<table>
<thead>
<tr>
<th><strong>Statistica Sinica Preprint</strong> No: SS-2021-0121</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
</tr>
<tr>
<td><strong>Manuscript ID</strong></td>
</tr>
<tr>
<td><strong>URL</strong></td>
</tr>
<tr>
<td><strong>DOI</strong></td>
</tr>
<tr>
<td><strong>Complete List of Authors</strong></td>
</tr>
<tr>
<td><strong>Corresponding Author</strong></td>
</tr>
<tr>
<td><strong>E-mail</strong></td>
</tr>
</tbody>
</table>

Notice: Accepted version subject to English editing.
STATISTICAL INFERENCE FOR FUNCTIONAL TIME SERIES:
AUTOCOVARIANCE FUNCTION

Chen Zhong and Lijian Yang

Tsinghua University

Abstract: Statistical inference for functional time series is investigated by extending the classic concept of autocovariance function (ACF) to functional ACF (FACF). It is established that for functional moving average (FMA) data, the FMA order can be determined as the highest nonvanishing order of FACF, just as in classic time series analysis. A two-step estimator is proposed for FACF, the first step involving simultaneous B-spline estimation of each time trajectory and the second step plug-in estimation of FACF by using the estimated trajectories in place of the latent true curves. Under simple and mild assumptions, the proposed tensor product spline FACF estimator is asymptotically equivalent to the oracle estimator with all known trajectories, leading to asymptotic correct simultaneous confidence envelope (SCE) for the true FACF. Simulation experiments validate the asymptotic correctness of the SCE and data-driven FMA order selection. The proposed SCEs are computed for the FACFs of an ElectroEncephalogram (EEG) functional time series with interesting discovery of finite FMA lag and Fourier form functional principal components.

Key words and phrases: Confidence envelope, Functional autocovariance func-
1. Introduction

During the last two decades, functional data analysis (FDA) has become a very important statistics research area due to its wide applications, see Ramsay and Silverman (2002), Ramsay and Silverman (2005), Ferraty and Vieu (2006), Hsing and Eubank (2015), Guo et al. (2019) for the development of FDA applications and theory.

A functional random variable is a square-integrable continuous stochastic process: namely, \( \eta(\cdot) \in C[0,1] \) almost surely, \( \text{E}\sup_{x \in [0,1]} \eta^2(x) < +\infty \). For such \( \eta(\cdot) \), both mean function \( \text{E}\{\eta(x)\}, x \in [0,1] \) and covariance function \( \text{Cov}\{\eta(x), \eta(x')\}, x, x' \in [0,1] \) exist and are continuous. Functional data consist of a sequence \( \{\eta_i(\cdot)\}_{i=1}^n \) of stochastic processes called trajectories, each having the same distribution as \( \eta(\cdot) \), thus also the same mean and covariance functions. These trajectories are decomposed as \( \eta_i(\cdot) = m(\cdot) + \chi_i(\cdot) \), where the centered trajectories \( \chi_i(\cdot) \) are small-scale variations of \( x \) on the \( i \)-th trajectory, and are continuous stochastic processes with \( \text{E}\chi_i(x) \equiv 0 \) and covariance function \( \text{Cov}\{\chi_i(x), \chi_i(x')\}, x, x' \in [0,1] \).

Estimation of functional mean \( \text{E}\{\eta(\cdot)\} \) with simultaneous confidence band (SCB) was investigated in Degras (2011), Cao et al. (2012), Ma et al. (2012), Gu et al. (2013), Zheng et al. (2013), Choi and Reimherr (2018) and Telschow.
Simultaneous confidence envelope (SCE) of bivariate covariance function was established in Cao et al. (2016) for dense functional data, and sharper SCB in Wang et al. (2020) for univariate stationary covariance function. See also related works on covariance estimation for stationary stochastic process over infinite domain in Hall et al. (1994), and for sparse longitudinal data in Meyer (1998) and Zhou et al. (2018).

All of the above are limited to i.i.d observations \( \{\eta_i(\cdot)\}_{i=1}^n \) of functional random variable \( \eta(\cdot) \). Many functional data from applied fields, however, are collected in timely order, and exhibit temporal dependence. These data can be regarded as functional dependent data and estimation of its functional mean \( \mathbb{E}\{\eta(\cdot)\} \) was investigated in Chen and Song (2015), Guo and Chen (2019), Horváth et al. (2013) and Li and Yang (2021). Take, for instance, the ElectroEncephalogram (EEG) of a human subject in eyes-closed resting state, studied in Li and Yang (2021). Recorded at sample rate 1000Hz (i.e., recording at every 0.001 second), the EEG series is divided into \( n = 300 \) consecutive segments \( \{\eta_t(\cdot)\}_{t=1}^{300} \) with 200 EEG signals in each segment. Figures 5 and 6 depict estimate (middle surface) of the autocovariance function \( \text{Cov}\{\eta_t(x), \eta_{t+1}(x')\}, x, x' \in [0, 1] \), which differs significantly from zero (flat surface in Figure 6), indicating strong dependence between \( \eta_t(\cdot) \) and \( \eta_{t+1}(\cdot) \), for more details, see Section 6.

To model timely ordered and dependent trajectories \( \{\eta_t(\cdot)\}_{t=1}^n \), the centered
trajectories $\chi_t(\cdot)$ are embedded into a strictly stationary functional infinite moving average, or FMA($\infty$) series $\{\chi_t(\cdot)\}_{t=-\infty}^{\infty}$ as in Li and Yang (2021)

$$\chi_t(\cdot) = \sum_{t' = 0}^{\infty} A_{t'} \zeta_{t-t'}(\cdot), t \in \mathbb{Z}$$

(1.1)

where bounded linear operators $A_{t'} : \mathcal{L}^2[0,1] \to \mathcal{L}^2[0,1]$ play the roles of scalar coefficients in classic MA($\infty$) (Definition 3.2.1, Brockwell and Davis (1991)).

The processes $\{\zeta_t(\cdot)\}_{t \in \mathbb{Z}}$ are strong functional white noises (Definition 3.1 of Bosq (2000)), corresponding to classic white noises in Brockwell and Davis (1991): they are square-integrable continuous stochastic processes, i.i.d over index $t \in \mathbb{Z}$, with $E\zeta_t(\cdot) \equiv 0$ and continuous covariance function $G(x,x') = E\zeta_t(x)\zeta_t(x')$. Mercer Lemma (Lemma 1.3, Bosq (2000)) entails that $G(x,x') = \sum_{k=1}^{\infty} \lambda_k \psi_k(x) \psi_k(x')$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ and corresponding eigenfunctions $\{\psi_k\}_{k=1}^{\infty}$, the latter an orthonormal basis of $\mathcal{L}^2[0,1]$, such that $\sum_{k=1}^{\infty} \lambda_k < \infty$, $\{\psi_k\}_{k=1}^{\infty} \subset C[0,1]$ and $\int G(x,x') \psi_k(x') dx' = \lambda_k \psi_k(x)$.

For any $t \in \mathbb{Z}$, $\zeta_t(\cdot)$ allows the well-known Karhunen-Loève $\mathcal{L}^2$ representation

$$\zeta_t(\cdot) = \sum_{k=1}^{\infty} \zeta_{tk} \phi_k(\cdot),$$

in which the rescaled eigenfunctions, $\phi_k$, called functional principal components (FPCs), satisfy that $\phi_k = \sqrt{\lambda_k} \psi_k$ and $\int \zeta_t(x) \phi_k(x) dx = \lambda_k \zeta_{tk}$, for $k \geq 1$; the random coefficients $\zeta_{tk}$, are therefore uncorrelated over $k \in \mathbb{N}_+$ and i.i.d over index $t \in \mathbb{Z}$, with mean 0 and variance 1. It is assumed that $\sum_{k=1}^{\infty} \|\phi_k\|_{\infty} < +\infty$ so that the Karhunen-Loève series converges absolutely uniformly almost surely by Dominated Convergence Theorem.
The FMA(∞) coefficient operators $A_t, t \in \mathbb{N}$ are defined relative to the orthonormal basis $\{\psi_k\}_{k=1}^{\infty}$ of $L^2[0,1]$

$$A_t \left\{ \sum_{k=1}^{\infty} c_k \psi_k(\cdot) \right\} = \sum_{k=1}^{\infty} a_{tk} c_k \psi_k(\cdot), \quad a_{tk} \in \mathbb{R}, \quad t \in \mathbb{N}$$

$$|a_{tk}| < C_a \rho_a^t, \quad C_a \in (0, \infty), \quad \rho_a \in (0,1), \quad t \in \mathbb{N}, k \in \mathbb{N}_+.$$  \hspace{1cm} (1.2)

Geometric decay of MA coefficients $\{a_{tk}\}_{t \in \mathbb{N}}$ in (1.2) holds for many causal time series such as ARMA, see equation (3.3.6) of Brockwell and Davis (1991).

