Journal of Multivariate Analysis 111 (2012) 310–334

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Detecting changes in functional linear models ☆

Lajos Horváth, Ron Reeder *

Department of Mathematics, University of Utah, 155 South 1440, 84112-0090, USA

ARTICLE INFO

Article history:
Received 21 December 2010
Available online 2 May 2012

AMS subject classifications:
primary 62H15
60F17
secondary 62M10

Keywords:
Functional data
Projections
Weak dependence
Change point
Weak convergence

ABSTRACT

We observe two sequences of curves which are connected via an integral operator. Our model includes linear models as well as autoregressive models in Hilbert spaces. We wish to test the null hypothesis that the operator did not change during the observation period. Our method is based on projecting the observations onto a suitably chosen finite dimensional space. The testing procedure is based on functionals of the weighted residuals of the projections. Since the quadratic form is based on estimating the long-term covariance matrix of the residuals, we also provide some results on Bartlett-type estimators.

1. Introduction

Suppose $\{X_n(t), n = 1, 2, \ldots, N\}$ and $\{Y_n(t), n = 1, 2, \ldots, N\}$ are sequences of random functions on $[0, 1]$ that satisfy the linear relationship

$$Y_n(t) = \int_0^1 \Psi_n(s, t)X_n(s)ds + \epsilon_n(t). \tag{1.1}$$

For example, $X_n(t)$ and $Y_n(t)$ may be the exchange rates of two currencies on day n at time t, where the trading day is normalized so that t ranges between 0 and 1. In other applications, X_n can be the temperature and Y_n the pollution level at a given location. If $\Psi_1 = \Psi_2 = \cdots = \Psi_N$, we say that the model is stable. However, as the underlying conditions change, the Ψ's may also change. Our estimates for the assumed common Ψ as well as our predictions and inferences based on the model would be flawed if we falsely assume that the Ψ's have not changed. To test the applicability of this model with an unchanging Ψ, we will test the null hypothesis,

$$H_0 : \Psi_1 = \Psi_2 = \cdots = \Psi_N, \tag{1.2}$$

against the alternative

$$H_A : \Psi_1 = \Psi_2 = \cdots = \Psi_{k_1^*} \neq \Psi_{k_1^*+1} = \cdots = \Psi_{k_r^*} \neq \Psi_{k_r^*+1} = \cdots = \Psi_N$$

with some unknown integers $k_1^*, \ldots, k_r^*$. The k_i^*'s are called change-points, and the alternative, H_A, is that there are exactly r change-points. We assume that (1.1) and H_0 hold and that both $\{X_n\}$ and $\{\epsilon_n\}$ are stationary sequences. The model with

☆ Research partially supported by NSF grant DMS 0905400.

* Corresponding author.
E-mail address: reeder@math.utah.edu (R. Reeder).

0047-259X/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2012.04.007

non-changing (stable) Ψ has received considerable attention in the literature. If X_n and ϵ_n are independent sequences of independent processes, then (1.1) is a functional version of the classical linear model (cf. [7–10]). If $X_n = Y_{n-1}$, then we have the functional AR(1) model in (1.1) (cf. [5,15,18]). Aue et al. [3] investigated the stability of high-frequency portfolio betas in the capital asset pricing model (CAPM). CAPM is a version of the model in (1.1) where, in our notation, a vector valued Y_n is a linear combination of vector valued X_n's and an additional error term.

Let $C(s,t) = \text{var}(X_n(t), X_n(s))$ and $D(s,t) = \text{var}(Y_n(t), Y_n(s))$. Let $\{(v_j(s), \lambda_j), 1 \le j \le \infty\}$ and $\{(w_i(t), \tau_i), 1 \le i \le \infty\}$ be eigenfunction–eigenvalue pairs associated with $C(s,t)$ and $D(s,t)$ respectively. This means that $\tau_i w_i(t) = \int_0^1 D(t,s)w_i(s)ds$ and $\lambda_j v_j(s) = \int_0^1 C(s,t)v_j(t)dt$. Assume that λ_j is the jth largest eigenvalue of $C(s,t)$ and that τ_i is the ith largest eigenvalue of $D(s,t)$. It can be assumed that the eigenfunctions of $C(s,t)$ are orthonormal and also that the eigenfunctions of $D(s,t)$ are orthonormal. We assume that $\Psi \in L^2[0,1]^2$ and can therefore be expressed as

$$\Psi(s,t) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \psi_{i,j} v_j(s) w_i(t). \tag{1.3}$$

Using (1.3) we can write the model (1.1) as

$$\begin{aligned} Y_n(t) &= \int_0^1 \Psi_n(s,t)X_n(s)ds + \epsilon_n(t) \\ &= \int_0^1 \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \psi_{i,j} w_i(t) v_j(s) X_n(s)ds + \epsilon_n(t) \\ &= \sum_{i=1}^{q}\sum_{j=1}^{p} \psi_{i,j} w_i(t) \int_0^1 v_j(s)X_n(s)ds + \epsilon_n^*(t), \end{aligned} \tag{1.4}$$

where

$$\epsilon_n^*(t) = \epsilon_n(t) + \sum_{i=1}^{q}\sum_{j=p+1}^{\infty} \psi_{i,j} w_i(t) \int_0^1 v_j(s)X_n(s)ds + \sum_{i=q+1}^{\infty}\sum_{j=1}^{\infty} \psi_{i,j} w_i(t) \int_0^1 v_j(s)X_n(s)ds.$$

Eq. (1.4) means that we keep the parts of Y_n and X_n which are explained by the first q and p principle components.

To reduce the dimensionality of the model we will project both sides of (1.4) onto the space spanned by the functions $\{w_i(t), 1 \le i \le q\}$. Doing this we obtain the linear model

$$\begin{pmatrix} \langle Y_n, w_1\rangle \\ \langle Y_n, w_2\rangle \\ \vdots \\ \langle Y_n, w_q\rangle \end{pmatrix} = \begin{pmatrix} \psi_{1,1} & \psi_{1,2} & \cdots & \psi_{1,p} \\ \psi_{2,1} & \psi_{2,2} & \cdots & \psi_{2,p} \\ \vdots & \vdots & \cdots & \vdots \\ \psi_{q,1} & \psi_{q,2} & \cdots & \psi_{q,p} \end{pmatrix} \begin{pmatrix} \langle X_n, v_1\rangle \\ \langle X_n, v_2\rangle \\ \vdots \\ \langle X_n, v_p\rangle \end{pmatrix} + \begin{pmatrix} \langle \epsilon_n^*, w_1\rangle \\ \langle \epsilon_n^*, w_2\rangle \\ \vdots \\ \langle \epsilon_n^*, w_q\rangle \end{pmatrix}. \tag{1.5}$$

Instead of testing the null hypothesis, (1.2), exactly as it is stated, we would like to test if the coefficients $\{\psi_{i,j}, 1 \le i \le q, 1 \le j \le p\}$ remained constant during the observation period. Essentially, we are testing the stability of $\Psi(s,t)$ over the space spanned by the most important principle components of the X_n's and the Y_n's. Eq. (1.5) has the form of a linear model, but it is not a classical linear model because the regressors are random variables and are correlated with the errors. Unfortunately, we cannot use (1.5) directly, since the covariance functions, $D(s,t)$ and $C(s,t)$, and hence the eigenfunctions, $\{w_i(t), i = 1,2,\ldots,q\}$ and $\{v_j(t), j = 1,2,\ldots,p\}$, are unknown. Instead, we will use the estimates $\hat{D}_N(s,t)$ and $\hat{C}_N(s,t)$ and their corresponding eigenfunctions, $\{\hat{w}_{i,N}(t), i = 1,2,\ldots,q\}$ and $\{\hat{v}_{j,N}(s), j = 1,2,\ldots,p\}$, where

$$\hat{D}_N(s,t) = \frac{1}{N}\sum_{k=1}^{N}(Y_k(t) - \bar{Y}_N(t))(Y_k(s) - \bar{Y}_N(s)) \quad \text{with } \bar{Y}_N(t) = \frac{1}{N}\sum_{i=1}^{N} Y_i(t),$$

$$\hat{C}_N(s,t) = \frac{1}{N}\sum_{k=1}^{N}(X_k(t) - \bar{X}_N(t))(X_k(s) - \bar{X}_N(s)) \quad \text{with } \bar{X}_N(t) = \frac{1}{N}\sum_{i=1}^{N} X_i(t).$$

Eigenfunctions corresponding to unique eigenvalues are uniquely determined up to signs. For this reason, we cannot expect more than to have $\hat{w}_{i,N}$ be close to $\hat{d}_{i,N} w_i$ and $\hat{v}_{j,N}$ be close to $\hat{c}_{j,N} v_j$, where $\hat{d}_{i,N}$, $\hat{c}_{i,N}$ are random signs (cf. Theorem 5.2). In order to obtain a linear model similar to Eq. (1.5) that is usable, we must use our estimates for the eigenfunctions. We replace Eq. (1.4) with

$$Y_n(t) = \sum_{i=1}^{q}\sum_{j=1}^{p} \hat{d}_{i,N}\psi_{i,j}\hat{c}_{j,N}\hat{w}_{i,N}(t) \int_0^1 \hat{v}_{j,N}(s)X_n(s)ds + \epsilon_n^{**}(t), \tag{1.6}$$

where

$$\epsilon_n^{**}(t) = \epsilon_n(t) + \sum_{i=1}^{q}\sum_{j=p+1}^{\infty}\psi_{i,j}w_i(t)\int_0^1 v_j(s)X_n(s)ds + \sum_{i=q+1}^{\infty}\sum_{j=1}^{\infty}\psi_{i,j}w_i(t)\int_0^1 v_j(s)X_n(s)ds$$
$$- \sum_{i=1}^{q}\sum_{j=1}^{p}\hat{d}_{i,N}\psi_{i,j}\hat{c}_{j,N}\hat{w}_{i,N}(t)\int_0^1 \hat{v}_{j,N}(s)X_n(s)ds + \sum_{i=1}^{q}\sum_{j=1}^{p}\psi_{i,j}w_i(t)\int_0^1 v_j(s)X_n(s)ds.$$

By projecting both sides of (1.6) onto the space spanned by the functions $\{\hat{w}_{j,N}(t),\ 1 \le j \le q\}$, we can replace the linear model (1.5) with the empirical linear model

$$\begin{pmatrix}\langle Y_n, \hat{w}_{1,N}\rangle\\ \langle Y_n, \hat{w}_{2,N}\rangle\\ \vdots\\ \langle Y_n, \hat{w}_{q,N}\rangle\end{pmatrix} = \begin{pmatrix}\hat{d}_{1,N}\psi_{1,1}\hat{c}_{1,N} & \hat{d}_{1,N}\psi_{1,2}\hat{c}_{2,N} & \cdots & \hat{d}_{1,N}\psi_{1,p}\hat{c}_{p,N}\\ \hat{d}_{2,N}\psi_{2,1}\hat{c}_{1,N} & \hat{d}_{2,N}\psi_{2,2}\hat{c}_{2,N} & \cdots & \hat{d}_{2,N}\psi_{2,p}\hat{c}_{p,N}\\ \vdots & \vdots & \cdots & \vdots\\ \hat{d}_{q,N}\psi_{q,1}\hat{c}_{1,N} & \hat{d}_{q,N}\psi_{q,2}\hat{c}_{2,N} & \cdots & \hat{d}_{q,N}\psi_{q,p}\hat{c}_{p,N}\end{pmatrix}\begin{pmatrix}\langle X_n, \hat{v}_{1,N}\rangle\\ \langle X_n, \hat{v}_{2,N}\rangle\\ \vdots\\ \langle X_n, \hat{v}_{p,N}\rangle\end{pmatrix} + \begin{pmatrix}\langle \epsilon_n^{**}, \hat{w}_{1,N}\rangle\\ \langle \epsilon_n^{**}, \hat{w}_{2,N}\rangle\\ \vdots\\ \langle \epsilon_n^{**}, \hat{w}_{q,N}\rangle\end{pmatrix}. \tag{1.7}$$

The signs $\{\hat{d}_{i,N},\ 1 \le i \le q\}$ and $\{\hat{c}_{j,N},\ 1 \le j \le p\}$ are computed from $X_1, X_2, \ldots, X_N$ and $Y_1, Y_2, \ldots, Y_N$ and they will not change during the testing procedure. Therefore, testing the stability of $\{\hat{d}_{i,N}\psi_{i,j}\hat{c}_{j,N},\ 1 \le i \le q,\ 1 \le j \le p\}$ is equivalent to testing the stability of $\{\psi_{i,j},\ 1 \le i \le q,\ 1 \le j \le p\}$.

Letting $\otimes$ be the Kronecker product, we can express Eq. (1.7) in a more condensed form:

$$\hat{\mathbf{Y}}(n) = \hat{\mathbf{Z}}(n)\boldsymbol{\beta} + \hat{\boldsymbol{\Delta}}(n), \quad 1 \le n \le N, \tag{1.8}$$

where

$$\hat{\mathbf{Y}}(n) = \begin{pmatrix}\langle Y_n, \hat{w}_{1,N}\rangle\\ \langle Y_n, \hat{w}_{2,N}\rangle\\ \vdots\\ \langle Y_n, \hat{w}_{q,N}\rangle\end{pmatrix}, \qquad \hat{\boldsymbol{\Delta}}(n) = \begin{pmatrix}\langle \epsilon_n^{**}, \hat{w}_{1,N}\rangle\\ \langle \epsilon_n^{**}, \hat{w}_{2,N}\rangle\\ \vdots\\ \langle \epsilon_n^{**}, \hat{w}_{q,N}\rangle\end{pmatrix},$$

$$\boldsymbol{\beta} = \begin{pmatrix}\hat{d}_{1,N}\psi_{1,1}\hat{c}_{1,N}\\ \vdots\\ \hat{d}_{1,N}\psi_{1,p}\hat{c}_{p,N}\\ \hat{d}_{2,N}\psi_{2,1}\hat{c}_{1,N}\\ \vdots\\ \hat{d}_{q,N}\psi_{q,p}\hat{c}_{p,N}\end{pmatrix} = \mathrm{vec}(\{\hat{d}_{i,N}\psi_{i,j}\hat{c}_{j,N},\ 1 \le i \le q,\ 1 \le j \le p\}^T),$$

and

$$\hat{\mathbf{Z}}(n) = \mathbf{I}_q \otimes \hat{\mathbf{M}}(n) \quad \text{with } \hat{\mathbf{M}}(n) = \left(\langle X_n, \hat{v}_{1,N}\rangle, \ldots, \langle X_n, \hat{v}_{p,N}\rangle\right).$$

The least squares estimator for $\boldsymbol{\beta}$ is defined by

$$\hat{\boldsymbol{\beta}}_N = \left(\hat{\mathbf{Z}}_N^T\hat{\mathbf{Z}}_N\right)^{-1}\hat{\mathbf{Z}}_N^T\hat{\mathbf{Y}}_N,$$

where the vectors $\hat{\mathbf{Y}}_{\lfloor Nt\rfloor}$ and the matrices $\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}$ for each $t \in [0, 1]$ are defined by

$$\hat{\mathbf{Y}}_{\lfloor Nt\rfloor} = \begin{pmatrix}\hat{\mathbf{Y}}(1)\\ \hat{\mathbf{Y}}(2)\\ \vdots\\ \hat{\mathbf{Y}}(\lfloor Nt\rfloor)\end{pmatrix} \quad \text{and} \quad \hat{\mathbf{Z}}_{\lfloor Nt\rfloor} = \begin{pmatrix}\hat{\mathbf{Z}}(1)\\ \hat{\mathbf{Z}}(2)\\ \vdots\\ \hat{\mathbf{Z}}(\lfloor Nt\rfloor)\end{pmatrix}.$$

Our testing procedure is based on the cumulative sums process of the weighted residuals,

$$\tilde{\mathbf{V}}_N(t) = N^{-1/2}\left[\sum_{n=1}^{\lfloor Nt\rfloor}\hat{\mathbf{Z}}^T(n)\tilde{\mathbf{Y}}(n) - t\sum_{n=1}^{N}\hat{\mathbf{Z}}^T(n)\tilde{\mathbf{Y}}(n)\right], \quad t \in [0, 1], \tag{1.9}$$

where $\tilde{\mathbf{Y}}(n) = \hat{\mathbf{Y}}(n) - \hat{\mathbf{Z}}(n)\hat{\boldsymbol{\beta}}_N$, $1 \le n \le N$ stands for the residuals.

