)$ intersect at some x, and let $\Pi_1(x), \ldots, \Pi_N(x)$ be the corresponding eigenfunctions. The eigencurves can be ordered, but, due to changing the order at each x, the ensuing trajectories $\ell_1(x) \geqslant \cdots \geqslant \ell_N(x)$ may not inherit the degree of smoothness of matrix $A(x)$. In spite of this, it turns out that the degree of smoothness of components of the spectral decomposition is influenced by the so-called eigengap or degree of separation of the eigencurves. In fact, Dieci & Eirola (1999) showed that the condition

$$\liminf_{\delta \to 0} \frac{|\lambda_i(x+\delta) - \lambda_j(x+\delta)|}{|\delta^d|} > 0 \quad \text{for some } d \leqslant k \text{ and for all } x \text{ and } i \neq j \tag{1}$$

guarantees that $\Pi(x)$ is in $\mathcal{C}^{k-d}$, where d is the order of contact of the eigencurves. The inequality in (1) suggests that, when two eigencurves approach a coalescing point with different slopes, there exists an orthonormal matrix $\Pi(x)$ that satisfies $\Pi(x)^* A(x) \Pi(x) = \mathrm{diag}\{\lambda_1(x), \ldots, \lambda_N(x)\}$ and is continuous in x. In order to estimate the continuous time-varying loadings, Motta et al. (2011) used the condition in (1) with $k = d = 1$. In this paper we go beyond continuity and estimate smooth eigenfunctions by matching them with the corresponding eigencurves around the estimated coelescing points; see § 2.3. We change the sign of the jth eigenvector $\{P_j(x), x > x_c\}$ after a coalescing point x_c if $\|P_j(x_c + \delta) - P_j(x_c - \delta)\| > \sqrt{2}$.

Prominent examples of dealing with eigendecompositions of matrix-valued functions in the context of dynamic factor models are the lag-zero covariance matrix, a function on rescaled time $[0, 1]$, of a locally stationary multivariate time series as in Motta et al. (2011), and the spectral density matrix, a function on frequency $[-\pi, \pi]$, of a multivariate stationary time series as in Forni et al. (2017). The phenomenon of coalescing eigencurves and the ensuing identification problem of the eigenfunctions have been overlooked in the statistics literature and has not received the attention it deserves, with the exception of Motta et al. (2011). Not surprisingly, the identification problem has a long history dating back to the work of von Neumann & Wigner (1929) in quantum mechanics, and is receiving growing interest in various areas of applied mathematics (Dieci & Eirola, 1999; Uhlig, 2020), where the focus is on coalescing eigencurves, but not on the corresponding eigenfunctions. In contrast to the recent work in mathematics, our focus is on estimating smooth eigenfunctions when the eigencurves are either well separated or intersect/coalesce. In the latter and more complicated case we investigate, depending on the number of coalescing points of the eigencurves, the smoothness or lack thereof of the corresponding eigenfunctions.

In this paper we deal with the smooth estimation of eigenvectors of two matrix-valued functions that are often employed in multivariate time series: the spectral density matrix and the time-varying covariance matrix.

In the first and stationary case, the matrix-valued function of interest is the spectral density $A(\omega) = \Sigma(\omega)$ defined in $\omega \in [-\pi, \pi]$. The importance of eigenvectors of the spectral density matrix has been emphasized by several authors; see, among others, Tiao & Box (1985, p. 815). N. R. Goodman, in his 1967 Air Force Technical Applications Center technical report, was the first to interpret and use eigenvalues and eigenvectors of spectral and sample-spectral density matrices of a multivariate stationary time series. R. R. Kneiper, E. S. Eby and H. S. Newman, in their 1970 Navy Underwater Sound Laboratory technical report, used eigenvalues and eigenvectors of the spectral density matrix to understand the beam patterns generated by the adaptive search and track array processor. Pignon (1981) estimated eigenvalues and eigenvectors of the spectral density matrix to localize sources and characterize propagating media with a short array of hydrophones. More recently, the resolving capability of passive array processing has been greatly improved by so-called high-resolution methods (Bienvenu, 1983). They are based on the eigenvalue-eigenvector decomposition of the spectral density matrix of the signals received on the sensors. The estimation of the dynamic factor model in econometrics relies on the continuity of eigenvalues and eigenvectors of the estimated frequency-varying spectral matrix; see Forni et al. (2000, p. 545).

In the second and locally stationary case, the matrix-valued function of interest is the time-varying covariance matrix $A(u) = \Gamma(u)$ defined in rescaled time $u \in [0, 1]$ by Dahlhaus (2000). Rodríguez-Poo & Linton (2001) used the eigenvalues of the estimated $\Gamma(u)$ to test whether the covariance matrix is of reduced rank. Motta et al. (2011) used the eigenvectors of the same matrix to estimate the loadings of a locally stationary factor model. Reiss & Winkelmann (2021) studied the rank of the time-varying instantaneous covariance matrix of a multi-dimensional continuous semimartingale.

The entries of the local covariance $\Gamma(u)$ are real-valued curves. In the case of the spectral matrix $\Sigma(\omega)$, with spectral decomposition

$$\Sigma(\omega) = P(\omega)L(\omega)P(\omega)^*, \qquad \omega \in (0, 1), \tag{2}$$

the identification problem is even more complicated because the entries of the eigenvectors $P_j(\omega)$ are complex valued. To see this, write $P(\omega) = P_R(\omega) + \mathrm{i}\, P_I(\omega)$, P_R and P_I being the real and imaginary parts of P, respectively, and $\mathrm{i} = \sqrt{-1}$. Thus, $\Sigma(\omega) = \Sigma_R(\omega) + \mathrm{i}\, \Sigma_I(\omega)$, where

$$\Sigma_R(\omega) = P_R(\omega)L(\omega)P_R(\omega)^{\mathrm{T}} + P_I(\omega)L(\omega)P_I(\omega)^{\mathrm{T}}$$
$$\text{and} \quad \Sigma_I(\omega) = P_I(\omega)L(\omega)P_R(\omega)^{\mathrm{T}} - \{P_I(\omega)L(\omega)P_R(\omega)^{\mathrm{T}}\}^{\mathrm{T}}$$

with M^{T} denoting the transpose of matrix M. Matrix $\Sigma_R(\omega)$ is invariant to changes in the sign of the eigenvectors $\{P_j(\omega), j = 1, \ldots, N\}$. However, if we change the sign of a column of $P_R(\omega)$, then we have to change the sign of the corresponding column of $P_I(\omega)$ in order to satisfy the spectral decomposition.

Throughout the paper, we denote by I_m the identity matrix of size m, by O the null matrix, by $\mathrm{tr}(A)$ the trace of A, by $\|A\|$ the Frobenius norm $\|A\| = \{\mathrm{tr}(A^{\mathrm{T}}A)\}^{1/2}$ and by $\rho(A)$ the spectral norm, that is, the largest singular value of A. We denote by A_j the jth columns of matrix $A = [A_1|A_2|\cdots|A_j|\cdots]$, where $A_j|A_k$ means that the columns A_j and A_k are contiguous. For matrix A of size $M \times N$, we denote by $A_{p\times q}$ the $p \times q$ matrix collecting the first p rows, from top to bottom, and the first q columns, from left to right, of matrix A, where $p \leqslant M$ and $q \leqslant N$. The `R` code of our approach is available at `https://github.com/giovanni-motta/`.

2. Nonparametric estimation and $\sqrt{2}$-smoothing

2.1. *Two issues in estimating smooth eigenfunctions*

Let $\Gamma(u)$ be the $N \times N$ contemporaneous time-varying covariance matrix and consider the spectral decompositions

$$\Gamma(u) = \Pi(u)\Lambda(u)\Pi(u)^{\mathrm{T}} = \sum_{j=1}^{N} \lambda_j(u)\Pi_j(u)\Pi_j(u)^{\mathrm{T}} \tag{3}$$

$$= P(u)L(u)P(u)^{\mathrm{T}} = \sum_{j=1}^{N} \ell_j(u)P_j(u)P_j(u)^{\mathrm{T}}, \tag{4}$$

where $\Lambda(u) = \mathrm{diag}\{\lambda_1(u), \ldots, \lambda_N(u)\}$, $L(u) = \mathrm{diag}\{\ell_1(u), \ldots, \ell_N(u)\}$, $P(u) = [P_1(u)|\cdots|P_N(u)]$ and $\Pi(u) = [\Pi_1(u)|\cdots|\Pi_N(u)]$ are orthonormal matrices for each u, $\lambda_j(\cdot)$ and $\Pi_j(\cdot)$ are smooth $\mathbb{C}^2[0,1]$ functions and $\ell_1(u) \geqslant \cdots \geqslant \ell_N(u)$ are ordered values of the latent eigencurves $\lambda_1(u), \ldots, \lambda_N(u)$. Note that $\ell_j(u)$ and $P_j(u)$ can be obtained from the conventional eigendecomposition of $\Gamma(u)$, but they are possibly not differentiable in u. Thus, one cannot directly obtain smooth eigenvalue estimates for $\ell(u)$, respectively $\lambda_j(u)$, and $P_j(u)$, respectively $\Pi_j(u)$, from the conventional decomposition of a local estimate $\widetilde{\Gamma}(u)$ of $\Gamma(u)$.

To estimate $\Gamma(u)$, we consider the locally weighted cross-products

$$\widetilde{\Gamma}(u) = \sum_{t=1}^{T} Y_t Y_t^{\mathrm{T}} w_t(u), \qquad w_t(u) = K_h(u - t/T) \Big/ \sum_{i=1}^{T} K_h(u - i/T), \tag{5}$$

where $K_h(x) = K(x/h)/h$, K is a symmetric kernel function with support $[-1, 1]$ and the bandwidth sequence $h = h_T \to 0$, satisfying $Th \to \infty$ as $T \to \infty$. Zhang & Wu (2021) discussed convergence properties of (5) for high-dimensional locally stationary processes. The conventional eigendecomposition of $\widetilde{\Gamma}(u)$ yields

$$\widetilde{\Gamma}(u) = \widetilde{P}(u)\widetilde{L}(u)\widetilde{P}(u)^{\mathrm{T}} \quad \text{with} \quad \widetilde{L}(u) = \mathrm{diag}\{\widetilde{\ell}_1(u), \ldots, \widetilde{\ell}_N(u)\}, \tag{6}$$

where $\widetilde{\ell}_1(u) \geqslant \cdots \geqslant \widetilde{\ell}_N(u)$ and $\widetilde{P}(u) = [\widetilde{P}_1(u)|\cdots|\widetilde{P}_N(u)]$ is an orthonormal matrix. In this paper we modify $\widetilde{L}(u)$ and $\widetilde{P}(u)$ in (6) and obtain smooth estimates for $\lambda_j(\cdot)$ and $\Pi_j(\cdot)$, respectively. For $j = 1, \ldots, r$, our goal is to obtain estimates $\widetilde{\lambda}_j(\cdot)$ smooth and close to $\lambda_j(\cdot)$, as well as estimates $\widetilde{\Pi}_j(\cdot)$ that are smooth, close to $\Pi_j(\cdot)$ and orthonormal: $\widetilde{\Pi}_j(u)^{\mathrm{T}}\widetilde{\Pi}_k(u) = \mathbb{1}_{\{j=k\}}$.

