

Article

Cross-Validated Functional Generalized Partially Linear Single-Functional Index Model

Mustapha Rachdi ^{1,†}, Mohamed Alahiane ^{2,*,†}, Idir Ouassou ^{2,†}, Abdelaziz Alahiane ^{3,†} and Lahoucine Hobbad ^{2,†}

¹ Laboratory AGEIS, Grenoble Alps University, UFR SHS, BP. 47, Cedex 09, 38040 Grenoble, France; mustapha.rachdi@univ-grenoble-alpes.fr

² Complex Systems Modeling Laboratory, National School of Applied Sciences, Cadi Ayyad University, Av. Abdelkrim Khattabi, BP. 575, Marrakesh 40000, Morocco; i.ouassou@uca.ma (I.O.); lahoucine.hobbad@ced.uca.ma (L.H.)

³ SmartICT Lab, ENSAO, Mohamed Premier University, Oujda 60000, Morocco; abdelaziz.alahiane@ump.ac.ma

* Correspondence: mohamed.alahiane@ced.uca.ma

† These authors contributed equally to this work.

Abstract: In this paper, we have introduced a functional approach for approximating nonparametric functions and coefficients in the presence of multivariate and functional predictors. By utilizing the Fisher scoring algorithm and the cross-validation technique, we derived the necessary components that allow us to explain scalar responses, including the functional index, the nonlinear regression operator, the single-index component, and the systematic component. This approach effectively addresses the curse of dimensionality and can be applied to the analysis of multivariate and functional random variables in a separable Hilbert space. We employed an iterative Fisher scoring procedure with normalized B-splines to estimate the parameters, and both the theoretical and practical evaluations demonstrated its favorable performance. The results indicate that the nonparametric functions, the coefficients, and the regression operators can be estimated accurately, and our method exhibits strong predictive capabilities when applied to real or simulated data.

Keywords: single-index coefficient; cross-validation; asymptotic normality; Fisher scoring algorithm; functional data analysis (FDA); functional index; quasi-likelihood

MSC: 62R10



Citation: Rachdi, M.; Alahiane, M.; Ouassou, I.; Alahiane, A.; Hobbad, L. Cross-Validated Functional Generalized Partially Linear Single-Functional Index Model. *Mathematics* **2024**, *12*, 2649. <https://doi.org/10.3390/math12172649>

Academic Editors: Heng Lian and Manuel Alberto M. Ferreira

Received: 3 July 2024

Revised: 5 August 2024

Accepted: 20 August 2024

Published: 26 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Parametric regression models are a common tool in generalized linear models (GLMs), where the relationship between the average response and covariates is explored through a chosen link function, often with a canonical or predefined form, as noted by McCullagh [1] and Nelder [2]. However, in certain cases, this approach may not be suitable due to the lack of knowledge or the presence of a more intricate link function.

In order to address this challenge, various models have been developed, including nonparametric and semiparametric regression models. However, these models are often limited in their application due to the curse of dimensionality, which poses significant challenges when dealing with high-dimensional data.

Efforts have been made to overcome this limitation by employing two key approaches: (i) the approximation of the link functions and (ii) dimension reduction techniques. One approach that has been proposed is the generalized additive model (GAM), which was introduced by Hastie [3] and extensively discussed by Wood [4]. In the GAM, the nonparametric component is represented as a sum of univariate functions, allowing for the flexible modeling of the complex relationships. However, it is important to note that a potential limitation of the GAM is its inability to explicitly account for interactions between explanatory variables. This means that the model may not fully capture the effects of interactions, which could be important in certain contexts.

The single-index model (SIM) was introduced by Härdle et al. [5] and Hristache et al. [6] to address certain cases where dimensionality reduction and relaxation of restrictive parametric assumptions are needed. The SIM approach achieves this by transforming multiple covariates into a linear combination of the covariates. Building upon the SIM, Ait-Saïdi et al. [7] investigated the functional single-index model (FSIM), which extends the single-index model usefulness framework to incorporate functional predictors.