2. Main results

In this section we formally state all of the assumptions that we need and then we state our main theorem. Throughout this paper we use $|\cdot|$ to mean the absolute value of a scalar or the largest of the absolute values of the elements of a vector or matrix. It will always be clear from the context which is meant.

Our first condition means that the processes X_n and ϵ_n are Bernoulli shifts:

Assumption 2.1. $X_n(t)$ and $\epsilon_n(t)$ can be expressed as

$$X_n(t) = a(\boldsymbol{\eta}_n(t), \boldsymbol{\eta}_{n-1}(t), \ldots) \quad \text{and} \quad \epsilon_n(t) = b(\boldsymbol{\eta}_n(t), \boldsymbol{\eta}_{n-1}(t), \ldots),$$

for some functionals a and b where $\{\boldsymbol{\eta}_k, -\infty < k < \infty\}$ are iid vector-valued random functions.

Assumption 2.1 implies immediately that the vector-valued process (X_n, ϵ_n), $1 \le n < \infty$ is stationary and ergodic. If H_0 holds, then (X_n, ϵ_n, Y_n), $1 \le n < \infty$ is also stationary and ergodic. We also require that the processes have at least 4 moments:

Assumption 2.2.

$$EX_n(t) = 0 \quad \text{and} \quad E\epsilon_n(t) = 0, \tag{2.1}$$

$$\int_0^1 EX_n^4(t)dt < \infty \quad \text{and} \quad \int_0^1 E\epsilon_n^4(t)dt < \infty. \tag{2.2}$$

Assumption 2.3. $X_n(t)$ and $\epsilon_n(s)$ are uncorrelated, i.e. $EX_n(t)\epsilon_n(s) = 0$ for all $0 \le t, s \le 1$.

Under Assumption 2.1 one can even have long-range dependence among the observations. However, in this paper we are only interested in weakly dependent sequences which is stated in the next assumption:

Assumption 2.4. We assume that

$$\sum_{1\le k<\infty} \left(E\int_0^1 \left(X_n(t) - X_n^{(k)}(t)\right)^4 dt \right)^{1/4} < \infty \tag{2.3}$$

and

$$\sum_{1\le k<\infty} \left(E\int_0^1 \left(\epsilon_n(t) - \epsilon_n^{(k)}(t)\right)^4 dt \right)^{1/4} < \infty \tag{2.4}$$

with

$$X_n^{(k)}(t) = a(\boldsymbol{\eta}_n(t), \boldsymbol{\eta}_{n-1}(t), \ldots, \boldsymbol{\eta}_{n-k+1}(t), \boldsymbol{\eta}_{n,n-k}^{(k)}(t), \boldsymbol{\eta}_{n,n-k-1}^{(k)}(t), \ldots)$$

and

$$\epsilon_n^{(k)}(t) = b(\boldsymbol{\eta}_n(t), \boldsymbol{\eta}_{n-1}(t), \ldots, \boldsymbol{\eta}_{n-k+1}(t), \boldsymbol{\eta}_{n,n-k}^{(k)}(t), \boldsymbol{\eta}_{n,n-k-1}^{(k)}(t), \ldots),$$

where $\{\boldsymbol{\eta}_{n,\ell}^{(k)}, -\infty < k, \ell, n < \infty\}$ are iid copies of $\boldsymbol{\eta}_0$.

We note that, due to stationarity required by Assumption 2.1, it is enough to assume that (2.3) and (2.4) hold for at least one n. Hörmann and Kokoszka [13] call the processes satisfying Assumption 2.4 L^4-k-decomposable processes. This property appeared first in [16] and is used several times in [4] in case of random variables on the line. Aue et al. [3] use an analogue of Assumption 2.4 for random vectors when they derive tests to detect a change in the covariance structure of the observations. Wied et al. [28] investigate the change in the correlation under the same assumptions as in [3]. Aue et al. [2] provide several examples when Assumptions 2.1 and 2.4 hold. For example, autoregressive, moving-average, linear processes in Hilbert spaces satisfy this condition. Also, the non-linear functional ARCH(1) model (cf. [14]) and bilinear models (cf. [13]) also satisfy Assumption 2.4.

Our next assumption ensures that the p and q largest eigenvalues of C and D, respectively, are unique.

Assumption 2.5.

$$\lambda_1 > \lambda_2 > \cdots > \lambda_{p+1}$$

and

$$\tau_1 > \tau_2 > \cdots > \tau_{q+1}.$$

Assumption 2.6.

$$\int_0^1 \int_0^1 \Psi^4(s,t)dtds < \infty.$$

We note that under Assumptions 2.2 and 2.6 we also have that $EY_n(t) = 0$ and $\int_0^1 EY_n^4(t)dt < \infty$. Let

$$\boldsymbol{\gamma}_\ell = \text{vec}\left(\{\gamma_\ell(i,j), 1 \le i \le q, 1 \le j \le p\}^T\right),$$

where

$$\gamma_\ell(i,j) = \langle X_\ell, v_j \rangle \langle \epsilon_\ell, w_i \rangle + \langle X_\ell, v_j \rangle \langle X_\ell, u_i \rangle,$$

and

$$u_i(s) = \sum_{r=p+1}^{\infty} \psi_{i,r} v_r(s), \quad 1 \le i \le q.$$

Define $\boldsymbol{\Sigma}$ as

$$\boldsymbol{\Sigma} = E\boldsymbol{\gamma}_0\boldsymbol{\gamma}_0^T + \sum_{l=1}^{\infty} E\boldsymbol{\gamma}_0\boldsymbol{\gamma}_\ell^T + \sum_{l=1}^{\infty} E\boldsymbol{\gamma}_\ell\boldsymbol{\gamma}_0^T.$$

We now define our detector as

$$V_N(t) = \tilde{\mathbf{V}}_N^T(t)\breve{\boldsymbol{\Sigma}}_N^{-1}\tilde{\mathbf{V}}_N(t),$$

where $\tilde{\mathbf{V}}_N(t)$ is defined in (1.9) and $\breve{\boldsymbol{\Sigma}}_N$ is an estimator (up to random signs) for $\boldsymbol{\Sigma}$. The Bartlett-type estimator that we propose for $\breve{\boldsymbol{\Sigma}}_N$ is a function of the estimators $\hat{v}_{j,N}(t)$ and $\hat{w}_{i,N}(t)$, which are estimators for $v(t)$ and $w(t)$ up to random signs. For this reason, we cannot expect that $\breve{\boldsymbol{\Sigma}}_N$ will be close to $\boldsymbol{\Sigma}$. The best we can expect is that $\boldsymbol{\zeta}_N\breve{\boldsymbol{\Sigma}}_N\boldsymbol{\zeta}_N$ will be close to $\boldsymbol{\Sigma}$, where $\boldsymbol{\zeta}_N$ is a matrix corresponding to the random signs, $\hat{c}_{j,N}$ and $\hat{d}_{i,N}$. This is described in Assumption 2.7.

Next we introduce the diagonal matrices $\hat{\mathbf{C}}_N$ and $\hat{\mathbf{D}}_N$ which consists of the random signs, i.e. $\hat{\mathbf{C}}_N = \text{diag}(\hat{c}_{1,N}, \dots, \hat{c}_{p,N})$, $\hat{\mathbf{D}}_N = \text{diag}(\hat{d}_{1,N}, \dots, \hat{d}_{q,N})$ and $\boldsymbol{\zeta}_N = \hat{\mathbf{D}}_N \otimes \hat{\mathbf{C}}_N$.

Assumption 2.7. $\hat{\boldsymbol{\Sigma}}_N = \boldsymbol{\zeta}_N\breve{\boldsymbol{\Sigma}}_N\boldsymbol{\zeta}_N$ is an estimator for $\boldsymbol{\Sigma}$ such that

$$\left|\hat{\boldsymbol{\Sigma}}_N - \boldsymbol{\Sigma}\right| = o_P(1).$$

Note in particular that

$$\boldsymbol{\zeta}_N\boldsymbol{\gamma}_\ell = \text{vec}\left(\{\hat{c}_{j,N}\hat{d}_{i,N}\gamma_\ell(i,j), 1 \le i \le q, 1 \le j \le p\}^T\right).$$

Note also that Assumption 2.7 and the continuous mapping theorem combined imply that $\hat{\boldsymbol{\Sigma}}_N^{-1} = \boldsymbol{\zeta}_N\breve{\boldsymbol{\Sigma}}_N^{-1}\boldsymbol{\zeta}_N \xrightarrow{\mathcal{P}} \boldsymbol{\Sigma}^{-1}$.

Although any estimator satisfying Assumption 2.7 can be used, we recommend using a Bartlett-type estimator as $\breve{\boldsymbol{\Sigma}}_N$, which we will describe in Section 3.

Theorem 2.1. *If Assumptions 2.1–2.7 hold, then we have*

$$V_N(t) \xrightarrow{\mathcal{D}} \sum_{\ell=1}^{pq} \mathcal{B}_\ell^2(t),$$

where $\{\mathcal{B}_\ell(t), \ell = 1, \dots, pq\}$ are iid standard Brownian bridges.

The testing procedure can be based on Theorem 2.1, using functionals of $V_N(t)$. The distribution of functionals of the limit was considered by Kiefer [19] who provided formulae for the distribution functions of the supremum and L^2 functionals of the limit. For tables, approximations and further discussion on the distribution of functionals of the limit we refer to [3].

3. Bartlett-type estimators

In this section we discuss the estimation of the long-run covariance matrix of the sums of weakly dependent vectors. We start with estimators based on the sequence γ_ℓ, $1 \le \ell \le N$. Since $\boldsymbol{\Sigma}$ is the spectral density at 0, the kernel-type estimators discussed in [1,6,12,24–26,29] can be used. The estimator is defined by

$$\tilde{\boldsymbol{\Sigma}}_N = \sum_{k=-(N-1)}^{N-1} K(k/B_N)\boldsymbol{\phi}_{k,N},$$

where

$$\boldsymbol{\phi}_{k,N} = \frac{1}{N}\sum_{\ell=\max(1,1-k)}^{\min(N,N-k)} \boldsymbol{\gamma}_\ell \boldsymbol{\gamma}_{\ell+k}^T.$$

The kernel K satisfies the following condition:

Assumption 3.1.

(i) $K(0) = 1$

(ii) K is a symmetric, Lipschitz function

(iii) K has a bounded support

(iv) $\hat{K}$, the Fourier transform of K, is also Lipschitz and integrable.

These conditions are mild, and they are satisfied by the most commonly used kernels, like the triangle of Bartlett and the polynomial kernel of [22,23]. Assumption 3.1(iii) makes the present proofs relatively technically simple and it could be replaced with the assumption that $K(x)$ decays sufficiently fast as $|x| \to \infty$. The next assumption is standard in the estimation of spectral densities and long term variances and covariances.

Assumption 3.2.

$$B_N \to \infty \quad \text{and} \quad B_N/N \to 0, \quad \text{as } N \to \infty.$$

Jansson [17] proved the consistency of covariance estimation for linear processes under the assumption $B_N = o(N^{1/2})$. Similarly, [13] obtained consistency results for the estimation of the long run covariance matrices of the projections of functional observations assuming $B_N = o(N^{1/2})$. Liu and Wu [20] established consistency results for estimation of spectral densities under Assumption 3.2.

Theorem 3.1. *If Assumptions 2.1–2.4, 2.6, 3.1 and 3.2 hold, then*

$$\tilde{\boldsymbol{\Sigma}}_N \xrightarrow{\mathcal{P}} \boldsymbol{\Sigma}.$$

We would like to point out that the proof of Theorem 3.1 only requires that γ_ℓ is a Bernoulli shift with zero mean and finite second moment for which (5.13) holds.

The estimator, $\tilde{\boldsymbol{\Sigma}}_N$, cannot be computed since the variables $\boldsymbol{\gamma}_\ell$ are not observed directly and we need to estimate them from the sample. We have estimators for v_j as well as for w_i, but we will also need an estimator for ϵ_ℓ. We use the residuals to get inference on ϵ_ℓ:

$$\hat{\epsilon}_\ell(t) = Y_\ell(t) - \sum_{i=1}^{q}\sum_{j=1}^{p} \hat{\psi}_{i,j}\hat{w}_{i,N}(t)\langle X_\ell, \hat{v}_{j,N}\rangle,$$

where $\hat{\psi}_{i,j}$ is the (i,j)th element of $\hat{\boldsymbol{\beta}}_N$ when it is written in the matrix form, i.e. $\{\hat{\psi}_{i,j}, 1 \le i \le q, 1 \le j \le p\} = \text{vec}^{-1}(\hat{\boldsymbol{\beta}}_N)$. Now $\boldsymbol{\gamma}_\ell$ will be replaced with

$$\hat{\boldsymbol{\gamma}}_\ell = \text{vec}\left(\{\hat{\gamma}_\ell(i,j), 1 \le i \le q, 1 \le j \le p\}^T\right),$$

where

$$\hat{\gamma}_\ell(i,j) = \langle X_\ell, \hat{v}_{j,N}\rangle\langle \hat{\epsilon}_\ell, \hat{w}_{i,N}\rangle.$$

Now the Bartlett-type estimator is defined as

$$\breve{\boldsymbol{\Sigma}}_N = \sum_{k=-(N-1)}^{N-1} K(k/B_N)\hat{\boldsymbol{\phi}}_{k,N}, \tag{3.1}$$

where

$$\hat{\boldsymbol{\phi}}_{k,N} = \frac{1}{N} \sum_{\ell=\max(1,1-k)}^{\min(N,N-k)} \hat{\boldsymbol{\gamma}}_\ell \hat{\boldsymbol{\gamma}}_{\ell+k}^T.$$

The next result states that the proposed estimator satisfies Assumption 2.7.