Our estimation approach and asymptotic theory cover both cases of fixed N and $N \to \infty$. For the latter case, we assume that only $r \ll N$ eigenvalues are diverging with N, whereas the remaining $N - r$ stay bounded. For the sake of readability and conciseness, we focus on estimating the r smooth eigenfunctions $\Pi_1(u), \ldots, \Pi_r(u)$ corresponding to the largest r eigencurves $\lambda_1(u), \ldots, \lambda_r(u)$. For this reason, we need to assume that the eigencurves $\lambda_j(u)$ are smooth in u only for $j \leqslant r$; see Condition 1 in § 3. We also assume that the largest r eigencurves are well separated from the remaining $N - r$; see Condition 2(iv) in § 3. The functions $\ell_1(u) \geqslant \cdots \geqslant \ell_r(u)$ in (4) are ordered versions of the possibly unordered $\lambda_1(u), \ldots, \lambda_r(u)$ in (3), and thus the $\ell_j(u)$ may not be differentiable. To establish the relationship between the $\lambda_j(u)$ and the $\ell_j(u)$, we introduce the permutation function $\psi(u) = \{\psi_1(u), \ldots, \psi_r(u)\}$,

which takes values as a permutation of $\{1, \ldots, r\}$, that is,

$$\lambda_{\psi_j(u)}(u) = \ell_j(u) \quad \text{for all } u \in (0,1). \tag{7}$$

Let $0 < u_1 < \cdots < u_c < 1$ be points where the r largest eigencurves coalesce. Namely,

$$\{u_1, u_2, \ldots, u_c\} = \{\text{ordered } u\colon \exists\, i < j \leqslant r\colon \ell_i(u) = \ell_j(u)\} \tag{8}$$

or, equivalently, since the ℓ_i are ordered,

$$\{u_1, u_2, \ldots, u_c\} = \{\text{ordered } u\colon \exists\, i \leqslant r-1\colon \ell_i(u) = \ell_{i+1}(u)\},$$

and therefore $\psi(u)$ is piecewise constant with $\psi(u) = \psi(u')$ when both $u, u' \in (u_k, u_{k+1})$.

Even in the absence of coalescing points, the sign of the eigenfunctions needs to be properly chosen to guarantee sufficiently smooth trajectories, that is, to preserve local smoothness. In the presence of a coalescing point, in addition to the sign selection the eigenfunctions need to be properly labelled and matched with the corresponding eigencurves. Moreover, another change of sign might be required to guarantee sufficiently smooth trajectories, that is, to preserve global smoothness. In § 2.2 below we estimate the eigenvectors in the case of distinct eigenvalues, whereas in § 2.3 below we estimate the eigenfunction when the eigencurves intersect.

The possibly unmatched eigenvectors corresponding to the ordered eigenvalues ℓ_j are denoted P_j, whereas matched eigenfunctions corresponding to the eigencurves λ_j are denoted Π_j. In the case where the eigenvalues are distinct for all u, $L(u) \equiv \Lambda(u)$ and we obtain smooth eigenfunctions $\widetilde{\Pi}_j$ by simply $\sqrt{2}$-signing the $\widetilde{P}_j$. In the case of coalescing eigenvalues, an initial ranking is necessary to obtain eigenfunctions matched with the corresponding unordered eigencurves. Then the $\sqrt{2}$-signing plus an additional $\sqrt{2}$-bridging around the coalescing points are needed to obtain smooth eigenfunctions $\widetilde{\Pi}_j$.

2.2. *Estimating smooth eigenvectors in the case of distinct eigenvalues*

Assume that, for all $u \in [0,1]$, $\lambda_1(u) > \cdots > \lambda_r(u) > \lambda_{1+r}(u)$. Then we have $\psi_j(u) = j$, $1 \leqslant j \leqslant r$, in view of (7) and $\lambda_j(u) = \ell_j(u)$. To overcome the sign indeterminacy of $P_j(u)$, by the smoothness assumption of $\Pi_j(u)$, we can recover the latter setting $\Pi_j(0) = P_j(0)$ and

$$\Pi_j(v_{g+1}) = P_j(v_{g+1}) \times \operatorname*{arg\,min}_{\iota=\pm 1} \|\iota P_j(v_{g+1}) - P_j(v_g)\| \qquad (g = 0, 1, \ldots, G-1), \tag{9}$$

where $v_g = g/G$ and $G \in \mathbb{N}$ is the number of grid points of the unit interval. In practice, we can choose $G = T$. For any two unit length vectors x and y, $\|x - y\|^2 = 2 - 2x^{\mathrm{T}}y$, so $\|x - y\| > \|x + y\|$ if and only if $\|x - y\| > \sqrt{2}$ or $x^{\mathrm{T}}y < 0$. Thus, $\iota = -1$, or $\|P_j(v_{g+1}) - P_j(v_g)\| > \| - P_j(v_{g+1}) - P_j(v_g)\|$ in (9), is equivalent to $\|P_j(v_{g+1}) - P_j(v_g)\| > \sqrt{2}$. That is, $\iota = \operatorname{sign}\{\cos\theta\} = \operatorname{sign}\{P_j(v_{g+1})^{\mathrm{T}} P_j(v_g)\}$, where θ is the angle between $P_j(v_{g+1})$ and $P_j(v_g)$. Consequently, the signing operation in (9) makes $\Pi_j(u)$ vary smoothly in u.

Given the multivariate N-dimensional time series $X_1, \ldots, X_T$, we estimate the localized covariance $\Gamma(u)$ in (3) by means of $\widetilde{\Gamma}(u)$ in (5). Based on (9), we propose the estimate

$$\widetilde{\Pi}_j(v_{g+1}) = \widetilde{P}_j(v_{g+1}) \times \operatorname*{arg\,min}_{\iota=\pm 1} \|\iota \widetilde{P}_j(v_{g+1}) - \widetilde{P}_j(v_g)\| \qquad (g = 0, 1, \ldots, G-1). \tag{10}$$

To illustrate how our novel criterion works, we define, for all $u \in (0,1)$, the $r \times r$ diagonal matrix $\Lambda(u) = \{\lambda_j(u)\}_{j=1}^r$ as

$$\lambda_j(u) = 0.2(2-j) + (3-j)(u^2 - u^3) \qquad (j = 1, \ldots, r). \tag{11}$$

Then we simulate the N-dimensional vectors of frequencies $\omega_N = [\omega_1, \ldots, \omega_N]^{\mathrm{T}}$ and phases $\varphi_N = [\varphi_1, \ldots, \varphi_N]^{\mathrm{T}}$ according to two independent uniform distributions $\omega_n \sim U(0.6, 0.9)$, $\varphi_n \sim U(0.1, 0.3)$, and we define the $N \times r$ matrix $Q(u)$ whose (n,j) entry is

$$Q_{nj}(u) = \begin{cases} \sqrt{2}\sin(j\pi\omega_n u - \phi_n)/N, & n \text{ odd}, \\ \sqrt{2}\cos(j\pi\omega_n u - \phi_n)/N, & n \text{ even}, \end{cases}$$

for all $n \leqslant N$, all $j \leqslant r$ and all $u \in (0,1)$. Then we define, for all u, the $N \times r$ matrix

$$\underset{N\times r}{\Pi}(u) = \underset{N\times r}{Q}(u)\,\{\underset{r\times r}{Q}(u)^{\mathrm{T}}\underset{r\times r}{Q}(u)\}^{-1/2}, \tag{12}$$

such that $\Pi_{N\times r}(u)^{\mathrm{T}}\Pi_{N\times r}(u) = \mathrm{I}_r$ for all u. Next, we define the $N \times 1$ vector $X_t = \Pi_{N\times r}(t/T)f_t$, where f_t is a zero-mean random vector of size r defined as $f_t = \Lambda_{r\times r}(t/T)^{1/2}\varepsilon_t$, where $\varepsilon_t \sim \mathcal{N}(0, \mathrm{I}_r)$ and $E(\varepsilon_t\varepsilon_s^{\mathrm{T}}) = \mathbb{1}_{\{s=t\}}\mathrm{I}_r$, with covariance matrix $\Gamma^f(t/T) = E(f_t f_t^{\mathrm{T}}) = \Lambda_{r\times r}(t/T)$, so that covariance matrix of X_t is

$$\Gamma_N^X\left(\frac{t}{T}\right) = E(X_t X_t^{\mathrm{T}}) = \Pi_{N\times r}\left(\frac{t}{T}\right)\Lambda_{r\times r}\left(\frac{t}{T}\right)\Pi_{N\times r}\left(\frac{t}{T}\right)^{\mathrm{T}}. \tag{13}$$

For all $u \in (0,1)$, the estimator $\widetilde{\Gamma}(u)$ of $\Gamma_N^X(u)$ in (13) is computed according to (5), and the matrices $\widetilde{P}(u)$ and $\widetilde{L}(u)$ of its spectral decomposition are defined according to (6). Since the eigenvalues are not coalescing, in this example we only need to fix the eigenvectors. From $\widetilde{P}(u)$ obtained according to (6), we estimate the smooth eigenvectors $\widetilde{\Pi}_{N\times r}(u)$ according to (10). We show an example with $N = 10$ and $r = 2$ in Fig. 1. The matrices $\widetilde{L}_{r\times r}(u)$, $\widetilde{\Gamma}_N^X(u)$, $\widetilde{P}_{N\times r}(u)$ and $\widetilde{\Pi}_{N\times r}(u)$ are shown in Figs. 1(a), 1(b), 1(c) and 1(d), respectively. The curves of $\Pi_{N\times r}(u)$ simulated according to (12) are smooth by construction, whereas the curves of matrix $\widetilde{P}_{N\times r}(u)$ computed according to (6) and obtained from $\widetilde{\Gamma}_N^X(u)$ are identifiable only up to sign. Figure 1(c) shows that the indeterminacy $P_j(u) = \pm\Pi_j(u)$ at each $u \in (0,1)$ makes the estimated loadings unsmooth. To remove this indeterminacy, we implement (10); see Fig. 1(d).

