Statistica Sinica Preprint No: SS-2024-0211	
Title	Sufficient Dimension Reduction for the Conditional Quantiles of Functional Data
Manuscript ID	SS-2024-0211
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202024.0211
Complete List of Authors	Eliana Christou, Eftychia Solea, Shanshan Wang and Jun Song
Corresponding Authors	Jun Song
E-mails	junsong@korea.ac.kr

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

Sufficient Dimension Reduction for the Conditional Quantiles of Functional Data

Eliana Christou, Eftychia Solea, Shanshan Wang, Jun Song

University of North Carolina at Charlotte, Queen Mary University of London, Korea University

Abstract: Functional data analysis holds transformative potential across fields but often relies on mean regression, with limited focus on quantile regression. Furthermore, the infinite-dimensional nature of the functional predictors necessitates the use of dimension reduction techniques. Therefore, in this work, we address this gap by developing dimension reduction techniques for the conditional quantiles of functional data. The idea is to replace the functional predictors with a few finite predictors without losing important information on the conditional quantile while maintaining a flexible nonparametric model. We derive the convergence rates of the proposed estimators and demonstrate their finite sample performance using simulations and a real dataset from fMRI studies.

Key words and phrases: conditional quantiles, dimension reduction, functional data analysis.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

1. Introduction

Functional data analysis (FDA) considers data as realizations of random functions, impacting fields like finance (Müller, Sen, and Stadmüller, 2011), natural language processing (NLP) (Gubian, Torreira, and Bovwes, 2015), electric grid stabilization (Elías, Jiménez, and Shang, 2022; Fontana, Tavoni, and Vantini, 2019), and notably medicine (Li and Luo, 2017; Pratt, Su, Hayes, Clancy, and Szczesniak, 2021). For example, neuroimaging studies on disorders like attention deficit hyperactivity (ADHD) or Alzheimer's, aim to understand neural development, substance use, and brain structure and function. These studies often collect functional data such as functional magnetic resonance imaging (fMRI) and electroencephalogram (EEG). While traditionally analyzed using conventional methods, FDA offers a more precise representation along the continuum.

FDA, first coined by Ramsay and Dalzell (1991), has been a very active research area. Key works on functional linear regression include Cardot, Ferraty and Sarda (2003), Yao, Müller and Wang (2005), Cai and Hall (2006), Hall and Horowitz (2007), Crambes, Kneip and Sarda (2009), Shin (2009), James, Wang and Zhu (2009), and Yuan and Cai (2010); see also the monographs of Ramsay and Silverman (2005) and Kokoszka and Reimherr (2017). Nonparametric models have been applied to chemometrics, meteo-

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

rology, speech recognition, and medicine; see Besse, Cardot and Stephenson (2000), Ferraty and Vieu (2002), Ferraty and Vieu (2003), Ferraty and Vieu (2004), and Aneiros Pérez, Cardot, Estévez Pérez and Vieu (2004). Other works include extensions to multiple functional predictors (Chiou, Yang and Chen, 2016; Happ and Greven, 2018), and to generalized linear models with functional predictors (Marx and Eilers, 1999; James, 2002; Müller and Stadmüller, 2005; James and Silverman, 2005).

Functional data, though infinite-dimensional, can be reduced to finite data using operators. Ferré and Yao (2003) extended sliced inverse regression (SIR) of Li (1991) to functional sliced inverse regression (FSIR), and later introduced functional inverse regression (FIR) using kernel smoothing (Ferré and Yao, 2005). FSIR was revisited by Lian and Li (2014), who also extended sliced average variance estimation (SAVE) of Cook and Weisberg (1991) to the functional context. Wang, Lin and Zhang (2013) extended contour regression (Li, Zha and Chiaromonte, 2005) to functional contour regression (FCR), while Amato, Antoniadis and De Feis (2006) proposed a wavelet based minimum average variance estimation (MAVE). Ait-Saïdi, Ferraty, Kassa and Vieu (2008) and Chen, Hall and Müller (2011) explored single and multiple index functional models, respectively. Li and Song (2022) developed dimension reduction for random response and pre-

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

dictor functions using weak conditional expectation, and Mahzarnia and Song (2022) addressed simultaneous functional predictor selection and coefficient estimation. For a review, see Song (2019).

Many functional data exhibit features like skewness and outliers, making *quantile regression (QR)* essential. However, QR for functional data is underexplored. For functional linear QR, Cardot, Crambes and Sarda (2005) used smoothing splines, while Kato (2012) applied functional principal component analysis (FPCA). Nonparametric approaches via inversion of the conditional distribution were considered by Ferraty, Rabhi and Vieu (2005) and Chen and Müller (2012). Other works include functional partially linear QR (Lu, Du and Sun, 2014; Yao, Sue-Chee and Wang, 2017; Ma, Li, Zhu and Zhu, 2019), functional quadratic QR (Shi, Xie and Zhang, 2020), functional linear semiparametric models (Qingguo and Kong, 2017), and generalized regression quantiles (Guo, Zhou, Huang and Härdle, 2015).

Existing QR methods for functional data are either parametric or overlook its infinite-dimensional nature. To address this gap, we propose a novel approach that reduces the functional predictors to a few finite ones, while preserving crucial information about the conditional quantiles, enabling the use of flexible, nonparametric QR models. Our contribution is twofold: it introduces the *first* supervised dimension reduction technique for functional

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

conditional quantiles, and it facilitates nonparametric QR modeling.

To enhance clarity, we review the basics of Hilbert spaces in Section 2 and sufficient dimension reduction in both classical and functional settings in Section 3. The proposed methodology on dimension reduction for the conditional quantiles of functional data is introduced in Section 4. However, Sections 2 and 3 are essential for establishing the foundational notation used in the paper. The remainder of the paper is organized as follows: Section 5 presents asymptotic results, Section 6 details the algorithm, Section 7 covers simulations, and Section 8 applies the method to neuroimaging data on ADHD. Proofs and additional simulations are in the supplementary file.

2. Basics on Hilbert space

Let $(\Omega, \mathcal{F}, P)$ be a probability space, and $\mathcal{H}$ a separable Hilbert space with inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and norm $\|\cdot\|_{\mathcal{H}}$. If $U : \Omega \to \mathcal{H}$ is measurable with respect to the Borel σ-field $\mathcal{B}$, it is called a *random element* in $\mathcal{H}$. For this paper, $\mathcal{H}$ is a space of functions of time, so U is viewed as a random function. If $E\|U\|_{\mathcal{H}} < \infty$, the linear functional $\mathcal{H} \to \mathbb{R}$, $f \to E\langle f, U\rangle_{\mathcal{H}}$ is bounded. By Riesz's representation theorem, there exists a unique *mean element* $\mu_U = E(U) \in \mathcal{H}$, such that

$$\langle \mu_U, f\rangle_{\mathcal{H}} = E\langle f, U\rangle_{\mathcal{H}}, \text{ for all } f \in \mathcal{H}. \tag{2.1}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

If $E\|U\|_{\mathcal{H}}^2 < \infty$, the covariance operator of U is $\Sigma_{UU} = E[\{U-E(U)\}\otimes\{U-E(U)\}] = E(U\otimes U) - \{E(U)\}\otimes\{E(U)\}$, where $\otimes$ is the tensor product on $\mathcal{H}$: for $f, g, h \in \mathcal{H}$, $(f\otimes g)(h) = f\langle g, h\rangle_{\mathcal{H}}$. Under the assumption $E\|U\|_{\mathcal{H}}^2 < \infty$, Σ_{UU} is a trace-class operator that satisfies the spectral decomposition $\sum_{r=1}^{\infty}\delta_r\phi_r\otimes\phi_r$, where $\{\delta_r\}_{r\geq 1}$ are the eigenvalues satisfying $\delta_1 \geq \delta_2 \geq \cdots \geq 0$, and $\{\phi_r\}_{r\geq 1}$ are the eigenfunctions forming an orthonormal basis in $\mathcal{H}$. Then, the Karhunen–Loève expansion of $U - \mu_U$ is

$$U - \mu_U = \sum_{r=1}^{\infty}\delta_r^{1/2}\xi_r\phi_r, \tag{2.2}$$

where $\xi_r = \delta_r^{-1/2}\langle U-\mu_U, \phi_r\rangle_{\mathcal{H}}$, $r = 1, 2, \ldots$ are zero mean, unit variance, and uncorrelated random variables, called the scores (Bosq, 2000).

In this study, we consider a multivariate functional predictor and extend the above to vector-valued random functions. Let $\mathscr{H} = \bigoplus_{i=1}^{p}\mathcal{H}_i$ be the direct sum of $\mathcal{H}_1, \ldots, \mathcal{H}_p$, i.e., the Cartesian product $\mathcal{H}_1 \times \cdots \times \mathcal{H}_p$. A member $f \in \bigoplus_{i=1}^{p}\mathcal{H}_i$ is $f = (f_1, \ldots, f_p)$, where $f_i \in \mathcal{H}_i$. Then, for $f, g \in \bigoplus_{i=1}^{p}\mathcal{H}_i$, the inner product is $\langle f, g\rangle_{\oplus\mathcal{H}} = \sum_{i=1}^{p}\langle f_i, g_i\rangle_{\mathcal{H}_i}$. Note that, this additivity assumption aids the mathematical derivations and asymptotic analysis and is a common practice for multivariate functional predictors; see, for example, Li and Song (2017), Song and Li (2021), Mahzarnia and Song (2022), Solea, Christou and Song (2026).