Theorem 3.2. *If Assumptions 2.1–2.6 and 3.1 hold and*

$$B_N \to \infty \quad \text{and} \quad B_N/N^{1/2} \to 0, \tag{3.2}$$

then Assumption 2.7 is satisfied.

The estimator $\breve{\boldsymbol{\Sigma}}_N$ is based on the empirical projections, $\hat{\gamma}_\ell(i,j)$; in the proofs, this will be replaced with $\hat{d}_{i,N}\hat{c}_{j,N}\gamma_\ell(i,j)$. We point out that the rates in (5.5) and (5.6) of Theorem 5.2, which are optimal, allow us to use the relaxed assumption that $B_N/N^{1/2} \to 0$ instead of the typical assumption that $B_N/N \to 0$.

4. A simulation study

In this section, we investigate the empirical size and power of a testing procedure using the integral of the detector, $\int |V_N(t)|dt$, as our test statistic. Seeking to obtain a test of size $\alpha = .01, .05$, or $.10$, a rejection region was chosen according to the limiting distribution of the test statistic. Simulated data was then used to compute the outcome of the test statistic. Iterating this procedure 10,000 times, we kept track of the proportion of times that the outcome fell in the predetermined rejection region. When simulations are done under H_0, this gives us the empirical size of the test, which we expect to be close to the nominal size, α, for large sample sizes. When simulations are done under the alternative, H_A, the proportion gives us the empirical power of the test.

The $X_n(t)$'s and $\varepsilon_n(t)$'s were generated according to the distribution of independent standard Brownian bridges. Then, using $\psi(s,t) = e^{-(s-t)^2}$, we obtained the first half of our sample according to (1.1). The second half of the sample was also obtained from (1.1) but used $\psi(s,t) = ce^{-(s-t)^2}$. Thus the power of the test is a function of the parameter c. In particular, when $c = 1$, the null hypothesis is true. The Bartlett estimator for Σ uses the flat-top kernel

$$K(t) = \begin{cases} 1 & 0 \le |t| < .1 \\ 1.1 - |t| & .1 \le |t| < 1.1 \\ 0 & |t| \ge 1.1. \end{cases}$$

The resulting empirical size and power are given in Tables 1–4 for various values of p and q.

5. Random processes in Hilbert spaces

In this section we summarize some basic results on random variables in Hilbert spaces which are used in the proofs. Let $\|\cdot\|$ denote the L^2-norm of functions defined on the unit interval, the unit square or the unit cube.

Theorem 5.1. *If Assumptions 2.1–2.4 hold, then we have*

$$\left\| \frac{1}{N^{1/2}} \sum_{n=1}^N X_n(t)\epsilon_n(s) \right\| = O_P(1), \tag{5.1}$$

$$\left\| \frac{1}{N^{1/2}} \sum_{n=1}^N (X_n(t)X_n(s) - C(t,s)) \right\| = O_P(1), \tag{5.2}$$

$$\left\| \frac{1}{N^{1/2}} \sum_{n=1}^N (\epsilon_n(t)\epsilon_n(s) - F(t,s)) \right\| = O_P(1), \tag{5.3}$$

with $F(t,s) = E(\epsilon_n(t)\epsilon_n(s))$. If in addition Assumption 2.6 is also satisfied, then

$$\left\| \frac{1}{N^{1/2}} \sum_{n=1}^N (Y_n(t)Y_n(s) - D(t,s)) \right\| = O_P(1). \tag{5.4}$$

Proof. It was pointed out in [13] that the k-approximable property in Assumption 2.4 implies (5.2) and (5.3). Using (1.1), we get that the sums of $X_n(t)\epsilon_n(s)$ and $Y_n(t)Y_n(s)$ are also k-approximable so the rest of the result again follows from Theorem 3.1 of [13]. □

Table 1
Empirical power of test (in %) using $p = 1, q = 1, B_N = N^{1/3}/4$, and a flat-top kernel for $K(t)$.

c	$N = 100$			$N = 500$			$N = 1000$		
	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$
1.0	0.8	5	10.2	0.9	5.1	10	1.1	5.1	10.2
1.2	2.5	10.1	18	15.1	34.9	46.9	35.8	60.1	71.7
1.4	8.9	26.5	38.9	70.5	88.5	93.3	96.9	99.4	99.8
1.6	24.1	52.2	65.5	98	99.7	99.9	100	100	100
1.8	46.5	75.1	85.1	100	100	100	100	100	100
2.0	69.7	90.7	95.3	100	100	100	100	100	100

Table 2
Empirical power of test (in %) using $p = 1, q = 2, B_N = N^{1/3}/4$, and a flat-top kernel for $K(t)$.

c	$N = 100$			$N = 500$			$N = 1000$		
	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$
1.0	0.6	4.5	9.4	1	5.1	10.4	1.1	5.3	10.3
1.2	1.5	7.9	15.8	10.1	26.7	38.7	25.9	50.2	63
1.4	5.7	19.5	30.9	58	80.9	88.6	93.6	98.5	99.4
1.6	15.3	40.5	55.2	95.8	99.2	99.6	100	100	100
1.8	35.	65.4	78.2	100	100	100	100	100	100
2.0	56.6	83.6	91.6	100	100	100	100	100	100

Table 3
Empirical power of test (in %) using $p = 1, q = 3, B_N = N^{1/3}/4$, and a flat-top kernel for $K(t)$.

c	$N = 100$			$N = 500$			$N = 1000$		
	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$
1.0	0.7	4.4	9.6	0.7	5.3	10.2	0.8	5.1	10.2
1.2	1.4	9.5	17.5	18.8	41.8	54.8	50	74.2	83.5
1.4	7.9	27.9	42.3	87.8	96.9	98.5	99.8	100	100
1.6	24.9	57.2	72.1	99.9	100	100	100	100	100
1.8	53.2	82.7	90.8	100	100	100	100	100	100
2.0	76	94.6	97.8	100	100	100	100	100	100

Table 4
Empirical power of test (in %) using $p = 2, q = 2, B_N = N^{1/3}/4$, and a flat-top kernel for $K(t)$.

c	$N = 100$			$N = 500$			$N = 1000$		
	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$	$\alpha = .01$	$\alpha = .05$	$\alpha = .10$
1.0	1.4	5.9	10.7	0.9	4.8	9.6	1	4.9	10
1.2	2.1	8	14.1	7.6	20.8	31.2	19	39.8	52.7
1.4	4.9	15.5	25.4	45.8	70.2	80.5	88	96.5	98.4
1.6	11.1	29.9	43.4	90.4	97.6	98.9	100	100	100
1.8	23.3	48.8	62.6	99.7	100	100	100	100	100
2.0	38.6	68	80.6	100	100	100	100	100	100

Theorem 5.2. *If Assumptions 2.1–2.6 hold, then we have*

$$\max_{1\le i\le q} \|\hat{w}_{i,N}(t) - \hat{d}_{i,N} w_i(t)\| = O_P\left(N^{-1/2}\right), \tag{5.5}$$

$$\max_{1\le j\le p} \|\hat{v}_{j,N}(t) - \hat{c}_{j,N} v_j(t)\| = O_P\left(N^{-1/2}\right) \tag{5.6}$$

and

$$\max_{1\le i\le q} |\hat{\tau}_{i,N} - \tau_i| = O_P\left(N^{-1/2}\right), \tag{5.7}$$

$$\max_{1\le j\le q} |\hat{\lambda}_{j,N} - \lambda_j| = O_P\left(N^{-1/2}\right). \tag{5.8}$$

Proof. Using Corollary 1.6 of [11, p. 99] we get that (5.5) follows from (5.4). According to Lemma 4.3 of [5], (5.4) implies (5.7). Similarly, (5.2) yields (5.6) and (5.8). □

The next result is a uniform version of Theorem 5.1.

Theorem 5.3. *If Assumptions 2.1–2.4 and 2.6 hold, then we have*

$$\max_{1\le k\le N}\left\|\frac{1}{N^{1/2}}\sum_{n=1}^{k}X_n(t)\epsilon_n(s)\right\| = O_P(\log N), \tag{5.9}$$

$$\max_{1\le k\le N}\left\|\frac{1}{N^{1/2}}\sum_{n=1}^{k}(X_n(t)X_n(s)-C(t,s))\right\| = O_P(\log N), \tag{5.10}$$

$$\max_{1\le k\le N}\left\|\frac{1}{N^{1/2}}\sum_{n=1}^{k}(\epsilon_n(t)\epsilon_n(s)-F(t,s))\right\| = O_P(\log N) \tag{5.11}$$

with $F(t,s)=E(\epsilon_n(t)\epsilon_n(s))$. If in addition Assumption 2.6 is also satisfied, then

$$\max_{1\le k\le N}\left\|\frac{1}{N^{1/2}}\sum_{n=1}^{k}(Y_n(t)Y_n(s)-D(t,s))\right\| = O_P(\log N). \tag{5.12}$$

Proof. Following the proof in Section A.1 in [13] one can easily verify that there is an integrable function $g(t,s)$ such that

$$E\left(\sum_{n=1}^{k}X_n(t)\epsilon_n(s)\right)^2 \le kg(t,s).$$

Hence by Menshov's inequality (cf. [21]) we have that

$$E\max_{1\le k\le N}\left(\sum_{n=1}^{k}X_n(t)\epsilon_n(s)\right)^2 \le (\log N)^2 Ng(t,s),$$

implying (5.9). Similar arguments yield (5.10)–(5.12). □

The next results establish the weak convergence of the sum of the γ_ℓ's.

Theorem 5.4. *If Assumptions 2.1–2.4 and 2.6 hold, then*

$$\frac{1}{N^{1/2}}\sum_{\ell=1}^{\lfloor Nt\rfloor}\boldsymbol{\gamma}_\ell \xrightarrow{\mathcal{D}^{pq}[0,1]} \mathbf{W}_{\boldsymbol{\Sigma}}(t),$$

where $\mathbf{W}_{\boldsymbol{\Sigma}}$ is a pq dimensional Brownian motion with zero mean and $E(\mathbf{W}_{\boldsymbol{\Sigma}}(t)\mathbf{W}_{\boldsymbol{\Sigma}}(s)^T)=\min(t,s)\boldsymbol{\Sigma}$.

Proof. First we note that Assumptions 2.1–2.4 imply that

$$\sum_{m=1}^{\infty}\left(E(\gamma_\ell(i)-\gamma_\ell^{(m)}(i))^2\right)^{1/2} < \infty, \tag{5.13}$$

where $\gamma_\ell(i)$ and $\gamma_\ell^{(m)}(i)$ are the ith coordinates of the vectors $\boldsymbol{\gamma}_\ell$ and $\boldsymbol{\gamma}_\ell^{(m)}$ with

$$\boldsymbol{\gamma}_\ell^{(m)} = \mathrm{vec}(\{\gamma_\ell^{(m)}(i,j),\ 1\le i\le q,\ 1\le j\le p\}^T),$$

and

$$\gamma_\ell^{(m)}(i,j) = \langle X_\ell^{(m)}, v_j\rangle\langle\epsilon_\ell^{(m)}, w_i\rangle + \langle X_\ell^{(m)}, v_j\rangle\langle X_\ell^{(m)}, u_i\rangle.$$

The result now follows immediately from Theorem A.1 of [3]. □

6. Proof of Theorem 2.1

First we outline the proof of Theorem 2.1. Using the definition of the residual vectors we can write that

$$\begin{aligned}\tilde{\mathbf{V}}_N(t) &= N^{-1/2}\left(\left(\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}^T\hat{\mathbf{Y}}_{\lfloor Nt\rfloor} - \hat{\mathbf{Z}}_{\lfloor Nt\rfloor}^T\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}\hat{\boldsymbol{\beta}}_N\right) - t\left(\hat{\mathbf{Z}}_N^T\hat{\mathbf{Y}}_N - \hat{\mathbf{Z}}_N^T\hat{\mathbf{Z}}_N\hat{\boldsymbol{\beta}}_N\right)\right)\\ &= N^{-1/2}\left(\left(\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}^T\hat{\boldsymbol{\Delta}}_{\lfloor Nt\rfloor} - t\hat{\mathbf{Z}}_N^T\hat{\boldsymbol{\Delta}}_N\right) + \left(\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}^T\hat{\mathbf{Z}}_{\lfloor Nt\rfloor} - t\hat{\mathbf{Z}}_N^T\hat{\mathbf{Z}}_N\right)\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}_N\right)\right)\\ &= N^{-1/2}\left(\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}^T\hat{\boldsymbol{\Delta}}_{\lfloor Nt\rfloor} - t\hat{\mathbf{Z}}_N^T\hat{\boldsymbol{\Delta}}_N\right) + \left(\frac{\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}^T\hat{\mathbf{Z}}_{\lfloor Nt\rfloor} - t\hat{\mathbf{Z}}_N^T\hat{\mathbf{Z}}_N}{N}\right)\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}_N\right)\sqrt{N},\end{aligned} \tag{6.1}$$

with

$$\hat{\boldsymbol{\Delta}}_{\lfloor Nt \rfloor} = \begin{pmatrix} \hat{\boldsymbol{\Delta}}(1) \\ \hat{\boldsymbol{\Delta}}(2) \\ \vdots \\ \hat{\boldsymbol{\Delta}}(\lfloor Nt \rfloor) \end{pmatrix}.$$

We show that

$$\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_N\right)\sqrt{N} = O_P(1), \tag{6.2}$$

(cf. Lemma 6.10) and we prove in Lemma 6.2 that

$$\sup_{t\in[0,1]} \left| \frac{\hat{\mathbf{Z}}^T_{\lfloor Nt \rfloor}\hat{\mathbf{Z}}_{\lfloor Nt \rfloor} - t\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N}{N} \right| = o_P(1). \tag{6.3}$$

Combining (6.2) and (6.3) we conclude that

$$\sup_{t\in[0,1]} \left| \left(\frac{\hat{\mathbf{Z}}^T_{\lfloor Nt \rfloor}\hat{\mathbf{Z}}_{\lfloor Nt \rfloor} - t\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N}{N} \right) \left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_N\right)\sqrt{N} \right| = o_P(1).$$

Thus we see that $N^{-1/2}\left(\hat{\mathbf{Z}}^T_{\lfloor Nt \rfloor}\hat{\boldsymbol{\Delta}}_{\lfloor Nt \rfloor} - t\hat{\mathbf{Z}}^T_N\hat{\boldsymbol{\Delta}}_N\right)$ is the leading term while the remainder can be disregarded when considering the limiting distribution of our cumulative sum process (1.9).

We now start with the proof of (6.3).