2.3. *Estimating smooth eigenfunctions when eigenvalues coalesce*

Compared to the case of distinct eigenvalues, when eigenvalues coalesce, we need to (i) estimate the coalescing points $u_1, u_2, \ldots, u_c$ in (8) and (ii) match the eigenvectors with suitable eigencurves. In § 2.4 below we deal with the problem of estimating the coalescing points. Let $\mathcal{I}(\delta) = \bigcup_{k=1}^c [u_k - \delta, u_k + \delta]$, where $u_1 < u_2 < \cdots < u_c$, and $\mathcal{I}^\circ(\delta) = [0,1] - \mathcal{I}(\delta)$. For all $r \leqslant N$, we define the registration functions $\psi^\circ(u) = \{\psi_1^\circ(u), \ldots, \psi_r^\circ(u)\}$ of $\{1, 2, \ldots, r\}$ for $u \in \mathcal{I}^\circ(\delta)$ as the stepwise functions whose values do not change on $[u_k + \delta, u_{k+1} - \delta]$ $(k = 1, \ldots, c)$. Here we use the convention that $u_{1+c} = 1 + \delta$ so that the right end point $u_{1+c} - \delta = 1$. Assume that $1/G \ll \delta$ and $\delta \to 0$. We first state our algorithm in the population setting with known $\Gamma(\cdot)$ to obtain $\lambda(\cdot)$ and $\Pi(\cdot)$, and then we compute the sample

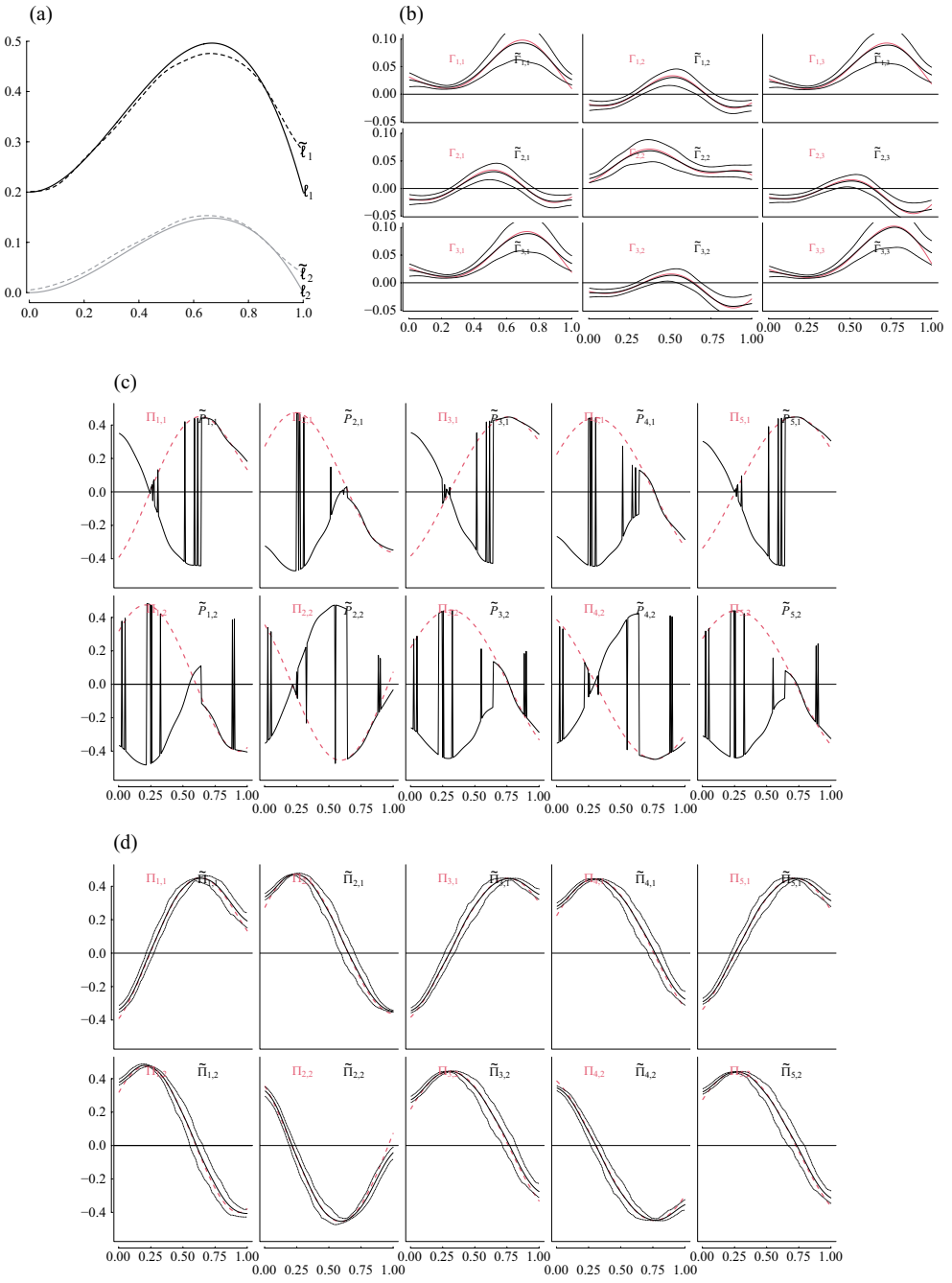

Fig. 1. Simulation scenario of §2.2 with $M = 100$ Monte Carlo replications, $N = 10$, $r = 2$ and $T = 400$. (a),(b) Eigenvalues and covariances. (c),(d) First five columns of the $r \times N$ matrices of eigenvectors; we only plot the first five among $N = 10$ columns. In (a) solid lines denote $\ell_1(u)$ and $\ell_2(u)$ defined in (11) and dashed lines denote averages, over the M replications, of $\tilde{\ell}_1(u)$ and $\tilde{\ell}_2(u)$ in (6). (b) Upper left 3×3 submatrix of the 10×10 time-varying covariance matrix. The dashed line denotes entries of $\Gamma(u)$ simulated according to (13). Solid lines denote averages, over the M replications, of the covariances estimated according to (5). Dotted lines denote 95% confidence bands obtained over the M replications. In (c) the dashed line denotes the entries of matrix $\Pi(u)$, simulated according to (12), and the solid line denotes one realization of matrix $\widetilde{P}(u)$ in (6), obtained with the statistical software R. In (d) the dashed line denotes the entries of matrix $\Pi(u)$, simulated according to (12), the solid line denotes the average, over the M replications, of estimator $\widetilde{\Pi}(u)$ obtained according to (10) from matrix $\widetilde{P}(u)$ in (6) and the dotted lines denote 95% quantiles, obtained from the M replications, of $\widetilde{\Pi}(u)$.

counterparts $\widetilde{\lambda}(\cdot)$ and $\widetilde{\Pi}(\cdot)$. The theoretical optimal choice of δ is given in Theorems 1 and 2. To compute the registration function, we proceed progressively in the following way.

Step 1. For all $u \in [0, u_1 - \delta]$ and all $j = 1, \ldots, r$, let $\psi_j^\circ(u) = j$, $\lambda_j(u) = \ell_j(u)$, and apply (9) to obtain $\Pi_j(v_g)$ for $0 \leqslant g \leqslant (u_1 - \delta)G$.

Step 2. For $j \leqslant r$ and $k = 1, \ldots, c-1$, define the registration function for the jth eigencurve

$$\psi_j^\circ(u_k + \delta) = \arg\min_{i \leqslant r} \min_{\iota = \pm 1} \|\iota P_j(u_k + \delta) - P_{\psi_i^\circ(u_k - \delta)}(u_k - \delta)\| \tag{14}$$

and, for all $u \in [u_k + \delta, u_{k+1} - \delta]$, define

$$\psi_j^\circ(u) = \psi_j^\circ(u_k + \delta), \tag{15}$$

$$\lambda_{\psi_j^\circ(u)}(u) = \ell_j(u), \tag{16}$$

$$\Pi_{\psi_j^\circ(u)}(u_k + \delta) = P_j(u_k + \delta) \times \arg\min_{\iota = \pm 1} \|\iota P_j(u_k + \delta) - \Pi_{\psi_j^\circ(u)}(u_k - \delta)\|, \tag{17}$$

and, for g with $u_k + \delta \leqslant g/G \leqslant u_{k+1} - \delta$,

$$\Pi_{\psi_j^\circ(u)}(v_{g+1}) = P_j(v_{g+1}) \times \arg\min_{\iota = \pm 1} \|\iota P_j(v_{g+1}) - \Pi_{\psi_j^\circ(u)}(v_g)\|. \tag{18}$$

Step 3. For $k = c$, we obtain $\psi_j(u_k + \delta)$ according to (14) and, for all $u \in [u_k + \delta, 1]$, we define $\psi_j^\circ(u)$, $\lambda_{\psi_j^\circ(u)}(u)$ and $\Pi_{\psi_j^\circ(u)}(u)$ according to (15), (16) and (18), respectively.

It is important to note that the permutation function in (7), which establishes the relationship between the eigencurves $\lambda_{\psi_j(u)}(\cdot)$ and the eigenvalues $\ell_j(\cdot)$, depends neither on the eigenvectors nor the eigenfunctions. However, we are using the smoothness of the eigenfunctions $\Pi_j(u)$ to determine the registration function in (16), and establish the relationship between the eigencurves $\lambda_{\psi_j^\circ(u)}(\cdot)$ and the eigenvalues $\ell_j(\cdot)$ via the eigenvectors $P_j(\cdot)$ by (14). In (9) we set initialization $\Pi_j(0) = P_j(0)$. Equation (17) determines the starting eigenfunction $\Pi_{\psi_j^\circ(u)}(u_k + \delta)$ over the interval $[u_k + \delta, u_{k+1} - \delta]$ by aligning with the ending eigenfunction $\Pi_{\psi_j^\circ(u)}(u_k - \delta)$ in the previous interval $[u_{k-1} + \delta, u_k - \delta]$, thus ensuring the closeness between them. Equation (17) is the $\sqrt{2}$-bridging, whereas (18) is the $\sqrt{2}$-signing.

In Fig. 2 we present an example of $N = 3$ eigenvalues that intersect at $c = 2$ coalescing points, and illustrate how eigencurves can be recovered according to (14) and (16). The eigenfunctions obtained according to (17) are correctly matched with the corresponding coalescing eigencurves, but not smooth. Therefore we apply the $\sqrt{2}$-signing in (18) to obtain smooth eigenfunctions $\Pi_{\psi_j^\circ}(u)$. Then, for all $j \leqslant r$ and all $u \in (0, 1)$, define $\Pi_j^\circ(u) = \Pi_{\psi_j^\circ(u)^{-1}}(u)$, and let

$$\Pi^\circ(u) = [\Pi_1(u) | \cdots | \Pi_r(u)], \qquad u \in (0, 1). \tag{19}$$

In Fig. 3 we use the same example as in Fig. 2, and illustrate how eigenfunctions can be recovered according to (17), (18) and (19).