Let $\mathbb{U} = (U^1, \ldots, U^p)$ be a random element in $\bigoplus_{i=1}^{p}\mathcal{H}_i$, where U^i, $i =$

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

$1, \ldots, p$, is the ith component of $\mathbb{U}$. The mean of $\mathbb{U}$ is $\mu_{\mathbb{U}} = (\mu_{U^1}, \ldots, \mu_{U^p})$, where μ_{U^i} is as defined in (2.1). Moreover, for $i, j = 1, \ldots, p$, the covariance operator between U^i and U^j is $\Sigma_{U^iU^j} = E\{(U^i - \mu_{U^i}) \otimes (U^j - \mu_{U^j})\}$. Note that, $\Sigma_{U^iU^j} \in \mathscr{B}(\mathcal{H}_j, \mathcal{H}_i)$, where $\mathscr{B}(\mathcal{H}_j, \mathcal{H}_i)$ denotes the set of all bounded operators from $\mathcal{H}_j$ to $\mathcal{H}_i$. Then, $\Sigma_{\mathbb{U}\mathbb{U}}$ is the $p \times p$ matrix whose (i, j)th entry is $\Sigma_{U^iU^j}$ and is linear, self-adjoint, positive semi-definite, and compact. Its decomposition is $\Sigma_{\mathbb{U}\mathbb{U}} = \sum_{r=1}^{\infty} \gamma_r \psi_r \otimes \psi_r$, where eigenvalues $\{\gamma_r\}_{r \geq 1}$ satisfying $\gamma_1 \geq \gamma_2 \geq \cdots \geq 0$, and eigenfunctions $\{\psi_r\}_{r \geq 1}$ form an orthonormal basis in $\bigoplus_{i=1}^{p} \mathcal{H}_i$. The Karhunen–Loève expansion of $\mathbb{U} - \mu_{\mathbb{U}}$ is

$$\mathbb{U} - \mu_{\mathbb{U}} = \sum_{r=1}^{\infty} \gamma_r^{1/2} \rho_r \psi_r, \tag{2.3}$$

where $\rho_r = \gamma_r^{-1/2} \langle \mathbb{U} - \mu_{\mathbb{U}}, \psi_r \rangle_{\bigoplus \mathcal{H}}$, $r = 1, 2, \ldots$, are zero mean, unit variance, and uncorrelated random variables (Happ and Greven, 2018).

3. Basics on Sufficient Dimension Reduction

3.1 Scalar-on-scalar

We start with the case where the response variable $Y : \Omega \to \mathbb{R}$ and the $p \times 1$ vector of predictors $\mathbf{X} : \Omega \to \mathbb{R}^p$ are scalars. Assume there exists a $p \times d$ matrix $\mathbf{A}$, $d \leq p$, such that $Y \perp\!\!\!\perp \mathbf{X} | \mathbf{A}^\top \mathbf{X}$, i.e., Y and $\mathbf{X}$ are independent given $\mathbf{A}^\top \mathbf{X}$. This allows replacing $\mathbf{X}$ with the reduced $d \times 1$ vector $\mathbf{A}^\top \mathbf{X}$

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

without loss of information on the regression. The space spanned by $\mathbf{A}$ is called the *dimension reduction subspace* for the regression of Y on $\mathbf{X}$; the smallest such subspace is called the *central subspace*. Techniques by Li (1991) (SIR) and Cook and Weisberg (1991) (SAVE) are commonly used.

For heteroscedastic data, focus is often on estimating conditional quantiles. Let $Q_\tau(Y|\mathbf{x}) = \inf\{y : P(Y \leq y|\mathbf{X} = \mathbf{x}) \geq \tau\}$ be the τth conditional quantile of Y given $\mathbf{X} = \mathbf{x}$, satisfying $Q_\tau(Y|\mathbf{x}) = \arg\min_q E\{\rho_\tau(Y-q)|\mathbf{X} = \mathbf{x}\}$, where $\rho_\tau(u) = \{\tau - I(u < 0)\}u$ is the loss function (check function). We seek a $p \times d_\tau$ matrix $\mathbf{B}_\tau$, $d_\tau \leq p$, such that $Y \perp\!\!\!\perp Q_\tau(Y|\mathbf{X})|\mathbf{B}_\tau^\top\mathbf{X}$. The space spanned by $\mathbf{B}_\tau$ is the *τth quantile dimension reduction subspace* for the regression of Y on $\mathbf{X}$, with the smallest being the *τth central quantile subspace* (τ-CQS). Methods for estimating the τ-CQS include Luo, Li and Yin (2014), Kong and Xia (2014), Christou (2020), Lee and Hilafu (2022).

3.2 Scalar-on-function

Continuing with X as a functional predictor, we introduce additional notation and terminology involving Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$. For a linear operator $A \in \mathscr{B}(\mathcal{H}, \mathcal{K})$, denote $\ker(A) = \{h : A(h) = 0\}$ the kernel of A, $\text{ran}(A) = \{A(h) : h \in \mathcal{H}\}$ the range of A, $\overline{\text{ran}}(A)$ the closure of $\text{ran}(A)$, and A^* the adjoint operator of A. For a self-adjoint operator $A \in \mathscr{B}(\mathcal{H})$,

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

let $A|_{\ker(A)^\perp} : \ker(A)^\perp \to \mathcal{H}$ be the restriction of A on $\ker(A)^\perp$. Then, the operator $A^\dagger : \mathrm{ran}(A) \to \ker(A)^\perp$ that maps each $g \in \mathrm{ran}(A)$ to the unique element $f \in \ker(A)^\perp$ such that $Af = g$ is called the Moore-Penrose inverse of A (Hsing and Eubank (2015), Definition 3.5.7). If, in addition, A is positive semi-definite, then, for any $\alpha > 0$, we define $A^{\dagger\alpha} = (A^\alpha)^\dagger$.

For $i = 1, \ldots, p$, let $\mathcal{H}_i$ be a separable Hilbert space of real-valued functions on T, a bounded closed interval in $\mathbb{R}$. Let $Y : \Omega \to \mathbb{R}$ be a univariate response and $X = (X^1, \ldots, X^p) : \Omega \to \bigoplus_{i=1}^p \mathcal{H}_i$ be a random element. Following Li and Song (2022)'s formulation, assume a finite rank linear operator $L \in \mathscr{B}(\bigoplus_{i=1}^p \mathcal{H}_i, \mathbb{R}^d)$, $d \in \mathbb{N}$, such that

$$Y \perp\!\!\!\perp X | L(X). \tag{3.1}$$

Then, $\overline{\mathrm{ran}}(L^*)$ is called the *functional dimension reduction subspace*.

Li and Song (2022) noted that (3.1) aligns with Ferré and Yao (2003) since, by Riesz's representation theorem, there exists $f_1, \ldots, f_d \in \bigoplus_{i=1}^p \mathcal{H}_i$ such that, $L(X) = (L_1(X), \ldots, L_d(X)) = (\langle f_1, X\rangle_{\oplus\mathcal{H}}, \ldots, \langle f_d, X\rangle_{\oplus\mathcal{H}})$, and, hence, for any $v \in \mathbb{R}^d$, $v^\top L(X) = \sum_{j=1}^d v_j \langle f_j, X\rangle_{\oplus\mathcal{H}} = \langle \sum_{j=1}^d v_j f_j, X\rangle_{\oplus\mathcal{H}}$. Then, $L^*(v) = \sum_{j=1}^d v_j f_j$ and $\overline{\mathrm{ran}}(L^*) = \mathrm{span}(f_1, \ldots, f_d)$.

Define $\cap\{\overline{\mathrm{ran}}(L^*) : L \text{ satisfies (3.1)}\}$, which we assume also satisfies (3.1). The assumption that this intersection is a functional dimension reduction subspace can be proved similarly to Proposition 6.4 of Cook (1998)

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

for the classical setting and Lee and Li (2022) for the functional data setting. The condition is mild and taken for granted without further development. We call this intersection the *functional central subspace* (FCS), denoted by $\mathscr{S}_{Y|X}$. We assume $\mathscr{S}_{Y|X}$ exists and satisfies (3.1), and that $L \in \mathscr{B}(\bigoplus_{i=1}^{p} \mathcal{H}_i, \mathbb{R}^d)$ denotes an operator such that $\mathscr{S}_{Y|X} = \overline{\text{ran}}(L^*)$.

Ferré and Yao (2003) proposed FSIR and proved that, under model (3.1) and Assumption 1 given below, $E(X|Y) - E(X)$ belongs to $\Sigma_{XX}\mathscr{S}_{Y|X}$, where Σ_{XX} is the covariance operator of X.

Assumption 1. There is a bounded linear operator $\Lambda : \text{ran}(L) \to \bigoplus_{i=1}^{p} \mathcal{H}_i$ such that $E\{X|L(X)\} = \Lambda L(X)$.

To our knowledge, no work focuses on supervised dimension reduction for conditional quantiles of functional predictors. This paper introduces the *τth functional central quantile subspace* (τ-FCQS) and its estimation.

4. The τth functional central quantile subspace

4.1 The Methodology

We formalize the setting of the paper as follows: For each $i = 1, \ldots, p$, $\mathcal{H}_i$ is a separable Hilbert space of real-valued functions on T, which is a bounded closed interval in $\mathbb{R}$. Let $Y : \Omega \to \mathbb{R}$ be a univariate response and

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

$X = (X^1, \ldots, X^p) : \Omega \to \bigoplus_{i=1}^p \mathcal{H}_i$ be a random element, such that

Assumption 2. $E\,\|X\|_{\mathscr{H}}^2 < \infty$,

where, from now on, $\mathscr{H}$ denotes $\bigoplus_{i=1}^p \mathcal{H}_i$, i.e., $\mathscr{H} = \bigoplus_{i=1}^p \mathcal{H}_i$.

Definition 1. For a finite rank linear operator $L_\tau \in \mathscr{B}(\mathscr{H}, \mathbb{R}^{d_\tau})$, $d_\tau \in \mathbb{N}$, such that

$$Y \perp\!\!\!\perp Q_\tau(Y|X)|L_\tau(X), \tag{4.1}$$

$\overline{\text{ran}}(L_\tau^*)$ is called the *τth functional quantile dimension reduction subspace*.

This formulation is equivalent to assuming that there exist functional parameters $\beta_{\tau,1}, \ldots, \beta_{\tau,d_\tau} \in \mathscr{H}$, such that $Y \perp\!\!\!\perp Q_\tau(Y|X)|\langle\beta_{\tau,1}, X\rangle_{\mathscr{H}}, \ldots, \langle\beta_{\tau,d_\tau}, X\rangle_{\mathscr{H}}$.

Definition 2. Under the assumption that $\cap\{\overline{\text{ran}}(L_\tau^*) : L_\tau \text{ satisfies (4.1)}\}$ satisfies (4.1), we call it the *τth functional central quantile subspace* (τ-FCQS) and denote it by $\mathscr{S}_{Q_\tau(Y|X)}$.

Remark 1. The assumption that the intersection of all τth functional quantile dimension reduction subspaces is itself a functional quantile dimension reduction subspace is mild and can be proven similarly to the functional dimension reduction subspace case. Henceforth, we assume that $\mathscr{S}_{Q_\tau(Y|X)}$ exists, satisfies (4.1), and that $L_\tau \in \mathscr{B}(\mathscr{H}, \mathbb{R}^{d_\tau})$ is an operator such that $\mathscr{S}_{Q_\tau(Y|X)} = \overline{\text{ran}}(L_\tau^*)$.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

Remark 2. It is evident that $\mathscr{S}_{Q_\tau(Y|X)} \subseteq \mathscr{S}_{Y|X}$, for any τ. Thus, the τ-FCQS can offer additional dimension reduction when these subspaces differ. For example, Model I from Section 7.1 is a case where $\mathscr{S}_{Q_\tau(Y|X)} = \mathscr{S}_{Y|X} = \text{span}\{\beta_1\}$, for all τ. However, in Model IV, $\mathscr{S}_{Q_\tau(Y|X)} = \text{span}\{\beta_1 + Q_\tau(\varepsilon)\beta_2\}$ for any τ, while $\mathscr{S}_{Y|X} = \text{span}\{\beta_1, \beta_2\}$; see Section 7.1 for details.