The following holds absolutely uniformly almost surely according to (1.1)

$$\chi_t(\cdot) = \sum_{t'=0}^{\infty} A_{t'} \left\{ \sum_{k=1}^{\infty} \zeta_{t-t',k} \phi_k(\cdot) \right\} = \sum_{t'=0}^{\infty} \sum_{k=1}^{\infty} a_{t',k} \zeta_{t-t',k} \phi_k(\cdot) = \sum_{k=1}^{\infty} \sum_{t'=0}^{\infty} a_{t',k} \zeta_{t-t',k} \phi_k(\cdot), \quad t \in \mathbb{Z}$$ \hspace{1cm} (1.3)

in which

$$\xi_{tk} = \sum_{t'=0}^{\infty} a_{t',k} \zeta_{t-t',k}, \quad t \in \mathbb{Z}, k \in \mathbb{N}_+.$$ \hspace{1cm} (1.4)

Thus for each fixed $k \in \mathbb{N}_+$, the time series $\{\xi_{tk}\}_{t \in \mathbb{Z}}$ is a classic MA(∞) expressed in terms of the i.i.d sequence $\{\zeta_{tk}\}_{t \in \mathbb{Z}}$. It should be pointed out that classic MA(∞) is sufficiently broad to include as special cases widely used causal ARMA($p,q$) and consequently AR($p$) and MA($q$).

Assume for convenience that for each fixed $k \in \mathbb{N}_+$, $\sum_{t=0}^{\infty} a_{tk}^2 \equiv 1$ then the mean and variance of $\xi_{tk}$ equal 0 and 1 as well, $t \in \mathbb{Z}, k \in \mathbb{N}_+$, and uncorrelated
over $k \in \mathbb{N}_+$. The covariance function of $\{\chi_t(\cdot)\}_{t=-\infty}^{\infty}$ equals that of $\{\zeta_t(\cdot)\}_{t=-\infty}^{\infty}$

$$E\chi_t(x)\chi_t(x') = E\sum_{k=1}^{\infty} \xi_{tk} \phi_k(x) \sum_{k=1}^{\infty} \xi_{tk} \phi_k(x') = \sum_{k=1}^{\infty} E\xi_{tk}^2 \phi_k(x) \phi_k(x') = \sum_{k=1}^{\infty} \phi_k(x) \phi_k(x') = G(x,x') \zeta_t(x) \zeta_t(x'), x,x' \in [0,1].$$

Hence $\chi_t(\cdot) = \sum_{k=1}^{\infty} \xi_{tk} \phi_k(\cdot)$ is in fact the Karhunen-Loève expansion of the strictly stationary FMA(∞) series $\{\chi_t(\cdot)\}_{t=-\infty}^{\infty}$, where the convergence is absolutely uniformly almost sure according to (1.3), and random coefficients $\{\xi_{tk}\}_{t \in \mathbb{Z}, k \in \mathbb{N}_+}$ are called FPC scores.

Raw data of the FMA(∞) are in the form

$$Y_{ij} = m(j/N) + \sum_{k=1}^{\infty} \xi_{tk} \phi_k(j/N) + \sigma(j/N) \varepsilon_{tj}, \quad 1 \leq t \leq n, \quad 1 \leq j \leq N, \quad (1.5)$$

with MA(∞) FPC scores $\xi_{tk}$ given in (1.4). The terms $\sigma(j/N) \varepsilon_{tj}$ represent measurement errors which occur with data collection, $\{\varepsilon_{tj}\}_{t=-\infty}^{N}$ are i.i.d with mean 0, variance 1, and the standard deviation function $\sigma(\cdot)$ satisfies Hölder continuity in Assumption (A2).

Autocovariance function (ACF) in classic time series extends naturally to the following functional autocovariance function (FACF)

$$C_h(x,x') = E\chi_t(x)\chi_{t+h}(x'), h \in \mathbb{N}, t \in \mathbb{Z}, x,x' \in [0,1]. \quad (1.6)$$

Obviously, $C_0(x,x') \equiv G(x,x')$. It is shown in Proposition II that if the highest order $h$ for which $C_h(x,x')$ is not identically zero is a finite integer $q \in \mathbb{N}_+$, then the FMA(∞) reduces to a simpler FMA(q).
This paper extends simultaneous confidence region for i.i.d. functional data covariance function in Cao et al. (2016) and Wang et al. (2020a) to the much more complicated FACFs $C_h(x, x')$, $h \in \mathbb{N}$. The FACFs are estimated in two steps: step one estimates by B spline all $n$ trajectories $\{\eta_t(\cdot)\}_{t=1}^n$ and their population mean $m(\cdot)$; step two estimates the FACF $C_h(x, x')$ using the estimated trajectories and population mean as if they were the true. The greatest technical difficulty is to establish in Proposition 2 under the assumption of moving average trajectories $\{\eta_t(\cdot)\}_{t=1}^n$ that the proposed FACF estimator is asymptotically equivalent to the infeasible FACF estimator when all trajectories $\{\chi_t(\cdot)\}_{t=1}^n$ are entirely observed. This ”oracle efficiency” allows for the construction of asymptotic SCE for each FACF $C_h(x, x')$, $h \in \mathbb{N}$. The second technical difficulty is to obtain in Theorem 4 the limiting distribution of FACF estimate under moving average dependence of trajectories $\{\chi_t(\cdot)\}_{t=1}^n$.

Nonparametric simultaneous confidence region is a versatile tool for global inference on functions, see for instance, Wang and Yang (2009a), Wang et al. (2014), Gu and Yang (2013), Zheng et al. (2016), Wang et al. (2020b), Gu et al. (2021), Yu et al. (2021), Zhong and Yang (2021) for theory and applications of SCB in diverse contexts. The SCE for $C_h(x, x')$ enables one to test against any null hypothesis such as $H_0 : C_h(x, x') \equiv 0$ for any positive integers $h$. For the EEG time series, this null hypothesis is rejected when $h = 9$ but retained for
$h = 10, 11, 12, 13, 14$, leading to the appropriate decision of FMA(9) for the data, see Section 6, especially the SCEs in Figures 5 and 6.

The paper is organized as follows. Section 2 examines theoretical properties of FACF $C_h(x, x')$ and define its infeasible and two-step estimators. Section 3 establishes limiting distribution of the infeasible FACF estimator and asymptotic equivalence of the infeasible and two-step estimators. Asymptotic SCEs for the FACF are proposed in Section 4 with implementation details. Section 5 presents the simulation studies, and analysis of EEG data is in Section 6. All technical proofs are in the supplement.

2. FACF and its estimation

We begin with deriving explicit formula of the FACF $C_h(x, x')$ in (1.4). One denotes by $\gamma_k(h) = E(\xi_{tk}\xi_{t+h,k}) = \sum_{m=0}^{\infty} a_{mk} a_{m+h,k}, h \in \mathbb{N}, t \in \mathbb{Z}, k \in \mathbb{N}_+$ the classic autocovariance function of $\{\xi_{tk}\}_{t \in \mathbb{Z}}$, with the convention that $\gamma_k(-h) \equiv \gamma_k(h), h \in \mathbb{N}$. Noting absolute almost sure convergence in (1.4), one derives with uniform absolute convergence

$$C_h(x, x') = \frac{1}{n-h} \sum_{t=1}^{n-h} \sum_{k,k'=1}^{\infty} E(\xi_{tk}\xi_{t+h,k'}) \phi_k(x) \phi_{k'}(x')$$

$$= \frac{1}{n-h} \sum_{t=1}^{n-h} \sum_{k=1}^{\infty} \gamma_k(h) \phi_k(x) \phi_k(x')$$

$$= \frac{1}{n-h} \sum_{t=1}^{n-h} \sum_{k,k'=1}^{\infty} \delta_{kk'} \gamma_k(h) \phi_k(x) \phi_{k'}(x'), x, x' \in [0, 1] \tag{2.1}$$

where $\delta_{kk'} = 1$ for $k = k'$ and 0 otherwise. The next Proposition plays the same
role of its classic analog, Proposition 3.2.1 in Brockwell and Davis (1991).

**Proposition 1.** If $C_h(x,x') \equiv 0, x,x' \in [0,1]$ for $h > q$ while $C_q(x,x') \neq 0$ for some $q \in \mathbb{N}_+$, and the MA($\infty$) equation in (1.4) is invertible for all $k \in \mathbb{N}_+$ (i.e., there exist $b_{t,k} \in \mathbb{R}$, such that $\zeta_{t,k} = \sum_{t'=0}^{\infty} b_{t',k}\xi_{t-t',k}, t \in \mathbb{Z}, k \in \mathbb{N}_+$), then $\chi_t(\cdot) = \sum_{t'=0}^{q} A_{t'}\zeta_{t-t'}(\cdot), t \in \mathbb{Z}$ absolutely uniformly almost surely hence $\{\chi_t(\cdot)\}_{t=-\infty}^{\infty}$ is an FMA($q$) series.