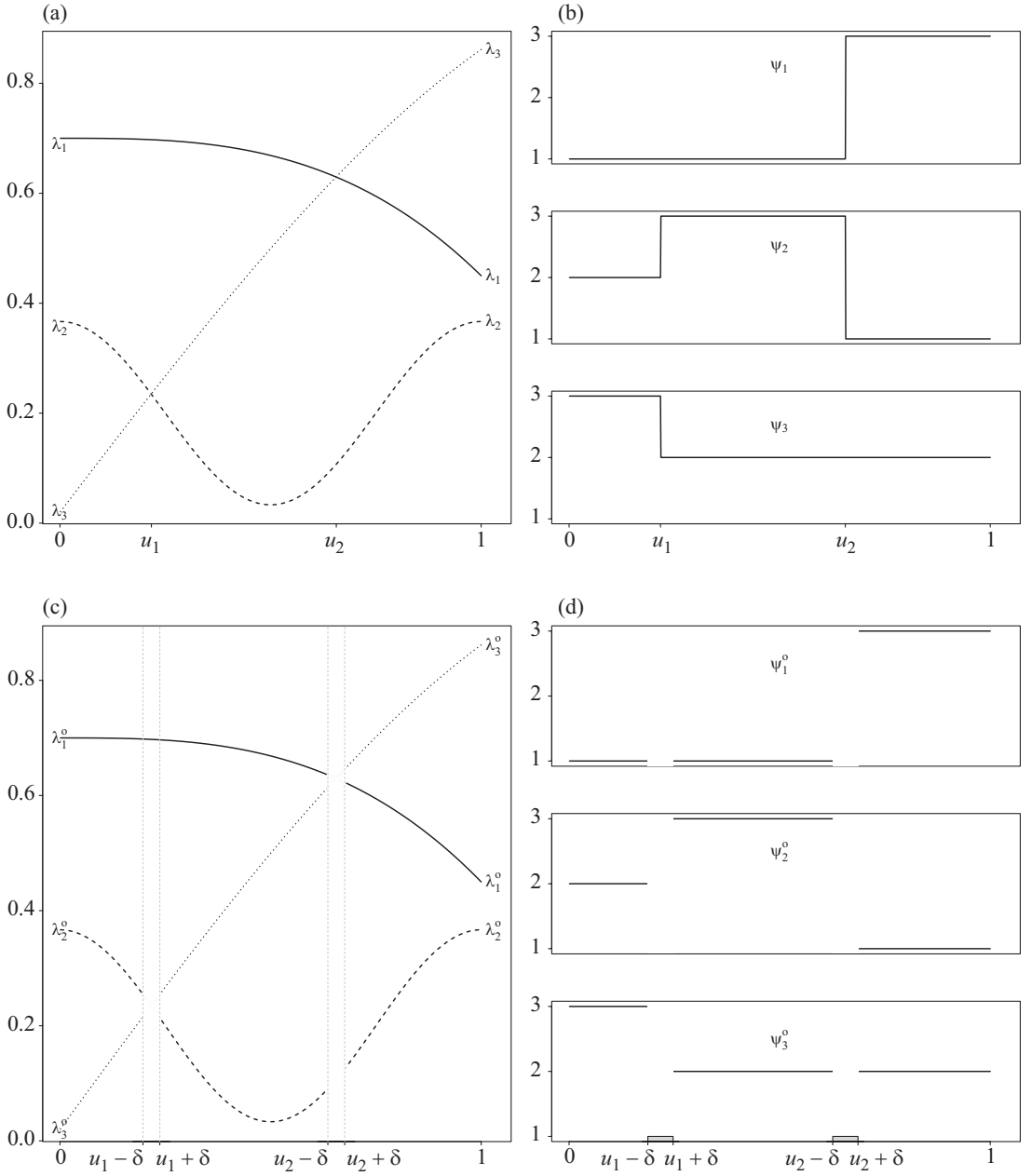

Fig. 2. Eigencurves defined according to (7) and recovered according to the definitions in (14), (15) and (16). A numerical example with $N = 3$ and $c = 2$. (a) An example of coalescing eigenvalues with $c = 2$, $u_1 = 0.21$ and $u_2 = 0.65$. (b) The corresponding registration functions defined in (7). (c) Eigencurves recovered according to (16). (d) Registration functions defined in (14) and (15).

We now compute the sample counterparts of eigencurves and eigenfunctions obtained according to Steps 1–3 above. Given $X_1, \ldots, X_T$, we compute the covariance matrix $\widetilde{\Gamma}(u)$ using (5). Recall that $\rho(\cdot)$ is the spectral norm, and write $\tilde{\lambda}^* = \max_{0 \leqslant u \leqslant 1} \rho\{\widetilde{\Gamma}(u)\}$. Denote by a_T the upper bound of $\max_u \rho\{\widetilde{\Gamma}(u) - \Gamma(u)\}/\tilde{\lambda}^*$, with $a_T \to 0$ as $T \to \infty$; see (23) in § 2.4 below.

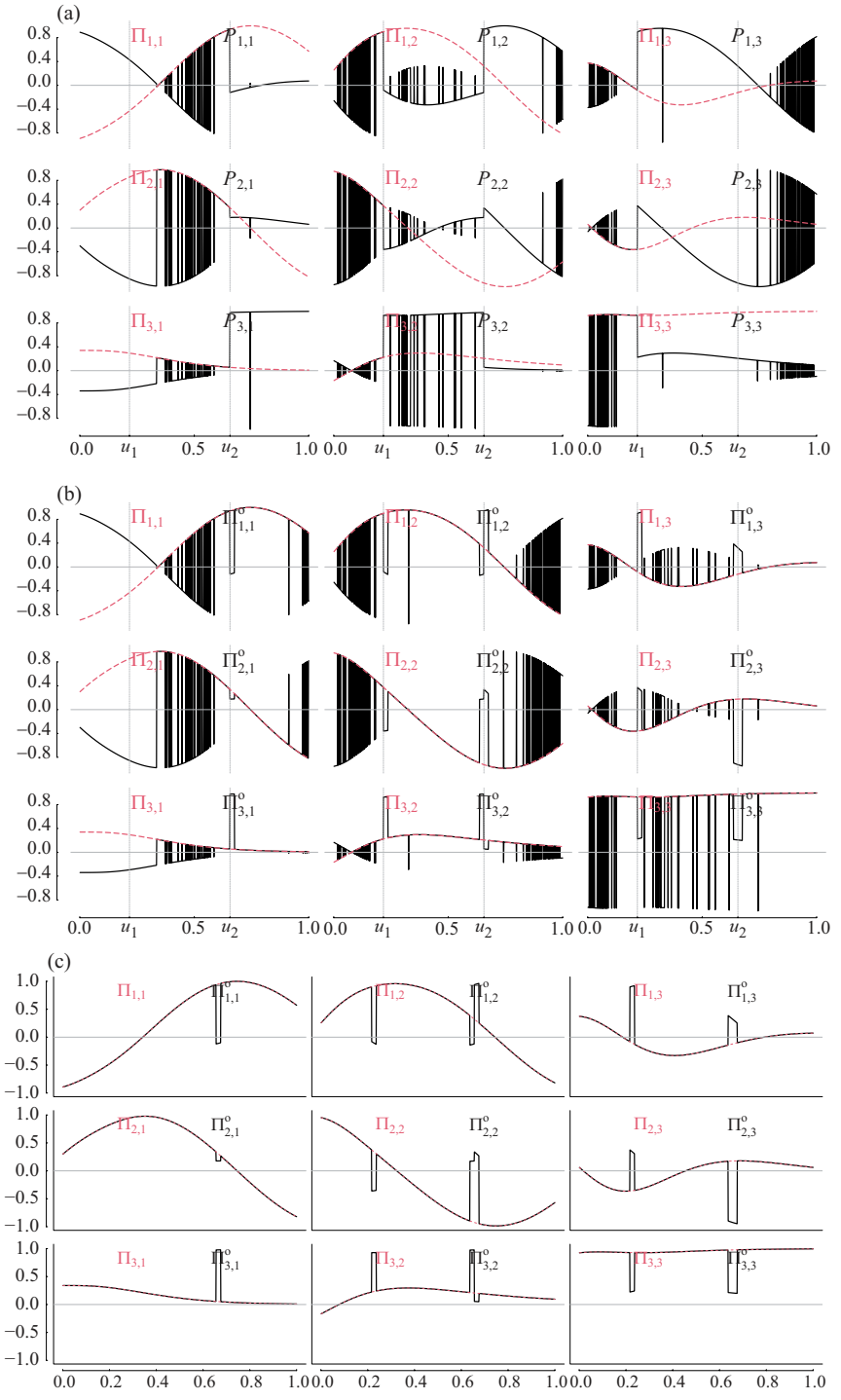

Fig. 3. Eigenfunctions recovered according to the definitions in (17), (18) and (19). A numerical example of the orthonormal matrix $\Pi(u) \in \mathbb{C}^2[0,1]$ in (3) with $N = 3$ and $c = 2$. In (a) dashed lines denote entries of the orthonormal matrix $\Pi(u) \in \mathbb{C}^2[0,1]$ in (3) and solid lines denote entries of $P(u)$ in (4) obtained with the statistical software R. In (b) dashed lines denote entries of the orthonormal matrix $\Pi(u) \in \mathbb{C}^2[0,1]$ in (3) and solid lines denote entries of $\Pi_{\psi_j^\circ(u)}(u)$ in (17): √2-bridging only. In (c) dashed lines denote entries of the orthonormal matrix $\Pi(u) \in \mathbb{C}^2[0,1]$ in (3) and solid lines denote entries of matrix $\Pi^\circ(u)$ in (19) obtained according to (17) and (18): √2-bridging and √2-signing.

Let $\hat{u}_1 < \cdots < \hat{u}_c$ be the estimated coalescing points, define $\hat{\mathcal{I}}(\delta) = \bigcup_{k=1}^{c}[\hat{u}_k - \delta, \hat{u}_k + \delta]$ and choose δ such that

$$\left\{u \in (0,1): \min_{j \leqslant r} \frac{\widetilde{\ell}_j(u) - \widetilde{\ell}_{j+1}(u)}{\tilde{\lambda}^*} \leqslant 2a_T\right\} \subset \hat{\mathcal{I}}(\delta). \tag{20}$$

Then we can carry out Steps 1 and 2 above using the sample version $\widetilde{P}(u)$ and $\widetilde{L}(u)$ in (6) and obtain estimates $\widetilde{\psi}_j^{\circ}(u)$ and $\widetilde{\Pi}_j(u)$ that are defined on $\widetilde{\mathcal{I}}^{\circ}(\delta) = [0,1] - \widetilde{\mathcal{I}}(\delta)$. To define $\widetilde{\psi}_j^{\circ}(u)$ and $\widetilde{\Pi}_j(u)$ on $\widetilde{\mathcal{I}}(\delta)$, we can use interpolation with

$$\widetilde{\lambda}_j(u) = \sum_{g:\, v_g \in \mathcal{I}^o(\delta)} w_g(u)\widetilde{\lambda}_j^{\circ}(v_g) \Big/ \sum_{g:\, v_g \in \mathcal{I}^o(\delta)} w_g(u), \tag{21}$$

$$\widetilde{\Pi}_j(u) = \sum_{g:\, v_g \in \mathcal{I}^o(\delta)} w_g(u)\widetilde{\Pi}_j^{\circ}(v_g) \Big/ \sum_{g:\, v_g \in \mathcal{I}^o(\delta)} w_g(u), \tag{22}$$

where the weights $w_g(u)$ can be chosen according to cubic-regression spline estimation, as in Claeskens et al. (2009, § 2.3), or can be defined as rescaled kernel functions $w_g(u) = K\{(v_g - u)/h\}/h$ according to local-constant or local polynomial smoothing, as in Fan & Gijbels (1996). In Fig. 4 we use the same example of Fig. 2, and illustrate how the eigenvectors in Fig. 3 can be estimated according to (20) and (22).

Note that $\widetilde{\Pi}_j(u)$, $j = 1, \ldots, r$, are not necessarily orthonormal. We can apply the Gram–Schmidt orthogonalization by recursively calculating $\widetilde{\Pi}_1^{\dagger}(u) = \widetilde{\Pi}_1(u)$, and

$$\widetilde{\Pi}_k^{\dagger}(u) = \widetilde{\Pi}_k(u) - \sum_{j=1}^{k-1} \widetilde{\Pi}_k(u)^{\mathrm{T}}\widetilde{\Pi}_j^{\dagger}(u)\widetilde{\Pi}_j^{\dagger}(u)/\|\widetilde{\Pi}_j^{\dagger}(u)\|^2 \quad \text{for } k = 2, \ldots, c,$$

and obtain the orthonormal vectors as $\widetilde{\Pi}_k(u) = \widetilde{\Pi}_k^{\dagger}(u)/\|\widetilde{\Pi}_k^{\dagger}(u)\|$.