Assumption 1 is necessary for the FCS. However, when focusing on a specific conditional quantile, the assumption must hold for each quantile level τ under consideration. This leads to the following modification.

Assumption 3. For a given τ, there is a bounded linear operator $\Lambda_\tau : \text{ran}(L_\tau) \to \mathscr{H}$ such that $E\{X|L_\tau(X)\} = \Lambda_\tau L_\tau(X)$.

Assumption 3 implies that, for any function $b_\tau \in \mathscr{H}$, there exist constants $c_{\tau,0}, c_{\tau,1}, \ldots, c_{\tau,d_\tau}$, such that $E(\langle b_\tau, X\rangle_{\mathscr{H}} | \langle \beta_{\tau,1}, X\rangle_{\mathscr{H}}, \ldots, \langle \beta_{\tau,d_\tau}, X\rangle_{\mathscr{H}}) = c_{\tau,0} + c_{\tau,1}\langle \beta_{\tau,1}, X\rangle_{\mathscr{H}} + \cdots + c_{\tau,d_\tau}\langle \beta_{\tau,d_\tau}, X\rangle_{\mathscr{H}}$, and it is satisfied by elliptical distributions Eaton (1986). Ferré and Yao (2003) provided a discussion and stated that this assumption 'is similar to the one used in the finite-dimensional case, where it holds when the explanatory variables have a symmetric elliptical distribution.' The proof of how the elliptically distributed Hilbertian variables satisfy the condition is available upon request from Ferré and Yao (2003).

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

The following theorem is analogous to a known result in the classical setting (see, e.g. Cook (1998), page 57) and is necessary to ensure that the functional parameters derived in Theorems 2 and 3 belong to $\mathscr{S}_{Q_\tau(Y|X)}$. Essentially, Theorem 1 states that $E\{X|L_\tau(X)\}$ is equal to the projection of X onto the subspace spanned by $L_\tau(X)$.

Theorem 1. *Under Assumptions 2, 3, and* $E(X) = 0$, *we have* $E\{X|L_\tau(X)\} = \Sigma_{XX}L_\tau^*(L_\tau\Sigma_{XX}L_\tau^*)^\dagger L_\tau(X)$.

We now focus on retrieving $\beta_{\tau,1}, \ldots, \beta_{\tau,d_\tau}$, such that $\mathscr{S}_{Q_\tau(Y|X)} = \text{span}\{\beta_{\tau,1}, \ldots, \beta_{\tau,d_\tau}\}$. Our methodology relies on two main theorems. Theorem 3 shows that given one functional parameter b_τ in $\mathscr{S}_{Q_\tau(Y|X)}$, we can construct another one using $E(\langle b_\tau, X\rangle_{\mathscr{H}} X)$. Since the τ-FCQS is spanned by d_τ functional parameters, we can use Theorem 3 to construct as many as needed. However, Theorem 3 requires an initial b_τ to start the iterative process. Theorem 2 provides a starting point by showing that one functional parameter can be extracted through minimizing $\arg\min_{(a_\tau,b_\tau)} E\{Q_\tau(Y|X) - a_\tau - \langle b_\tau, X\rangle_{\mathscr{H}}\}^2$, for $a_\tau \in \mathbb{R}$ and $b_\tau \in \mathscr{H}$. Note that, in practice, this minimization problem requires estimating $Q_\tau(Y|X)$, which is challenging due to the infinite-dimensional nature of X. Therefore, since $Y \perp\!\!\!\perp X|L(X)$, we perform an initial dimension reduction by replacing X with $L(X)$.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

Theorem 2. *For a given $\tau \in (0,1)$, assume that $Y \perp\!\!\!\perp Q_\tau(Y|X)|L_\tau(X)$, where $L_\tau \in \mathscr{B}(\mathscr{H}, \mathbb{R}^{d_\tau})$ is such that $\mathscr{S}_{Q_\tau(Y|X)} = \overline{\text{ran}}(L_\tau^*)$. Under Assumptions 2 and 3, and if*

$$(\alpha_\tau^*, \beta_\tau^*) = \arg\min_{(a_\tau, b_\tau)} E[Q_\tau\{Y|L(X)\} - a_\tau - \langle b_\tau, X\rangle_{\mathscr{H}}]^2, \tag{4.2}$$

where $L \in \mathscr{B}(\mathscr{H}, \mathbb{R}^d)$ is such that $\mathscr{S}_{Y|X} = \overline{\text{ran}}(L^*)$, then $\beta_\tau^* \in \mathscr{S}_{Q_\tau(Y|X)}$.

Theorem 2 retrieves one functional parameter such that $Y \perp\!\!\!\perp Q_\tau(Y|X)|\langle\beta_\tau^*, X\rangle_{\mathscr{H}}$. If a single-index functional QR model is assumed ($d_\tau = 1$), then $\beta_\tau^* = \beta_{\tau,1}$, and the process ends. However, if $d_\tau > 1$, Theorem 2 is insufficient, and more functional parameters are required, provided by the next theorem. We note that the idea for constructing these additional parameters originates from Cook and Li (2002); see their Theorem 3.

Theorem 3. *For a given $\tau \in (0,1)$, assume that $Y \perp\!\!\!\perp Q_\tau(Y|X)|L_\tau(X)$, where $L_\tau \in \mathscr{B}(\mathscr{H}, \mathbb{R}^{d_\tau})$ is such that $\mathscr{S}_{Q_\tau(Y|X)} = \overline{\text{ran}}(L_\tau^*)$. If Assumptions 2 and 3 hold, and $b_\tau \in \mathscr{S}_{Q_\tau(Y|X)}$, then $E(\langle b_\tau, X\rangle_{\mathscr{H}} X) \in \Sigma_{XX}\mathscr{S}_{Q_\tau(Y|X)}$.*

4.2 Population Level

Theorem 3 suggests a method for constructing the τ-FCQS. Specifically, if $b_{\tau,0} \in \mathscr{S}_{Q_\tau(Y|X)}$, then, for $j = 1, \ldots,$ we can construct additional functional

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

parameters in $\mathscr{S}_{Q_\tau(Y|X)}$ using

$$E(\langle b_{\tau,j-1}, X\rangle_{\mathscr{H}} X). \tag{4.3}$$

Hence, it is enough to find an initial $b_{\tau,0}$, and we can use $b_{\tau,0} = \beta^*_\tau$. Then:

1. Set $b_{\tau,0} = \beta^*_\tau$, where β^*_τ is given in (4.2).

2. Choose an integer m and, for $j = 1, \ldots, m$, sequentially form $b_{\tau,j} = E(\langle b_{\tau,j-1}, X\rangle_{\mathscr{H}} X)$.

3. Let $B_\tau = \sum_{j=0}^{m} b_{\tau,j} \otimes b_{\tau,j}$.

4. For $v_{\tau,1}, \ldots, v_{\tau,d_\tau}$ the set of solutions to the generalized eigenvalue problem $B_\tau v_{\tau,j} = \lambda \Sigma_{XX} v_{\tau,j}$, $j = 1, \ldots, d_\tau$, $v_{\tau,j} \in \mathscr{S}_{Q_\tau(Y|X)}$. This is because, for $j = 1, \ldots, d_\tau$, $v_{\tau,j} = \lambda^{-1} \Sigma^\dagger_{XX} B_\tau v_{\tau,j} \in \Sigma^\dagger_{XX} \text{span}\{b_{\tau,0}, \ldots, b_{\tau,m}\} \subseteq \mathscr{S}_{Q_\tau(Y|X)}$.

Note that, an eigenvector v of a generalized eigenvalue problem $Av = \lambda Bv$ implies that $v = B^{-1/2}u$, where u is an eigenvector of $B^{-1/2}AB^{-1/2}$. Thus, the eigenvectors $v_{\tau,j}$, $j = 1, \ldots, d_\tau$, of Step 4 that satisfy

$$\begin{aligned} &\text{argmax} \quad \langle v_\tau, B_\tau v_\tau\rangle_{\mathscr{H}} \\ &\text{subject to} \quad v_\tau \in \mathscr{H}, \langle v_\tau, \Sigma_{XX} v_\tau\rangle_{\mathscr{H}} = 1, \langle v_\tau, \Sigma_{XX} v_{\tau,j}\rangle_{\mathscr{H}} = 0, j = 1, \ldots, d_\tau - 1, \end{aligned}$$

can be expressed as $v_{\tau,j} = \Sigma^{\dagger 1/2}_{XX} \eta_{\tau,j}$, leading to the following problem

$$\begin{aligned} &\text{argmax} \quad \langle \eta_\tau, \Sigma^{\dagger 1/2}_{XX} B_\tau \Sigma^{\dagger 1/2}_{XX} \eta_\tau\rangle_{\mathscr{H}} \\ &\text{subject to} \quad \eta_\tau \in \mathscr{H}, \langle \eta_\tau, \eta_\tau\rangle_{\mathscr{H}} = 1, \langle \eta_\tau, \eta_{\tau,j}\rangle_{\mathscr{H}} = 0, j = 1, \ldots, d_\tau - 1. \end{aligned} \tag{4.4}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

Remark 3. Step 2 of the algorithm requires an integer m. Simulation studies show that the algorithm is robust across different values of m; see Supplementary S6. For this study, we chose $m = pq-1$, where p is the number of predictors and q is the number of basis functions used to approximate the functional predictors.

Remark 4. In this paper, we assume the structural dimension d_τ is known. However, in practice, it is unknown and must be estimated. One approach is the cross-validation Bayesian information criterion (CVBIC) from Li and Song (2017); details and simulations are provided in Supplementary S6.

4.3 Sample Level

We now derive the sample estimates of the expressions (4.2) and (4.3) when the functions are fully observed. Specifically, for $u = 1, \ldots, n$, let Y_u be an independent and identically distributed (iid) sample from Y, and $X_1, \ldots, X_n$ be an independent sample from the random element $X = (X^1, \ldots, X^p)$, with $X_u = (X_u^1, \ldots, X_u^p)^\top$.