According to Proposition 1, a finite moving average order $q$ is determined if the FACF $C_h(\cdot,\cdot)$ vanishes exactly beyond order $q$. Testing if an FACF is identically zero, is conveniently done by estimation of FACF with simultaneous confidence envelope (SCE).

If the $t$-th centered trajectory $\chi_t(\cdot) = \eta_t(\cdot) - m(\cdot), 1 \leq t \leq n$ were fully observed, one would estimate the ACF $C_h(x,x')$ with the following sample ACF as in (7.2.1) of Brockwell and Davis (1991):

\[
\tilde{C}_h(x,x') = n^{-1} \sum_{t=1}^{n-h} \chi_t(x) \chi_{t+h}(x'), \ x,x' \in [0,1];
\]

which can be written explicitly as

\[
\tilde{C}_h(x,x') = n^{-1} \sum_{t=1}^{n-h} \sum_{k,k'=1}^{\infty} \xi_{tk}\xi_{t+h,k'}\phi_k(x)\phi_{k'}(x').
\]

Since $\{\chi_t(\cdot)\}_{t=1}^{n}$ are unobservable, the above estimator function $\tilde{C}_h(x,x')$ is “infeasible”. It does suggest, however, the following plug-in sample covariance esti-
mator

\[ \hat{C}_h (x, x') = n^{-1} \sum_{t=1}^{n-h} \hat{\chi}_t (x) \hat{\chi}_{t+h} (x'), \quad x, x' \in [0, 1], \quad (2.4) \]

in which

\[ \hat{\chi}_t (\cdot) = \hat{\eta}_t (\cdot) - \hat{m} (\cdot), \quad 1 \leq t \leq n, \quad (2.5) \]

where \( \hat{\eta}_t (\cdot) \) and \( \hat{m} (\cdot) \) are some suitable estimators of \( \eta_t (\cdot) \) and \( m (\cdot) \), respectively.

In this work B spline estimators \( \hat{\eta}_t (x) \) and \( \hat{m} (x) \) are employed. To describe B spline estimator, denote by \( \{t_\ell\}_{\ell=1}^{J_s} \) a sequence of equally-spaced points. We call \( t_\ell = \ell / (J_s + 1) \), \( \ell \in \{1, \ldots, J_s\} \), \( 0 < t_1 < \cdots < t_{J_s} < 1 \) interior knots, which divide the interval \([0, 1]\) into \((J_s + 1)\) equal subintervals \( I_0 = [0, t_1) \), \( I_\ell = [t_\ell, t_{\ell+1}) \), \( \ell \in \{1, \ldots, J_s - 1\} \), \( I_{J_s} = [t_{J_s}, 1] \). For any positive integer \( p \), let \( t_{1-p} = \cdots = t_0 = 0 \) and \( 1 = t_{J_s+1} = \cdots = t_{J_s+p} \) be auxiliary knots. Let \( S^{(p-2)} = S^{(p-2)}[0, 1] \) be the polynomial spline space of order \( p \) on \( I_\ell \), \( \ell \in \{0, \ldots, J_s\} \), which consists of all \((p-2)\) times continuously differentiable functions on \([0, 1]\) that are polynomials of degree \((p-1)\) on subintervals \( I_\ell \), \( 0 \leq \ell \leq J_s \). Denote by \( \{B_{\ell,p} (x) \}, 1 \leq \ell \leq J_s + p \) the \( p \)-th order B-spline basis functions of \( S^{(p-2)} \) (Boor, C (2001)),

\[ S^{(p-2)} = \left\{ \sum_{\ell=1}^{J_s+p} \lambda_{\ell,p} B_{\ell,p} (\cdot) \bigg| \lambda_{\ell,p} \in \mathbb{R} \right\}. \]

In this paper, trajectories \( \eta_t (\cdot) \) and their population mean \( m (\cdot) \) are all estimated by spline regression

\[ \hat{m} (\cdot) = n^{-1} \sum_{t=1}^{n} \hat{\eta}_t (\cdot), \quad \hat{\eta}_t (\cdot) = \arg \min_{g (\cdot) \in S^{(p-2)}} \sum_{j=1}^{N} \{Y_{tj} - g (j/N)\}^2. \quad (2.6) \]
3. Asymptotic properties

This section studies the asymptotic properties for the proposed estimators.

3.1 Assumptions and the infeasible estimator

To study the asymptotic properties of the two-step spline estimator \( \hat{C}_h(x, x') \), one gives some mild assumptions. For sequences of real numbers \( a_n \) and \( b_n \), denote \( a_n \asymp b_n \) if \( |a_n| \leq C |b_n| \), \( |b_n| \leq C |a_n| \), \( n \in \mathbb{N}_+ \) for some constant \( C > 0 \). For any measurable function \( \varphi(\cdot) \) defined on \([0, 1]\), denote \( \|\varphi\|_\infty = \sup_{x \in [0, 1]} |\varphi(x)| \), and \( \varphi^{(q)}(x) \) its \( q \)-th order derivative with respect to \( x \) if it exists. For any functions \( \phi(\cdot), \varphi(\cdot) \in L^2[0, 1] \), define their theoretical and empirical inner products as \( \langle \phi, \varphi \rangle = \int_{[0, 1]} \phi(x) \varphi(x) \, dx \), \( \langle \phi, \varphi \rangle_N = N^{-1} \sum_{j=1}^{N} \phi(j/N) \varphi(j/N) \). Correspondingly and respectively, theoretical and empirical norms are \( \|\phi\|_2^2 = \langle \phi, \phi \rangle \), \( \|\phi\|_{2,N}^2 = \langle \phi, \phi \rangle_N \).

For a non-negative integer \( q \) and a real number \( \mu \in (0, 1] \), write \( \mathcal{H}^{(q, \mu)}[0, 1] \) as the space of \((q, \mu)\)-Hölder continuous functions, i.e.,

\[
\mathcal{H}^{(q, \mu)}[0, 1] = \left\{ \varphi : [0, 1] \to \mathbb{R} \mid \|\varphi\|_{q, \mu} = \sup_{x, y \in [0, 1], x \neq y} \left| \frac{\varphi^{(q)}(x) - \varphi^{(q)}(y)}{|x - y|^\mu} \right| < +\infty \right\}
\]

in which \( \|\cdot\|_{q, \mu} \) denotes the \((q, \mu)\)-Hölder semi-norm. Denote by \( I_n \in \mathbb{N}_+ \) a truncation index such that the FMA(\( \infty \)) \( \chi_t(\cdot) \) is well approximated by \( \sum_{t'=0}^{I_n} A_{t'} \zeta_{t-t'}(\cdot) \):

\[
I_n > -10 \log n / \log \rho_a, \quad I_n \asymp \log n, \quad (3.1)
\]
where \( \rho_a \) is the geometric decay parameter in \((1.2)\).

The following are some technical assumptions.

(A1) There exist \( q \in N_+ , \mu \in (0, 1] \), such that \( m(\cdot) \in H^{(q, \mu)} [0, 1] \). In the following, one denotes \( p^* = q + \mu \).

(A2) The standard deviation function \( \sigma(\cdot) \in H^{(0, \nu)} [0, 1] \) for \( \nu \in (0, 1] \) and for some constants \( M_{\sigma} > 0 \), \( \sup_{x \in [0, 1]} \sigma(x) \leq M_{\sigma} \).

(A3) There exists \( \theta > 0 \), such that as \( N \to \infty \), \( n = n(N) \to \infty \), \( n = O(N^\theta) \).

(A4) The rescaled FPCs \( \phi_k(\cdot) \in H^{(q, \mu)} [0, 1] \) with \( \sum_{k=1}^{\infty} \| \phi_k \|_{q, \mu} < +\infty \), \( \sum_{k=1}^{\infty} \| \phi_k \|_\infty < +\infty \); for increasing positive integers \( \{ k_n \}_{n=1}^{\infty} \), as \( n \to \infty \), \( \sum_{k_n+1}^{\infty} \| \phi_k \|_\infty = o(n^{-1/2}) \) and \( k_n = O(n^\omega) \) for some \( \omega > 0 \).