Longitudinal data scenarios involving discrete explanatory variables in the linear component were investigated by Liang et al. [8], and Chen et al. [9] extended the framework to develop the partially linear single-index models (PLSIMs). These models allow for the modeling of discrete variables within the single-index framework.

Partially generalized linear single-index models (PGLSIMs), introduced by Carroll et al. [10], employ kernel smoothing techniques to estimate the single-index link function. PGLSIMs offer increased flexibility and modeling capabilities. In a study by Wang et al. [11], a novel approach was proposed. This approach combines penalized spline smoothing of the quasi-likelihood with the Fisher scoring mechanism, resulting in improved theoretical robustness and suitability for PGLSIM modeling.

Overall, these models, including SIM, FSIM, PLSIM, and PGLSIM, provide valuable tools for addressing complex scenarios, accommodating functional predictors and discrete variables, and effectively handling high-dimensional data.

Several models have been developed to address the complexity of functional variables in the regression analysis. However, it is important to note that the models mentioned earlier may not fully capture the intricacies of the data when some covariates are of the functional kind. Researchers have dedicated their efforts to explore functional variables in the regression models, as evidenced by the works of Ramsay et al. [12] and Ferraty et al. [13].

Furthermore, the field has seen investigations into specific models such as semi-functional partial linear regression, as studied by Aneiros et al. [14], and partially linear modeling with multifunctional covariates, as explored by Aneiros et al. [15]. Various other works have contributed to our understanding of this topic, including studies on inference by Horváth et al. [16]; the introduction to this subject by Kokoszka et al. [17]; the spline approaches by Schumaker [18]; the functional principal component analysis (FPCA) by Cao et al. [19]; the lack-of-fit testing by Li et al. [20]; and the works by Ould-Saïd et al. [21], Laksaci et al. [22], and Ouassou et al. [23,24] on the regression analysis with functional covariates.

We can refer to specific studies concerned by different aspects of the regression models. For instance, Yu et al. [25] focus on the partially functional linear single-index regression model, while Yu et al. [26] provide a comprehensive review of the penalized spline-smoothing methodology for the partially linear single-index model (PLSIM), where the sub-regression function is assumed to be a spline function with a fixed number of knots. Regarding the FDA, Rachdi et al. [27] and Alahiane et al. [28] investigate partially linear generalized single-index models for functional data (PLGSIMF) using the B-spline expansion and quasi-likelihood function. In their studies, the functional model is assumed to be linear. On the other hand, Alahiane et al. [29] explore the high-dimensional case, where the functional model is nonlinear.

The projection pursuit regression, proposed by Friedman et al. [30], Hall [31], and Huber [32], introduces an additive model that operates on derived features rather than the original inputs. This approach aims to capture complex relationships between variables by projecting them onto a lower-dimensional space. In the context of FDA, Ferraty et al. [33] extended the projection pursuit regression to the functional framework. They achieved this by incorporating the cross-validation method in the Nadaraya–Watson estimation technique and the spline functions. This generalization allows for the extension of the FDA results to the case where both the covariates and the responses are curves.

In this paper, we present a novel model called cross-validated estimations for generalized partially linear single-functional index models (CVGPLSFIM). We therefore estimate: (i) the nonlinear regression operator comprising a functional index that is chosen by the

cross-validation method; (ii) the nonlinear regression operator combining a spline approximation and the one-dimensional systematic component with unknown link; and (iii) the single-index function using an iterative algorithm based on smoothing by spline functions and the maximization of the quasi-likelihood function. We also provide the convergence rates of our different estimators of the different CVGPLSFIM parameters. In fact, our contributions include examining the performance of the single and functional index, the nonparametric regression operator, the systematic component, and the optimal direction. Through numerical calculations on simulated data and then on real data, we demonstrate that our model outperforms the models mentioned throughout this introduction.

The structure of this article is as follows. In Sections 2.1, 2.2, and 3, we present our estimation methodology, discuss the asymptotic properties of the proposed estimators, and give an iterative algorithm that maximizes the quasi-likelihood function, allowing us to compute the estimators. Section 4 presents the results of a simulation study. Furthermore, in Section 5, we apply our methodology to a real dataset consisting of curves from the chemometrics field. The technical lemmas needed to prove Theorems 1–3 are provided in Appendix A. In order to save space, the detailed proofs of the various results have been compiled in an additional and available supplementary material.