To achieve (4.2), we apply the FSIR method from Ferré and Yao (2003) and replace X with the new d-dimensional predictor vector $\widehat{L}(X)$; see Ferré and Yao (2003) for details. Then, we use the data $\{Y_u, X_u\}_{u=1}^n$ to obtain

$$(\widehat{\alpha}_\tau, \widehat{\beta}_\tau) = \arg\min_{(a_\tau, b_\tau)} \sum_{u=1}^{n} [\widehat{Q}_\tau\{Y|\widehat{L}(X_u)\} - a_\tau - \langle b_\tau, X_u\rangle_{\mathscr{H}}]^2, \tag{4.5}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

where $\widehat{Q}_\tau\{Y|\widehat{L}(X_u)\}$ is a nonparametric estimate of $Q_\tau\{Y|\widehat{L}(X_u)\}$. For that we use the local linear conditional quantile estimation method of Guerre and Sabbah (2012), where $\widehat{Q}_\tau\{Y|\widehat{L}(X_u)\} = \widehat{q}_\tau\{\widehat{L}(X_u)\}$ and

$$\big(\widehat{q}_\tau\{\widehat{L}(X_u)\},\ \widehat{\mathbf{s}}_\tau\{\widehat{L}(X_u)\}\big) = \arg\min_{(q_\tau,\mathbf{s}_\tau)} \sum_{k=1}^{n} \rho_\tau\Big[Y_k - q_\tau - \mathbf{s}_\tau^\top\{\widehat{L}(X_k) - \widehat{L}(X_u)\}\Big] \times K\left(\frac{\widehat{L}(X_k) - \widehat{L}(X_u)}{h}\right). \tag{4.6}$$

Here, $K(\cdot)$ is a d-dimensional kernel function, and $h > 0$ is the bandwidth. We use a Gaussian kernel and select h based on the rule-of-thumb in Yu and Jones (1998). Specifically, $h = h_m[\tau(1-\tau)/[\phi\{\Phi^{-1}(\tau)\}]^2]^{1/5}$, where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and cumulative distribution functions, and h_m is the optimal bandwidth for local mean regression.

Next, to achieve (4.3), we set $\widehat{\beta}_{\tau,0} = \widehat{\beta}_\tau$, and, for $j = 1, \ldots, m$, we form

$$\widehat{\beta}_{\tau,j} = n^{-1}\sum_{u=1}^{n}\langle\widehat{\beta}_{\tau,j-1}, X_u\rangle_{\mathscr{H}} X_u. \tag{4.7}$$

After obtaining $\widehat{\beta}_{\tau,0}, \widehat{\beta}_{\tau,1}, \ldots, \widehat{\beta}_{\tau,m}$, set $\widehat{B}_\tau = \sum_{j=0}^{m}\widehat{\beta}_{\tau,j}\otimes\widehat{\beta}_{\tau,j}$ and solve problem (4.4), which implies solving $\Sigma_{XX}^{\dagger 1/2} B_\tau \Sigma_{XX}^{\dagger 1/2}\eta_\tau = \lambda\eta_\tau$ and $v_\tau = \Sigma_{XX}^{\dagger 1/2}\eta_\tau$.

At the sample level, we estimate Σ_{XX} using the $p\times p$ matrix $\widehat{\Sigma}_{XX}$, where $\widehat{\Sigma}_{X^iX^j} = E_n[\{X^i - E_n(X^i)\}\otimes\{X^j - E_n(X^j)\}]$, $i, j = 1, \ldots, p$, and $\Sigma_{XX}^{\dagger 1/2}$ using the regularized inverse $(\widehat{\Sigma}_{XX} + \epsilon_n I_p)^{-1/2}$, where $\{\epsilon_n\}_{n\geq 1}$ is a sequence of positive numbers approaching zero as $n\to\infty$, and I_p is the $p\times p$ identity

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

matrix. Thus, we aim to find the first d_τ eigenfunctions $\widehat{\eta}_{\tau,1}, \ldots, \widehat{\eta}_{\tau,d_\tau}$ of

$$\widehat{M}_\tau = (\widehat{\Sigma}_{XX} + \epsilon_n I)^{-1/2} \widehat{B}_\tau (\widehat{\Sigma}_{XX} + \epsilon_n I)^{-1/2} \tag{4.8}$$

and transform back to the eigenfunctions $\widehat{v}_{\tau,j} = (\widehat{\Sigma}_{XX} + \epsilon_n I)^{-1/2} \widehat{\eta}_{\tau,j}$ and the sufficient predictors $\langle \widehat{v}_{\tau,j}, X \rangle_{\mathscr{H}}$, $j = 1, \ldots, d_\tau$.

5. Asymptotic Theory

To derive the consistency and convergence rate of $\widehat{M}_\tau$, where the population counterpart is $M_\tau = \Sigma_{XX}^{\dagger 1/2} B_\tau \Sigma_{XX}^{\dagger 1/2}$, we introduce a commonly used assumption, as seen in Assumption 7 of Li and Song (2022).

Assumption 4. $\mathscr{S}_{Q_\tau(Y|X)} \subseteq \overline{\text{ran}}(\Sigma_{XX})$.

This assumption is not restrictive and suggests that the τ-FCQS is identifiable up to the range of Σ_{XX}. This is because if $\beta_\tau \perp \overline{\text{ran}}(\Sigma_{XX})$, then $Var(\langle \beta_\tau, X \rangle_{\mathscr{H}}) = 0$, implying β_τ is orthogonal to the support of $X - E(X)$.

Theorem 4. *Let Assumptions 2, 3, 4, and S1-S6 from the Supplementary hold. Then, if $E(X) = 0$, $\widehat{L}(X)$ is consistent of the directions of the FCS, and $n^{-1/4} \prec \epsilon_n \prec 1$, for a given $\tau \in (0, 1)$, $\widehat{M}_\tau$, given in (4.8), is a consistent estimate of M_τ, and*

$$\left\| \widehat{M}_\tau - M_\tau \right\| = O_p(n^{-1/2} \epsilon_n^{-2}) + O(\epsilon_n^{1/2}),$$

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

where $M_\tau = \Sigma_{XX}^{\dagger 1/2} B_\tau \Sigma_{XX}^{\dagger 1/2}$, *and* $\|\cdot\|$ *is the operator norm.*

Next, the first d_τ eigenvalues and eigenfunctions of $\widehat{M}_\tau$ converge to those of M_τ at the same rate, as in Corollary 2 of Li and Song (2022).

Corollary 1. *Under the assumptions of Theorem 4, and for a given* $\tau \in (0,1)$ *and* $j = 1, \ldots, d_\tau$,

$$\widehat{\lambda}_j - \lambda_j = O_p(n^{-1/2}\epsilon_n^{-2}) + O(\epsilon_n^{1/2}), \quad \text{and} \quad \|\widehat{\eta}_{\tau,j} - \eta_{\tau,j}\|_{\mathscr{H}} = O_p(n^{-1/2}\epsilon_n^{-2}) + O(\epsilon_n^{1/2}),$$

where $\widehat{\lambda}_j$ *and* $\widehat{\eta}_{\tau,j}$ *are the eigenvalues and eigenfunctions of* $\widehat{M}_\tau$, *respectively, and* λ_j *and* $\eta_{\tau,j}$ *are the eigenvalues and eigenfunctions of* M_τ, *respectively.*

The eigenfunctions of interest are $\widehat{v}_{\tau,j} = (\widehat{\Sigma}_{XX} + \epsilon_n I)^{-1/2}\widehat{\eta}_{\tau,j}$, and the next theorem gives the convergence rates of $\widehat{v}_{\tau,j}$ and of the predictors $\langle \widehat{v}_{\tau,j}, X\rangle_{\mathscr{H}}$, $j = 1, \ldots, d_\tau$.

Theorem 5. *Let Assumptions 2, 3, 4, and S1-S6 from the Supplementary hold. Then, if* $E(X) = 0$, $\widehat{L}(X)$ *is consistent of the directions of the FCS, and* $n^{-1/5} \prec \epsilon_n \prec 1$, *for a given* $\tau \in (0,1)$ *and for* $j = 1, \ldots, d_\tau$,

$$\|\widehat{v}_{\tau,j} - v_{\tau,j}\|_{\mathscr{H}} = O_p(n^{-1/2}\epsilon_n^{-5/2}) + O(\epsilon_n^{1/2}),$$
$$\langle \widehat{v}_{\tau,j}, X\rangle_{\mathscr{H}} - \langle v_{\tau,j}, X\rangle_{\mathscr{H}} = O_p(n^{-1/2}\epsilon_n^{-5/2}) + O(\epsilon_n^{1/2}),$$

where $\widehat{v}_{\tau,j} = (\widehat{\Sigma}_{XX} + \epsilon_n I)^{-1/2}\widehat{\eta}_{\tau,j}$ *and* $v_{\tau,j} = \Sigma_{XX}^{\dagger 1/2}\eta_{\tau,j}$.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

6. Implementation

In practice, the functions $X_u(t)$, $u = 1, \ldots, n$, are observed at a finite set of points, $t_{u1}, \ldots, t_{uN_u}$, and need to be estimated using the observed data $\{(t, X_u(t)) : t = t_{u1}, \ldots, t_{uN_u}\}$. Common methods for estimating $X_u(t)$ use coordinate representation with a finite number of basis functions, converting a function to a vector, a linear operator to a matrix, and an eigenfunction problem to an eigenvector problem. Preliminaries on coordinate representation are provided in Supplementary S5. We omit indices like ${}_{\mathcal{G}_j}[\cdot]_{\mathcal{G}_i}$ and $[\cdot]_{\mathcal{G}_i}$, as the bases for the coordinates can be identified by the operators' domain and range. In this section, square brackets are reserved for coordinates.