2.4. *Estimating the coalescing points*

Let $\lambda^* = \max_{u \in [0,1]} \ell_1(u)$, and let $a_T \to 0$ be a sequence such that

$$\mathrm{pr}(\mathcal{A}) \to 1, \quad \text{where} \quad \mathcal{A} = \left\{\max_{u \in [0,1]} \rho\{\widetilde{\Gamma}(u) - \Gamma(u)\}/\lambda^* \leqslant a_T\right\}. \tag{23}$$

In (23), the sequence a_T represents the uniform convergence rate of $\widetilde{\Gamma}(u)$ normalized by λ^*. For independent and individually distributed sub-Gaussian vectors $X_1, \ldots, X_T$, Koltchinskii & Lounici (2017) obtained a sharp bound for $\rho(\widetilde{\Gamma} - \Gamma)/\lambda^*$, where $\widetilde{\Gamma} = T^{-1}\sum_{i=1}^{T} X_i X_i^{\mathrm{T}}$ and $\Gamma = \mathbb{E}(X_i X_i^{\mathrm{T}})$. In such a special case a_T can be chosen as $C\max\{T^{-1}r(\Gamma), T^{-1/2}r(\Gamma)^{1/2}\}$ if the effective rank $r(\Gamma) = \mathrm{tr}(\Gamma)/\rho(\Gamma) \to \infty$. In the asymptotic theory in § 3 we establish (23) for dependent and nonstationary processes; see also the Supplementary Material, where an asymptotic expression for a_T is provided. Based on (23), we can estimate coalescing points. Let

$$\mathcal{W} := \left\{u \in (0,1): \min_{j \leqslant r} \frac{\widetilde{\ell}_j(u) - \widetilde{\ell}_{j+1}(u)}{\lambda^*} \leqslant 2a_T\right\}. \tag{24}$$

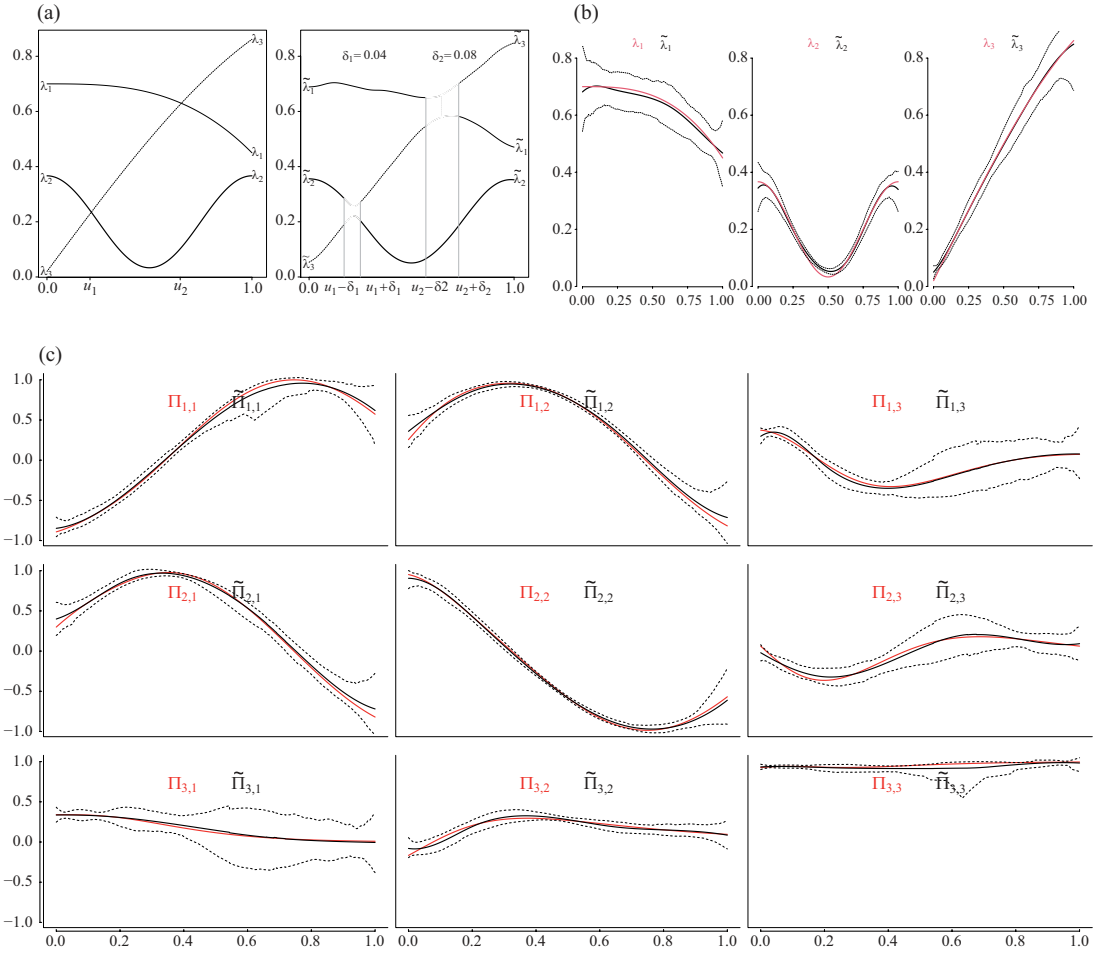

Fig. 4. Simulation scenario with $N = 3$, $c = 2$, $T = 1500$ and $M = 200$. In (a) the left panel shows an example of coalescing eigencurves with $c = 2$, $u_1 = 0.21$ and $u_2 = 0.65$ and the right panel shows the average, over the M replications, of the estimates $\widetilde{\lambda}_j^{\circ}$ $(j = 1, \ldots, r)$ obtained according to (16). In (b) the solid line denotes the average, over the M replications, of the estimates $\widetilde{\lambda}_j$ $(j = 1, \ldots, r)$ obtained according to (21) and the dashed lines denote 95% confidence intervals, computed over the M replications. (c) Comparing the simulated entries of matrix $\Pi(u)$ with the entries of the estimated matrix $\widetilde{\Pi}(u)$. Solid line denotes the average, over the M replications, of the estimates $\widetilde{\Pi}(u)$ obtained according to (22) and the dotted lines denote 95% quantiles obtained over M replications.

At a coalescing point u_i, we have $\ell_j(u_i) = \ell_{j+1}(u_i)$ for some $j \leqslant r$. By Weyl's theorem,

$$0 \leqslant \widetilde{\ell}_j(u_i) - \widetilde{\ell}_{1+j}(u_i) = \{\widetilde{\ell}_j(u_i) - \ell_j(u_i)\} - \{\widetilde{\ell}_{1+j}(u_i) - \ell_{1+j}(u_i)\} \leqslant 2\rho\{\widetilde{\Gamma}(u_i) - \Gamma(u_i)\}.$$

Then $\mathrm{pr}(u_i \in \mathcal{W}) \to 1$. We now consider noncoalescing points. Assume that Condition 2 in § 3 holds. Under (28), if $u \notin \mathcal{I}(\delta)$ is not a coalescing point with $\delta > 2a_T/\kappa$, we have, for all $j = 1, \ldots, r$,

$$\widetilde{\ell}_j(u) - \widetilde{\ell}_{j+1}(u) \geqslant \ell_j(u) - \ell_{j+1}(u) - 2a_T\lambda^* > \delta\kappa - 2a_T\lambda^* > 0.$$

Hence, $\mathrm{pr}\{\mathcal{W} \subset \mathcal{I}^{\circ}(\delta)\} = \mathrm{pr}\{\mathcal{I}(\delta) \subset [0,1] - \mathcal{W}\} \to 1$. Therefore, under event $\mathcal{A}$, there exist disjoint intervals $[s_i, t_i] \subset [0,1]$, $i = 1, \ldots, c$, with $m := \max_{i \leqslant c}(t_i - s_i) \to 0$ and $\mathcal{W} \subset$

$\bigcup_{i=1}^{c}[s_i, t_i]$. For a data-driven version, we can replace λ^* in (24) by $\tilde{\lambda}^* = \max_{0\leqslant u\leqslant 1}\rho\{\widetilde{\Gamma}(u)\}$, $\hat{u}_i = (t_i + s_i)/2$ and let $\delta = m/2$. A data-driven choice of a_T is given in the Supplementary Material, where we propose a bootstrap method based on (23) and (24).

3. Asymptotic estimation theory

3.1. *Spectral density and time-varying covariance*

In this section we provide an asymptotic theory for estimators of smooth eigenvector functions of spectral density matrices and time-varying covariance matrices. The key ingredient involves bounding the spectral norms of the errors of the spectral density matrix estimate and the time-varying covariance matrix estimate. We deal with them separately, since the former involves stationary processes while the latter is for locally stationary processes. Bai & Silverstein (2010), Tropp (2012), Vershynin (2012) and Koltchinskii & Lounici (2017) considered properties of sample covariance matrices for independent random vectors, and their results are therefore not directly applicable here. Fan et al. (2022) obtained a central limit theory for estimated eigenvalues and eigenvectors for random matrices with diverging spikes and independent noises. Recently, there have been some developments on dependent random vectors. Bhattacharjee & Bose (2016) and Wang et al. (2017) considered limiting behaviour of empirical spectral distributions for moving average processes. Heiny & Mikosch (2021) studied high-dimensional linear time series with infinite fourth moment, and showed that their spectral properties are very different from those of short-tailed processes. Bhattacharjee & Bose (2014) and Furmańczyk (2021) investigated spectral norm convergence of regularized covariance matrix estimates for high-dimensional linear time series, and Chen et al. (2013) considered high-dimensional nonlinear processes. Our problem differs from those approached so far because we are dealing with matrix-valued functions, and we need to establish a uniform convergence rate of $\max_u \rho\{\widetilde{\Gamma}^Y(u) - \Gamma^Y(u)\}$ or $\max_\omega \rho\{\widetilde{\Sigma}^Y(\omega) - \Sigma^Y(\omega)\}$, so that we can derive asymptotic properties of our smooth eigenvector function estimators.