2. Estimation Methodology

2.1. Preliminary Definitions

Let Y be a scalar response variable, Z be a functional random variable that is valued in $\mathcal{H}$ separable Hilbert space endowed with inner product $\langle f, g \rangle = \int_I f(t)g(t)dt$. Let $(X, Z) \in \mathbb{R}^d \times \mathcal{H}$ be the predictor vector where $X = (X_1, X_2, \dots, X_d)$ and d are a fixed integer. For a fixed $(x, z) \in \mathbb{R}^d \times \mathcal{H}$, we assume that the conditional density function of the response Y given $(X, Z) = (x, z)$ belongs to the following canonical exponential family

$$f_{Y|X=x, Z=z}(y) = \exp \left(y \zeta(x, z) - B(\zeta(x, z)) + C(y) \right), \quad (1)$$

where B and C are two known functions that are defined from $\mathbb{R}$ into $\mathbb{R}$, and $\zeta : \mathbb{R}^d \times \mathcal{H} \rightarrow \mathbb{R}$ is the parameter in the generalized linear model that is linked to the dependent variable

$$\mu(x, z) = \mathbb{E} \left[Y | X = x, Z = z \right] = B'(\zeta(x, z)). \quad (2)$$

where B' denotes the first derivative of the function B .

In what follows, we modelize the scalar response Y as a cross-validated generalized partially linear functional single-index model (CVGPLFSIM) by

$$g(\mu(X, Z)) = \eta_0(\alpha^\top X) + R(\langle \beta, Z \rangle) + \varepsilon. \quad (3)$$

where $\alpha = (\alpha_1, \dots, \alpha_d)^\top \in \mathbb{R}^d$ is the d -dimensional single-index coefficient vector; η_0 is the unknown single-index link function (the systematic nonlinear component), which will be assumed to be sufficiently smooth; and R is the nonlinear regression operator to be estimated, where $x^\top$ denotes the transpose vector of x and $\beta \in \mathcal{H}$ is the so-called the functional index, which is such that $\|\beta\|^2 = 1$ and $\langle \beta, Z \rangle$ is the functional single-index.

We suggest estimators for the unknown single-index vector α , the unknown systematic component $\eta_0(\cdot)$, the unknown functional index β , and the unknown nonlinear regression operator R . Then, we will derive their asymptotic distributions, and we also provide some illustrations of these models and their performances.

Remark 1.

- For identifiability purposes, we guess that $\|\alpha\|_d = 1$ and the first component of α is non-negative, i.e., $\alpha_1 > 0$, where $\|\cdot\|_d$ denotes the Euclidean norm on $\mathbb{R}^d$.

- In order to identify the function $\eta_0(\cdot)$, we define its support as $[a, b]$, where $a = \inf \alpha^\top X$ and $b = \sup \alpha^\top X$.
- In the definition of the real canonical link function g , we will assume that the functional random variable $Z = \{Z(t), t \in [0, 1]\}$ is valued in $\mathcal{H}$ such that:

$$\mathbb{E}[Z] = 0, \mathbb{E}(\varepsilon|X, Z) = 0 \quad \text{and} \quad \text{var}(\varepsilon|X, Z) = \sigma^2.$$

- If the conditional variance $\text{var}(Y|X = x, Z = z) = \sigma^2 V(\mu(x, z))$ where $V(\cdot)$ is an unknown positive function, then the estimation of the mean function $g(\mu)$ may be obtained by replacing the log-likelihood $f_{Y|X=x, Z=z}$ given by (1) by the quasi-likelihood $Q(u, v)$, which is given for any real numbers u and v by

$$\frac{\partial Q(u, v)}{\partial u} = \frac{v - u}{\sigma^2 V(u)} = \frac{v - u}{\text{var}(Y|X = x, Z = z)}.$$

2.2. Methodology

Let (X_i, Y_i, Z_i) for $i = 1, \dots, n$, be an independent and identically distributed (i.i.d.) n -sample of (X, Y, Z) and, for each $i = 1, \dots, n$,

$$g(\mu(X_i, Z_i)) = \eta_0(\alpha^\top X_i) + R(\langle \beta, Z_i \rangle) + \varepsilon_i. \quad (4)$$

Let $\{B_j(u), j = 1, \dots, N_n\}$ the B-spline basis functions of order r ; $h = \frac{(b-a)}{J_n+1}$ is the distance between the neighbors knots, where J_n denote the number of interior knots in $[a, b]$, as in Wang et al. [11], Rachdi et al. [27], Alahiane et al. [28] and Alahiane et al. [29].