For $i = 1, \ldots, p$, assume that $\mathcal{H}_i$ is spanned by a finite set of basis functions $\mathcal{G}_i = \{g^i_1, \ldots, g^i_{k_n}\}$, such that each X^i_u, $u = 1, \ldots, n$, can be approximated by $[X^i_u]^\top g^i_{1:k_n}$. Although $\widehat{X}^i_u$ could be used for clarity, we omit this distinction for simplicity. Let $Q_n = I_n - n^{-1}1_n 1_n^\top$ be the projection onto the orthogonal complement of the subspace spanned by 1_n, where I_n is the $n \times n$ identity matrix and 1_n is the n-dimensional vector of ones, and let $G = \mathrm{diag}(G_i : i = 1, \ldots, p) \in \mathbb{R}^{pk_n \times pk_n}$, where G_i is the Gram matrix of $\mathcal{H}_i$. Then, $[X_u] = ([X^1_u]^\top, \ldots, [X^p_u]^\top)^\top \in \mathbb{R}^{pk_n}$ is the coordinate representation of X_u, $u = 1, \ldots, n$, i.e., the ith block is the coordinate representation of X^i_u. Therefore, the matrix $[X_{1:n}]$ is the $pk_n \times n$ matrix $([X_1], \ldots, [X_n])$.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

The coordinates of (4.5) and (4.7) with respect to $\mathcal{G} = \bigoplus_{i=1}^{p} \mathcal{G}_i$ are

$$(\widehat{\alpha}_\tau, [\widehat{\beta}_\tau]) = \arg\min_{(a_\tau, b_\tau)} \sum_{u=1}^{n} \left[\widehat{Q}_\tau\{Y|\widehat{L}(X_u)\} - a_\tau - [b_\tau]^\top G[X_u] \right]^2, \tag{6.1}$$

where $\widehat{Q}_\tau\{Y|\widehat{L}(X_u)\} = \widehat{q}_\tau\{\widehat{L}(X_u)\}$ for

$$(\widehat{q}_\tau\{\widehat{L}(X_u)\}, \widehat{\mathbf{s}}_\tau\{\widehat{L}(X_u)\}) = \arg\min_{(q_\tau, \mathbf{s}_\tau)} \sum_{k=1}^{n} \rho_\tau \Big[Y_k - q_\tau - \mathbf{s}_\tau^\top \{\widehat{L}(X_k) - \widehat{L}(X_u)\} \Big] \times K\left(\frac{\widehat{L}(X_k) - \widehat{L}(X_u)}{h} \right), \tag{6.2}$$

and, for $j = 1, \ldots, m$,

$$[\widehat{\beta}_{\tau,j}] = n^{-1} \sum_{u=1}^{n} [\widehat{\beta}_{\tau,j-1}]^\top G[X_u][X_u], \tag{6.3}$$

where $\widehat{\beta}_{\tau,0} = \widehat{\beta}_\tau$ from (6.1). Then, for $\widehat{B}_\tau = \sum_{j=0}^{m} \widehat{\beta}_{\tau,j} \otimes \widehat{\beta}_{\tau,j}$, we have $[\widehat{B}_\tau] = \sum_{j=0}^{m} [\widehat{\beta}_{\tau,j} \otimes \widehat{\beta}_{\tau,j}] = \sum_{j=0}^{m} [\widehat{\beta}_{\tau,j}][\widehat{\beta}_{\tau,j}]^\top G = [\widehat{\beta}_{\tau,0:m}][\widehat{\beta}_{\tau,0:m}]^\top G$ and the eigenvalue problem (4.4) becomes

$$\begin{aligned} &\text{argmax} \quad [\eta_\tau]^\top G[\widehat{\Sigma}_{XX}^{\dagger 1/2}][\widehat{\beta}_{\tau,0:m}][\widehat{\beta}_{\tau,0:m}]^\top G[\widehat{\Sigma}_{XX}^{\dagger 1/2}][\eta_\tau] \\ &\text{subject to} \quad [\eta_\tau]^\top G[\eta_\tau] = 1, [\eta_\tau]^\top G[\eta_{\tau,j}] = 0, j = 1, \ldots, d_{\tau-1}. \end{aligned}$$

Let $\omega_\tau = G^{1/2}[\eta_\tau]$, then we get the eigenvector problem

$$\begin{aligned} &\text{argmax} \quad \omega_\tau^\top G^{1/2}[\widehat{\Sigma}_{XX}^{\dagger 1/2}][\widehat{\beta}_{\tau,0:m}][\widehat{\beta}_{\tau,0:m}]^\top G[\widehat{\Sigma}_{XX}^{\dagger 1/2}] G^{\dagger 1/2} \omega_\tau \\ &\text{subject to} \quad \omega_\tau^\top \omega_\tau = 1, \omega_\tau^\top \omega_{\tau,j} = 0, j = 1, \ldots, d_{\tau-1}. \end{aligned}$$

Using $[\widehat{\Sigma}_{XX}^{\dagger 1/2}] = G^{\dagger 1/2}\{n^{-1} G^{1/2}([X_{1:n}] Q_n [X_{1:n}]^\top) G^{1/2}\}^{\dagger 1/2} G^{1/2}$, proven in Supplementary S4.6, and $A = \{n^{-1} G^{1/2}([X_{1:n}] Q_n [X_{1:n}]^\top) G^{1/2}\}^{\dagger 1/2}$, we get

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

$$\begin{aligned} &\text{argmax} \quad \omega_\tau^\top A G^{1/2} [\widehat{\beta}_{\tau,0:m}] [\widehat{\beta}_{\tau,0:m}]^\top G^{1/2} A \omega_\tau \\ &\text{subject to} \quad \omega_\tau^\top \omega_\tau = 1, \omega_\tau^\top \omega_{\tau,j} = 0, j = 1, \ldots, d_{\tau-1}. \end{aligned} \tag{6.4}$$

Then, $[\widehat{v}_\tau] = [\widehat{\Sigma}_{XX}^{\dagger 1/2}] G^{\dagger 1/2} \omega_\tau = G^{\dagger 1/2} A \omega_\tau$ and $\langle X_u, \widehat{v}_\tau \rangle_{\mathscr{H}} = [X_u]^\top G [\widehat{v}_\tau] = [X_u]^\top G^{1/2} A \omega_\tau$. See Algorithm 1 for the details.

7. Simulation studies

7.1 Computational Remarks

Algorithm and parameters. Step 1 computes the coordinates of X_u using a B-spline basis with $q = 4$. Step 2 applies FSIR (Ferré and Yao, 2003) with $H = 10$ slices. Step 3 performs local linear conditional quantile estimation, selecting the kernel and bandwidth as in Section 4.3. Finally, Step 5 generates additional directions with $m = pq - 1$, as in Remark 3.

Simulation Setting. The simulation setting is an extension of Wang, Liu, Han and Di (2023) to multivariate data. First, create 101 equally spaced time points in $[0, 1]$, separated by 0.01. Then, for $i = 1, \ldots, p$ and $u = 1, \ldots, n$, generate $X_u^i(t)$ using the Karhunen–Loève expansion (2.2), where $\mu_{X^i} = 0$ and $X_u^i(t) = \sum_{r=1}^4 \xi_{ur}^i \phi_r^i(t)$, $t \in [0, 1]$, with $\phi_1^i(t) = \sqrt{2}\sin(2\pi t)$, $\phi_2^i(t) = \sqrt{2}\cos(2\pi t)$, $\phi_3^i(t) = \sqrt{2}\sin(4\pi t)$, $\phi_4^i(t) = \sqrt{2}\cos(4\pi t)$, and ξ_{ur}^i independent random variables with zero mean and variance $var(\xi_{ur}^i) = \lambda_r$,

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

Algorithm 1

1: Let $\{Y_u, (t, X_u(t)) : t = t_{u1}, \ldots, t_{uN_u}\}_{u=1}^n$ iid sample.

2: For each $u = 1, \ldots, n$, obtain the coordinates of X_u relative to the basis $\mathcal{G} = \bigoplus_{i=1}^p \mathcal{G}_i$ of $\mathscr{H}$ and derive the gram matrix G.

3: Use FSIR of Ferré and Yao (2003) to compute the d-dimensional predictors $\widehat{L}(X_u)$.

4: For each $u = 1, \ldots, n$, estimate $Q_\tau\{Y|\widehat{L}(X_u)\}$ using the local linear conditional quantile estimation method of Guerre and Sabbah (2012). That is, take $\widehat{Q}_\tau\{Y|\widehat{L}(X_u)\} = \widehat{q}_\tau\{\widehat{L}(X_u)\}$, where $\widehat{q}_\tau\{\widehat{L}(X_u)\}$ is in (6.2).

5: Compute $[\widehat{\beta}_\tau]$ according to (6.1) and set $[\widehat{\beta}_{\tau,0}] = [\widehat{\beta}_\tau]$.

6: For $j = 1, \ldots, m$, compute $[\widehat{\beta}_{\tau,j}]$ according to (6.3).

7: Compute the matrix $A = \{n^{-1}G^{1/2}([X_{1:n}]Q_n[X_{1:n}]^\top)G^{1/2}\}^{\dagger 1/2}$.

8: Solve the eigenvalue problem (6.4) and obtain $\omega_{\tau,j}$, $j = 1, \ldots, d_\tau$.

9: Obtain $[\widehat{v}_{\tau,j}] = G^{\dagger 1/2}A\omega_{\tau,j}$ and the sufficient predictors $\langle X_u, \widehat{v}_{\tau,j}\rangle_{\mathscr{H}} = [X_u]^\top G^{1/2}A\omega_{\tau,j}$, $j = 1, \ldots, d_\tau$, $u = 1, \ldots, n$.

$r = 1, \ldots, 4$, with $\lambda_1 = 2, \lambda_2 = 1, \lambda_3 = 1/2, \lambda_4 = 1/4$.

To simulate multivariate functional data $X_u(t) = (X_u^1(t), \ldots, X_u^p(t))$ using the Karhunen–Loève expansion (2.3), we follow Proposition 5 of Happ and Greven (2018) and use the multivariate FPCA eigenfunctions through an orthogonalization of the univariate eigenfunctions. Specifi-

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

cally, for $i = 1, \ldots, p$ and $u = 1, \ldots, n$, let $\boldsymbol{\xi}_u^i = (\xi_{u1}^i, \ldots, \xi_{u4}^i)^\top$ and $\boldsymbol{\xi}_u = ((\boldsymbol{\xi}_u^1)^\top, \ldots, (\boldsymbol{\xi}_u^p)^\top)^\top$. Moreover, define $\mathbb{Z} \in \mathbb{R}^{4p \times 4p}$ to be the covariance matrix of the univariate FPCA scores $\boldsymbol{\xi}_u$ with (j, j')th entry the matrix $\mathbb{Z}^{jj'} = cov(\boldsymbol{\xi}_u^j, \boldsymbol{\xi}_u^{j'}) \in \mathbb{R}^{4\times 4}$. Then, the kth eigenfunction $\boldsymbol{\psi}_k(t) = (\psi_k^1(t), \ldots, \psi_k^p(t))^\top$ of Σ_{XX} is defined by $\psi_k^i(t) = \boldsymbol{\phi}^i(t)^\top \boldsymbol{z}_k^i$, $k = 1, \ldots, 4p$, where $\boldsymbol{\phi}^i(t) = (\phi_1^i(t), \ldots, \phi_4^i(t))^\top$ and $\boldsymbol{z}_k^i = (z_{k1}^i, \ldots, z_{k4}^i)^\top$ denotes the ith block of the eigenvector $\boldsymbol{z}_k$ of $\mathbb{Z}$. Finally, the scores are $\rho_{uk} = \sum_{i=1}^p \sum_{r=1}^4 z_{kr}^i \xi_{ur}^i$, $k = 1, \ldots, 4p$ and $u = 1, \ldots, n$, where the coordinate-wise scores ξ_{ur}^i are standard normal random variables.