3.2. *Spectral density matrix functions*

Consider the stationary factor model

$$\underset{N\times 1}{Y_t} = \underset{N\times q}{\Psi^X(B)}\,\underset{q\times 1}{f_t} + \underset{N\times N}{\Psi^Z(B)}\,\underset{N\times 1}{\eta_t} = X_t + Z_t, \tag{25}$$

where $\Psi^X(B) = \sum_{j=0}^{J}\Psi_j^X B^j$ and $\Psi^Z(B) = \sum_{k=0}^{K}\Psi_k^Z B^k$. Model (25) allows for dimension reduction as long as q is assumed to be smaller than N. Factors f_t and errors η_t are both zero mean and stationary, with spectral density matrices $\Sigma^f(\omega)$ and $\Sigma^\eta(\omega)$, respectively. If the common components X_t are orthogonal to the idiosyncratic components Z_t, that is, $Ef_t\eta_s^{\mathrm{T}} = \mathrm{O}$, $s(t = 1, \ldots, T)$, the spectral density matrix of Y_t is

$$\Sigma^Y(\omega) = \Sigma^X(\omega) + \Sigma^Z(\omega) = \Psi^X(\omega)\Sigma^f(\omega)\Psi^X(\omega)^* + \Psi^Z(\omega)\Sigma^\eta(\omega)\Psi^Z(\omega)^*, \qquad \omega \in (0, 1).$$

In general, the number q of latent factors in the dynamic factor model (25) is different, and typically smaller, than the number r of latent factors in a static factor model, such as the nonstationary model (31) we consider in the next section. Nevertheless, for the sake of readability and conciseness, in this paper we do not distinguish between q and r, and choose the letter r to denote the number of smooth eigenvectors we aim to estimate. This is

without loss of generality, since the methodology we propose in this paper works for both stationary-dynamic and nonstationary-static factor models. We emphasize that our method becomes useful whenever we are interested in estimating smooth eigenvectors, and it is not limited to the case where the underlying multivariate dataset follows a factor model. Recall (3) and (4), and write

$$\Sigma^Y(\omega) = \Pi(\omega)\Lambda(\omega)\Pi(\omega)^* = \sum_{j=1}^{N} \lambda_j(\omega)\Pi_j(\omega)\Pi_j(\omega)^*$$

$$= P(\omega)L(\omega)P(\omega)^* = \sum_{j=1}^{N} \ell_j(\omega)P_j(\omega)P_j(\omega)^*,$$

where $\ell_1(\omega) \geqslant \cdots \geqslant \ell_N(\omega)$ are ordered values of $\lambda_1(\omega), \ldots, \lambda_N(\omega)$. We are interested in the estimation of $\Pi_j(\omega)$ only for $j \leqslant r$, the eigenfunctions corresponding to the r largest eigencurves; see Conditions 1 and 2 below. Under the latter condition, $\ell_1(\omega) \geqslant \cdots \geqslant \ell_r(\omega)$ are ordered values of $\lambda_1(\omega), \ldots, \lambda_r(\omega)$. We estimate the permutation function $\psi(\omega) = \{\psi_1(\omega), \ldots, \psi_r(\omega)\}$, where $\psi(\omega)$ is a permutation of $\{1, 2, \ldots, r\}$ in (7), namely, $\{\ell_j(\omega) = \lambda_{\psi_j(\omega)}(\omega), j = 1, \ldots, r\}$, which implies the eigenvalue registration (Uhlig, 2020).

Assume that (Y_t) is a stationary zero-mean Gaussian vector process. Our asymptotic theory is applicable for bounded N as well as the large-dimensional case with $N \to \infty$ at an appropriate rate. Given the data $Y_1, \ldots, Y_T$, we define the sample counterpart of the spectral density matrix in (2) as the lag-window estimate

$$\widetilde{\Sigma}^Y(\omega) = \sum_{k=-B}^{B} w(|k|/B)\widetilde{\Gamma}_k\, e^{-2\pi\imath k\omega}, \tag{26}$$

where $\widetilde{\Gamma}_k = T^{-1}\sum_{t=1+k}^{T} Y_{t-k}Y_t^{\mathrm{T}}$ is the sample autocovariance matrix, with $\widetilde{\Gamma}_{-k} = \widetilde{\Gamma}_k^{\mathrm{T}}$ and $0 \leqslant k \leqslant T-1$, $w(x) = \max(0, 1-|x|)$ is the Bartlett kernel, $B = B_T = 1.1447\,(\alpha\, T)^{1/3}$ and α can be estimated according to Andrews (1991). To state our asymptotic results, we introduce the following technical conditions. They are not mathematically weakest. For the sake of readability and conciseness, we choose these forms. Condition 1 asserts smoothness of the eigenvector functions $\Pi_j(\cdot)$, while Condition 2 concerns separation of eigenvalue functions.

Condition 1. Assume that $\Pi_j(\omega) \in \mathbb{C}^2[0,1]$, $1 \leqslant j \leqslant r$.

Condition 2. Let $\lambda^* = \max_{j\leqslant r}\max_{0\leqslant\omega\leqslant 1}\lambda_j(\omega)$, and assume that the following conditions hold.

(i) The eigencurves are such that $\lambda_j(\omega) \in \mathbb{C}^2[0,1]$ $(j = 1, \ldots, r)$ and there exists a constant $\phi > 0$ such that $\max_{0\leqslant\omega\leqslant 1}|\lambda_j'(\omega)| \leqslant \lambda^*\phi$ for all $j = 1, \ldots, r$.

(ii) There exists a constant $\kappa > 0$ such that

$$\min_{0\leqslant\omega\leqslant 1}\liminf_{\epsilon\to 0}\min_{j<j'\leqslant r}\frac{|\lambda_j(\omega+\epsilon) - \lambda_{j'}(\omega+\epsilon)|}{\lambda^*|\epsilon|} \geqslant \kappa. \tag{27}$$

(iii) Let $\mathcal{I}(\delta) = \bigcup_{k=1}^{c}[\omega_k - \delta, \omega_k + \delta]$, where $\omega_1 < \omega_2 < \cdots < \omega_c$ are coalescing points defined in (8), and $\mathcal{I}^{\circ}(\delta) = [0,1] - \mathcal{I}(\delta)$. Assume that

$$\liminf_{\delta \downarrow 0} \min_{\omega \in \mathcal{I}^{\circ}(\delta)} \min_{j<j' \leqslant r} \frac{|\lambda_j(\omega) - \lambda_{j'}(\omega)|}{\lambda^* \delta} \geqslant \kappa. \tag{28}$$

(iv) There exists a constant $g > 0$ such that, for all $\omega \in [0,1]$, $\min_{j \leqslant r} \lambda_j(\omega) - \max_{j>r} \lambda_j(\omega) \geqslant g\lambda^*$. That is, the eigencurves $\{\lambda_j(\cdot)\}_{j=1}^{r}$ and $\{\lambda_j(\cdot)\}_{j=1+r}^{N}$ are separated by $g\lambda^*$.

We now discuss Condition 2. Note that (27) can be viewed as a normalized version, by the maximum eigenvalue λ^*, of the condition in (1) with $d = 1$. It describes the well separateness of the eigencurves $\lambda_j(\omega)$ around the coalescing points. If ω is not a coalescing point then the value of the $\liminf_{\delta \to 0}$ in (27) is ∞. If ω is a coalescing point of the eigencurves $\lambda_j(\cdot)$ and $\lambda_{j'}(\cdot)$ so that $\lambda_j(\omega) = \lambda_{j'}(\omega)$, then (27) implies that the normalized difference of derivatives $|\lambda_j'(\omega) - \lambda_{j'}'(\omega)|/\lambda^*$ is bounded below by κ. On the other hand, (28) indicates the well separateness of the eigencurves on the region outside the coalescing points. Condition 2(iv) asserts that the r largest eigencurve $\lambda_j(\omega)$ is well separated from the remaining $N - r$ eigencurves.

Under Condition 2, we define, according to § 2.3, the stepwise permutation functions $\psi(\omega) = \{\psi_1(\omega), \ldots, \psi_r(\omega)\}$ of $\{1, 2, \ldots, r\}$ for $\omega \in \mathcal{I}^{\circ}(\delta)$, whose values do not change on $[\omega_k + \delta, \omega_{k+1} - \delta]$, $k = 1, \ldots, c$. Let $\ell_1(\omega) \geqslant \cdots \geqslant \ell_r(\omega)$ be eigenvalues of $\Sigma(\omega)$, with eigenvectors $P_1(\omega), \ldots, P_r(\omega)$. Since $\ell_1(\omega) > \cdots > \ell_r(\omega)$ are mutually different, the gap is actually $\geqslant \lambda^* \kappa \delta$; for $\omega \in \mathcal{I}^{\circ}(\delta)$, the vectors $P_1(\omega), \ldots, P_r(\omega)$ are uniquely determined up to ± 1 signs.

Condition 3. For some $\alpha > 0$, $C > 0$, the bias of the lag-window estimator satisfies

$$\max_{0 \leqslant \omega \leqslant 1} \rho[E\{\widetilde{\Sigma}^{Y}(\omega)\} - \Sigma^{Y}(\omega)]/\lambda^* \leqslant CB^{-\alpha}.$$

Condition 3 concerns the bias of the lag-window estimate (26). Note that $E\{\widetilde{\Sigma}^{Y}(\omega)\} - \Sigma^{Y}(\omega) = \sum_{k=-B}^{B} r_k \Gamma_k e^{-2\pi \imath k\omega} - \sum_{|k|>B} \Gamma_k e^{-2\pi \imath k\omega}$, where $r_k = w(|k|/B)(1 - |k|/B) - 1$. As a sufficient condition, if $\rho(\Gamma_m)/\lambda^* = O(m^{-1-\alpha})$, since $r_k = O(|k|/B)$, Condition 3 holds in view of $\sum_{k=-B}^{B} r_k O(k^{-1-\alpha}) = O(B^{-\alpha})$ and $\sum_{|k|>B} O(k^{-1-\alpha}) = O(B^{-\alpha})$. For stationary causal vector autoregressive moving average processes, $\rho(\Gamma_m)$ decays to 0 geometrically quickly. For example, consider the var(1) process $Y_t = AY_{t-1} + \varepsilon_t$, where the ε_t are independent and individually distributed mean 0 vectors with covariance matrix Σ_ε and A is a coefficient matrix with spectral radius $r_A < 1$, guaranteeing the causality of Y_t with $Y_t = \sum_{b=0}^{\infty} A^b \varepsilon_{t-b}$. By Gelfand's formula (Dunford & Schwartz, 1963), $\lim_{m \to \infty} \rho(A^m)^{1/m} = r_A$. Thus, $\rho(A^m) \leqslant r^m$ for all sufficiently large m, where $r \in (r_A, 1)$. Consequently, $\Gamma_m = \sum_{b=0}^{\infty} A^b \Sigma_\varepsilon A^{m+b}$ has the property that $\rho(\Gamma_m)$ decays to 0 geometrically quickly.

Theorem 1. *Let* (Y_t) *in* (25) *be a mean* 0 *stationary Gaussian process. Assume that Conditions* 1, 2 *and* 3 *hold. Define* $\chi = \max_{j \leqslant r} \max_{0 \leqslant \omega \leqslant 1} \|\Pi_j'(\omega)\|$, *where* $\Pi_j'(\omega) = d\Pi_j(\omega)/d\omega$. *Let*

$$\theta = \{BT^{-1}(N + \log T)\}^{1/2} + B^{-\alpha},$$

and assume that $\delta = \delta_T$ *satisfies* $\delta \max(1, \chi) \to 0$ *and* $\theta/(\kappa\delta) \to 0$. *Then the estimator* $\widetilde{\Pi}_j(\omega)$ *converges to the orthonormal matrix* $\Pi_j(\omega)$ *defined according to* (3)*:*

$$\max_{j \leqslant r} \max_{0 \leqslant \omega \leqslant 1} \|\widetilde{\Pi}_j(\omega) - \Pi_j(\omega)\| = \mathcal{O}_{\mathbb{P}}\{\delta\chi + \theta/(\kappa\delta)\} \to 0. \tag{29}$$

In particular, choosing $B \asymp \{T/(N + \log T)\}^{1/(1+2\alpha)}$ *and* $\delta^2 \asymp \theta/(\chi\kappa) \to 0$, *we have*

$$\max_{j \leqslant r} \max_{0 \leqslant \omega \leqslant 1} \|\widetilde{\Pi}_j(\omega) - \Pi_j(\omega)\| = \mathcal{O}_{\mathbb{P}}[(\chi/\kappa)^{1/2}\{T^{-1}(N + \log T)\}^{\alpha/(2+4\alpha)}].$$

The Gaussian process assumption in Theorem 1 can be relaxed. In the Supplementary Material we show that Theorem 1 still holds for linear processes with sub-Gaussian innovations. In the special case that N is fixed and the lower-dimensional case with $N = O(\log T)$ and $\chi/\kappa = O(1)$, Theorem 1 gives the bound

$$\max_{j \leqslant r} \max_{0 \leqslant \omega \leqslant 1} \|\widetilde{\Pi}_j(\omega) - \Pi_j(\omega)\| = \mathcal{O}_{\mathbb{P}}\{(T^{-1} \log T)^{\alpha/(2+4\alpha)}\},$$

which is also new for eigenvector function estimation for spectral density matrix function for multivariate (finite)-dimensional time series.