(A5) The i.i.d variables \( \{ \varepsilon_{tj} \}_{t \geq 1, j \geq 1} \) are independent of \( \{ \zeta_{tk} \}_{t \in \mathbb{Z}, k \in N_+} \). The random coefficients \( \{ \zeta_{tk} \}_{t \in \mathbb{Z}, k \in N_+} \) are independent over \( k \in N_+ \) and i.i.d over \( t \in \mathbb{Z} \). The number of distinct distributions for all \( \{ \zeta_{tk} \}_{t \in \mathbb{Z}, k \in N_+} \) is finite and \( \sup_{k \geq 1} E \zeta_{tk}^{r_0} < \infty \) for some \( r_0 \geq 4 \). There exist \( C_1, C_2 \in (0, +\infty) \), \( \gamma_1, \gamma_2 \in (1, +\infty) \), \( \beta_1, \beta_2 \in (0, 1/2) \), i.i.d \( N(0, 1) \) variables \( \{ Z_{tj, \varepsilon} \}_{t=1,j=1}^{n,N} \), \( \{ Z_{tk, \zeta} \}_{t=-J_n+1,k=1}^{n,k_n} \) such that

\[
\mathbb{P} \left\{ \max_{1 \leq k \leq k_n} \max_{-J_n+1 \leq \tau \leq n} \left| \sum_{t=-J_n+1}^{\tau} \zeta_{tk} - \sum_{t=-J_n+1}^{\tau} Z_{tk, \zeta} \right| > n^{\beta_1} \right\} < C_1 n^{-\gamma_1},
\]

\[
\mathbb{P} \left\{ \max_{1 \leq \tau \leq n} \max_{1 \leq \tau \leq N} \left| \sum_{j=1}^{\tau} \varepsilon_{tj} - \sum_{j=1}^{\tau} Z_{tj, \varepsilon} \right| > N^{\beta_2} \right\} < C_2 N^{-\gamma_2}.
\]
(A6) The spline order \( p \geq p^* \), the number of interior knots \( J_s \asymp N^\tau d_N \) for some \( \tau > 0 \) with \( d_N + d_N^{-1} = O(\log^\tau N) \) as \( N \to \infty \), and for \( p^* \) in Assumption (A1), \( \nu \) in Assumption (A2), \( \theta \) in Assumption (A3) and \( r_0, \beta_1, \beta_2 \) in Assumption (A5)

\[
\max \left\{ \frac{4\theta}{r_0 p^*} + \frac{3\theta}{2p^*}, 1 - \nu \right\} < \gamma < 1 - \frac{\theta}{2} - \beta_2 - \theta \beta_1.
\]

Assumptions (A1)–(A2) are regular conditions for the spline smoother. In particular, Assumption (A1) controls the size of the bias of the spline smoother for \( m(\cdot) \) and Assumption (A2) ensures the variance function should be uniformly bounded. Assumption (A3) requires that sample size \( n \) grows in sync with the number \( N \) of observations per curve, and not faster than \( N^\theta \), hence all asymptotics in this section are stated with \( N \to \infty \) only, not with \( N \to \infty, n \to \infty \). Assumption (A4) guarantees the bounded smoothness of the principal components.

The independence of latent FPC scores \( \{\zeta_{tk}\}_{t \geq 1, k \geq 1} \) over \( k \in \mathbb{N}_+ \) in Assumption (A5) is common in existing works on functional data analysis, see \cite{Cao2012}, \cite{Ma2012}, \cite{Gu2014}, \cite{Zheng2014}. The probability inequalities in Assumption (A5) provide strong Gaussian approximations of measurement errors \( \{\varepsilon_{tj}\}_{t \geq 1, j \geq 1} \) and the latent FPC scores \( \{\zeta_{tk}\}_{t \geq 1, k \geq 1} \), so that all the main results of this section hold without data itself being Gaussian. Assumption (A5) is a high level assumption which can be guaranteed by the elementary Assumption (A5') below together with Assumption (A4), see Lemma
S.3 in the Supplement. Assumption (A6) on choice of knot number satisfies the requirements for the B-spline smoothing.

\begin{itemize}
\item[(A5')] The i.i.d variables \(\{\varepsilon_{tj}\}_{t \geq 1, j \geq 1}\) are independent of \(\{\zeta_{tk}\}_{t \in \mathbb{Z}, k \in \mathbb{N}_+}\). The random coefficients \(\{\zeta_{tk}\}_{t \in \mathbb{Z}, k \in \mathbb{N}_+}\) are independent over \(k \in \mathbb{N}_+\) and i.i.d over \(t \in \mathbb{Z}\). The number of distinct distributions for all \(\{\zeta_{tk}\}_{t \in \mathbb{Z}, k \in \mathbb{N}_+}\) is finite.
\end{itemize}

**Remark 1.** The aforementioned Assumptions are easily satisfied in practice.

One simple and reasonable setup for the above parameters \(q, \mu, \theta, p, \gamma\) can be the following: \(q + \mu = p^* = 4, \nu = 1, r_0 > 4, \theta = 1, p = 4\) (cubic spline), \(\gamma = 3/8\), \(d_N \asymp \log \log N\). These constants are used as defaults for implementation in Section 4.

We first examine the infeasible estimator \(\tilde{C}_h (x, x')\). Denote \(\Delta_h (x, x') = \tilde{C}_h (x, x') - n^{-h} C_h (x, x'), x, x' \in [0, 1]\). Then (2.1) and (2.3) imply that

\[
\Delta_h (x, x') = n^{-1} \sum_{t=1}^{n-h} \sum_{k,k'=1}^{\infty} \{\xi_{tk} \xi_{t+h,k'} - \delta_{kk'} \gamma_k (h)\} \phi_k (x) \phi_{k'} (x').
\]

Let \(\varphi_h (x, x')\) be a zero mean Gaussian random field defined on \([0, 1]^2\) with co-
variance function $\Omega_h(x, x', y, y')$ defined as

$$
\Omega_h(x, x', y, y') = \text{Cov}\{\varphi_h(x, x'), \varphi_h(y, y')\}
= \sum_{l=-\infty}^{\infty} C_l(x, y) C_l(x', y') + \sum_{l=-\infty}^{\infty} C_{l+h}(x, x') C_{l-h}(y, y')
+ \sum_{m=1}^{\infty} (E\zeta_0^4 - 3) \gamma_m^2(h) \phi_m(x) \phi_m(x') \phi_m(y) \phi_m(y')
$$

for $x, x', y, y' \in [0, 1]$. Then $E\varphi_h^2(x, x') = \Omega_h(x, x', x, x') = \Xi_h(x, x')$ satisfies a functional version of Bartlett’s formula

$$
\Xi_h(x, x') = \sum_{l=-\infty}^{\infty} \left\{C_l(x, x) C_l(x', x') + C_{l+h}(x, x') C_{l-h}(x, x')\right\}
+ \sum_{m=1}^{\infty} (E\zeta_0^4 - 3) \gamma_m^2(h) \phi_m^2(x) \phi_m^2(x').
$$

(3.2)

Theorem 1 below gives the asymptotics of $\Delta_h(x, x')$.

**Theorem 1.** Under Assumptions (A3)-(A5), for $h \in \mathbb{N}$, as $N \to \infty$

$$
\|nE\Delta_h(x, x') \Delta_h(y, y') - \Omega_h(x, x', y, y')\|_{\infty} \to 0,
$$

$$
\sqrt{n}\Delta_h(x, x') \to_D \varphi_h(x, x'), \sqrt{n}\left\{\tilde{C}_h(x, x') - C_h(x, x')\right\} \to_D \varphi_h(x, x') .
$$

**Remark 2.** The Bartlett’s formula (Chapter 7, Brockwell and Davis (1991)) for sample ACF of numerical time series is extended to functional Bartlett’s formula (3.2) for infeasible FACF $\tilde{C}_h(x, x')$. If one takes the degenerate FMA($\infty$) with $\phi_1(x) \equiv 1, \phi_k(x) \equiv 0, k \geq 2$, the infeasible ACF $\tilde{C}_h(x, x')$ simplifies to the sample ACF $\gamma^*(h) = n^{-1} \sum_{t=-h}^{n-h} \xi_{t+1} \xi_{t+h+1}$ of numerical time series $\{\xi_{t+1}\}_{t=-\infty}^\infty$,
and $\Xi (x, x')$ becomes the asymptotic covariance of $\gamma^* (h)$,

$$(E_{C_0^4} - 3) \gamma_1^2 (h) + \sum_{l=-\infty}^{\infty} \left\{ \gamma_1^2 (l) + \gamma_1 (l + h) \gamma_1 (l - h) \right\},$$

same as equation (7.3.3) in Brockwell and Davis (1991) by setting $p = q = h$.

Theorem 1 is proved in Section S3 of the Supplement. While it provides desirable asymptotics of the infeasible estimator $\hat{C}_h (x, x')$, it is not a statistic since the centered trajectories $\chi_t (\cdot)$ are unobservable.