Let $\mathcal{S}_n$ be the space of polynomial splines on $[a, b]$ of order $r \geq 1$, $v \in \mathbb{N}^*$ and $e \in (0, 1]$ such that $p = v + e > 1.5$. We denote by $\mathcal{H}(p)$ the collection of functions g that are defined on $[a, b]$ whose v -th order derivative, $g^{(v)}$, exists and satisfies the following e -th order Lipschitz condition

$$|g^{(v)}(m_1) - g^{(v)}(m_2)| \leq C|m_1 - m_2|^e, \quad \text{for all } a \leq m_1, m_2 \leq b.$$

Using the method by De Boor [34], we can approximate η , which is assumed in $\mathcal{H}(p)$, by the function $\tilde{\eta} \in \mathcal{S}_n$. So, we can write $\tilde{\eta}(u) = \tilde{\gamma}^\top B(u)$, where $B(u)$ is the spline basis and $\tilde{\gamma} \in \mathbb{R}^{N_n}$ is the spline coefficient vector.

We introduce a new knot sequence $t_0 < t_1 < \dots < t_{k+1}$ in the range of R . Then, there exists $N' = k + r + 1$ functions in the B-splines basis, which are normalized and of order r , such that $g(\cdot) \approx \delta^\top B_1(\cdot)$ where $B_1(\cdot) = (B_{11}(\cdot), B_{12}(\cdot), \dots, B_{1N'}(\cdot))^\top$ and $\delta \in \mathbb{R}^{N'}$.

Using the setting $W_i = \langle \beta, Z_i \rangle$, the mean function estimator $\hat{m}(x, z)$ is therefore obtained by assessing the parameter $\theta = (\alpha^\top, \gamma^\top, \delta^\top)^\top$ and inverting the subsequent equation $g(\mu(X_i, Z_i)) = \hat{\gamma}^\top B(\hat{\alpha}^\top x) + \hat{\delta}^\top B_1(W_i)$.

We may notice that the parameter $\theta = (\alpha^\top, \gamma^\top, \delta^\top)^\top$ is deduced by maximizing the following requirement of quasi-likelihood

$$\hat{\theta} = (\hat{\alpha}^\top, \hat{\gamma}^\top, \hat{\delta}^\top)^\top = \arg \max_{\theta = (\alpha, \gamma, \delta) \in \mathbb{R}^d \times \mathbb{R}^{N_n} \times \mathbb{R}^{N'}} L(\theta), \quad (5)$$

where $L(\theta) := L(\alpha, \gamma, \delta) = \frac{1}{n} \sum_{i=1}^n Q(g^{-1}(m_i), Y_i)$ and $m_i := \gamma^\top B(\alpha^\top X_i) + \delta^\top B_1(W_i)$.

In order to overcome the constraints $\|\alpha\| = 1$ and $\alpha_1 > 0$ that are imposed on the d -dimensional index α , we adopt a reparametrization approach that is inspired by the methodology employed by Yu [26]

$$\alpha(\tau) = \left(\sqrt{1 - \|\tau\|^2}, \tau^\top \right)^\top \text{ for } \tau \in \mathbb{R}^{d-1}.$$

The true value τ_0 of τ must be $\|\tau_0\| \leq 1$. Subsequently, we assume that $\|\tau_0\| < 1$. The Jacobian matrix of $\alpha \rightarrow \alpha(\tau)$ of dimension $d \times (d-1)$ is

$$J(\tau) = \begin{pmatrix} -\frac{1}{\sqrt{1 - \|\tau\|^2}} & \tau^\top \\ I_{(d-1) \times (d-1)} \end{pmatrix}$$