Estimation accuracy is measured using the multiple correlation between true and estimated predictors (Li and Song, 2022): $mcorr(U, V) = tr(C_{VV}^{-1/2} C_{VU} C_{UU}^{-1} C_{UV} C_{VV}^{-1/2})$, where U and V are random vectors of dimension d. It ranges from 0 to d, with values near d indicating better performance. Unless stated otherwise, simulations use $N = 100$ iterations, $n = 400$, $p = 5$, $\tau = 0.1, 0.25, 0.5, 0.75, 0.9$, and assume d_τ to be known.

Models under consideration. We consider the following models.

$$\text{M-I:} \quad Y = \frac{1}{0.5 + (\langle \beta_1, X \rangle_{\mathscr{H}} + 1)^2} + 0.2\varepsilon,$$

$$\text{M-II:} \quad Y = \arctan(\pi \langle \beta_1, X \rangle_{\mathscr{H}} / 2) + \varepsilon,$$

$$\text{M-III:} \quad Y = \exp(\langle \beta_1, X \rangle_{\mathscr{H}}) \varepsilon,$$

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

$$\text{M-IV:} \quad Y = \langle \beta_1, X \rangle_{\mathscr{H}} + \langle \beta_2, X \rangle_{\mathscr{H}} \varepsilon,$$

$$\text{M-V:} \quad Y = \arctan(\pi \langle \beta_1, X \rangle_{\mathscr{H}}) + 0.5 \sin(\pi \langle \beta_2, X \rangle_{\mathscr{H}} / 6) + 0.1\varepsilon,$$

$$\text{M-VI:} \quad Y = \langle \beta_1, X \rangle_{\mathscr{H}} + \langle \beta_2, X \rangle_{\mathscr{H}} + \sqrt{0.5 + \langle \beta_1, X \rangle_{\mathscr{H}}^2 + \langle \beta_2, X \rangle_{\mathscr{H}}^2}\varepsilon,$$

$$\text{M-VII:} \quad Y = \langle \beta_1, X \rangle_{\mathscr{H}}^3 + \exp(\langle \beta_2, X \rangle_{\mathscr{H}}) + \langle \beta_3, X \rangle_{\mathscr{H}} \varepsilon,$$

$$\text{M-VIII:} \quad Y = \langle \beta_1, X \rangle_{\mathscr{H}}^3 + \langle \beta_2, X \rangle_{\mathscr{H}} + \langle \beta_3, X \rangle_{\mathscr{H}} \varepsilon,$$

where $\beta_1(t) = \psi_1(t)$, $\beta_2(t) = \psi_2(t)$, and $\beta_3(t) = \psi_3(t)$ are the first three eigenfunctions of Σ_{XX}, X is simulated as described above, and the error ε is generated from standard normal ($\mathcal{N}$), chi-square with three degrees of freedom ($\mathcal{X}_3^2$), and exponential with rate of 0.5 (Exp(0.5)) distributions. Note that $\mathscr{S}_{Q_\tau(Y|X)} = \text{span}\{\beta_1\}$ for Models I-III, $\mathscr{S}_{Q_\tau(Y|X)} = \text{span}\{\beta_1, \beta_2\}$ for Models V-VI, $\mathscr{S}_{Q_\tau(Y|X)} = \text{span}\{\beta_1, \beta_2, \beta_3\}$ for Models VII, $\mathscr{S}_{Q_\tau(Y|X)} = \text{span}\{\beta_1 + Q_\tau(\varepsilon)\beta_2\}$ for Model IV, and $\mathscr{S}_{Q_\tau(Y|X)} = \text{span}\{\beta_1, \beta_2 + Q_\tau(\varepsilon)\beta_3\}$ for Model VIII. Results for all models are available, but we report selective ones for brevity and since they exhibit similar patterns.

7.2 Results

Example 1 - Effect of n and p. We evaluate the algorithm for sample sizes $n = 200, 400, 1000$, and number of predictors $p = 5, 10, 20, 40$, primarily focusing on Model I. Table 1 reports the means and standard deviations

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

Table 1: *Mean (and standard deviation) of multiple correlation for Model I, when the error follows a standard normal distribution.*

n	p	0.1	0.25	0.5	0.75	0.9
200	5	0.9957 (0.0039)	0.9956 (0.0042)	0.9957 (0.0040)	0.9957 (0.0039)	0.9957 (0.0038)
	10	0.9912 (0.0105)	0.9911 (0.0108)	0.9913 (0.0106)	0.9913 (0.0105)	0.9910 (0.0105)
	20	0.9848 (0.0542)	0.9849 (0.0565)	0.9911 (0.0146)	0.9907 (0.0173)	0.9908 (0.0155)
	40	0.9692 (0.0827)	0.9680 (0.0825)	0.9722 (0.0657)	0.9726 (0.0698)	0.9538 (0.1210)
400	5	0.9964 (0.0045)	0.9965 (0.0039)	0.9965 (0.0035)	0.9966 (0.0034)	0.9966 (0.0032)
	10	0.9948 (0.0048)	0.9948 (0.0047)	0.9949 (0.0047)	0.9949 (0.0046)	0.9949 (0.0047)
	20	0.9926 (0.0055)	0.9927 (0.0053)	0.9926 (0.0054)	0.9926 (0.0054)	0.9927 (0.0054)
	40	0.9898 (0.0092)	0.9897 (0.0095)	0.9892 (0.0137)	0.9898 (0.0092)	0.9897 (0.0097)
1000	5	0.9974 (0.0018)	0.9974 (0.0017)	0.9974 (0.0017)	0.9974 (0.0017)	0.9974 (0.0016)
	10	0.9966 (0.0027)	0.9966 (0.0027)	0.9966 (0.0026)	0.9966 (0.0026)	0.9966 (0.0026)
	20	0.9952 (0.0035)	0.9950 (0.0035)	0.9951 (0.0035)	0.9950 (0.0034)	0.9950 (0.0034)
	40	0.9935 (0.0045)	0.9936 (0.0044)	0.9935 (0.0044)	0.9935 (0.0045)	0.9934 (0.0046)

of the multiple correlation for various τ values when the error follows a normal distribution; Tables S1 and S2 in the Supplementary S6 present the results for chi-square and exponential error distributions, respectively. The efficiency of the methodology increases with n and decreases with p, and is consistent across different quantile levels and error distributions.

Example 2 - Performance of the algorithm. We evaluate the performance of the algorithm across all models. Table 2 reports the means

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

and standard deviations of the multiple correlation for various τ and error distributions. The mean multiple correlation is close to d_τ.

Table 2: *Mean (and standard deviation) of multiple correlation for Models I-VIII.*

M	error	0.1	0.25	0.5	0.75	0.9
I	$\mathcal{N}$	0.9967 (0.0033)	0.9967 (0.0035)	0.9967 (0.0035)	0.9967 (0.0035)	0.9968 (0.0031)
	$\mathcal{X}_3^2$	0.9952 (0.0056	0.9953 (0.0054)	0.9952 (0.0053)	0.9953 (0.0053)	0.9954 (0.0052)
	Exp	0.9965 (0.0031)	0.9966 (0.0030)	0.9966 (0.0031)	0.9966 (0.0031)	0.9966 (0.0031)
II	$\mathcal{N}$	0.9975 (0.0016)	0.9975 (0.0016)	0.9975 (0.0017)	0.9975 (0.0016)	0.9975 (0.0016)
	$\mathcal{X}_3^2$	0.9961 (0.0048)	0.9961 (0.0047)	0.9961 (0.0045)	0.9961 (0.0046)	0.9961 (0.0043)
	Exp	0.9967 (0.0032)	0.9967 (0.0032)	0.9967 (0.0032)	0.9967 (0.0032)	0.9965 (0.0037)
III	$\mathcal{N}$	0.9975 (0.0020)	0.9973 (0.0031)	0.9955 (0.0082)	0.9971 (0.0038)	0.9975 (0.0018)
	$\mathcal{X}_3^2$	0.9970 (0.0024)	0.9970 (0.0022)	0.9971 (0.0020)	0.9971 (0.0020)	0.9971 (0.0021)
	Exp	0.9969 (0.0023)	0.9969 (0.0022)	0.9969 (0.0022)	0.9969 (0.0023)	0.9969 (0.0023)
IV	$\mathcal{N}$	0.6305 (0.0393)	0.8380 (0.0270)	0.9969 (0.0035)	0.8373 (0.0292)	0.6297 (0.0413)
	$\mathcal{X}_3^2$	0.9426 (0.0660)	0.9176 (0.0828)	0.7953 (0.1576)	0.7018 (0.1972)	0.6628 (0.2151)
	Exp	0.9340 (0.0991)	0.9604 (0.0346)	0.8205 (0.1053)	0.6547 (0.1619)	0.5882 (0.1979)
V	$\mathcal{N}$	1.9692 (0.0852)	1.9591 (0.1186)	1.9538 (0.1353)	1.9637 (0.0971)	1.9690 (0.0881)
	$\mathcal{X}_3^2$	1.9492 (0.1383)	1.9448 (0.1537)	1.9352 (0.1763)	1.9423 (0.1489)	1.9425 (0.1561)
	Exp	1.9679 (0.0869)	1.9629 (0.0967)	1.9625 (0.0829)	1.9619 (0.0931)	1.9688 (0.0813)
VI	$\mathcal{N}$	1.9802 (0.0425)	1.9808 (0.0404)	1.9805 (0.0410)	1.9793 (0.0438)	1.9751 (0.0496)
	$\mathcal{X}_3^2$	1.9549 (0.1063)	1.9569 (0.1006)	1.9549 (0.0884)	1.9269 (0.1210)	1.8588 (0.2145)
	Exp	1.9676 (0.1039)	1.9680 (0.1034)	1.9659 (0.1053)	1.9571 (0.1197)	1.9054 (0.1754)
VII	$\mathcal{N}$	2.3663 (0.3282)	2.4085 (0.3093)	2.4119 (0.3117)	2.4280 (0.3098)	2.4434 (0.3042)
	$\mathcal{X}_3^2$	2.8988 (0.1798)	2.9270 (0.1569)	2.9400 (0.1270)	2.9380 (0.1144)	2.9234 (0.1185)
	Exp	2.8644 (0.2028)	2.8925 (0.1689)	2.9173 (0.1231)	2.9084 (0.1369)	2.8836 (0.1693)
VIII	$\mathcal{N}$	1.3976 (0.1730)	1.6780 (0.1561)	1.9568 (0.0816)	1.6614 (0.1481)	1.3793 (0.1805)
	$\mathcal{X}_3^2$	1.604 (0.2456)	1.8709 (0.1626)	1.9040 (0.1342)	1.8619 (0.1563)	1.8475 (0.1684)
	Exp	1.4798 (0.2692)	1.7974 (0.1893)	1.9124 (0.1116)	1.8273 (0.1653)	1.7864 (0.2003)

Example 3 - Methods to compare. We compare our method with FSIR (Ferré and Yao, 2003) and robust FSIR (R-FSIR; Solea, Christou and Song

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

(2026). Since FSIR and R-FSIR target FCS while we focus on FCQS, we limit comparisons to Models I-III and V-VII, where subspaces coincide, and fix $\tau = 0.5$. Table 3 reports the means and standard deviations of the multiple correlation. Our method consistently outperforms the others, except in Model III, where FSIR performs slightly better with normal errors, excelling especially in multi-index models.