Lemma 6.1. *If Assumptions 2.1–2.5 hold, then we have*

$$\frac{1}{k}\sum_{n=1}^{k}\langle X_n, v_i\rangle\langle X_n, v_j\rangle \xrightarrow{a.s.} \lambda_i\, 1\{i=j\} \quad \text{as } k\to\infty.$$

Proof. We recall that $X_n(t)$ is stationary and ergodic. Thus the ergodic theorem shows us that as $k \to \infty$

$$\begin{aligned}
\frac{1}{k}\sum_{n=1}^{k}\langle X_n, v_i\rangle\langle X_n, v_j\rangle &\xrightarrow{a.s.} E\int_0^1 X_n(s)v_i(s)ds\int_0^1 X_n(t)v_j(t)dt \\
&= E\int_0^1 v_j(t)\int_0^1 v_i(s)X_n(t)X_n(s)dsdt \\
&= \int_0^1 v_j(t)\int_0^1 v_i(s)E\left(X_n(t)X_n(s)\right)dsdt \\
&= \int_0^1 v_j(t)\int_0^1 v_i(s)C(s,t)dsdt \\
&= \int_0^1 v_j(t)\lambda_i v_i(t)dt \\
&= \lambda_i\, 1\{i=j\},
\end{aligned}$$

completing the proof. □

Lemma 6.2. *If Assumptions 2.1–2.5 hold, then we have*

$$\frac{1}{N}\sup_{t\in[0,1]}\left|\hat{\mathbf{Z}}^T_{\lfloor Nt \rfloor}\hat{\mathbf{Z}}_{\lfloor Nt \rfloor} - t\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N\right| = o_P(1) \tag{6.4}$$

and

$$\frac{1}{N}\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N \xrightarrow{\mathcal{P}} \mathbf{C} = \mathbf{I}_q \otimes \boldsymbol{\Lambda}, \tag{6.5}$$

where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$.

Proof. First we show that for $\delta > 0$ and $\gamma > 0$ there are K_0 and N_0 such that

$$P\left(\sup_{K_0\le k\le N}\left|\frac{1}{k}\sum_{n=1}^{k}\langle X_n,\hat{v}_{i,N}\rangle\langle X_n,\hat{v}_{j,N}\rangle-\lambda_i 1\{i=j\}\right|>\delta\right)\le\gamma, \tag{6.6}$$

if $N \ge N_0$. Note that by the Cauchy–Schwarz inequality we have

$$\left|\frac{1}{k}\sum_{n=1}^{k}\big(\langle X_n,\hat{v}_{i,N}\rangle\langle X_n,\hat{v}_{j,N}\rangle-\langle X_n,\hat{c}_{i,N}v_i\rangle\langle X_n,\hat{c}_{j,N}v_j\rangle\big)\right|\le\frac{1}{k}\sum_{n=1}^{k}\|X_n\|^2(\|\hat{v}_{i,N}-\hat{c}_{i,N}v_i\|+\|\hat{v}_{j,N}-\hat{c}_{j,N}v_j\|).$$

Using the ergodic theorem we get that

$$\sup_{1\le k<\infty}\frac{1}{k}\sum_{n=1}^{k}\|X_n\|^2<\infty\quad\text{a.s.,}$$

so (6.6) follows from Theorem 5.2 and Lemma 6.1.

Assume $N > N_0$. It now follows that

$$\begin{aligned}
&P\left(\sup_{t\in[0,1]}\left|\hat{\mathbf{Z}}^T_{\lfloor Nt\rfloor}\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}-t\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N\right|>4N\delta\right)\\
&\quad\le P\left(\sup_{0\le t\le K_0/N}\left|\hat{\mathbf{Z}}^T_{\lfloor Nt\rfloor}\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}-t\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N\right|+\sup_{K_0/N\le t\le 1}\left|\hat{\mathbf{Z}}^T_{\lfloor Nt\rfloor}\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}-t\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N\right|>4N\delta\right)\\
&\quad\le P\left(\sup_{0\le t\le K_0/N}\left|\frac{\hat{\mathbf{Z}}^T_{\lfloor Nt\rfloor}\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}-t\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N}{N}\right|+\sup_{K_0/N\le t\le 1}\left|\frac{\hat{\mathbf{Z}}^T_{\lfloor Nt\rfloor}\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}}{Nt}-\frac{\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N}{N}\right|>4\delta\right)\\
&\quad\le P\left(\sup_{0\le t\le K_0/N}\left|\frac{\hat{\mathbf{Z}}^T_{\lfloor Nt\rfloor}\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}-t\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N}{N}\right|+\sup_{K_0/N\le t\le 1}\left|\frac{\hat{\mathbf{Z}}^T_{\lfloor Nt\rfloor}\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}}{Nt}-\mathbf{C}\right|+\left|\frac{\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N}{N}-\mathbf{C}\right|>4\delta\right)\\
&\quad\le P\left(\max_{1\le k\le K_0}\left|\frac{\hat{\mathbf{Z}}^T_k\hat{\mathbf{Z}}_k}{N}\right|+\left|\frac{K_0\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N}{N^2}\right|+\max_{K_0\le k\le N}\left|\frac{\hat{\mathbf{Z}}^T_k\hat{\mathbf{Z}}_k}{k}-\mathbf{C}\right|+\left|\frac{\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N}{N}-\mathbf{C}\right|>4\delta\right).
\end{aligned}$$

For every K_0 we have that $P(\max_{1\le k\le K_0}|\hat{\mathbf{Z}}^T_k\hat{\mathbf{Z}}_k|/N>\delta)\to 0$ and by (6.6) $P(|K_0\hat{\mathbf{Z}}^T_N\hat{\mathbf{Z}}_N|/N^2>\delta)\to 0$ as $N\to\infty$. Using (6.6) again we conclude $P(\max_{K_0\le k\le N}|\hat{\mathbf{Z}}^T_k\hat{\mathbf{Z}}_k/k-\mathbf{C}|>\delta)\le\gamma$. Since γ and δ can be chosen as small as we wish, Lemma 6.2 is established. □

We continue with the properties of $\hat{\mathbf{Z}}^T_{\lfloor Nt\rfloor}\hat{\boldsymbol{\Delta}}_{\lfloor Nt\rfloor}$. First we observe that

$$\begin{aligned}
\hat{\mathbf{Z}}^T_{\lfloor Nt\rfloor}\hat{\boldsymbol{\Delta}}_{\lfloor Nt\rfloor}&=\sum_{\ell=1}^{\lfloor Nt\rfloor}\hat{\mathbf{Z}}^T(\ell)\hat{\boldsymbol{\Delta}}(\ell)\\
&=\sum_{\ell=1}^{\lfloor Nt\rfloor}\mathrm{vec}\left(\{\langle X_\ell,\hat{v}_{j,N}\rangle\langle\epsilon^{**}_\ell,\hat{w}_{i,N}\rangle,\ 1\le i\le q,\ 1\le j\le p\}^T\right).
\end{aligned} \tag{6.7}$$

We note that

$$\epsilon^{**}_\ell(t)=\epsilon_\ell(t)+\eta_{\ell,1}(t)+\eta_{\ell,2}(t)+\eta_{\ell,3}(t)+\eta_{\ell,4}(t)+\eta_{\ell,5}(t),$$

with

$$\begin{aligned}
\eta_{n,1}(t)&=\sum_{i=q+1}^{\infty}\sum_{j=1}^{\infty}\psi_{i,j}w_i(t)\int_0^1 v_j(s)X_n(s)ds=\sum_{i=q+1}^{\infty}\sum_{j=1}^{\infty}\psi_{i,j}w_i(t)\langle v_j,X_n\rangle,\\
\eta_{n,2}(t)&=\sum_{i=1}^{q}\sum_{j=p+1}^{\infty}\psi_{i,j}w_i(t)\int_0^1 v_j(s)X_n(s)ds=\sum_{i=1}^{q}\sum_{j=p+1}^{\infty}\psi_{i,j}w_i(t)\langle v_j,X_n\rangle,
\end{aligned}$$

$$\begin{aligned}
\eta_{n,3}(t) &= \sum_{i=1}^{q}\sum_{j=1}^{p}\hat{d}_{i,N}\psi_{i,j}\hat{c}_{j,N}\hat{d}_{i,N}w_i(t)\int_0^1\left(\hat{c}_{j,N}v_j(s)-\hat{v}_{j,N}(s)\right)X_n(s)ds\\
&= \sum_{i=1}^{q}\sum_{j=1}^{p}\hat{d}_{i,N}\psi_{i,j}\hat{c}_{j,N}\hat{d}_{i,N}w_i(t)\left\langle\left(\hat{c}_{j,N}v_j-\hat{v}_{j,N}\right),X_n\right\rangle,\\
\eta_{n,4}(t) &= \sum_{i=1}^{q}\sum_{j=1}^{p}\hat{d}_{i,N}\psi_{i,j}\hat{c}_{j,N}\left(\hat{d}_{i,N}w_i(t)-\hat{w}_{i,N}(t)\right)\int_0^1\hat{c}_{j,N}v_j(s)X_n(s)ds\\
&= \sum_{i=1}^{q}\sum_{j=1}^{p}\hat{d}_{i,N}\psi_{i,j}\hat{c}_{j,N}\left(\hat{d}_{i,N}w_i(t)-\hat{w}_{i,N}(t)\right)\langle\hat{c}_{j,N}v_j,X_n\rangle,\\
\eta_{n,5}(t) &= \sum_{i=1}^{q}\sum_{j=1}^{p}\hat{d}_{i,N}\psi_{i,j}\hat{c}_{j,N}\left(\hat{w}_{i,N}(t)-\hat{d}_{i,N}w_i(t)\right)\int_0^1\left(\hat{c}_{j,N}v_j(s)-\hat{v}_{j,N}(s)\right)X_n(s)ds\\
&= \sum_{i=1}^{q}\sum_{j=1}^{p}\hat{d}_{i,N}\psi_{i,j}\hat{c}_{j,N}\left(\hat{w}_{i,N}(t)-\hat{d}_{i,N}w_i(t)\right)\left\langle\left(\hat{c}_{j,N}v_j-\hat{v}_{j,N}\right),X_n\right\rangle.
\end{aligned}$$

In particular, we can write

$$\langle\epsilon_\ell^{**},\hat{w}_{i,N}\rangle=\langle\epsilon_\ell,\hat{w}_{i,N}\rangle+\langle\eta_{\ell,1},\hat{w}_{i,N}\rangle+\langle\eta_{\ell,2},\hat{w}_{i,N}\rangle+\langle\eta_{\ell,3},\hat{w}_{i,N}\rangle+\langle\eta_{\ell,4},\hat{w}_{i,N}\rangle+\langle\eta_{\ell,5},\hat{w}_{i,N}\rangle. \tag{6.8}$$

We show that $\hat{\mathbf{Z}}^T_{\lfloor Nt\rfloor}\hat{\boldsymbol{\Delta}}_{\lfloor Nt\rfloor}$ can be written as the sum of weakly dependent variables and an additional term which is just t times a random variable matrix. The additional term reflects the replacement of Ψ with a finite sum and the estimation of the eigenfunctions $\{w_i,\ 1\le i\le q\}$ and $\{v_j,\ 1\le i\le p\}$. The drift term is given by

$$\mathbf{R}_N=\mathrm{vec}\left(\{R_N(i,j),\ 1\le i\le q,\ 1\le j\le p\}^T\right),$$

where

$$\begin{aligned}
&R_N(i,j)=R_N^{(1)}(i,j)+R_N^{(2)}(i,j)+R_N^{(3)}(i,j)+R_N^{(4)}(i,j),\\
&R_N^{(1)}(i,j)=\hat{c}_{j,N}\lambda_j\sum_{r=q+1}^{\infty}\psi_{r,j}\int_0^1 w_r(x)\left(\hat{w}_{i,N}(x)-\hat{d}_{i,N}w_i(x)\right)dx,\\
&R_N^{(2)}(i,j)=\hat{d}_{i,N}\int_0^1\left(\hat{v}_{j,N}(z)-\hat{c}_{j,N}v_j(z)\right)\sum_{n=p+1}^{\infty}\psi_{i,n}\lambda_n v_n(z)dz,\\
&R_N^{(3)}(i,j)=\hat{c}_{j,N}\hat{d}_{i,N}\lambda_j\sum_{n=1}^{p}\psi_{i,n}\hat{c}_{n,N}\int_0^1\left(\hat{c}_{n,N}v_n(s)-\hat{v}_{n,N}(s)\right)v_j(s)ds,\\
&R_N^{(4)}(i,j)=\hat{c}_{j,N}\hat{d}_{i,N}\lambda_j\sum_{r=1}^{q}\hat{d}_{r,N}\psi_{r,j}\int_0^1 w_i(x)\left(\hat{d}_{r,N}w_r(x)-\hat{w}_{r,N}(x)\right)dx.
\end{aligned}$$

Lemma 6.3. *If Assumptions 2.1–2.5 hold, then we have*

$$\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}\rangle\langle\epsilon_\ell,\hat{w}_{i,N}\rangle-\hat{c}_{j,N}\hat{d}_{i,N}T^{(1)}_{\lfloor Nt\rfloor}(i,j)\right|=O_P(\log N),$$

where

$$T^{(1)}_{\lfloor Nt\rfloor}(i,j)=\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,v_j\rangle\langle\epsilon_\ell,w_i\rangle.$$

Proof. We note that

$$\begin{aligned}
&\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}\rangle\langle\epsilon_\ell,\hat{w}_{i,N}\rangle-\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{c}_{j,N}v_j\rangle\langle\epsilon_\ell,\hat{d}_{i,N}w_i\rangle\right|\\
&\quad\le\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}-\hat{c}_{j,N}v_j\rangle\langle\epsilon_\ell,\hat{w}_{i,N}\rangle\right|+\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{c}_{j,N}v_j\rangle\langle\epsilon_\ell,\hat{w}_{i,N}-\hat{d}_{i,N}w_i\rangle\right|.
\end{aligned}$$

Using the Cauchy–Schwarz inequality we get that

$$\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell, \hat{v}_{j,N}-\hat{c}_{j,N}v_j\rangle\langle\epsilon_\ell, \hat{w}_{i,N}\rangle\right| \le \sup_{t\in[0,1]}\left\|\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(x)\epsilon_\ell(s)\right\|\left(\|\hat{v}_{j,N}-\hat{c}_{j,N}v_j\|\,\|\hat{w}_{i,N}\|\right) = O_P(\log N),$$

on account of (5.6), Theorem 5.3 and $\|\hat{w}_{i,N}\| = 1$. Similar arguments give that

$$\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell, \hat{c}_{j,N}v_j\rangle\langle\epsilon_\ell, \hat{w}_{i,N}-\hat{d}_{i,N}w_i\rangle\right| = O_P(\log N),$$

completing the proof of the lemma. □

Lemma 6.4. *If Assumptions 2.1–2.6 hold, then we have*

$$\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell, \hat{v}_{j,N}\rangle\langle\eta_{\ell,1}, \hat{w}_{i,N}\rangle - \lfloor Nt\rfloor R_N^{(1)}(i,j)\right| = O_P(\log N).$$

Proof. Using the orthogonality of the w_i's we get that

$$\begin{aligned}\langle\eta_{\ell,1}, w_i\rangle &= \int_0^1\sum_{r=q+1}^\infty\sum_{n=1}^\infty\psi_{r,n}w_r(x)\left\{\int_0^1 v_n(s)X_\ell(s)ds\right\}w_i(x)dx\\ &= \sum_{r=q+1}^\infty\sum_{n=1}^\infty\psi_{r,n}\int_0^1 v_n(s)X_\ell(s)ds\left(\int_0^1 w_i(x)w_r(x)dx\right)\\ &= 0.\end{aligned}$$

Therefore we have

$$\langle X_\ell, \hat{v}_{j,N}\rangle\langle\eta_{\ell,1}, \hat{w}_{i,N}\rangle = \langle X_\ell, \hat{v}_{j,N}-\hat{c}_{j,N}v_j\rangle\langle\eta_{\ell,1}, \hat{w}_{i,N}-\hat{d}_{i,N}w_i\rangle + \langle X_\ell, \hat{c}_{j,N}v_j\rangle\langle\eta_{\ell,1}, \hat{w}_{i,N}-\hat{d}_{i,N}w_i\rangle.$$

Now,

$$\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell, \hat{c}_{j,N}v_j\rangle\langle\eta_{\ell,1}, \hat{w}_{i,N}-\hat{d}_{i,N}w_i\rangle = A^{(1)}_{\lfloor Nt\rfloor} + A^{(2)}_{\lfloor Nt\rfloor},$$

where

$$A^{(1)}_{\lfloor Nt\rfloor} = \hat{c}_{j,N}\int_0^1\int_0^1\int_0^1 v_j(z)\sum_{r=q+1}^\infty\sum_{n=1}^\infty\psi_{r,n}w_r(x)v_n(s)\left(\hat{w}_{i,N}(x)-\hat{d}_{i,N}w_i(x)\right)\left(\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(z)X_\ell(s)-\lfloor Nt\rfloor C(z,s)\right)dzdsdx$$

and

$$\begin{aligned}A^{(2)}_{\lfloor Nt\rfloor} &= \lfloor Nt\rfloor\hat{c}_{j,N}\int_0^1\int_0^1\int_0^1 v_j(z)\sum_{r=q+1}^\infty\sum_{n=1}^\infty\psi_{r,n}w_r(x)v_n(s)\left(\hat{w}_{i,N}(x)-\hat{d}_{i,N}w_i(x)\right)C(z,s)dzdsdx\\ &= \lfloor Nt\rfloor\hat{c}_{j,N}\int_0^1\lambda_j\sum_{r=q+1}^\infty\psi_{r,j}w_r(x)\left(\hat{w}_{i,N}(x)-\hat{d}_{i,N}w_i(x)\right)dx\\ &= \lfloor Nt\rfloor R_N^{(1)}(i,j),\end{aligned}$$

where we used that the v_j's are orthonormal eigenfunctions of C.