3.3. *Locally stationary factor models*

Consider the reduced rank model

$$X_t = \underset{N \times r}{\Pi_{N \times r}\left(\frac{t}{T}\right)} \underset{r \times 1}{f_t} = \Pi_{N \times r}\left(\frac{t}{T}\right) \Lambda_{r \times r}\left(\frac{t}{T}\right)^{1/2} \underset{r \times 1}{\varepsilon_t}, \tag{30}$$

where $\Pi_{N \times r}(u) = [\Pi_1(u)|\cdots|\Pi_r(u)]$ collects the first r columns of the $N \times N$ orthonormal matrix $\Pi(u) \in \mathbb{C}^2[0, 1]$ in (3), $\Lambda_{r \times r}(u) = \operatorname{diag}\{\lambda_1(u), \ldots, \lambda_r(u)\}$ is the upper left $r \times r$ submatrix of $\Lambda(u)$ in (3) and $\varepsilon_t \sim \mathcal{N}(0, \mathrm{I}_r)$. If there exists a matrix-valued function $\Pi_{N \times r}(u)$ that is smooth in rescaled time $u \in (0, 1)$ then X_t is a multivariate locally stationary process; see Dahlhaus (2000). Extending the reduced rank model (30), we now consider the locally stationary factor model

$$Y_t = X_t + Z_t = \Pi_{N \times r}\left(\frac{t}{T}\right) f_t + Z_t, \tag{31}$$

where (Z_t) is a mean 0 Gaussian process. We are interested in estimating time-varying loadings $\Pi(u)$, $u \in [0, 1]$. Note that

$$\Gamma^X(u) = \sum_{k=1}^{r} \lambda_k(u) \Pi_k(u) \Pi_k(u)^{\mathrm{T}}. \tag{32}$$

Assume that the factor process (ε_t) and the error process (Z_t) are independent, and note that we can allow temporal dependencies in (ε_t) and (Z_t). Since (ε_t) and (Z_t) are independent, $\Gamma_N^Y(t/T) = E(Y_t Y_t^{\mathrm{T}}) = \Gamma^X(t/T) + \Gamma_0^Z$, where $\Gamma_k^Z = E(E_0 E_k^{\mathrm{T}})$. Eigenvalues and eigenvectors of $\Gamma(u) = \Gamma^X(u) + \Gamma_0^Z$ may not be $\lambda_k(u)$ and $\Pi_k(u)$ given in (32) due to the presence of Γ_0^Z. Nonetheless, we can use the local weighted estimate (33) below and study the distance (34),

thus obtaining a convergence result for estimates of $\Pi_k(u)$ based on the algorithms proposed in § 2.

Motta et al. (2011) proved that the first r columns of $\widetilde{P}(u)$ in (6), the sample version of $P(u)$ in (4), converge to a possibly unsmooth linear tranformation of the loadings $\Pi_{N\times r}(u)$ in (31). In this paper we estimate the smooth time-varying loadings $\Pi_1(\cdot), \ldots, \Pi_r(\cdot)$ from the data $Y_1, \ldots, Y_T$. We exploit the well-known fact that the r largest eigenvalues of the covariance matrix of N times series grow unboundedly as N increases, while the remaining $N-r$ eigenvalues remain bounded; see Chamberlain & Rothschild (1983) and Bai (2003). To determine the number r of factors of the approximate static factor model in (31), we adopt the estimator of Ahn & Horenstein (2013), which is obtained by maximizing the ratio of two adjacent eigenvalues.

In our theoretical result, we consider estimator (5) with the rectangle kernel, namely,

$$\widetilde{\Gamma}^{Y}(u) = \frac{1}{D_T(u)} \sum_{t=\lceil T(u-h)\rceil}^{\lfloor T(u+h)\rfloor} Y_t Y_t^{\mathrm{T}}, \quad \text{where} \quad D_T(u) = \lceil T(u+h)\rceil - \lfloor T(u-h)\rfloor + 1, \quad (33)$$

and $h = h_T$ is the bandwidth sequence satisfying $h_T \to 0$ and $Th_T \to \infty$. The above rectangle kernel setting can simplify the presentation of our asymptotic theory. We estimate $\Gamma_N^Y(u)$ by $\widetilde{\Gamma}^Y(u)$. Then we can estimate $\Lambda(u)$ and $\Pi(u)$, eigencurves and eigenfunctions of $\Gamma^X(u)$, from those of $\widetilde{\Gamma}^Y(u)$ via swapping and signing according to the methodology presented in § 2. To develop an asymptotic theory for the estimates, we use a similar argument as in the proof of Theorem 1. The key step is to establish relations (S26) and (S27) in the Supplementary Material for

$$\Delta = \max_{h\leqslant u\leqslant 1-h} \rho\{\widetilde{\Gamma}^Y(u) - \Gamma^X(u)\}. \quad (34)$$

Because of nonstationarity and the presence of (ε_t) and (Z_t) and the temporal dependencies in those processes, it is quite technical to obtain a tail probability bound for Δ. Because of the noise process (Z_t), eigenvectors of $\Gamma^X(u)$, namely $\Pi(u)$, are typically different from those of $\Gamma_N^Y(u)$. Additionally, due to the smoothing procedure, eigenvectors of $E\{\widetilde{\Gamma}^Y(u)\}$ and $\Gamma_N^Y(u)$ are also different. An upper bound of this bias is given in the Supplementary Material.

As in the treatment of the spectral density matrix case in § 3.2, the key issue is to bound

$$\Delta_0 = \max_u \rho[\,\widetilde{\Gamma}^Y(u) - E\{\widetilde{\Gamma}^Y(u)\}].$$

In the case of independent and identically distributed sub-Gaussian random vectors, Koltchinskii & Lounici (2017) considered the magnitude of the operator norms of the difference in the sample and the population covariance operators. Our setting here is more complicated since it involves dependent as well as nonstationary vectors. To account for dependence, let

$$\lambda_Z^* = \max_{0\leqslant\omega\leqslant 2\pi} \rho\{F_Z(\omega)\}, \quad \text{where} \quad F_Z(\omega) = \sum_{k=-\infty}^{\infty} \Gamma_k^Z e^{\iota k\omega},$$

and

$$\lambda_\varepsilon^* = \max_{0\leqslant\omega\leqslant 2\pi} \rho\{F_\varepsilon(\omega)\}, \quad \text{where} \quad F_\varepsilon(\omega) = \sum_{k=-\infty}^{\infty} \operatorname{cov}(\varepsilon_0, \varepsilon_k)\, e^{\imath k\omega}.$$

Define $\eta = \lambda_Z^*/(\lambda^*\lambda_\varepsilon^*)$, $\upsilon(m,h) = \max[\{m/(Th)\}^{1/2}, m/(Th)]$ and

$$\zeta(h) = [\eta + \sqrt{\eta} + \min\{1, (\chi+\phi)h\}]\upsilon(N + \log T, h) + \upsilon(\log T, h).$$

Intuitively, $\lambda^*\lambda_\varepsilon^*$ and λ_Z^* respectively quantify the magnitudes of the factor and error processes, and η quantifies the relative magnitude of the error process. Smaller η implies strong factors and $\eta = 0$ corresponds to the noiseless case with $Z_t = 0$.

THEOREM 2. *Assume that the factor process (ε_t) and the error process (Z_t) are independent Gaussian processes, and that Conditions 1 and 2 hold. Let $\theta = \rho(\Gamma_0^Z)/\lambda^* + \min\{1, h(\chi+\phi)\} + \zeta(h)\lambda_\varepsilon^*$, and assume that $\delta = \delta_T$ satisfies $\delta\max(1,\chi) \to 0$ and $\theta/(\kappa\delta) \to 0$. Then (29) holds, namely, as $T \to \infty$,*

$$\max_{j\leqslant r} \max_{0\leqslant u\leqslant 1} \|\widetilde{\Pi}_j(u) - \Pi_j(u)\| = \mathcal{O}_{\mathbb{P}}\{\delta\chi + \theta/(\kappa\delta)\} \to 0 \quad \textit{in probability.} \tag{35}$$

In particular, choosing $\delta \asymp \{\theta/(\chi\kappa)\}^{1/2}$, then the bound in (35) becomes $\mathcal{O}_{\mathbb{P}}\{(\theta\chi/\kappa)^{1/2}\}$.

Proposition 1 below concerns the optimal h for minimizing the function

$$\mu(h) = \min\{1, h(\chi+\phi)\} + \zeta(h)\lambda_\varepsilon^*$$

in Theorem 2 for the lower-dimensional case with $N \leqslant \log T$ and the larger-dimensional case with $N > \log T$, respectively. In the noiseless case with $\eta = 0$ or the weak noise case with $\eta = O(N^{-1}\log T)$, (36) and (37) become $(\chi+\phi)T^{-1}\log T \to 0$.

PROPOSITION 1. *Assume that $\lambda_\varepsilon^* \asymp 1$.*

(i) *If $N \leqslant \log T$, assume that*

$$b = [\{(1+\eta)\lambda_\varepsilon^*\}^2(\chi+\phi)T^{-1}\log T]^{1/3} + \{(1+\eta)\lambda_\varepsilon^*(\chi+\phi)T^{-1}\log T\}^{1/2} \to 0. \tag{36}$$

Then $h = h_0 \asymp b/(\chi+\phi)$ minimizes $\mu(h)$ and $\mu(h_0) \asymp b$.

(ii) *If $N > \log T$, assume that*

$$f = T^{-1}(\eta N + \log T)(\chi+\phi) \to 0. \tag{37}$$

Then the minimizer $h_0 \asymp f^{1/3}/(\chi+\phi)$ and $\mu(h_0) \asymp f^{1/3} + T^{-1}N(\chi+\phi)$.

Similar to Theorem 1, Theorem 2 can be extended to linear processes with sub-Gaussian innovations; see the Supplementary Material.