8. Application

We apply our method to the ADHD-200 resting-state fMRI (rs-fMRI) data from 222 ADHD patients (2 predominantly hyperactive/impulsive, 44 predominantly inattentive, and 77 combined), and 99 controls, obtained from the NYU Child Study Center. The data are publicly available from the ADHD-200 Consortium (`http://fcon_1000.projects.nitrc.org/indi/adhd200/index.html`). We focus on the ADHD-C group, removing five subjects due to missing information. Pre-processing was done by the Neuro Bureau organization using the Athena pipeline (`http://www.theneurobureau.org/`). The 116 regions of interest were defined by the automated anatomical labeling atlas (AAL; Craddock, James, Holtzheimer, Hu and Mayberg (2012)). Time series were extracted by averaging voxel signals within each region, yielding 116 regional fMRI time series observed at 172 time points.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

Table 3: *Mean (and standard deviation) of multiple correlation for Models I-VII, for FSIR, R-FSIR, and* 0.5*-FCQS and various error distributions.*

M	error	FSIR	R-FSIR	0.5-FCQS
I	$\mathcal{N}$	0.9652 (0.0138)	0.9623 (0.0156)	**0.9968** (0.0033)
	$\mathcal{X}_3^2$	0.9414 (0.0263)	0.9151 (0.0458)	**0.9964** (0.0027)
	Exp	0.9532 (0.0207)	0.9373 (0.0315)	**0.9961** (0.0059)
II	$\mathcal{N}$	0.9821 (0.0071)	0.9829 (0.0070)	**0.9972** (0.0019)
	$\mathcal{X}_3^2$	0.9519 (0.0207)	0.8963 (0.0803)	**0.9967** (0.0030)
	Exp	0.9703 (0.0124)	0.9568 (0.0217)	**0.9968** (0.0032)
III	$\mathcal{N}$	**0.9900** (0.0039)	0.9898 (0.0036)	0.9872 (0.0555)
	$\mathcal{X}_3^2$	0.9926 (0.0028)	0.9926 (0.0031)	**0.9972** (0.0020)
	Exp	0.9887 (0.0046)	0.9874 (0.0057)	**0.9969** (0.0032)
V	$\mathcal{N}$	1.9287 (0.0235)	1.9300 (0.0242)	**1.9666** (0.0606)
	$\mathcal{X}_3^2$	1.9009 (0.0414)	1.8945 (0.0428)	**1.9476** (0.1281)
	Exp	1.9141 (0.0298)	1.9122 (0.0336)	**1.9626** (0.0685)
VI	$\mathcal{N}$	1.0634 (0.1199)	1.0213 (0.0735)	**1.9856** (0.0243)
	$\mathcal{X}_3^2$	1.0116 (0.1265)	0.9403 (0.1256)	**1.9563** (0.0778)
	Exp	1.0603 (0.1446)	1.0003 (0.0959)	**1.9713** (0.0765)
VII	$\mathcal{N}$	1.9413 (0.1792)	1.7187 (0.2106)	**2.5043** (0.3515)
	$\mathcal{X}_3^2$	2.1001 (0.2947)	1.7029 (0.2785)	**2.9454** (0.0856)
	Exp	2.1188 (0.2686)	1.7358 (0.2709)	**2.9289** (0.1420)

We analyze the association between rs-fMRI brain activity, $X_u(t) = (X_u^1(t), \ldots, X_u^{116}(t))$, and ADHD scores, Y_u, for 72 subjects. The right-skewed distribution and extreme ADHD scores highlight the suitability of QR and our method; see Figure S1 in Supplementary S6.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

We apply our method to derive the first d_τ sufficient predictors $\langle \widehat{v}_{\tau 1}, X\rangle_{\mathscr{H}}$, $\ldots, \langle \widehat{v}_{\tau d_\tau}, X\rangle_{\mathscr{H}}$ for $\tau = 0.1, 0.25, 0.5, 0.75, 0.9$. For all τ, we choose $d_\tau = 5$ and estimate the fMRI data using 15 B-splines basis functions of order 4. Figure 1 shows scatterplots of the first two sufficient predictors by quantile, with different colors indicating different ADHD scores. The distinct grouping of lighter and darker colors demonstrate a clear separation of subject based on their ADHD scores. These groupings effectively highlight the separation of subjects with low, moderate, and high severity ADHD scores.

To compare our method with FSIR, we split the data into training (80%) and test (20%) sets and fit the local linear QR (Guerre and Sabbah, 2012) for various quantiles ($\tau = 0.1, 0.25, 0.5, 0.75, 0.9$). We use the first five sufficient predictors from FSIR and τ-FCQS, aligning τ with the quantile of the local linear model. Table 4 reports the average mean square error, indicting that our methodology outperforms FSIR, except for $\tau = 0.25$. Our results agree with Joshi, Li, Akrami and Leahy (2019), demonstrating that dimension reduction in rs-fMRI data can predict ADHD scores.

Supplementary Material

The online Supplementary Material contains additional assumptions, preliminary results and lemmas, proofs, and additional simulation results.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

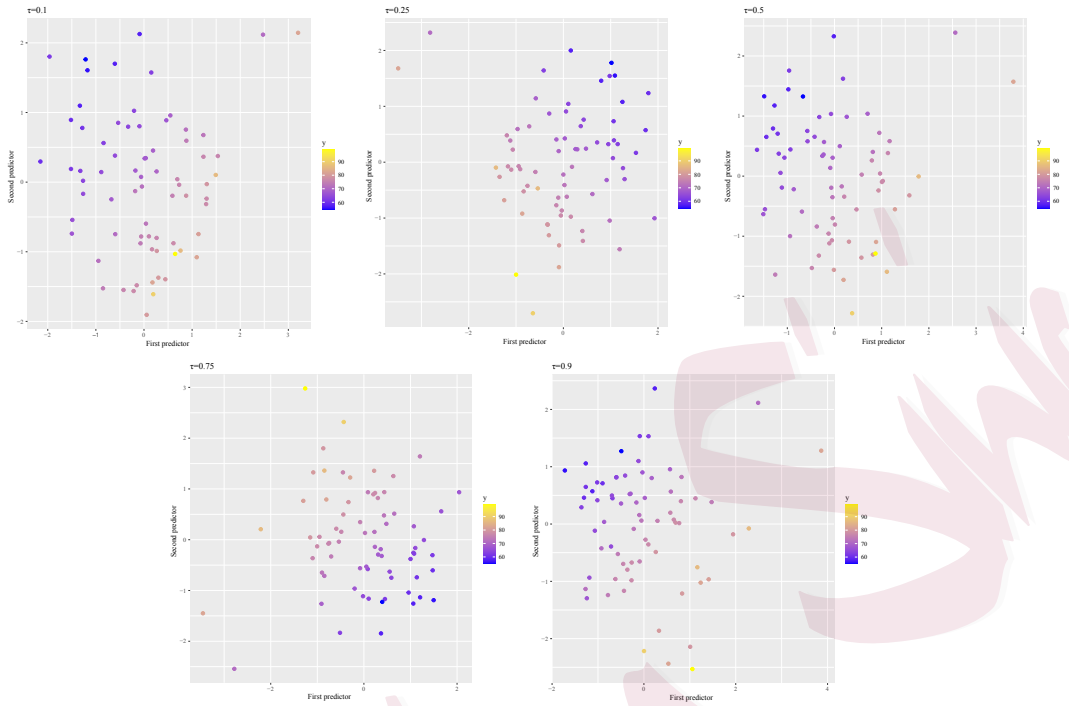

Figure 1: Scatterplots of the first two sufficient predictors for the fMRI data set across quantile levels $\tau = 0.1, 0.25, 0.5, 0.75, 0.9$. Each point represents an observation, with colors indicating ADHD scores. Lighter colors correspond to higher ADHD scores. The axes represent the values of the first and second sufficient predictors.

Acknowledgments

Eliana Christou's research was supported by the National Science Foundation under Grant No. DMS-2213140. Jun Song's research was supported

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

Table 4: *Average mean square error of local linear QR model using the first d_τ sufficient predictors constructed by FSIR and τ-FCQS.*

Method	0.1	0.25	0.5	0.75	0.9
FSIR	5.18	4.75	4.53	5.89	7.95
τ-FCQS	4.26	5.21	3.84	5.42	6.93

by the National Research Foundation of Korea grants funded by the Korea government (MSIT) (No. 2022R1C1C1003647, 2022M3J6A1063595, and RS-2023-00219212). The authors would like to thank the associate editor and the two anonymous referees, whose comments lead to improvements in the presentation of this paper.

References

Ait-Saïdi, A., Ferraty, F., Kassa, R., and Vieu, P. (2008). Cross-validated estimations in the single-functional index model. *Statistics* **42**, 475–494.

Amato, U., Antoniadis, A., and De Feis, I. (2006). Dimension reduction in functional regression with applications. *Computational Statistics and Data Analysis* **50**, 2422–2446.

Aneiros Pérez, G., Cardot, H., Estévez Pérez, G., and Vieu, P. (2004). Maximum ozone forecasting by functional nonparametric approach. *Environmetrics* **15**, 675–685.

Besse, P., Cardot, J., and Stephenson, D. (2000). Autoregressive forecasting of some functional

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

climatic variations. *Scandinavian Journal of Statistics* **27**, 673–687.

Bosq, D. (2000). *Linear processes in function spaces: theory and applications*, Volume 149. Springer Science & Business Media.

Christou, E. (2020). Central quantile subspace. *Statistics and Computing* **30**, 677–695.