Denote $l = 1, 2$, $q_l(m, y) = \frac{\partial^l}{\partial m^l} Q(m, y)$ and $\rho_l = \frac{1}{\sigma^2 V(g^{-1}(m))} \left[\frac{d}{dm} (g^{-1}(m)) \right]^l$. So,

$$q_1(m, y) = (y - m)\rho_1(m) \quad \text{and} \quad q_2(m, y) = (y - m)\rho'_1(m) - \rho_2(m).$$

The score vector is

$$S(\theta_\tau) = \frac{\partial l}{\partial \theta_\tau}(\theta_\tau) = \frac{1}{n} \sum_{i=1}^n q_1(m_i, Y_i) \xi_i(\tau, \gamma, \delta),$$

where

$$\xi_i(\tau, \gamma, \delta) = \begin{pmatrix} \gamma^\top B'(\alpha^\top(\tau) X_i) J^\top(\tau) X_i \\ B(\alpha^\top(\tau) X_i) \\ B_1(W_i) \end{pmatrix}.$$

The conditional expectation of the Hessian matrix given X and Z is

$$H(\theta_\tau) = -\frac{1}{n} \sum_{i=1}^n \rho_2(m_i) \xi_i(\tau, \gamma, \delta) \xi_i^\top(\tau, \gamma, \delta).$$

The Fisher scoring update equations $\theta_\tau^{(k+1)} = \theta_\tau^{(k)} - [H(\theta_\tau^{(k)})]^{-1} S(\theta_\tau^{(k)})$ becomes

$$\begin{aligned} \theta_\tau^{(k+1)} &= \theta_\tau^{(k)} + \left[\sum_{i=1}^n \rho_2(m_i^{(k)}) \xi_i(\tau^{(k)}, \gamma^{(k)}, \delta^{(k)}) \xi_i^\top(\tau^{(k)}, \gamma^{(k)}, \delta^{(k)}) \right]^{-1} \\ &\quad \times \left[\sum_{i=1}^n (Y_i - \mu_i^{(k)}) \rho_1(m_i^{(k)}) \xi_i(\tau^{(k)}, \gamma^{(k)}, \delta^{(k)}) \right], \end{aligned}$$

where $m_i^{(k)} = \gamma^{(k)\top} B(\alpha^{(k)\top}(\tau^{(k)}) X_i) + \delta^{(k)\top} B_1(W_i)$ and $\mu_i^{(k)} = g^{-1}(m_i^{(k)})$, for $1 \leq i \leq n$.

So, we obtain

$$\begin{aligned} \hat{\eta}(t) &= \hat{\gamma}^\top B(t) = \gamma^{(k)\top} B(t), \\ \hat{m}_i &= \hat{\gamma}^\top B(\alpha^\top(\hat{\tau}) X_i) + \hat{\delta}^\top B_1(W_i) = \gamma^{(k)\top} B(\alpha^\top(\tau^k) X_i) + \delta^{(k)\top} B_1(W_i), \\ \hat{R}(Z_i) &= \hat{\delta}^\top B_1(W_i) = \delta^{(k)\top} B_1(W_i), \end{aligned}$$

where $\hat{\alpha} = \alpha(\tau^{(k)})$ is the estimator of the single-index coefficient vector α and $\hat{\mu}_i = g^{-1}(\hat{m}_i)$.

Through the utilization of this procedure, we will employ the cross-validation method (see the Section 2.3) to determine the optimal direction $\beta(\cdot)$. Additionally, we will estimate in Section 2.4 the various components of our CVGPLSFIM model using a plug-in approach.