### 3.2 Oracle efficiency

The next Proposition 2 states that the proposed two-step FCAF estimator $\hat{C}_h (x, x')$ in (2.3) is oracally efficient: up to order $n^{1/2}$, it is asymptotically equivalent to, or as efficient as $\check{C}_h (x, x')$, the infeasible FCAF estimator if all centered trajectories $\chi_t (\cdot)$ were fully known by “oracle”. Thus $\hat{C}_h (x, x')$ enjoys all the same asymptotic properties as $\check{C}_h (x, x')$.

**Proposition 2.** Under Assumptions (A1)-(A6), for $h \in \mathbb{N}$, as $N \to \infty$,

$$\sup_{(x, x') \in [0,1]^2} \left| \hat{C}_h (x, x') - \check{C}_h (x, x') \right| = o_p \left( n^{-1/2} \right).$$

Proposition 2 and Theorem 1 lead to the following.

**Theorem 2.** Under Assumptions (A1)-(A6), for $h \in \mathbb{N}$, as $N \to \infty$,

$$\sup_{(x, x') \in [0,1]^2} \left| \hat{C}_h (x, x') - C_h (x, x') - \Delta_h (x, x') \right| = o_p \left( n^{-1/2} \right),$$

Statistica Sinica: Newly accepted Paper
(accepted author-version subject to English editing)
\[
\sqrt{n} \left\{ \hat{C}_h (x, x') - C_h (x, x') \right\} \rightarrow_D \varphi_h (x, x').
\] (3.3)

Therefore the limiting distribution of \( \hat{C}_h (x, x') - C_h (x, x') \) is the same as \( \hat{C}_h (x, x') - C_h (x, x') \), thus the term “oracle efficiency”.

4. Simultaneous confidence envelope

In this section, asymptotic simultaneous confidence envelope (SCE) is constructed for the functional autocovariance function \( C_h (x, x') \).

4.1 Asymptotic SCE

The main theorem in this section establishes asymptotic behavior of the normalized maximal deviation of the FACF estimator \( \hat{C}_h (x, x') \). Making use of (3.3) in Theorem 2, one obtains the following standardized limiting distribution

\[
\sqrt{n} \left\{ \hat{C}_h (x, x') - C_h (x, x') \right\} \Xi_h^{-1/2} (x, x') \rightarrow_D \varphi_h (x, x') \Xi_h^{-1/2} (x, x'),
\] (4.1)

in which \( \varphi_h (x, x') \) is the mean zero Gaussian random field defined in Theorem 1 with pointwise variance function \( \Xi_h (x, x') = \text{E} \varphi_h^2 (x, x') \). The standardized random field \( \varphi_h (x, x') \Xi_h^{-1/2} (x, x') \) therefore has mean zero and variance one.

Denote by \( Q_{1-\alpha} \) the 100 \((1 - \alpha)\)-th percentile of the absolute maxima distribution of \( \varphi_h (x, x') \Xi_h^{-1/2} (x, x') \), i.e., for \( \alpha \in (0, 1), h \in \mathbb{N} \),

\[
P \left\{ \sup_{(x, x') \in [0, 1]^2} |\varphi_h (x, x')| \Xi_h^{-1/2} (x, x') \leq Q_{1-\alpha} \right\} = 1 - \alpha.
\] (4.2)
The next theorem follows directly from (4.1) and (4.2):

**Theorem 3.** Under Assumptions (A1)-(A6), for \( \alpha \in (0, 1) \), \( h \in \mathbb{N} \),

\[
\lim_{N \to \infty} \mathbb{P} \left\{ \sup_{(x,x') \in [0,1]^2} n^{1/2} \left| \hat{C}_h (x, x') - C_h (x, x') \right| \Xi_h^{-1/2} (x, x') \leq Q_{1-\alpha} \right\} = 1 - \alpha.
\]

**Corollary 1.** Under Assumptions (A1)-(A6), as \( N \to \infty \), an asymptotic 100 \((1-\alpha)\)% SCE for \( C_h (x, x') \) is \( \hat{C}_h (x, x') \pm n^{-1/2} Q_{1-\alpha} \Xi_h^{-1/2} (x, x') \), \( x, x' \in [0,1] \).

4.2 Knot selection

The number of knots \( J_s \) for spline smoothing is selected subject to the constraints of Assumption (A6). The smoothness order \((q, \mu)\) of mean function \( m (\cdot) \) and eigenfunctions \( \phi_k (\cdot) \) is taken as \((3, 1)\) or \((4, 0)\) by default with a matching spline order \( p = 4 \) (cubic spline). Therefore, \( J_s = \lfloor c N^\gamma \log \log (N) \rfloor \) is recommended with constant \( c \), where \( \lfloor a \rfloor \) denotes the integer part of \( a \). The default values of parameters \( \gamma = 3/8 \) and \( c = 0.8 \) are adequate in simulation.

4.3 FPC analysis

We now describe covariance function estimator \( \hat{C}_0 (\cdot, \cdot) = \hat{G} (\cdot, \cdot) \), eigenfunction estimators \( \hat{\phi}_k (\cdot) \), and eigenvalue estimators \( \hat{\lambda}_k \) in the FPC analysis. One estimates \( C_0 (\cdot, \cdot) \) by

\[
\hat{C}_0 (x, x') = n^{-1} \sum_{t=1}^{n} \hat{\chi}_t (x) \hat{\chi}_t (x') = \sum_{s=1}^{J_s+p} \sum_{s'=1}^{J_s+p} \hat{\beta}_{ss'} B_{s,p} (x) B_{s',p} (x'),
\]

where \( \hat{\chi}_t (\cdot) \) is defined in (2.5) and \( \hat{\beta}_{ss'} \)'s are the coefficients.
FUNCTIONAL TIME SERIES: AUTOCOVARIANCE FUNCTION

Denote \( B(x) = \{B_{1,p}(x), \ldots, B_{J_s+p,p}(x)\}^\top \), and \( N \times (J_s + p) \) design matrix \( B \) for spline regression is \( B = \{B(1/N), \ldots, B(N/N)\}^\top \). Then for any \( k \in \{1, \ldots, \kappa\} \), one considers the following spline approximation for \( \psi_k(\cdot) \): 

\[
\hat{\psi}_k(x') = \sum_{\ell=1}^{J_s+p} \hat{\gamma}_{\ell k} B_{\ell,p}(x'),
\]

where \( \hat{\gamma}_{\ell k} \)'s are coefficients of B-spline estimator subject to 

\[
\hat{\lambda}_k B^\top \hat{B} \hat{\gamma}_k = N \quad \text{with} \quad \hat{\gamma}_k = (\hat{\gamma}_{1,k}, \ldots, \hat{\gamma}_{J_s+p,k})^\top.
\]

The estimates of eigenfunctions and eigenvalues correspond \( \psi_k \) and \( \lambda_k \) can be obtained by solving the eigenequations,

\[
\int \hat{C}_0(x, x') \hat{\psi}_k(x') \, dx' = \hat{\lambda}_k \hat{\psi}_k(x), \quad 1 \leq k \leq J_s + p.
\]

According to (4.3), solving (4.4) is equivalent to solving 

\[
N^{-1}B^\top(x) \hat{\beta} B^\top \hat{\gamma}_k = \hat{\lambda}_k B^\top(x) \hat{\gamma}_k, \quad 1 \leq k \leq J_s + p.
\]

Consider the following Cholesky decomposition: 

\[
B^\top B = L_B L_B^\top,
\]

solving (4.4) is equivalent to solving 

\[
\hat{\lambda}_k L_B^\top \hat{\gamma}_k = N^{-1}L_B^\top \hat{\beta} L_B L_B^\top \hat{\gamma}_k \quad \text{subject to} \quad \hat{\gamma}_k L_B L_B^\top \hat{\gamma}_k = N,
\]

that is, \( \hat{\lambda}_k \) and \( L_B^\top \hat{\gamma}_k \), 

\( 1 \leq k \leq J_s + p \) are the eigenvalues and unit eigenvectors of 

\( N^{-1}L_B^\top \hat{\beta} L_B \). In other words, \( \hat{\gamma}_k \) is obtained by left multiplying \( N^{1/2} (L_B^\top)^{-1} \) to unit eigenvectors of 

\( N^{-1}L_B^\top \hat{\beta} L_B \), hence \( \hat{\psi}_k(\cdot) \) is obtained. Consequently, 

\[
\hat{\phi}_k(x') = \hat{\lambda}_k^{1/2} \hat{\psi}_k(x').
\]

The \( k \)-th FPC score of the \( t \)-th curve is estimated by numerical integration

\[
\hat{\xi}_{tk} = N^{-1} \sum_{j=1}^N \hat{\lambda}_k^{-1} \hat{\chi}_t \left( \frac{j}{N} \right) \hat{\phi}_k \left( \frac{j}{N} \right), 1 \leq t \leq n, 1 \leq k \leq J_s + p.
\]

Instead of computing all \( J_s + p \) eigenvalues, it is typical to truncate the spectral decomposition at a much smaller number of eigenvalues that account for 95%
of the variation in the data, \( \kappa = \arg \min_{1 \leq l \leq J_s+p} \left\{ \sum_{k=1}^{J_s+p} \tilde{\lambda}_k / \sum_{k=1}^{J_s+p} \lambda_k \geq 0.95 \right\} \), see Cao et al. (2016), Wang et al. (2020a).