Cai, T. T., and Hall, P. (2006). Prediction in functional linear regression. *Annals of Statistics* **34**, 2159–2179.

Cardot, H., Crambes, C., and Sarda, P. (2005). Quantile regression when the covariates are functions. *Journal of Nonparametric Statistics* **17**, 841–856.

Cardot, H., Ferraty, F., and Sarda, P. (2003). Spline estimators for the functional linear model. *Statistica Sinica* **13**, 571–591.

Chen, D., Hall, P., and Müller, H.-G. (2011). Single and multiple index functional regression models with nonparametric link. *The Annals of Statistics* **39**, 1720–1747.

Chen, K., and Müller, H.-G. (2012). Conditional quantile analysis when covariates are functions, with application to growth data. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **74(1)**, 67–89.

Chiou, J., Yang, Y., and Chen, Y. (2016). Multivariate functional linear regression and prediction. *Journal of Multivariate Analysis* **146**, 301–312.

Cook, D. (1998). Regression Graphics: Ideas for Studying Regressions Through Graphics. *John Wiley & Sons*.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

REFERENCES

Cook, R.D., and Li, B. (2002). Dimension reduction for conditional mean in regression. *The Annals of Statistics* **30(2)**, 455–474.

Cook, R.D., and Weisberg, S. (1991). Sliced Inverse Regression for Dimension Reduction: Comment. *Journal of the American Statistical Association* **86(414)**, 328–332.

Craddock, R.C., James, G.A., Holtzheimer, P.E., Hu, X.P., and Mayberg, H.S. (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human brain mapping* **33**, 1914–1928.

Crambes, C., Kneip, A., and Sarda, P. (2009). Smoothing splines estimators for functional linear regression. *Annals of Statistics* **37**, 35–72.

Eaton, M. L. (1986). A characterization of spherical distribution. *Journal of Multivariate Analysis* **20(2)**, 272–276.

Elías, A., Jiménez, R., and Shang, H. L. (2022). On projection methods for functional time series forecasting. *Journal of Multivariate Analysis* **189**, 104890.

Ferraty, F., Rabhi, A., and Vieu, P. (2005). Conditional quantiles for dependent functional data with application to the climatic El Nino Phenomenon. *Sankhyã: The Indian Journal of Statistics* **67**, 378–398.

Ferraty, F., and Vieu, P. (2002). The functional nonparametric model and application to spectometric data. *Computational Statistics* **17**, 545–564.

Ferraty, F., and Vieu, P. (2003). Curves discrimination: a nonparametric functional approach.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

Computational Statistics and Data Analysis **44**, 161–173.

Ferraty, F., and Vieu. P. (2004). Nonparametric models for functional data, with application in regression, time-series prediction and curve discrimination. *Journal of Nonparametric Statistics* **16**, 111–125.

Ferré, L., and Yao, A. F. (2003). Functional sliced inverse regression analysis. *Statistics* **37**, 475–488.

Ferré, L., and Yao, A. F. (2005). Smoothed functional inverse regression. *Statistica Sinica* **15**, 665–683.

Fontana, M., Tavoni, M., and Vantini, S. (2019). Functional Data Analysis of high-frequency load curves reveals drivers of residential electricity consumption. *PLoS One* **14(6)**, e0218702.

Gubian, M., Torreira, F., and Boves, L. (2015). Using Functional Data Analysis for investigating multidimensional dynamic phonetic contrasts. *Journal of Phonetics* **49**, 16–40.

Guerre, E., and Sabbah, C. (2012). Uniform bias study and Bahadur representation for local polynomial estimators of the conditional quantile function. *Econometric Theory* **28**, 87–129.

Guo, M., Zhou, L., Huang, J. Z., and Härdle, W. K. (2015). Functional data analysis of generalized regression quantiles. *Statistics and Computing* **25**, 189–202.

Hall, P., and Horowitz, J. L. (2007). Methodology and convergence rates for functional linear

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

regression. *Annals of Statistics* **35**, 70–91.

Happ, C., and Greven S. (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. *Journal of the American Statistical Association* **13**, 649–659.

Hsing, T., and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. Wiley.

James, G. M. (2002). Generalized Linear Models With Functional Predictors. *Journal of the Royal Statistical Society, Series B* **64**, 411–432.

James, G. M., and Silverman, B. W. (2005). Functional Adaptive Model Estimation. *Journal of the American Statistical Association* **100**, 565–576.

James, G. M., Wang, J., and Zhu, J. (2009). Functional linear regression that's interpretable. *Annals of Statistics* **37**, 2083–2108.

Joshi, A. A., Li, J., Akrami, H., and Leahy, R. M. (2019). Predicting cognitive scores from resting fMRI data and geometric features of the brain. *Medical Imaging 2019: Image Processing* 10949, 619-625.

Kato, K. (2012). Estimation in functional linear quantile regression. *The Annals of Statistics* **40**, 3108–3136.

Kokoszka, P., and Reimherr, M. (2017). *Introduction to Functional Data Analysis*. CRC Press.

Kong, E., and Xia. Y. (2014). An adaptive composite quantile approach to dimension reduction.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

The Annals of Statistics **42(4)**, 1657–1688.

Lee, C.E., and Hilafu, H. (2022). Quantile martingale difference divergence for dimension reduction. *Statistica Sinica* **32**, 65–87.

Lee, K-Y, and Li, L. (2022). Functional sufficient dimension reduction through average Fréchet derivatives. *The Annals of Statistics* **50(2)**, 904–929.

Li, K.-C. (1991). Sliced inverse regression for dimension reduction. *Journal of the American Statistical Association* **86**, 316–327.

Li, K., and Luo, S. (2017). Functional Joint Model for Longitudinal and Time-to-Event Data: An Application to Alzheimer's Disease. *Stat Med* **36(22)**, 3560–3572.

Li, B., and Song, J. (2017). Nonlinear sufficient dimension reduction for functional data. *The Annals of Statistics* **45(3)**, 1059–1095.

Li, B, and Song, J. (2022). Dimension reduction for functional data based on weak conditional moments. *The Annals of Statistics* **50(1)**, 107–128.

Li, B., Zha, H., and Chiaromonte, F. (2005). Contour regression: A general approach to dimension reduction. *The Annals of Statistics* **33**, 1580–1616.

Lian, H., and Li, G. (2014). Series expansion for functional sufficient dimension reduction. *Journal of Multivariate analysis* **124**, 150–165.

Lu, Y., Du., J., and Sun, Z. (2014). Functional partially linear quantile regression model. *Metrika* **77**, 317–332.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

REFERENCES

Luo, W., Li, B., and Yin, X. (2014). On efficient dimension reduction with respect to a statistical functional of interest. *The Annals of Statistics* **42(1)**, 382–412.

Ma, H., Li, T., Zhu, H., and Zhu, Z. (2019). Quantile regression for functional partially linear model in ultra-high dimensions. *Computational Statistics and Data Analysis* **129**, 135–147.

Mahzarnia, A., and Song, J. (2022). Multivariate functional group sparse regression: Functional predictor selection. *PLoS ONE* **17(4)**, e0265940.

Marx, B. D., and Eilers, P. H. (1999). Generalized Linear Regression on Sampled Signals and Curves: A P-Spline Approach. *Technometrics* **41**, 1–13.

Müller, H. G., Sen, R., and Stadmüller, U. (2011). Functional data analysis for volatility. *Journal of Econometrics* **165**, 233–245.

Müller, H. G., and Stadmüller, U. (2005). Generalized Functional Linear Models. *The Annals of Statistics* **33**, 774–805.

Pratt, J., Su, W., Hayes, D. Jr., Clancy, J.P., and Szczesniak, R.D. (2021). An Animated Functional Data Analysis Interface to Cluster Rapid Lung Function Decline and Enhance Center-Level Care in Cystic Fibrosis. *Journal of Healthcare Engineering* **ID 6671833**

Qingguo, T., and Kong, L. (2017). Quantile regression in functional linear semiparametric model. *Statistics* **51**, 1342–1358.

Ramsay, J. O., and Dalzell, C. J. (1991). Some tools for functional data analysis. *Journal of the Royal Statistical Society: Series B (Methodological)* **53**, 539–572.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

REFERENCES

Ramsay, J. O., and Silverman, B. W. (2005). *Functional Data Analysis*. New York: Springer.

Shi, G., Xie, T., and Zhang, Z. (2020). Statistical inference for the functional quadratic quantile regression model. *Metrika* **83**, 937–960.

Shin, H. (2009). Partial functional linear regression. *Journal of Statistical Planning and Inference* **139**, 3405–3418.

Solea, E., Christou, E., and Song, J. (2026). Robust Inverse Regression for Multivariate Elliptical Functional Data. *Statistica Sinica* https://www3.stat.sinica.edu.tw/preprint/SS-2023-0341_Preprint.pdf

Song, J. (2019). On sufficient dimension reduction for functional data: Inverse moment-based methods. *WIREs Computational Statistics* **11**, e1459.

Song, J., and Li, B. (2021). Nonlinear and additive principal component analysis for functional data. *Journal of Multivariate Analysis* **181**, 104675.

Wang, G., Lin, N., and Zhang, B. (2013). Functional contour regression. *Journal of Multivariate Analysis* **116**, 1–13.

Wang, G., Liu, S., Han, F., and Di, C.-Z. (2023). Robust functional principal component analysis via a functional pairwise spatial sign operator. *Biometrics* **79(2)**, 1239–1253.

Yao, F., Müller, H.-G., and Wang, J.-L. (2005). Functional linear regression analysis for longitudinal data. *Annals of Statistics* **33**, 2873–2903.

Yao, F., Sue-Chee, S., and Wang, F. (2017). Regularized partially functional quantile regression.

Statistica Sinica: Preprint
doi:10.5705/ss.202024.0211

Journal of Multivariate Analysis **156**, 39–56.

Yu, K., and Jones, M.C. (1998). Local linear quantile regression. *Journal of the American Statistical Association* **93**, 228–238.

Yuan, M., and Cai, T. T. (2010). A reproducing kernel Hilbert space approach to functional linear regression. *Annals of Statistics* **38**, 3412–3444.

Eliana Christou

Department of Mathematics and Statistics, University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, NC 28223, USA. E-mail: echris15@charlotte.edu

Eftychia Solea

School of Mathematical Sciences, Queen Mary University of London, London, E1 4NS, United Kingdom. E-mail: e.solea@qmul.ac.uk

Shanshan Wang

Department of Mathematics and Statistics, University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, NC 28223, USA. E-mail: swang47@charlotte.edu

Jun Song

Department of Statistics, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul, 02841, South Korea. E-mail: junsong@korea.ac.kr