Applying (5.5) and (5.10) again we conclude

$$\sup_{t\in[0,1]}\left|A^{(1)}_{\lfloor Nt\rfloor}\right| = O_P(\log N).$$

Finally, using Theorems 5.2 and 5.3, we obtain that

$$\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}-\hat{c}_{j,N}v_j\rangle\langle\eta_{\ell,1},\hat{w}_{i,N}-\hat{d}_{i,N}w_i\rangle\right|$$
$$\leq\left|\left|\hat{v}_{j,N}(z)-\hat{c}_{j,N}v_j(z)\right|\right|\left|\left|\sum_{r=q+1}^{\infty}\sum_{n=1}^{\infty}\psi_{r,n}w_r(x)v_n(s)\right|\right|\left|\left|\hat{w}_{i,N}(x)-\hat{d}_{i,N}w_i(x)\right|\right|\sup_{t\in[0,1]}\left|\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(z)X_\ell(s)\right|\right|$$
$$=O_P\left(N^{-1/2}\right)O(1)O_P\left(N^{-1/2}\right)O_P\left(N\right). \quad\square$$

Lemma 6.5. *If Assumptions 2.1–2.6 hold, then we have*

$$\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}\rangle\langle\eta_{\ell,2},\hat{w}_{i,N}\rangle-\left(\hat{c}_{j,N}\hat{d}_{i,N}T^{(2)}_{\lfloor Nt\rfloor}(i,j)+\lfloor Nt\rfloor R^{(2)}_N(i,j)\right)\right|=O_P\left(\log N\right),$$

where

$$T^{(2)}_{\lfloor Nt\rfloor}(i,j)=\sum_{\ell=1}^{\lfloor Nt\rfloor}\int_0^1\int_0^1\left(X_\ell(s)X_\ell(z)-C(z,s)\right)\sum_{r=p+1}^{\infty}\psi_{ir}v_r(s)v_j(z)dzds.$$

Proof. First we write

$$\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}\rangle\langle\eta_{\ell,2},\hat{w}_{i,N}\rangle=A^{(3)}_{\lfloor Nt\rfloor}+A^{(4)}_{\lfloor Nt\rfloor}+A^{(5)}_{\lfloor Nt\rfloor}+A^{(6)}_{\lfloor Nt\rfloor},$$

where

$$A^{(3)}_{\lfloor Nt\rfloor}=\hat{c}_{j,N}\hat{d}_{i,N}\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,v_j\rangle\langle\eta_{\ell,2},w_i\rangle,$$
$$A^{(4)}_{\lfloor Nt\rfloor}=\hat{c}_{j,N}\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,v_j\rangle\langle\eta_{\ell,2},\hat{w}_{i,N}-\hat{d}_{i,N}w_i\rangle,$$
$$A^{(5)}_{\lfloor Nt\rfloor}=\hat{d}_{i,N}\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}-\hat{c}_{j,N}v_j\rangle\langle\eta_{\ell,2},w_i\rangle,$$
$$A^{(6)}_{\lfloor Nt\rfloor}=\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}-\hat{c}_{j,N}v_j\rangle\langle\eta_{\ell,2},\hat{w}_{i,N}-\hat{d}_{i,N}w_i\rangle.$$

The orthonormality of $\{w_i,\ 1\leq i<\infty\}$ shows that for all $1\leq i\leq q$

$$\langle\eta_{\ell,2},w_i\rangle=\int_0^1\sum_{r=1}^{q}\sum_{n=p+1}^{\infty}\psi_{r,n}w_r(x)\left\{\int_0^1v_n(s)X_\ell(s)ds\right\}w_i(x)dx$$
$$=\sum_{n=p+1}^{\infty}\psi_{i,n}\int_0^1v_n(s)X_\ell(s)ds.$$

Therefore, using again that the v_j's are orthonormal eigenfunctions of C we have

$$A^{(3)}_{\lfloor Nt\rfloor}==\hat{c}_{j,N}\hat{d}_{i,N}\sum_{n=p+1}^{\infty}\psi_{i,n}\int_0^1\int_0^1v_j(z)v_n(s)\left(\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(z)X_\ell(s)-\lfloor Nt\rfloor C(s,z)\right)dsdz$$
$$=\hat{c}_{j,N}\hat{d}_{i,N}T^{(2)}_{\lfloor Nt\rfloor}(i,j).$$

We decompose $A^{(4)}_{\lfloor Nt\rfloor}$ as

$$A^{(4)}_{\lfloor Nt\rfloor}=\hat{c}_{j,N}\sum_{\ell=1}^{\lfloor Nt\rfloor}\int_0^1\int_0^1X_\ell(z)v_j(z)\left(\hat{w}_{i,N}(x)-\hat{d}_{i,N}w_i(x)\right)\sum_{r=1}^{q}\sum_{n=p+1}^{\infty}\psi_{r,n}w_r(x)\int_0^1v_n(s)X_\ell(s)dsdzdx$$
$$=\hat{c}_{j,N}\int_0^1\int_0^1\int_0^1v_j(z)\left(\hat{w}_{i,N}(x)-\hat{d}_{i,N}w_i(x)\right)\sum_{r=1}^{q}\sum_{n=p+1}^{\infty}\psi_{r,n}w_r(x)v_n(s)\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(s)X_\ell(z)dsdzdx$$
$$=A^{(4)}_{\lfloor Nt\rfloor,1}+A^{(4)}_{\lfloor Nt\rfloor,2},$$

where

$$A^{(4)}_{\lfloor Nt\rfloor,1} = \hat{c}_{j,N}\int_0^1\int_0^1\int_0^1 v_j(z)\left(\hat{w}_{i,N}(x)-\hat{d}_{i,N}w_i(x)\right)\sum_{r=1}^{q}\sum_{n=p+1}^{\infty}\psi_{r,n}w_r(x)v_n(s) \times\left(\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(s)X_\ell(z)-\lfloor Nt\rfloor C(s,z)\right)dsdzdx$$

and

$$\begin{aligned}A^{(4)}_{\lfloor Nt\rfloor,2} &= \hat{c}_{j,N}\lfloor Nt\rfloor\int_0^1\int_0^1\int_0^1 v_j(z)\left(\hat{w}_{i,N}(x)-\hat{d}_{i,N}w_i(x)\right)\sum_{r=1}^{q}\sum_{n=p+1}^{\infty}\psi_{r,n}w_r(x)v_n(s)C(s,z)dsdzdx\\ &= \hat{c}_{j,N}\lfloor Nt\rfloor\int_0^1\int_0^1 \lambda_j v_j(s)\left(\hat{w}_{i,N}(x)-\hat{d}_{i,N}w_i(x)\right)\sum_{r=1}^{q}\sum_{n=p+1}^{\infty}\psi_{r,n}w_r(x)v_n(s)dsdx\\ &= 0,\end{aligned}$$

using again that the v_j's are eigenfunctions of C. Therefore we obtain

$$\begin{aligned}\sup_{t\in[0,1]}\left|A^{(4)}_{\lfloor Nt\rfloor}\right| &= \sup_{t\in[0,1]}\left|A^{(4)}_{\lfloor Nt\rfloor,1}\right|\\ &\le \left\|\hat{w}_{i,N}(x)-\hat{d}_{i,N}w_i(x)\right\|\left\|v_j(z)\sum_{r=1}^{q}\sum_{n=p+1}^{\infty}\psi_{r,n}w_r(x)v_n(s)\right\|\sup_{t\in[0,1]}\left\|\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(s)X_\ell(z)-\lfloor Nt\rfloor C(s,z)\right\|\\ &= O_P\left(N^{-1/2}\right)O_P(1)\,O_P\left(N^{1/2}\log N\right).\end{aligned}$$

Similar arguments give

$$\begin{aligned}A^{(5)}_{\lfloor Nt\rfloor} &= \hat{d}_{i,N}\sum_{\ell=1}^{\lfloor Nt\rfloor}\int_0^1\int_0^1 X_\ell(z)\left(\hat{v}_{j,N}(z)-\hat{c}_{j,N}v_j(z)\right)w_i(x)\sum_{r=1}^{q}\sum_{n=p+1}^{\infty}\psi_{r,n}w_r(x)\int_0^1 v_n(s)X_\ell(s)dsdzdx\\ &= A^{(5)}_{\lfloor Nt\rfloor,1}+A^{(5)}_{\lfloor Nt\rfloor,2},\end{aligned}$$

where

$$A^{(5)}_{\lfloor Nt\rfloor,1} = \hat{d}_{i,N}\int_0^1\int_0^1\left(\hat{v}_{j,N}(z)-\hat{c}_{j,N}v_j(z)\right)\sum_{n=p+1}^{\infty}\psi_{i,n}v_n(s)\left(\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(s)X_\ell(z)-\lfloor Nt\rfloor C(s,z)\right)dsdz$$

and

$$\begin{aligned}A^{(5)}_{\lfloor Nt\rfloor,2} &= \hat{d}_{i,N}\lfloor Nt\rfloor\int_0^1\int_0^1\int_0^1\left(\hat{v}_{j,N}(z)-\hat{c}_{j,N}v_j(z)\right)w_i(x)\sum_{r=1}^{q}\sum_{n=p+1}^{\infty}\psi_{r,n}w_r(x)v_n(s)C(s,z)dsdzdx\\ &= \lfloor Nt\rfloor R_N^{(2)}(i,j).\end{aligned}$$

Repeating our previous arguments we get that

$$\begin{aligned}\sup_{t\in[0,1]}\left|A^{(5)}_{\lfloor Nt\rfloor,1}\right| &\le \left\|\hat{v}_{j,N}(z)-\hat{c}_{j,N}v_j(z)\right\|\left\|\sum_{n=p+1}^{\infty}\psi_{i,n}v_n(s)\right\|\sup_{t\in[0,1]}\left\|\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(s)X_\ell(z)-\lfloor Nt\rfloor C(s,z)\right\|\\ &= O_P\left(N^{-1/2}\right)O(1)O_P\left(N^{1/2}\log N\right).\end{aligned}$$

Similarly, using the Cauchy–Schwarz inequality with (5.2) and Theorem 5.2, we conclude that

$$\sup_{t\in[0,1]}\left|A^{(6)}_{\lfloor Nt\rfloor}\right|O_P(1),$$

completing the proof of the lemma. □

Lemma 6.6. *If Assumptions 2.1–2.6 hold, then we have*

$$\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}\rangle\langle\eta_{\ell,3},\hat{w}_{i,N}\rangle-\lfloor Nt\rfloor R_N^{(3)}(i,j)\right| = O_P(\log N).$$

Proof. We write

$$\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}\rangle\langle\eta_{\ell,3},\hat{w}_{i,N}\rangle = A^{(7)}_{\lfloor Nt\rfloor}+A^{(8)}_{\lfloor Nt\rfloor}+A^{(9)}_{\lfloor Nt\rfloor},$$

where

$$A^{(7)}_{\lfloor Nt\rfloor} = \sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}-\hat{c}_{j,N}v_j\rangle\langle\eta_{\ell,3},\hat{w}_{i,N}\rangle,$$

$$A^{(8)}_{\lfloor Nt\rfloor} = \hat{c}_{j,N}\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,v_j\rangle\langle\eta_{\ell,3},\hat{w}_{i,N}-\hat{d}_{i,N}w_i\rangle,$$

$$A^{(9)}_{\lfloor Nt\rfloor} = \hat{c}_{j,N}\hat{d}_{i,N}\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,v_j\rangle\langle\eta_{\ell,3},w_i\rangle.$$