4.4 Estimating the variance function \( \Xi_h \) and the percentile \( Q_{1-\alpha} \)

Strong approximation property of \( \xi_{tk} \), guaranteed by Assumption (A5) and Lemma S.4, makes possible a procedure to approximate \( \Xi_h \) and \( Q_{1-\alpha} \). Denote the matrices \( \Gamma^{(n)}_k = \{ \gamma_k(|i-j|) \}_{i,j=1}^n \), \( k = 1, \ldots, \kappa \), where \( \gamma_k(|i-j|) \) is the \((i,j)\)-th entry of \( \Gamma^{(n)}_k \). It is easily shown that \( \Gamma^{(n)}_k \) is the covariance matrix of random vector \((\xi_{1k}, \xi_{2k}, \ldots, \xi_{nk})\). Useful estimates of \( \gamma_k(h) \) can only be made if \( n \geq 50 \) and \( 0 \leq h \leq n/4 \), p. 221, Chapter 7 in Brockwell and Davis (1991). One therefore considers the \( r \times r \) submatrix \( \Gamma^{(r)}_k \) of \( \Gamma^{(n)}_k \), where \( r \) can be chosen as \( \lceil \sqrt{n} \rceil \) or \( \lfloor n/4 \rfloor \). Then according to the definition of \( \gamma_k(h) \), \( h \in \mathbb{N} \), one can construct the estimation \( \hat{\Gamma}^{(r)}_k \) of \( \Gamma^{(r)}_k \) by

\[
\hat{\gamma}^{(r)}_k(h) = n^{-1} \sum_{t=1}^{n-h} \left( \hat{\xi}_{tk} - \hat{\xi}. \right) \left( \hat{\xi}_{t+h,k} - \hat{\xi}. \right)
\]

for \( 0 \leq h \leq r \), where \( \hat{\xi}. = n^{-1} \sum_{t=1}^{n-h} \hat{\xi}_{tk} \), \( 1 \leq k \leq \kappa \).

For \( 1 \leq k \leq \kappa \), one generates \( B \) independent replications of \( r \)-dimensional normal random vectors \( \{ (Z_{1k}^b, \ldots, Z_{rk}^b) \}_{b=1}^B \) with mean zero and covariance matrix

\[
\hat{\Gamma}^{(r)}_k \), where \( B \) is a preset large integer. Let \( \hat{\xi}_h(x, x') = r^{-1} \sum_{t=1}^{r-h} \left\{ \sum_{k=1}^{\kappa} Z_{tk+k}^b \hat{\phi}_k(x) \right\} \left\{ \sum_{k=1}^{\kappa} Z_{tk+h,k}^b \hat{\phi}_k(x) \right\}, \) and one can estimate \( \Xi_h(x, x') \) by

\[
\hat{\Xi}_h^B (x, x') = \frac{r}{B-1} \sum_{b=1}^B \left\{ \hat{C}_h^b (x, x') - \bar{C}_h^* (x, x') \right\}^2 ,
\]

where \( \bar{C}_h^* (x, x') = B^{-1} \sum_{b=1}^B \hat{C}_h^b (x, x') \).
Lastly, define the empirical quantile
\[ \hat{Q}_{1-\alpha}^B \equiv (1-\alpha) \text{-th quantile of } \left\{ \sup_{x,x' \in [0,1]} \frac{n^{1/2} | \hat{C}_h^B (x,x') - \bar{C}_h^* (x,x') |}{\hat{\Xi}_h^B (x,x')^{1/2}} , b = 1, \ldots B \right\}, \]
where the supremum is computed over \( N \times N \) equal distance grid points on \([0,1]^2\) with \( B = 500 \). The SCEs are then computed as
\[ \hat{C}_h (x,x') \pm n^{-1/2} \hat{Q}_{1-\alpha}^B (x,x')^{1/2} , x,x' \in [0,1]. \] (4.5)

While consistency of \( \hat{Q}_{1-\alpha}^B \hat{\Xi}_h^B (x,x') \) as an estimate of \( Q_{1-\alpha} \Xi_h (x,x') \) is conceivable, its full theoretical proof is beyond the scope of this work.

5. Simulation studies

In this section, we examine finite-sample performance of the SCEs. The data are generated from the following model:
\[ Y_{tj} = m (j/N) + \sum_{k=1}^{2} \xi_{tk} \phi_k (j/N) + \sigma \varepsilon_{tj} , 1 \leq j \leq N, 1 \leq t \leq n. \]

**Case 1:** \( m (x) = 10 + \sin \{ 2 \pi (x - 1/2) \} , \varepsilon_{tj} \sim N (0,1) , 1 \leq t \leq n, 1 \leq j \leq N, \phi_1 (x) = -2 \cos \{ \pi (x - 1/2) \} \) and \( \phi_2 (x) = \sin \{ \pi (x - 1/2) \} \). \( \{ \xi_{tk} \}_{t=1,k=1}^{n,2} \) are generated from (1.4), where \( \{ \xi_{tk} \}_{t=1,k=1}^{n,2} \) are i.i.d \( N (0,1) \)
\[ a_{0k} = 0.8, a_{1k} = 0.6, a_{tk} = 0, \forall t \geq 2, k = 1,2, \]
so \( \{ \xi_{tk} \}_{t=1}^{n} \) is an MA(1) process. The number \( n \) of trajectories is taken to be 160, 400, 900, 1600 and the number \( N \) of observations per trajectory 200, 500, 1000, 2000 respectively, the noise level \( \sigma = 0.1 \).
Case 2: \( \{\xi_{tk}\}_{t=1,k=1}^{n,2} \) are generated from (1.4), where \( \{\zeta_{tk}\}_{t=1,k=1}^{n,2} \) are i.i.d \( N(0,1) \) variables and
\[
a_{0k} = 0.3, a_{1k} = 0.6, a_{2k} = 0.741, a_{tk} = 0, \forall t \geq 3, k = 1, 2,
\]
so \( \{\xi_{tk}\}_{t=1}^{n} \) is an MA(2) process, while \( m(x), \varepsilon_{tj}, \phi_{1}(x) \) and \( \phi_{2}(x) \) are the same as those in Case 1. The number \( n \) of curve is taken to be 400 and the number \( N \) of observations per curve is taken to be 500.

Case 3: \( \{\xi_{tk}\}_{t=1,k=1}^{n,2} \) are generated from (1.4), where \( \{\zeta_{tk}\}_{t=1,k=1}^{n,2} \) are i.i.d \( N(0,1) \) variables and
\[
a_{tk} = 2^{-\left(\frac{t+1}{2}\right)}, \forall t \in \mathbb{N}, k = 1, 2.
\]
so \( \{\xi_{tk}\}_{t=1}^{n} \) is an MA(\( \infty \)) process, while \( m(x), \phi_{1}(x) \) and \( \phi_{2}(x) \) are the same as those in Case 1 and \( \varepsilon_{tj} \) follows the standardized \( t \)-distribution with degree of freedom 40, \( \varepsilon_{tj} \sim \sqrt{\frac{19}{20}}t_{40} \). The number \( n \) of trajectories is taken to be 160, 400, 900, 1600 and the number \( N \) of observations per trajectory 200, 500, 1000, 2000 respectively, the noise level \( \sigma = 0.1 \).

Tables 1-3 display the empirical coverage frequencies out of 1000 replications of the true surface \( C_{h}(\cdot,\cdot), h = 0,1,2 \) in Case 1 being covered by the SCE in (1.3) at all \( N \times N \) grid points \( \{(j/N,j'/N) , 1 \leq j,j' \leq N \} \) on \([0,1]^{2} \). Overall, the empirical coverage frequency approaches the nominal level as \( n,N \) increase, slightly higher than nominal in Table 3 due to the true FACF \( C_{2}(\cdot,\cdot) \equiv 0 \). For
Case 3 with genuine FMA(∞) trajectories, Tables 4-6 exhibit similar patterns.

Figure 1 depicts the FACF estimate $\hat{C}_1(\cdot, \cdot)$ and 95\% SCEs for the true FACF for one replication of Case 1 with $\sigma = 0.1, r = [n/4]$ and various $n, N$ combinations. As expected, the SCEs become narrower as the sample size $n$ increases. Figure 2 with different lag $h$ exhibit similar patterns as Figure 1, forming positive confirmation of the asymptotic theory.