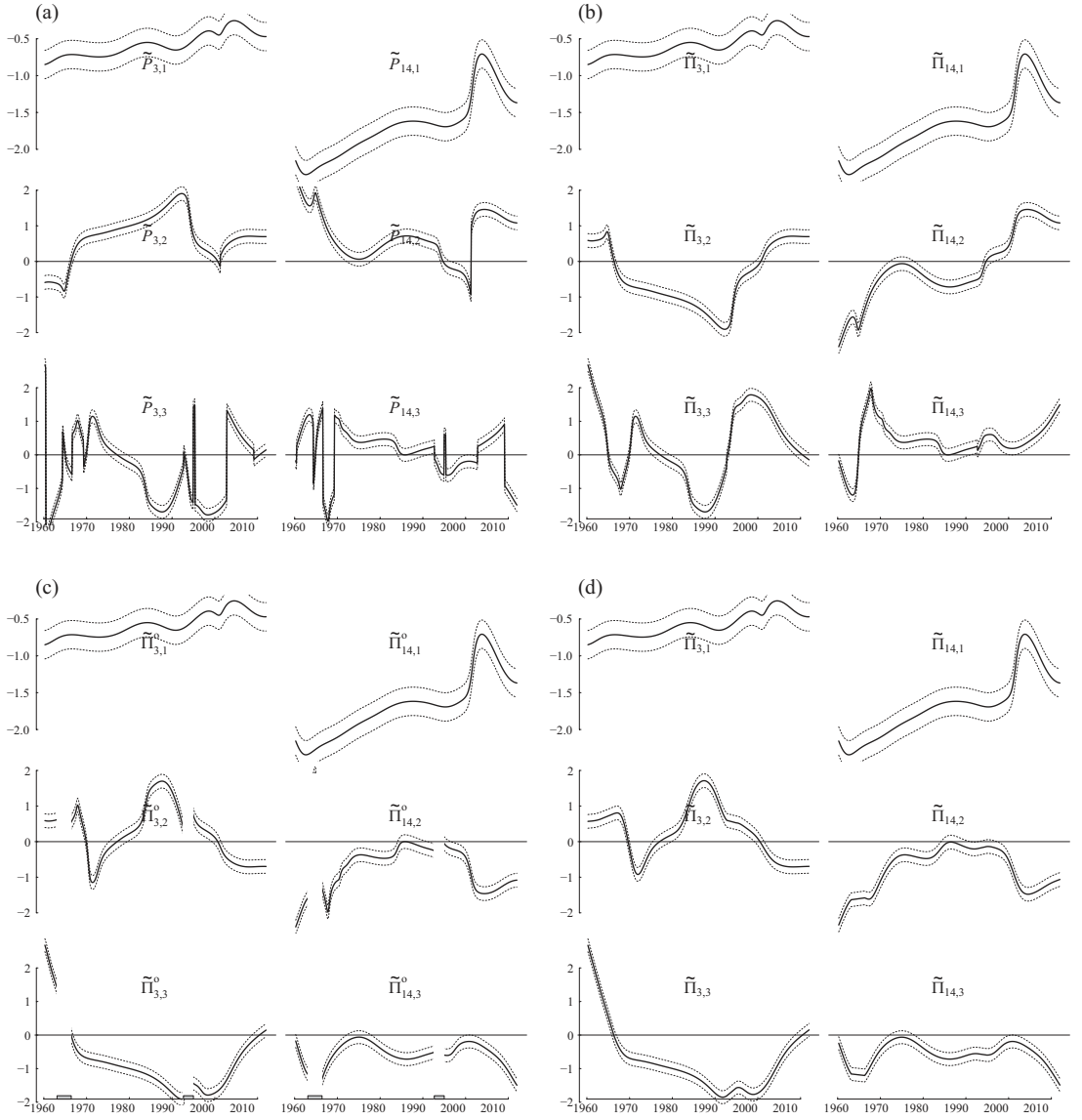

Fig. 5. A real data example with $T = 622$, $N = 132$, $r = 3$ and $c = 2$. The first and second columns respectively correspond to the third and fourteenth rows of the estimated 132×3 eigenfunctions matrix $\Pi(t)$. (a) Eigenvectors obtained by forcing the sign of the first row of $\widetilde{P}(t)$ in (6) to be positive, as suggested by Lawley & Maxwell (1971). (b) Eigenvectors estimated according to (10): $\sqrt{2}$-signing only. (c) Eigenvectors estimated according to (17) and (18): $\sqrt{2}$-bridging and $\sqrt{2}$-signing. (d) Eigenvectors estimated according to (22).

4. Application to real data

In this section we use real data examples to illustrate the phenomenon of unsmooth eigenfunctions of the time-varying lag-0 covariance matrix $\Gamma(t)$. We use the US macroeconomic time series in Jurado et al. (2015), with $N = 132$ monthly times series available from January 1959 to December 2011, for a total of $T = 636$ observations. In Fig. 5 we plot the estimated time-varying factor loadings for real personal consumption expenditures (left column) and the industrial production index of durable goods (right panel) corresponding to the $r = 3$ common factors.

In order to fit the locally stationary factor model (31) and estimate the factor loadings $\{\Pi_{ij}(t),\ i = 3, 14,\ j = 1, 2, 3\}$, we first compute the conventional matrix of eigenvectors $\widetilde{P}(t)$ in (6), the sample version of $P(t)$ in (4). Figure 5(a) is obtained by forcing the sign of the nonzero entries of the first row of $\widetilde{P}(t)$ in (6) to be positive. As explained in § 1, this convention leads to unsmooth eigenvectors. Figure 5(b) is obtained according to our novel estimator in (10), that is, assuming distinct eigenvalues. However, for $j = 2, 3$, the trajectories $\widetilde{\Pi}_{ij}$ in Fig. 5(b) seem to vary abruptly around 1965 and 1994. Using the approach of § 2.4, we detect $c = 2$ coalescing points, $\widetilde{\lambda}_2(t_i) = \widetilde{\lambda}_3(t_i)$, $i = 1, 2$, with $t_1 =$ September 1964 and $t_2 =$ October 1993, and estimate the smooth eigenfunctions according to the methodology in § 2.3; see Figs. 5(c)–5(d).

Supplementary material

The Supplementary Material contains information clarifying the combinatorial nature of the curse of indeterminacy, or misidentification, of smooth eigenvector functions, proofs of our theoretical results, details on estimation of coalescing points and additional simulation results.

Acknowledgement

We are grateful to the editor, an associate editor and two referees for their helpful comments and suggestions that led to a much improved version of this article. This research was partially supported by the US National Science Foundation (NSF-DMS-1916351 and NSF-DMS-2027723).

References

Ahn, S. C. & Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica* **81**, 1203–27.

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59**, 817–58.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica* **71**, 135–71.

Bai, Z. & Silverstein, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, vol. 20. New York: Springer.

Bhattacharjee, M. & Bose, A. (2014). Estimation of autocovariance matrices for infinite dimensional vector linear process. *J. Time Ser. Anal.* **35**, 262–81.

Bhattacharjee, M. & Bose, A. (2016). Large sample behaviour of high dimensional autocovariance matrices. *Ann. Statist.* **44**, 598–628.

Bienvenu, G. (1983). Eigensystem properties of the sampled space correlation matrix. In *ICASSP '83. IEEE Int. Conf. Acoust. Speech Sig. Proces.*, pp. 332–5. Piscataway, NJ: IEEE.

Bunse-Gerstner, A., Byers, R., Mehrmann, V. & Nichols, N. (1991). Numerical computation of an analytic singular value decomposition of a matrix valued function. *Numer. Math.* **60**, 1–39.

Chamberlain, G. & Rothschild, M. (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica* **51**, 1281–304.

Chen, X., Xu, M. & Wu, W. B. (2013). Covariance and precision matrix estimation for high-dimensional time series. *Ann. Statist.* **41**, 2994–3021.

Claeskens, G., Krivobokova, T. & Opsomer, J. D. (2009). Asymptotic properties of penalized spline estimators. *Biometrika* **96**, 529–44.

Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes. *Ann. Statist.* **28**, 1762–94.

Dieci, L. & Eirola, T. (1999). On smooth decompositions of matrices. *SIAM J. Matrix Anal. Applic.* **20**, 800–19.

Dunford, N. & Schwartz, J. T. (1963). *Linear Operators. Part II: Spectral Theory, Self Adjoint Operators in Hilbert Space* (with the assistance of William G. Bade and Robert G. Bartle). New York: Interscience.

Fan, J., Fan, Y., Han, X. & Lv, J. (2022). Asymptotic theory of eigenvectors for random matrices with diverging spikes. *J. Am. Statist. Assoc.* **117**, 996–1009.

Fan, J. & Gijbels, I. (1996). *Local Polynomial Modeling and Its Applications*. Boca Raton, FL: Chapman and Hall.

Forni, M., Hallin, M., Lippi, M. & Reichlin, L. (2000). The generalized dynamic factor model: identification and estimation. *Rev. Econ. Statist.* **82**, 540–54.

Forni, M., Hallin, M., Lippi, M. & Zaffaroni, P. (2017). Dynamic factor models with infinite-dimensional factor space: asymptotic analysis. *J. Economet.* **199**, 74–92.

Furmańczyk, K. (2021). Estimation of autocovariance matrices for high dimensional linear processes. *Metrika* **84**, 595–613.

Heiny, J. & Mikosch, T. (2021). Large sample autocovariance matrices of linear processes with heavy tails. *Stoch. Proces. Appl.* **141**, 344–75.

Jurado, K., Ludvigson, S. C. & Ng, S. (2015). Measuring uncertainty. *Am. Econ. Rev.* **105**, 1177–216.

Koltchinskii, V. & Lounici, K. (2017). Concentration inequalities and moment bounds for sample covariance operators. *Bernoulli* **23**, 110–33.

Lawley, D. & Maxwell, A. (1971). *Factor Analysis as a Statistical Method*. London: Butterworths.

Motta, G., Hafner, C. & von Sachs, R. (2011). Locally stationary factor models: Identification and nonparametric estimation. *Economet. Theory* **27**, 1279–319.

Pignon, J. P. (1981). Eigenvector directions of spectral density matrix. In *Underwater Acoustics and Signal Processing*, L. Bjørnø, ed. Dordrecht: Springer, 571–75.

Reiss, M. & Winkelmann, L. (2021). Inference on the maximal rank of time-varying covariance matrices using high-frequency data. *arXiv:* 2110.00363v1.

Rodríguez-Poo, J. & Linton, O. (2001). Nonparametric factor analysis of residual time series. *Test* **10**, 161–82.

Tiao, G. C. & Box, G. E. (1985). Modeling multiple times series with applications. *J. Am. Statist. Assoc.* **76**, 802–16.

Tropp, J. A. (2012). User-friendly tail bounds for sums of random matrices. *Foundat. Comp. Math.* **12**, 389–434.

Uhlig, F. (2020). Coalescing eigenvalues and crossing eigencurves of 1-parameter matrix flows. *SIAM J. Matrix Anal. Applic.* **41**, 1528–45.

Vershynin, R. (2012). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing: Theory and Applications*, Y. Eldar & G. Kutyniok, eds. Cambridge: Cambridge University Press, pp. 210–68.

von Neumann, J. & Wigner, E. (1929). Über merkwürdige diskrete eigenwerte. *Z. Physik* **30**, 465–7.

Wang, L., Aue, A. & Paul, D. (2017). Spectral analysis of sample autocovariance matrices of a class of linear time series in moderately high dimensions. *Bernoulli* **23**, 2181–209.

Wicklin, R. (2017). The curse of non-unique eigenvectors. SAS Blog, https://blogs.sas.com/content/iml/2017/08/07/eigenvectors-not-unique.html [last accessed 4 January 2023].

Zhang, D. & Wu, W. B. (2021). Convergence of covariance and spectral density estimates for high-dimensional locally stationary processes. *Ann. Statist.* **49**, 233–54.

[*Received on* 13 *December* 2021. *Editorial decision on* 22 *February* 2023]