Theorems 5.1 and 5.2 imply that

$$\begin{aligned}\sup_{t\in[0,1]}\left|A^{(7)}_{\lfloor Nt\rfloor}\right| &= \sup_{t\in[0,1]}\left|\int_0^1\int_0^1\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(z)\left(\hat{v}_{j,N}(z)-\hat{c}_{j,N}v_j(z)\right)\hat{w}_{i,N}(x)\sum_{r=1}^{q}\sum_{n=1}^{p}\psi_{r,n}\hat{c}_{n,N}w_r(x)\right.\\ &\qquad\left.\times\int_0^1\left(\hat{c}_{n,N}v_n(s)-\hat{v}_{n,N}(s)\right)X_\ell(s)dsdzdx\right|\\ &\le \sum_{r=1}^{q}\sum_{n=1}^{p}\sup_{t\in[0,1]}\left\|\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(s)X_\ell(z)\right\|\left\|\hat{v}_{j,N}(z)-\hat{c}_{j,N}v_j(z)\right\|\left\|\hat{w}_{i,N}(x)\psi_{r,n}\hat{c}_{n,N}w_r(x)\right\|\left\|\hat{c}_{n,N}v_n(s)-\hat{v}_{n,N}(s)\right\|\\ &= O_P(N)\,O_P\left(N^{-1/2}\right)O(1)O_P\left(N^{-1/2}\right),\end{aligned}$$

and similarly

$$\sup_{t\in[0,1]}\left|A^{(8)}_{\lfloor Nt\rfloor}\right| = O_P(1).$$

Next we observe that

$$\sup_{t\in[0,1]}\left|A^{(9)}_{\lfloor Nt\rfloor}\right| = A^{(9)}_{\lfloor Nt\rfloor,1}+A^{(9)}_{\lfloor Nt\rfloor,2},$$

where

$$\begin{aligned}A^{(9)}_{\lfloor Nt\rfloor,1} &= \hat{c}_{j,N}\hat{d}_{i,N}\lfloor Nt\rfloor\int_0^1\int_0^1\lambda_j v_j(s)w_i(x)\sum_{r=1}^{q}\sum_{n=1}^{p}\psi_{r,n}\hat{c}_{n,N}w_r(x)\left(\hat{c}_{n,N}v_n(s)-\hat{v}_{n,N}(s)\right)dsdx\\ &= \lfloor Nt\rfloor R^{(3)}_N(i,j)\end{aligned}$$

and

$$\begin{aligned}A^{(9)}_{\lfloor Nt\rfloor,2} &= \hat{c}_{j,N}\hat{d}_{i,N}\int_0^1\int_0^1\int_0^1\left(\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(z)X_\ell(s)-\lfloor Nt\rfloor C(z,s)\right)v_j(z)w_i(x)\\ &\quad\times\sum_{r=1}^{q}\sum_{n=1}^{p}\psi_{r,n}\hat{c}_{n,N}w_r(x)\left(\hat{c}_{n,N}v_n(s)-\hat{v}_{n,N}(s)\right)dsdzdx.\end{aligned}$$

Using Theorems 5.2 and 5.3 again, we obtain that

$$\sup_{t\in[0,1]}\left|A^{(9)}_{\lfloor Nt\rfloor,2}\right| = O_P(\log N).$$

This completes the proof. □

Lemma 6.7. *If Assumptions 2.1–2.6 hold, then we have*

$$\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}\rangle\langle\eta_{\ell,4},\hat{w}_{i,N}\rangle-\lfloor Nt\rfloor R^{(4)}_N(i,j)\right| = O_P(\log N).$$

Proof. Following the proofs of the previous lemmas we write

$$\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}\rangle\langle\eta_{\ell,4},\hat{w}_{i,N}\rangle=A^{(10)}_{\lfloor Nt\rfloor}+A^{(11)}_{\lfloor Nt\rfloor}+A^{(12)}_{\lfloor Nt\rfloor}+A^{(13)}_{\lfloor Nt\rfloor},$$

where

$$A^{(10)}_{\lfloor Nt\rfloor}=\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}-\hat{c}_{j,N}v_j\rangle\langle\eta_{\ell,4},\hat{w}_{i,N}-\hat{d}_{i,N}w_i\rangle,$$

$$A^{(11)}_{\lfloor Nt\rfloor}=\hat{c}_{j,N}\hat{d}_{i,N}\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,v_j\rangle\langle\eta_{\ell,4},w_i\rangle,$$

$$A^{(12)}_{\lfloor Nt\rfloor}=\hat{d}_{i,N}\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}-\hat{c}_{j,N}v_j\rangle\langle\eta_{\ell,4},w_i\rangle,$$

$$A^{(13)}_{\lfloor Nt\rfloor}=\hat{c}_{j,N}\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,v_j\rangle\langle\eta_{\ell,4},\hat{w}_{i,N}-\hat{d}_{i,N}w_i\rangle.$$

Repeating the arguments used in the proofs of Lemmas 6.4 and 6.5, one can show that

$$\sup_{t\in[0,1]}\left|A^{(10)}_{\lfloor Nt\rfloor}\right|=O_P(1),$$

$$\sup_{t\in[0,1]}\left|A^{(12)}_{\lfloor Nt\rfloor}\right|=O_P(1),$$

$$\sup_{t\in[0,1]}\left|A^{(13)}_{\lfloor Nt\rfloor}\right|=O_P(1).$$

Elementary arguments give

$$A^{(11)}_{\lfloor Nt\rfloor}=A^{(11)}_{\lfloor Nt\rfloor,1}+A^{(11)}_{\lfloor Nt\rfloor,2},$$

where

$$A^{(11)}_{\lfloor Nt\rfloor,2}=\hat{c}_{j,N}\hat{d}_{i,N}\int_0^1\int_0^1\int_0^1\left(\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(z)X_\ell(s)-\lfloor Nt\rfloor C(z,s)\right)v_j(z)w_i(x)$$
$$\times\sum_{r=1}^{q}\sum_{n=1}^{p}\hat{d}_{r,N}\psi_{r,n}\hat{c}_{n,N}\left(\hat{d}_{r,N}w_r(x)-\hat{w}_{r,N}(x)\right)\hat{c}_{n,N}v_n(s)dsdzdx,$$

and

$$A^{(11)}_{\lfloor Nt\rfloor,1}=\lfloor Nt\rfloor\hat{c}_{j,N}\hat{d}_{i,N}\lambda_j\int_0^1 w_i(x)\sum_{r=1}^{q}\hat{d}_{r,N}\psi_{r,j}\hat{c}_{j,N}\left(\hat{d}_{r,N}w_r(x)-\hat{w}_{r,N}(x)\right)\hat{c}_{j,N}dx$$
$$=\lfloor Nt\rfloor R^{(4)}_N(i,j).$$

Using Theorems 5.2 and 5.3 again, we conclude that

$$\sup_{t\in[0,1]}\left|A^{(11)}_{\lfloor Nt\rfloor,2}\right|=O_P(\log N),$$

completing the proof. □

Lemma 6.8. *If Assumptions 2.1–2.5 hold, then we have*

$$\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}\rangle\langle\eta_{\ell,5},\hat{w}_{i,N}\rangle\right|=O_P(1).$$

Proof. It follows from Theorems 5.1 and 5.2 that

$$\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}\rangle\langle\eta_{\ell,5},\hat{w}_{i,N}\rangle\right|$$
$$=\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\int_0^1\int_0^1 X_\ell(z)\hat{v}_{j,N}(z)\hat{w}_{i,N}(x)\sum_{r=1}^{q}\sum_{n=1}^{p}\hat{d}_{r,N}\psi_{r,n}\hat{c}_{n,N}\left(\hat{w}_{r,N}(x)-\hat{d}_{r,N}w_r(x)\right)\right.$$
$$\left.\times\int_0^1\left(\hat{c}_{n,N}v_n(s)-\hat{v}_{n,N}(s)\right)X_\ell(s)dsdzdx\right|$$
$$\le\sum_{r=1}^{q}\sum_{n=1}^{p}|\psi_{r,n}|\sup_{t\in[0,1]}\left\|\sum_{\ell=1}^{\lfloor Nt\rfloor}X_\ell(z)X_\ell(s)\right\|\|\hat{w}_{i,N}(x)\|\|\hat{v}_{j,N}(z)\|\left\|\hat{w}_{r,N}(x)-\hat{d}_{r,N}w_r(x)\right\|\|\hat{c}_{n,N}v_n(s)-\hat{v}_{n,N}(s)\|$$
$$=O_P(N)\,O_P\left(N^{-1/2}\right)O_P\left(N^{-1/2}\right).\quad\square$$

Lemma 6.9. *If Assumptions 2.1–2.6 hold, then we have*

$$\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}\rangle\langle\epsilon_\ell^{**},\hat{w}_{i,N}\rangle-\left(\lfloor Nt\rfloor R_N(i,j)+\hat{c}_{j,N}\hat{d}_{i,N}\sum_{\ell=1}^{\lfloor Nt\rfloor}\gamma_\ell(i,j)\right)\right|=O_P(\log N).$$

Proof. Combining Lemmas 6.3–6.8, we immediately see that

$$\sup_{t\in[0,1]}\left|\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,\hat{v}_{j,N}\rangle\langle\epsilon_\ell^{**},\hat{w}_{i,N}\rangle-\left(\lfloor Nt\rfloor R_N(i,j)+\hat{c}_{j,N}\hat{d}_{i,N}T_{\lfloor Nt\rfloor}(i,j)\right)\right|=O_P(\log N),$$

where

$$T_{\lfloor Nt\rfloor}(i,j)=T^{(1)}_{\lfloor Nt\rfloor}(i,j)+T^{(2)}_{\lfloor Nt\rfloor}(i,j).$$

Thus we need only to show that

$$T^{(2)}_{\lfloor Nt\rfloor}(i,j)=\sum_{\ell=1}^{\lfloor Nt\rfloor}\langle X_\ell,v_j\rangle\langle X_\ell,u_i\rangle.$$

However, using that the v_j's are orthogonal eigenfunctions of C, we get that

$$\int_0^1\int_0^1 C(z,s)v_j(z)\sum_{r=p+1}^{\infty}\psi_{i,r}v_r(s)dsdz=\lambda_j\int_0^1 v_j(s)\sum_{r=p+1}^{\infty}\psi_{i,r}v_r(s)dsdz=0,$$

completing the proof. $\square$

Lemma 6.10. *If Assumptions 2.1–2.6 hold, then we have*

$$\left|\sqrt{N}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}_N)\right|=O_P(1).$$

Proof. It is easy to see that

$$\sqrt{N}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}_N)=-N^{-1/2}\left(\frac{\hat{\mathbf{Z}}_N^T\hat{\mathbf{Z}}_N}{N}\right)^{-1}\hat{\mathbf{Z}}_N^T\hat{\boldsymbol{\Delta}}_N.$$

It follows from (6.5) that

$$\left|\left(\frac{\hat{\mathbf{Z}}_N^T\hat{\mathbf{Z}}_N}{N}\right)^{-1}\right|=O_P(1).$$

Lemma 6.9 and (6.7) yield that

$$\left|\hat{\mathbf{Z}}_N^T\hat{\boldsymbol{\Delta}}_N\right|\le\max_{1\le i\le q,1\le j\le p}\left\{N|R_N(i,j)|+\left|\sum_{\ell=1}^{N}\gamma_\ell(i,j)\right|\right\}+O_P(\log N).$$

It follows from Theorem 5.2 that for all $1 \le i \le q$, $1 \le j \le p$

$$N|R_N(i,j)| = O_P\left(N^{1/2}\right)$$

while Theorem 5.4 implies that

$$\left|\sum_{\ell=1}^{N} \gamma_\ell(i,j)\right| = O_P\left(N^{1/2}\right). \quad \square$$

Lemma 6.11. *If Assumptions 2.1–2.6 hold, then we have*

$$\sup_{t\in[0,1]} \left|\tilde{\mathbf{V}}_N(t) - \boldsymbol{\zeta}_N \frac{1}{N^{1/2}} \left(\sum_{\ell=1}^{\lfloor Nt\rfloor} \boldsymbol{\gamma}_\ell - t\sum_{\ell=1}^{N} \boldsymbol{\gamma}_\ell\right)\right| = o_P(1).$$

Proof. Lemmas 6.2 and 6.10 and (6.1) imply that

$$\sup_{t\in[0,1]} \left|\tilde{\mathbf{V}}_N(t) - \frac{1}{N^{1/2}} \left(\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}^T \hat{\boldsymbol{\Delta}}_{\lfloor Nt\rfloor} - t\hat{\mathbf{Z}}_N^T \hat{\boldsymbol{\Delta}}_N\right)\right| = o_P(1).$$

It also follows from Lemma 6.9 and (6.7)

$$\sup_{t\in[0,1]} \left|\hat{\mathbf{Z}}_{\lfloor Nt\rfloor}^T \hat{\boldsymbol{\Delta}}_{\lfloor Nt\rfloor} - \mathrm{vec}\left(\left\{\lfloor Nt\rfloor R_N(i,j) + \hat{c}_{j,N}\hat{d}_{i,N} \sum_{\ell=1}^{\lfloor Nt\rfloor} \gamma_\ell(i,j)\right\}^T\right)\right| = O_P(\log N),$$

and therefore the proof is complete. $\square$

Now we have all the necessary tools to prove the main result.

Proof of Theorem 2.1. It follows from Lemma 6.11 and Theorem 5.4 that

$$\boldsymbol{\zeta}_N \tilde{\mathbf{V}}_N(t) \overset{\mathcal{D}^{pq}[0,1]}{\longrightarrow} \mathbf{W}_{\boldsymbol{\Sigma}}(t) - t\mathbf{W}_{\boldsymbol{\Sigma}}(1).$$

Next we observe that

$$\left\{\boldsymbol{\Sigma}^{-1/2}(\mathbf{W}_{\boldsymbol{\Sigma}}(t) - t\mathbf{W}_{\boldsymbol{\Sigma}}(1)), 0 \le t \le 1\right\} \overset{\mathcal{D}}{=} \{\mathbf{B}(t), 0 \le t \le 1\},$$

where $\mathbf{B}(t) = (\mathcal{B}_1(t), \dots, \mathcal{B}_{pq}(t))^T$ and $\mathcal{B}_1, \dots, \mathcal{B}_{pq}$ are independent, identically distributed Brownian bridges. Hence

$$(\boldsymbol{\zeta}_N \tilde{\mathbf{V}}_N(t))^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\zeta}_N \tilde{\mathbf{V}}_N(t)) \overset{\mathcal{D}[0,1]}{\longrightarrow} \sum_{\ell=1}^{pq} \mathcal{B}_\ell^2(t).$$

Now, using Assumption 2.7 with Slutsky's lemma, the proof is complete. $\square$

7. Proof of Theorems 3.1 and 3.2

We can assume without loss of generality that $K(u) = 0$ if $|u| > 1$. Let m be a positive integer and define

$$\boldsymbol{\gamma}_\ell^{(m)} = \mathrm{vec}(\{\gamma_\ell^{(m)}(i,j), 1 \le i \le q, 1 \le j \le p\}^T),$$

where

$$\gamma_\ell^{(m)}(i,j) = \langle X_\ell^{(m)}, v_j\rangle\langle \epsilon_\ell^{(m)}, w_i\rangle + \langle X_\ell^{(m)}, v_j\rangle\langle X_\ell^{(m)}, u_i\rangle.$$

The long term covariance matrix associated with the stationary sequence $\{\boldsymbol{\gamma}_\ell^{(m)}, 1 \le \ell < \infty\}$ is given by

$$\boldsymbol{\Sigma}^{(m)} = E\boldsymbol{\gamma}_1^{(m)}(\boldsymbol{\gamma}_1^{(m)})^T + \sum_{\ell=1}^{\infty} E\boldsymbol{\gamma}_1^{(m)}(\boldsymbol{\gamma}_{\ell+1}^{(m)})^T + \sum_{\ell=1}^{\infty} E\boldsymbol{\gamma}_{\ell+1}^{(m)}(\boldsymbol{\gamma}_1^{(m)})^T.$$

The corresponding Bartlett estimator is defined as

$$\tilde{\boldsymbol{\Sigma}}_N^{(m)} = \sum_{k=-(N-1)}^{N-1} K(k/B_N)\boldsymbol{\phi}_{k,N}^{(m)},$$

where

$$\boldsymbol{\phi}_{k,N}^{(m)} = \frac{1}{N} \sum_{\ell=\max(1,1-k)}^{\min(N,N-k)} \boldsymbol{\gamma}_\ell^{(m)} (\boldsymbol{\gamma}_{\ell+k}^{(m)})^T$$

are the sample covariances of lag k. Since K is symmetric, $K(0) = 1$ and $K(u) = 0$ outside $[-1, 1]$ we have that

$$\tilde{\boldsymbol{\Sigma}}_N^{(m)} = \boldsymbol{\phi}_{0,N}^{(m)} + \sum_{k=1}^{B_N} K(k/B_N)\boldsymbol{\phi}_{k,N}^{(m)} + \sum_{k=1}^{B_N} K(k/B_N)(\boldsymbol{\phi}_{k,N}^{(m)})^T$$

for all sufficiently large N.