Figure 3 shows the FACF estimate and 95\% SCE of the true FACF for one replication of Case 2 with $\sigma = 0.1$ and $(n, N) = (400, 500), h = 0, 1, 2, 3, p = 4$. It is seen that the 95\% SCE fails to cover entirely the zero plane for lag $h = 0, 1, 2$, whereas it does for $h = 3$. If null hypothesis $H_0 : C_h(x, x') \equiv 0$ would be tested for $h = 0, 1, 2, 3$, one would retain the null hypothesis for $h = 3$ with $p$-value $> 0.05$, and reject the null hypothesis for $h = 0, 1, 2$ at significant level $\alpha = 0.05$. Thus one could infer that the FMA(∞) to be actually FMA(2) according to Proposition 1, exactly the model of Case 2.

For the genuine FMA(∞) Case 3, however, Figure 4 shows that for one replication, the 95\% SCE fails to cover entirely the zero plane for lag $h = 0, 1, 2, 3, 4$, whereas it does for $h = 5$. Then under the same testing procedure, one could infer that the FMA(∞) to be actually FMA(4) according to Proposition 1. It is unclear at the moment what the asymptotics of FMA order by Proposition 1 should look like when the true model is a genuine FMA(∞). Further research is also
needed on how Proposition should be applied to multiple h’s simultaneously or sequentially.

6. Real data analysis

In this section, we study more closely the ElectroEncephalogram (EEG) data discussed in Section 1. The data was collected by research group of Prof. Linhong Ji at Tsinghua University Department of Mechanical Engineering, from 142 University student participants. The data used in our study is recorded by electrodes placed on scalp surface of an individual subject A at the 6-th of 32 scalp locations, at 1000Hz sample rate. The mid portion of 60000 EEG recordings are used as 300 consecutive segments, each consisting of 200 EEG signals. Hence there are \( n = 300 \) unobserved curves with \( N = 200 \) signals recorded in each curve (see Figure 8 for the raw data). The data range is from -22 to 22. Inspection of the EEG data and its analysis code can be arranged with the authors.

Multiple SCEs are used for testing null hypothesis \( H_0 : C_h (\cdot, \cdot) \equiv 0 \) with preset lag \( h \). For \( h = 1 \), the estimated FACF and 95% SCE computed by \( (4.5) \) are shown on Figure 5. Figure 6 (a) shows that the 95% SCEs (upper and lower surfaces) and the null hypothesis zero plane. One rejects the null hypothesis at significant level \( \alpha = 0.05 \) since the zero plane is not entirely covered by 95% SCE. Moreover, Figure 6 (b) shows that even the 99% SCE does not contain the zero plane hence one rejects the null hypothesis \( H_0 : C_1 (\cdot, \cdot) \equiv 0 \) with p-
value < 0.01. Similar for \( h = 2, 3, \ldots, 9 \), one rejects the null hypothesis with 
\( p \)-values less than 0.05. For \( h = 10 \), the 95% SCE covers the null hypothesis 
zero plane and the lowest confidence levels of SCE covering the entire zero plane 
is 0.896, hence one retains the null hypothesis \( H_0 : C_{10}(x, x') \equiv 0 \) with \( p \)-value 
0.104. Similar null hypotheses are retained for \( h = 11, 12, 13, 14 \) with \( p \)-values 
= 0.150, 0.220, 0.204, 0.192 respectively. Thus Proposition II points to an FMA(9) 
model for the data.

We have also investigated specific form of covariance function \( C_0(x, x') \) for 
the EEG data. Figure II (a) depicts the estimated FACF, which hints strongly of 
trigonometric form. The null hypothesis \( H_0 : C_0(x, x') = a_0 + \sum_{k=1}^{2} \{ 2a_k \cos (2k\pi x) 
\cos (2k\pi x') + 2b_k \sin (2k\pi x) \sin (2k\pi x') \} \) is tested in which parameters \( \{ a_k \}_{k=0}^{2} \), \( \{ b_k \}_{k=1}^{2} \) 
are estimated by linear least squares. Figure II (b) shows that the 95% SCE (upper 
and lower surfaces) cover the null surface completely, and the coefficient of 
determination for the two-step estimate \( \hat{C}_0(x, x') \) by the trigonometric form is 
\( R^2 = 0.9661 \). The lowest confidence level by which the SCE covers the entire null 
surface (Figure II (c)) is 67.4%, so one retains the null hypothesis with \( p \)-value 
= 0.326.

Since all estimated parameters \( \hat{a}_0, \hat{a}_1, \hat{a}_2, \hat{b}_1, \hat{b}_2 \) are significantly positive, the 
above null hypothesis is a Mercer Lemma expansion of the FACF with positive
eigenvalues in descending order

\[ \lambda_{0,1} = \hat{b}_2 = 19.6, \lambda_{0,2} = \hat{a}_2 = 17.1, \lambda_{0,3} = \hat{a}_0 = 14.6, \lambda_{0,4} = \hat{a}_1 = 3.5, \lambda_{0,5} = \hat{b}_1 = 2.9, \]

and orthonormal eigenfunctions \( \psi_k(x) \) being Fourier basis:

\[
\psi_{0,1}(x) \equiv \sqrt{2} \cos(4\pi x), \psi_{0,2}(x) \equiv \sqrt{2} \sin(4\pi x), \psi_{0,3}(x) \equiv 1, \\
\psi_{0,4}(x) \equiv \sqrt{2} \sin(2\pi x), \psi_{0,5}(x) \equiv \sqrt{2} \cos(2\pi x). \tag{4.1}
\]

This strongly suggests a Karhunen-Loève expansion with Fourier form FPCs

\[
\phi_{0,1}(x) \equiv \left(2\hat{b}_2\right)^{1/2} \cos(4\pi x), \phi_{0,2}(x) \equiv \left(2\hat{a}_2\right)^{1/2} \sin(4\pi x), \phi_{0,3}(x) \equiv \hat{a}_0^{1/2}, \\
\phi_{0,4}(x) \equiv \left(2\hat{a}_1\right)^{1/2} \sin(2\pi x), \phi_{0,5}(x) \equiv \left(2\hat{b}_1\right)^{1/2} \cos(2\pi x).
\]

The centered trajectories \( \hat{\chi}_t(\cdot) \) in (2.5) are then well approximated by

\[
\sum_{k=1}^{5} \hat{\xi}_{tk} \phi_{0,k}(\cdot), \\hat{\xi}_{tk} = \lambda_{0,k}^{-1} \int_{t}^{\hat{\chi}_t(x) \phi_{0,k}(x) \, dx, 1 \leq k \leq 5, 1 \leq t \leq 300.}
\]

Figure 8 depicts for 2 randomly chosen integers \( t \) between 1 and 300, both the spline trajectory \( \hat{\eta}_t(\cdot) = \hat{m}(\cdot) + \hat{\chi}_t(\cdot) \) (solid), and the trajectory with hypothesized FPCs \( \{ \phi_{0,k}(\cdot) \}_{k=1}^{5} : \hat{m}(\cdot) + \sum_{k=1}^{5} \hat{\xi}_{tk} \phi_{0,k}(\cdot) \) (dash). Both appear to be faithful representations of the raw EEG data (crosses), with \( R^2 \) values (0.985, 0.959), (0.969, 0.948) respectively for spline and Fourier basis trajectories and the 2 raw data segments. This further corroborates that for this particular EEG data, the Fourier FPCs in (4.1) are appropriate.
7. Conclusions

Properties of FACF are studied for FMA(∞), and a two-step tensor-product spline estimator is proposed for the FACF, which is asymptotically equivalent to an infeasible estimator at the rate of $o_p\left(n^{-1/2}\right)$. Asymptotic SCE is established which can be used to test any hypothesis on the FACF such as equality to zero.

For EEG time series, strong evidence points to FMA of finite lag and Fourier form FPCs. The theoretically justified SCE is a powerful inference tool which is expected to find more applications in various scientific fields.

Acknowledgments

This research has been supported by National Natural Science Foundation of China awards 12171269, 12026242, 11771240. The authors are grateful to Professor Linhong Ji and his research group for providing the EEG data, and to two Referees whose comments have helped to significantly improve our manuscript.

Online Supplementary Materials

The online supplement contains the detailed proofs for the main results.

References


Gu, L. and Yang, L. (2015). Oracally efficient estimation for single-index link function with

Guo, J. and Chen, Y. (2019). An $L^2$-norm based ANOVA test for the equality of weakly depen-


Introduction to Linear Operators*. Wiley, Chichester.

online** DOI: 10.5705/ss.202021.0107.