We start with the consistency of $\tilde{\boldsymbol{\Sigma}}_N^{(m)}$.

Lemma 7.1. *If Assumptions 3.1 and 3.2 are satisfied, then we have for every m*

$$\tilde{\boldsymbol{\Sigma}}_N^{(m)} \xrightarrow{\mathcal{P}} \boldsymbol{\Sigma}^{(m)},$$

as $N \to \infty$.

Proof. Since the sequence $\boldsymbol{\gamma}_\ell^{(m)}$ is m-dependent we have that

$$\boldsymbol{\Sigma}^{(m)} = E\boldsymbol{\gamma}_1\boldsymbol{\gamma}_1^T + \sum_{\ell=1}^{m} E\boldsymbol{\gamma}_1\boldsymbol{\gamma}_{\ell+1}^T + \sum_{\ell=1}^{m} E\boldsymbol{\gamma}_{\ell+1}\boldsymbol{\gamma}_1^T.$$

It follows from the ergodic theorem that for any fixed k and m

$$\boldsymbol{\phi}_{k,N}^{(m)} \xrightarrow{\mathcal{P}} E\boldsymbol{\gamma}_1^{(m)}(\boldsymbol{\gamma}_{1+k}^{(m)})^T.$$

So using Assumption 3.1(i), (ii) and 3.2 we get that

$$\boldsymbol{\phi}_{0,N}^{(m)} + \sum_{k=1}^{m} K(k/B_N)\boldsymbol{\phi}_{k,N}^{(m)} + \sum_{k=1}^{m} K(k/B_N)(\boldsymbol{\phi}_{k,N}^{(m)})^T \xrightarrow{\mathcal{P}} E\boldsymbol{\gamma}_1\boldsymbol{\gamma}_1^T + \sum_{\ell=1}^{m} E\boldsymbol{\gamma}_1\boldsymbol{\gamma}_{\ell+1}^T + \sum_{\ell=1}^{m} E\boldsymbol{\gamma}_{\ell+1}\boldsymbol{\gamma}_1^T.$$

Lemma 7.1 is proven if we show that

$$\sum_{k=m+1}^{B_N} K(k/B_N)\boldsymbol{\phi}_{k,N}^{(m)} \xrightarrow{\mathcal{P}} 0 \tag{7.1}$$

and

$$\sum_{k=m+1}^{B_N} K(k/B_N)(\boldsymbol{\phi}_{k,N}^{(m)})^T \xrightarrow{\mathcal{P}} 0. \tag{7.2}$$

Clearly, it is enough to prove (7.1).

Let

$$\mathbf{G}_N^{(m)} = \sum_{k=m+1}^{B_N} K(k/B_N)\boldsymbol{\phi}_{k,N}^{(m)}.$$

Elementary arguments show that

$$\begin{aligned} \mathbf{G}_N^{(m)} &= \sum_{k=m+1}^{B_N} K(k/B_N)\boldsymbol{\phi}_{k,N}^{(m)} \\ &= \sum_{k=m+1}^{B_N} K(k/B_N)\frac{1}{N}\sum_{\ell=1}^{N-k} \boldsymbol{\gamma}_\ell^{(m)} \left(\boldsymbol{\gamma}_{\ell+k}^{(m)}\right)^T \\ &= \sum_{\ell=1}^{N-(m+1)} \boldsymbol{\gamma}_\ell^{(m)} \mathbf{H}_{\ell,N}^{(m)}, \end{aligned}$$

where

$$\mathbf{H}_{\ell,N}^{(m)} = \sum_{k=m+1}^{\min(N-\ell,B_N)} \frac{K(k/B_N)}{N} \left(\boldsymbol{\gamma}_{\ell+k}^{(m)}\right)^T.$$

Let

$$G_N^{(m)}(i,j) = \sum_{\ell=1}^{N-(m+1)} \gamma_\ell^{(m)}(i) H_{\ell,N}^{(m)}(j), \quad 1 \le i,j \le pq,$$

where $\gamma_\ell^{(m)}(i)$ and $H_{\ell,N}^{(m)}(j)$ are the ith and the jth coordinates of the vectors $\boldsymbol{\gamma}_{\ell,N}^{(m)}$ and $\mathbf{H}_{\ell,N}^{(m)}$, respectively. Next we write

$$\begin{aligned} E\left(G_N^{(m)}(i,j)\right)^2 &= E\left(\sum_{\ell=1}^{N-(m+1)} \gamma_\ell^{(m)}(i) H_{\ell,N}^{(m)}(j)\right)^2 \\ &= \sum_{\substack{1\le r\le N-(m+1) \\ 1\le \ell \le N-(m+1)}}\sum E\left(H_{\ell,N}^{(m)}(j)\gamma_\ell^{(m)}(i)\gamma_r^{(m)}(i)H_{r,N}^{(m)}(j)\right) \\ &= G_{1,N}^{(m)}(i,j) + G_{2,N}^{(m)}(i,j), \end{aligned}$$

where

$$G_{1,N}^{(m)}(i,j) = \sum_{\substack{1\le r\le N-(m+1) \\ 1\le \ell \le N-(m+1) \\ |r-\ell|\le m}}\sum E\left(H_{\ell,N}^{(m)}(j)\gamma_\ell^{(m)}(i)\gamma_r^{(m)}(i)H_{r,N}^{(m)}(j)\right),$$

and

$$G_{2,N}^{(m)}(i,j) = \sum_{\substack{1\le r\le N-(m+1) \\ 1\le \ell \le N-(m+1) \\ |r-\ell|> m}}\sum E\left(H_{\ell,N}^{(m)}(j)\gamma_\ell^{(m)}(i)\gamma_r^{(m)}(i)H_{r,N}^{(m)}(j)\right).$$

Notice that $\boldsymbol{\gamma}_\ell^{(m)}$ is independent of $\mathbf{H}_{\ell,N}^{(m)}$, $\mathbf{H}_{r,N}^{(m)}$ and $\boldsymbol{\gamma}_r^{(m)}$, if $r > m + \ell$. Hence

$$\begin{aligned} E\left(H_{\ell,N}^{(m)}(j)\gamma_\ell^{(m)}(i)\gamma_r^{(m)}(i)H_{r,N}^{(m)}(j)\right) &= \begin{cases} E\gamma_\ell^{(m)}(i)E\left(H_{\ell,N}^{(m)}(j)\gamma_r^{(m)}(i)H_{r,N}^{(m)}(j)\right) & r > m+\ell, \\ E\gamma_r^{(m)}(i)E\left(H_{\ell,N}^{(m)}(j)\gamma_\ell^{(m)}(i)H_{r,N}^{(m)}(j)\right) & \ell > m+r, \\ E\left(H_{\ell,N}^{(m)}(j)\gamma_\ell^{(m)}(i)\gamma_r^{(m)}(i)H_{r,N}^{(m)}(j)\right) & |\ell - r| \le m, \end{cases} \\ &= \begin{cases} 0 & |\ell - r| > m, \\ E\left(H_{\ell,N}^{(m)}(j)\gamma_\ell^{(m)}(i)\gamma_r^{(m)}(i)H_{r,N}^{(m)}(j)\right) & |\ell - r| \le m. \end{cases} \end{aligned}$$

Thus we have

$$EG_{2,N}^{(m)}(i,j) = 0.$$

Let M be an upper bound on $|K(t)|$. Using the fact that $\boldsymbol{\gamma}_\ell^{(m)}$ is an m-dependent sequence, we now obtain the following:

$$\begin{aligned} E(H_{\ell,N}^{(m)}(j))^2 &= \sum_{k=m+1}^{\min(N-\ell,B_N)} \sum_{v=m+1}^{\min(N-\ell,B_N)} \frac{K(k/B_N)}{N}\frac{K(v/B_N)}{N} E\left(\gamma_{\ell+k}^{(m)}(j)\gamma_{\ell+v}^{(m)}(j)\right) \\ &\le \frac{M^2}{N^2} \sum_{k=m+1}^{\min(N-\ell,B_N)} \sum_{v=m+1}^{\min(N-\ell,B_N)} E\left(\gamma_{\ell+k}^{(m)}(j)\gamma_{\ell+v}^{(m)}(j)\right) \\ &\le \frac{M^2}{N^2} B_N \sum_{r=-m}^{m} E\left|\gamma_0^{(m)}(j)\gamma_r^{(m)}(j)\right| \\ &= O\left(\frac{B_N}{N^2}\right). \end{aligned} \tag{7.3}$$

In the next step we will first use the Cauchy–Schwarz inequality, then the independence of $H^{(m)}_{\ell,N}(j)$ and $\gamma^{(m)}_{\ell}(i)$ and the independence of $H^{(m)}_{r,N}(j)$ and $\gamma^{(m)}_{r}(i)$ to get

$$\begin{aligned}
\left|G^{(m)}_{2,N}(i,j)\right| &\le \sum\sum_{\substack{1\le r\le N-(m+1)\\ 1\le \ell\le N-(m+1)\\ |r-\ell|\le m}} E\left|H^{(m)}_{\ell,N}(j)\gamma^{(m)}_{\ell}(i)\gamma^{(m)}_{r}(i)H^{(m)}_{r,N}(j)\right| \\
&\le \sum\sum_{\substack{1\le r\le N-(m+1)\\ 1\le \ell\le N-(m+1)\\ |r-\ell|\le m}} \left(E\left(H^{(m)}_{\ell,N}(j)\gamma^{(m)}_{\ell}(i)\right)^2\right)^{1/2}\left(E\left(\gamma^{(m)}_{r}(i)H^{(m)}_{r,N}(j)\right)^2\right)^{1/2} \\
&\le \sum\sum_{\substack{1\le r\le N-(m+1)\\ 1\le \ell\le N-(m+1)\\ |r-\ell|\le m}} \left(E\left(H^{(m)}_{\ell,N}(j)\right)^2\right)^{1/2}\left(E\left(\gamma^{(m)}_{\ell}(i)\right)^2\right)^{1/2}\left(E\left(\gamma^{(m)}_{r}(i)\right)^2\right)^{1/2}\left(E\left(H^{(m)}_{r,N}(j)\right)^2\right)^{1/2} \\
&\le 2mNO\left(\frac{B_N^{1/2}}{N}\right)O(1)O(1)O\left(\frac{B_N^{1/2}}{N}\right) \\
&= O\left(\frac{B_N}{N}\right) \\
&= o(1),
\end{aligned}$$

where we also used (7.3) and Assumption 3.2. This completes the proof of Lemma 7.1. □

Let $\mathbf{i}^2 = -1$.

Lemma 7.2. *If Assumptions 2.1–2.4, 3.1 and 3.2 are satisfied, then for all $1 \le j \le pq$ we have*

$$\limsup_{N\to\infty}\limsup_{m\to\infty}\sup_{-\infty<t<\infty} E\left(\frac{1}{N^{1/2}}\sum_{k=1}^{N}(\gamma_k(j)-\gamma^{(m)}_k(j))e^{\mathbf{i}kt}\right)^2 = 0, \tag{7.4}$$

$$\limsup_{N\to\infty}\limsup_{m\to\infty}\sup_{-\infty<t<\infty} E\left(\frac{1}{N^{1/2}}\sum_{k=1}^{N}\gamma_k(j)e^{\mathbf{i}kt}\right)^2 < \infty \tag{7.5}$$

and

$$\limsup_{N\to\infty}\limsup_{m\to\infty}\sup_{-\infty<t<\infty} E\left(\frac{1}{N^{1/2}}\sum_{k=1}^{N}\gamma^{(m)}_k(j)e^{\mathbf{i}kt}\right)^2 < \infty. \tag{7.6}$$

Proof. First we note that

$$\begin{aligned}
E\left(\sum_{k=1}^{N}(\gamma_k(j)-\gamma^{(m)}_k(j))e^{\mathbf{i}kt}\right)^2 &= \sum_{1\le k\le N} E((\gamma_k(j)-\gamma^{(m)}_k(j))e^{\mathbf{i}kt})^2 \\
&\quad + 2\sum_{1\le k<\ell\le N} E\left[(\gamma_k(j)-\gamma^{(m)}_k(j))(\gamma_\ell(j)-\gamma^{(m)}_\ell(j))\right]e^{\mathbf{i}(k+\ell)t}.
\end{aligned}$$

It follows from (5.13) that there is a sequence $c_1(m) \to 0$ such that

$$\left|\sum_{1\le k\le N} E(\gamma_k(j)-\gamma^{(m)}_k(j))^2 e^{\mathbf{i}2kt}\right| \le Nc_1(m).$$

Next we write

$$\sum_{1\le k<\ell\le N} E\left[(\gamma_k(j)-\gamma^{(m)}_k(j))\gamma_\ell(j)\right]e^{\mathbf{i}(k+\ell)t} = \sum_{1\le k<\ell\le N} E\left[(\gamma_k(j)-\gamma^{(m)}_k(j))(\gamma_\ell(j)-\gamma^{(\ell-k)}_\ell(j))\right]e^{\mathbf{i}(k+\ell)t},$$

since $(\boldsymbol{\gamma}_k, \boldsymbol{\gamma}_k^{(m)})$ and $\boldsymbol{\gamma}_\ell^{(\ell-k)}$ are independent. Using the Cauchy–Schwarz inequality first, then (5.13) again we get that

$$\begin{aligned}
&\sum_{1\le k<\ell\le N} \left| E\left[(\gamma_k(j) - \gamma_k^{(m)}(j))(\gamma_\ell(j) - \gamma_\ell^{(\ell-k)}(j)) \right] e^{\mathbf{i}(k+\ell)t} \right| \\
&\quad \le \sum_{1\le k<\ell\le N} \left[E(\gamma_k(j) - \gamma_k^{(m)}(j))^2 \right]^{1/2} \left[E(\gamma_\ell(j) - \gamma_\ell^{(\ell-k)}(j))^2 \right]^{1/2} \\
&\quad \le N \left[E(\gamma_1(j) - \gamma_1^{(m)}(j))^2 \right]^{1/2} \sum_{1\le k<\infty} \left[E(\gamma_1(j) - \gamma_1^{(k)}(j))^2 \right]^{1/2} \\
&\quad = Nc_2(m)
\end{aligned}$$

with some sequence $c_2(m) \to 0$. Similar arguments show that

$$\sum_{1\le k<\ell\le N} \left| E\left[(\gamma_k(j) - \gamma_k^{(m)}(j))\gamma_\ell^{(m)}(j) \right] e^{\mathbf{i}(k+\ell)t} \right| = Nc_3(m)$$

with some sequence $c_3(m) \to 0$, completing the proof of (7.4).

Similar to the proof of (7.4), we write

$$\begin{aligned}
E\left(\sum_{k=1}^N \gamma_k(j) e^{\mathbf{i}kt} \right)^2 &= \sum_{k=1}^N \sum_{\ell=1}^N E\gamma_k(j)\gamma_\ell(j) e^{\mathbf{i}(k+\ell)t} \\
&= \sum_{k=1}^N E\gamma_k^2(j) e^{2\mathbf{i}kt} + 2 \sum_{1\le k<\ell\le N} E\gamma_k(j)\gamma_\ell(j) e^{\mathbf{i}(k+\ell)t} \\
&= E\gamma_1^2(j) \sum_{k=1}^N e^{2\mathbf{i}kt} + 2 \sum_{1\le k<\ell\le N} E\gamma_k(j)(\gamma_\ell(j) - \gamma_\ell^{(\ell-k)}(j)) e^{\mathbf{i}(k+\ell)t},
\end{aligned}$$

since by the independence of $\gamma_k(j)$ and $\gamma_\ell^{(\ell-k)}(j)$ we have that $E\gamma_k(j)\gamma_\ell^{(\ell-k)}(j) = 0$. Using the Cauchy–Schwarz inequality with (5.13) we get that

$$\left| \sum_{1\le k<\ell\le N} E\gamma_k(j)(\gamma_\ell(j) - \gamma_\ell^{(\ell-k)}(j)) e^{\mathbf{i}(k+\ell)t} \right| \le cN$$

with some constant c, completing the proof of (7.5). The same arguments can be used to prove (7.6). □

Following [27] we define $\mathbf{S}_N(t) = \sum_{k=1}^N \boldsymbol{\gamma}_{k,N} e^{\mathbf{i}kt}$ and $\mathbf{S}_N^{(m)}(t) = \sum_{k=1}^N \boldsymbol{\gamma}_{k,N}^{(m)} e^{\mathbf{i}kt}$. Let $\mathbf{S}_N^*(t)$ be the conjugate transpose of $\mathbf{S}_N(t)$ and introduce

$$\begin{aligned}
\mathbf{I}_N(t) &= \frac{1}{N} \mathbf{S}_N(t) \mathbf{S}_N^*(t) \\
&= \frac{1}{N} \sum_{k=1}^N \boldsymbol{\gamma}_k e^{\mathbf{i}kt} \sum_{\ell=1}^N \boldsymbol{\gamma}_\ell^T e^{-\mathbf{i}\ell t} \\
&= \frac{1}{N} \sum_{\ell=1}^N \sum_{k=1}^N e^{\mathbf{i}t(k-\ell)} \boldsymbol{\gamma}_k \boldsymbol{\gamma}_\ell^T \\
&= \sum_{k=1-N}^{N-1} e^{-\mathbf{i}tk} \frac{1}{N} \sum_{\ell=\max(1,1-k)}^{\min(N,N-k)} \boldsymbol{\gamma}_k \boldsymbol{\gamma}_{\ell+k}^T \\
&= \sum_{k=1-N}^{N-1} e^{-\mathbf{i}tk} \boldsymbol{\phi}_{k,N}.
\end{aligned}$$

Similarly we define

$$\mathbf{I}_N^{(m)}(t) = \tfrac{1}{N} \mathbf{S}_N^{(m)}(t) \left(\mathbf{S}_N^{(m)}(t) \right)^* = \sum_{k=1-N}^{N-1} e^{-\mathbf{i}tk} \boldsymbol{\phi}_{k,N}^{(m)}.$$

Lemma 7.3. *If Assumptions 2.1–2.4, 3.1 and 3.2 are satisfied, then we have*

$$\limsup_{N\to\infty} \limsup_{m\to\infty} \sup_{-\infty<t<\infty} E\left| \mathbf{I}_N(t) - \mathbf{I}_N^{(m)}(t) \right| = 0.$$

Proof. By the triangle inequality we have

$$\left|\mathbf{I}_N(t)-\mathbf{I}_N^{(m)}(t)\right| = \left|\frac{1}{N}\mathbf{S}_N(t)\mathbf{S}_N^*(t)-\frac{1}{N}\mathbf{S}_N^{(m)}(t)\left(\mathbf{S}_N^{(m)}(t)\right)^*\right|$$
$$\leq \frac{1}{N}\left|\mathbf{S}_N(t)(\mathbf{S}_N^*(t)-(\mathbf{S}_N^{(m)}(t))^*)\right|+\frac{1}{N}\left|(\mathbf{S}_N(t)-\mathbf{S}_N^{(m)}(t))(\mathbf{S}_N^{(m)}(t))^*\right|.$$

Now the result follows from Lemma 7.2 via the Cauchy–Schwartz inequality. □

Proof of Theorem 3.1. Define the Fourier transform, $\hat{K}(u)$, of the kernel K as $\hat{K}(u)=\frac{1}{2\pi}\int_{-\infty}^{\infty}K(s)e^{-\mathbf{i}su}ds$. Since K and $\hat{K}$ are in L^1 and both are Lipschitz functions, the inversion formula gives $K(s)=\int_{-\infty}^{\infty}\hat{K}(u)e^{\mathbf{i}su}du$. From the relationship between K and $\hat{K}$ and from the fact that K is supported on the interval $[-1, 1]$, we obtain:

$$\begin{aligned}
\tilde{\boldsymbol{\Sigma}}_N &= \sum_{k=-B_N}^{B_N}K(k/B_N)\boldsymbol{\phi}_{k,N}\\
&= \sum_{k=1-N}^{N-1}K(k/B_N)\boldsymbol{\phi}_{k,N}\\
&= \sum_{k=1-N}^{N-1}\left(\int_{-\infty}^{\infty}\hat{K}(u)e^{\mathbf{i}(k/B_N)u}du\right)\boldsymbol{\phi}_{k,N}\\
&= \int_{-\infty}^{\infty}\hat{K}(u)\sum_{k=1-N}^{N-1}e^{-\mathbf{i}(-u/B_N)k}\boldsymbol{\phi}_{k,N}du\\
&= \int_{-\infty}^{\infty}\hat{K}(u)\mathbf{I}_N(-u/B_N)du.
\end{aligned}$$

Similarly,

$$\tilde{\boldsymbol{\Sigma}}_N^{(m)} = \int_{-\infty}^{\infty}\hat{K}(u)\mathbf{I}_N^{(m)}(-u/B_N)du.$$

Hence we have

$$\begin{aligned}
E\left|\tilde{\boldsymbol{\Sigma}}_N-\tilde{\boldsymbol{\Sigma}}_N^{(m)}\right| &= E\left|\int_{-\infty}^{\infty}\hat{K}(u)\left(\mathbf{I}_N(u/B_N)-\mathbf{I}_N^{(m)}(u/B_N)\right)du\right|\\
&\leq \int_{-\infty}^{\infty}\left|\hat{K}(u)\right|E\left|\left(\mathbf{I}_N(u/B_N)-\mathbf{I}_N^{(m)}(u/B_N)\right)\right|du\\
&\leq \sup_{-\infty<t<\infty}\left|\left|\mathbf{I}_N(t)-\mathbf{I}_N^{(m)}(t)\right|\right|_1\int_{-\infty}^{\infty}\left|\hat{K}(u)\right|du.
\end{aligned}$$

Applying Lemma 7.3 we conclude that

$$\left|\tilde{\boldsymbol{\Sigma}}_N-\tilde{\boldsymbol{\Sigma}}_N^{(m)}\right| \xrightarrow{\mathcal{P}} 0,$$

as $\min(N, m)\to\infty$. On the other hand, by Lemma 7.1, for every fixed m

$$\tilde{\boldsymbol{\Sigma}}_N^{(m)} \xrightarrow{\mathcal{P}} \boldsymbol{\Sigma}^{(m)}.$$

Since

$$\boldsymbol{\Sigma}^{(m)}\to\boldsymbol{\Sigma},$$

as $m\to\infty$, the proof of the theorem is complete. □

Proof of Theorem 3.2. It follows from the definition of $\hat{\epsilon}_\ell$, (1.4) and the orthonormality of $\{w_j, 1\leq j<\infty\}$ that

$$\langle\hat{\epsilon}_\ell, w_i\rangle = \langle\epsilon_\ell, w_i\rangle + \langle X_\ell, u_i\rangle + \langle\nu_\ell, w_i\rangle,$$

where

$$\nu_\ell(t) = \sum_{i=1}^{q}\sum_{j=1}^{p}\psi_{i,j}w_i(t)\langle X_\ell, v_j\rangle - \sum_{i=1}^{q}\sum_{j=1}^{p}\hat{\psi}_{i,j}\hat{w}_{i,N}(t)\langle X_\ell, \hat{v}_{j,N}\rangle.$$

Following the proof of Theorem 3.1 one can show that the estimates in (5.5) and (5.6) yield

$$\left|\breve{\Sigma}_N(i,j,i',j') - \hat{d}_{i,N}\hat{c}_{j,N}\hat{d}_{i',N}\hat{c}_{j',N}\Sigma_N^*(i,j,i',j')\right| = o_P(1), \tag{7.7}$$

where

$$\breve{\Sigma}_N(i,j,i',j') = \sum_{k=-(N-1)}^{N-1} K(k/B_N)\hat{\phi}_{k,N}(i,j,i',j')$$

and

$$\Sigma_N^*(i,j,i',j') = \sum_{k=-(N-1)}^{N-1} K(k/B_N)\phi_{k,N}^*(i,j,i',j')$$

with

$$\hat{\phi}_{k,N}(i,j,i',j') = \frac{1}{N}\sum_{\ell=\max(1,1-k)}^{\min(N,N-k)} \hat{\gamma}_\ell(i,j)\hat{\gamma}_{\ell+k}(i',j'),$$

$$\phi_{k,N}^*(i,j,i',j') = \frac{1}{N}\sum_{\ell=\max(1,1-k)}^{\min(N,N-k)} \gamma_\ell^*(i,j)\gamma_{\ell+k}^*(i',j'),$$

and

$$\gamma_\ell^*(i,j) = \langle X_\ell, v_j\rangle\langle\hat{\epsilon}_\ell, w_i\rangle.$$

Since

$$\langle X_\ell, v_j\rangle\langle\hat{\epsilon}_\ell, w_i\rangle = \gamma_\ell(i,j) + \langle X_\ell, v_j\rangle\langle\nu_\ell, w_i\rangle,$$

(5.5) and (5.6) and Lemma 6.10 imply that

$$\left|\tilde{\Sigma}_N - \Sigma_N^*\right| = o_P(1). \quad \Box \tag{7.8}$$

We have seen in Theorem 3.1 that $\left|\tilde{\boldsymbol{\Sigma}}_N - \boldsymbol{\Sigma}\right| = o_P(1)$. In (7.7) and (7.8) we have seen that $\left|\breve{\boldsymbol{\Sigma}}_N - \boldsymbol{\zeta}_N\boldsymbol{\Sigma}_N^*\boldsymbol{\zeta}_N\right| = o_P(1)$ and $\left|\tilde{\Sigma}_N - \Sigma_N^*\right| = o_P(1)$. Therefore, $\left|\hat{\boldsymbol{\Sigma}}_N - \boldsymbol{\Sigma}\right| = o_P(1)$, completing the proof.

References

[1] T.W. Anderson, The Statistical Analysis of Time Series, Wiley, New York, 1971.
[2] A. Aue, S. Hörmann, L. Horváth, M. Hušková, J. Steinebach J., Sequential testing for the stability of portfolio betas, Econometric Theory, (in press).
[3] A. Aue, S. Hörmann, L. Horváth, M. Reimherr, Break detection in the covariance structure of multivariate time series models, Annals of Statistics 37 (2009) 4046–4047.
[4] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.
[5] D. Bosq, Linear Processes in Function Spaces, Springer, New York, 2000.
[6] D.R. Brillinger, Time Series: Data Analysis and Theory, Holden-Day, 1969.
[7] H. Cardot, F. Ferraty, A. Mas, P. Sarda, Testing hypothesis in the functional linear model, Scandinavian Journal of Statistics 30 (2003) 241–255.
[8] J.-M. Chiou, H.-G. Müller, J.-L. Wang, Functional response models, Statistica Sinica 14 (2004) 675–693.
[9] T. Cai, P. Hall, Prediction in functional linear regression, The Annals of Statistics 34 (2006) 2159–2179.
[10] F. Ferraty, P. Vieu, Nonparametric Functional Data Analysis: Theory and Practice, Springer, New York, 2006.
[11] I. Gohberg, S. Golberg, M.A. Kaashoek, Classes of Linear Operators, Vol. I, in: Operator Theory: Advances and Applications, vol. 49, Birkhäuser, 1990, Basel.
[12] U. Grenander, M. Rosenblatt, Statistical Analysis of Stationary Time Series, Wiley, New York, 1957.
[13] S. Hörmann, P. Kokoszka, Weakly dependent functional data, Annals of Statistics 38 (2010) 1845–1884.
[14] S. Hörmann, L. Horváth, R. Reeder, A functional version of the ARCH model.
[15] L. Horváth, M. Hušková, P. Kokoszka, Testing the stability of the functional autoregressive process, Journal of Multivariate Analysis 101 (2010) 352–367.
[16] I.A. Ibragimov, Some limit theorems for stationary processes, Theory of Probability and Its Applications 7 (1962) 349–382.
[17] M. Jansson, Consistent covariance matrix estimation for linear processes, Econometric Theory 18 (2002) 1449–1459.
[18] V. Kargin, A. Onatski, Curve forecasting by functional autoregression, Journal of Multivariate Analysis 99 (2008) 2508–2526.
[19] J. Kiefer, K-sample analogues of the Kolmogorov–Smirnov and Cramér–v. Mises tests, Annals of Mathematical Statistics 30 (1959) 420–447.
[20] W. Liu, W.B. Wu, Asymptotics of spectral density estimates, Econometric Theory 26 (2010) 1218–1245.
[21] F. Móricz, Moment inequalities and the strong laws of large number, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 35 (1976) 299–314.
[22] E. Parzen, Mathematical considerations in the estimation of spectra, Technometrics 3 (1961) 167–190.
[23] E. Parzen, Notes on Fourier analysis and spectral windows, in: E. Parzen (Ed.), Time Series Analysis Papers, Holden-Day, 1967, pp. 190–250.
[24] M.B. Priestley, Spectral Analysis and Time Series I, Academic Press, New York, 1981.
[25] M.B. Priestley, Nonlinear and Nonstationary Time Series Analysis, Academic Press, New York, 1988.
[26] M. Rosenblatt, Stationary Sequences and Random Fields, Birkhäuser, Basel, 1985.
[27] M. Taniguchi, Y. Kakizawa, Asymptotic Theory of Statistical Inference for Time Series, Springer, New York, 2000.
[28] D. Wied, W. Krämer, H. Dehling, Testing for a change in correlation at an unknown point in time using an extended functional delta method, Econometric Theory 28 (2012) 570–589.
[29] N. Wiener, Nonlinear Problems in Random Theory, MIT Press, Cambridge, MA, 1958.