Ma, S., Yang, L. and Carroll, R. (2012). A simultaneous confidence band for sparse longitudinal


### Table 1: Coverage frequencies for $C_0(\cdot, \cdot)$ in Case 1 by SCE (1.5) with $p = 4$

<table>
<thead>
<tr>
<th>$(n, N)$</th>
<th>$r$</th>
<th>$1 - \alpha = 0.95$</th>
<th>$1 - \alpha = 0.99$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(160, 200)$</td>
<td>$\sqrt{n}$</td>
<td>0.912</td>
<td>0.986</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.912</td>
<td>0.975</td>
</tr>
<tr>
<td>$(400, 500)$</td>
<td>$\sqrt{n}$</td>
<td>0.953</td>
<td>0.991</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.952</td>
<td>0.991</td>
</tr>
<tr>
<td>$(900, 1000)$</td>
<td>$\sqrt{n}$</td>
<td>0.946</td>
<td>0.992</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.952</td>
<td>0.990</td>
</tr>
<tr>
<td>$(1600, 2000)$</td>
<td>$\sqrt{n}$</td>
<td>0.952</td>
<td>0.992</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.956</td>
<td>0.992</td>
</tr>
</tbody>
</table>

### Table 2: Coverage frequencies for $C_1(\cdot, \cdot)$ in Case 1 by SCE (1.5) with $p = 4$

<table>
<thead>
<tr>
<th>$(n, N)$</th>
<th>$r$</th>
<th>$1 - \alpha = 0.95$</th>
<th>$1 - \alpha = 0.99$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(160, 200)$</td>
<td>$\sqrt{n}$</td>
<td>0.914</td>
<td>0.988</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.906</td>
<td>0.982</td>
</tr>
<tr>
<td>$(400, 500)$</td>
<td>$\sqrt{n}$</td>
<td>0.939</td>
<td>0.995</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.946</td>
<td>0.992</td>
</tr>
<tr>
<td>$(900, 1000)$</td>
<td>$\sqrt{n}$</td>
<td>0.954</td>
<td>0.994</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.954</td>
<td>0.987</td>
</tr>
<tr>
<td>$(1600, 2000)$</td>
<td>$\sqrt{n}$</td>
<td>0.954</td>
<td>0.994</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.950</td>
<td>0.994</td>
</tr>
</tbody>
</table>
Table 3: Coverage frequencies for $C_2 (\cdot , \cdot )$ in Case 1 by SCE (1.5) with $p = 4$

<table>
<thead>
<tr>
<th>$(n, N)$</th>
<th>$r$</th>
<th>$1 - \alpha = 0.95$</th>
<th>$1 - \alpha = 0.99$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(160, 200)</td>
<td>$\sqrt{n}$</td>
<td>0.954</td>
<td>0.997</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.966</td>
<td>0.997</td>
</tr>
<tr>
<td>(400, 500)</td>
<td>$\sqrt{n}$</td>
<td>0.958</td>
<td>0.998</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.963</td>
<td>0.995</td>
</tr>
<tr>
<td>(900, 1000)</td>
<td>$\sqrt{n}$</td>
<td>0.952</td>
<td>0.994</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.956</td>
<td>0.991</td>
</tr>
<tr>
<td>(1600, 2000)</td>
<td>$\sqrt{n}$</td>
<td>0.950</td>
<td>0.994</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.963</td>
<td>0.995</td>
</tr>
</tbody>
</table>

Table 4: Coverage frequencies for $C_0 (\cdot , \cdot )$ in Case 3 by SCE (1.5) with $p = 4$

<table>
<thead>
<tr>
<th>$(n, N)$</th>
<th>$r$</th>
<th>$1 - \alpha = 0.95$</th>
<th>$1 - \alpha = 0.99$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(160, 200)</td>
<td>$\sqrt{n}$</td>
<td>0.898</td>
<td>0.954</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.904</td>
<td>0.951</td>
</tr>
<tr>
<td>(400, 500)</td>
<td>$\sqrt{n}$</td>
<td>0.939</td>
<td>0.981</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.942</td>
<td>0.977</td>
</tr>
<tr>
<td>(900, 1000)</td>
<td>$\sqrt{n}$</td>
<td>0.939</td>
<td>0.993</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.945</td>
<td>0.982</td>
</tr>
<tr>
<td>(1600, 2000)</td>
<td>$\sqrt{n}$</td>
<td>0.953</td>
<td>0.995</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.951</td>
<td>0.992</td>
</tr>
</tbody>
</table>
Table 5: Coverage frequencies for $C_1 (\cdot, \cdot)$ in Case 3 by SCE (4.5) with $p = 4$

<table>
<thead>
<tr>
<th>$(n, N)$</th>
<th>$r$</th>
<th>$1 - \alpha = 0.95$</th>
<th>$1 - \alpha = 0.99$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(160, 200)$</td>
<td>$\sqrt{n}$</td>
<td>0.902</td>
<td>0.963</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.909</td>
<td>0.950</td>
</tr>
<tr>
<td>$(400, 500)$</td>
<td>$\sqrt{n}$</td>
<td>0.948</td>
<td>0.986</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.941</td>
<td>0.982</td>
</tr>
<tr>
<td>$(900, 1000)$</td>
<td>$\sqrt{n}$</td>
<td>0.949</td>
<td>0.996</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.944</td>
<td>0.986</td>
</tr>
<tr>
<td>$(1600, 2000)$</td>
<td>$\sqrt{n}$</td>
<td>0.951</td>
<td>0.995</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.944</td>
<td>0.986</td>
</tr>
</tbody>
</table>

Table 6: Coverage frequencies for $C_2 (\cdot, \cdot)$ in Case 3 by SCE (4.5) with $p = 4$

<table>
<thead>
<tr>
<th>$(n, N)$</th>
<th>$r$</th>
<th>$1 - \alpha = 0.95$</th>
<th>$1 - \alpha = 0.99$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(160, 200)$</td>
<td>$\sqrt{n}$</td>
<td>0.921</td>
<td>0.979</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.913</td>
<td>0.982</td>
</tr>
<tr>
<td>$(400, 500)$</td>
<td>$\sqrt{n}$</td>
<td>0.946</td>
<td>0.985</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.942</td>
<td>0.982</td>
</tr>
<tr>
<td>$(900, 1000)$</td>
<td>$\sqrt{n}$</td>
<td>0.940</td>
<td>0.992</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.945</td>
<td>0.986</td>
</tr>
<tr>
<td>$(1600, 2000)$</td>
<td>$\sqrt{n}$</td>
<td>0.943</td>
<td>0.993</td>
</tr>
<tr>
<td></td>
<td>$n/4$</td>
<td>0.955</td>
<td>0.994</td>
</tr>
</tbody>
</table>
Figure 1: Plots of the true $C_1(\cdot, \cdot)$ in Case 1 and two-step estimate $\hat{C}_1(\cdot, \cdot)$ (middle surfaces), with 95% SCEs (upper and lower surfaces), $r = \lceil n/4 \rceil$. (a)-(d) correspond to $(n, N) = (160, 200), (400, 500), (900, 1000), (1600, 2000)$ respectively.
Figure 2: Plots of the true $C_2(\cdot,\cdot)$ in Case 1 and two-step estimate $\hat{C}_2(\cdot,\cdot)$ (middle surfaces), with 95% SCEs (upper and lower surfaces), $r = [n/4]$. (a)-(d) correspond to $(n, N) = (160, 200), (400, 500), (900, 1000), (1600, 2000)$ respectively.
Figure 3: Plots of the true $C_h(\cdot, \cdot)$ in Case 2 and two-step estimate $\hat{C}_h(\cdot, \cdot)$ (middle surfaces), with 95% SCEs (upper and lower surfaces), $r = [n/4]$. (a)-(d) correspond to $h = 0, 1, 2, 3$ respectively.
Figure 4: Plots of the true $C_h(\cdot, \cdot)$ in Case 3 and two-step estimate $\hat{C}_h(\cdot, \cdot)$ (middle surfaces), with 95% SCEs (upper and lower surfaces), $r = [n/4]$. (a)-(f) correspond to $h = 0, 1, 2, 3, 4, 5$ respectively.
Figure 5: Two-step estimate $\hat{C}_1(\cdot, \cdot)$ of EEG data with $r = [n/4]$ (middle surface) and its 95% SCE (upper and lower surfaces).

Figure 6: Null hypothesis zero plane and SCEs (upper and lower surfaces) for $C_1(\cdot, \cdot)$ of EEG data with (a) $\alpha = 0.05$, (b) $\alpha = 0.01$. 
Figure 7: Null hypothesis middle surface $a_0 + \sum_{k=1}^{2} \{2a_k \cos(2k\pi x) \cos(2k\pi x') + 2b_k \sin(2k\pi x) \sin(2k\pi x')\}$ and SCEs (upper and lower surfaces) of $C_0(x, x')$ of EEG data with (a) $\alpha = 0.05$, (b) $\alpha = 0.326$.

Figure 8: Randomly selected segments of raw EEG data (crosses), spline trajectories (solid) and trajectories based on trigonometric FPCs (dash).