2.3. Cross-Validated Estimation of the Functional Index

We use the cross-validation principle based on the leave-one-out statistical sample $\{(X_j, Z_j), j = 1, \dots, n, j \neq i\}$ to estimate the functional index β by:

$$\hat{\beta} = \arg \min_{\beta \in \Xi} CV(\beta),$$

where

$$CV(\beta) = \frac{1}{n} \sum_{i=1}^n \left(Y_i - g^{-1}(\hat{m}_i^{(-i)}) \right)^2 I_{Z_i \in \mathcal{G}}, \quad (6)$$

with $\hat{m}_i^{(-i)} = \hat{\gamma}^\top B(\alpha^\top(\hat{\tau})X_i^{(-i)}) + \hat{\delta}^\top B_1(W_i^{(-i)})$; $\mathcal{G}$ is a subset of $\mathcal{H}$ introduced for usual technical boundedness reasons, and $\Xi = \Xi_n$ with $\Xi \subset \mathcal{H} \cap \{\beta : \|\beta\|^2 = 1\}$ is constructed in a similar way as in Ait-Saïdi et al. [7]. Specifically:

- Each direction $\beta \in \Xi_n$ is obtained from an l_n -dimensional space generated by the B-spline basis functions $\{e_1, \dots, e_{l_n}\}$. Therefore, we focus on directions

$$\beta(\cdot) = \sum_{k=1}^{l_n} \lambda_k e_k, \text{ where } (\lambda_1, \dots, \lambda_{l_n}) \in \mathcal{V}. \quad (7)$$

- The set of coefficients vectors $\mathcal{V}$ in (7) is obtained by the following procedure:
 - Step 1: For each $(b_1, \dots, b_{l_n}) \in \mathcal{C}^{l_n}$, where $\mathcal{C} = \{c_1, \dots, c_K\} \subset \mathbb{R}^K$ denotes a set of K “seed-coefficients”, we construct the initial functional direction $\beta_{\text{init}}(\cdot) = \sum_{k=1}^{l_n} b_k e_k(\cdot)$.
 - Step 2: For each β_{init} selected in Step 1 that verifies the condition $\beta_{\text{init}}(t_0) > 0$, we construct $(\lambda_1, \dots, \lambda_{l_n}) = (b_1, \dots, b_{l_n}) / \langle \beta_{\text{init}}, \beta_{\text{init}} \rangle^{1/2}$.
 - Step 3: Construct $\mathcal{V}$ as the set of vectors $(b_1, \dots, b_{l_n})$ obtained in Step 2. Therefore, the final set of eligible functional direction is

$$\Xi_n = \left\{ \beta(\cdot) = \sum_{k=1}^{l_n} \lambda_k e_k(\cdot) \text{ such that } (\lambda_1, \dots, \lambda_{l_n}) \in \mathcal{V} \right\}.$$

Using the method by De Boor [34], if β_0 is sufficiently smooth, then it is well approximated by some function in the l_n -dimensional space, which is generated by the B-spline basis. From the construction (see Step 2), each $\beta \in \Xi_n$ satisfies $\langle \beta, \beta \rangle = 1$ and $\beta(t_0) > 0$. So, the identifiability of the CVGPLFSIM model is guaranteed. Like in Ait-Saïdi et al. [7], we consider the cubic B-spline functions and $\mathcal{V} = \{-1, 0, 1\}$.

2.4. The CVGPLFSIM Model

By plug-in, first, the functional index $\hat{\beta}$ in the model is

$$g(\mu(X_i, Z_i)) = \eta(\alpha^\top X_i) + R(W_i) \text{ for } i = 1, \dots, n, \quad (8)$$

where $W_i = \langle \hat{\beta}, Z_i \rangle$ denotes the functional index component. We seek a function $\eta \in \mathcal{S}_n$ along with a value of α that minimizes the following quasi-likelihood function

$$L(\eta, \alpha) = \frac{1}{n} \sum_{i=1}^n Q\left(g^{-1}\left\{\eta\left(\alpha^\top X_i\right) + \delta^\top B_1\left(W_i\right)\right\}, Y_i\right). \quad (9)$$

By denoting $\theta = (\alpha^\top, \gamma^\top, \delta^\top)^\top$, the maximization problem (9) is equivalent to find a value θ maximizing

$$l(\theta) = \frac{1}{n} \sum_{i=1}^n Q\left(g^{-1}\left\{\gamma^\top B\left(\alpha^\top X_i\right) + \delta^\top B_1(W_i)\right\}, Y_i\right), \quad (10)$$

where

$$g(\mu(X_i, Z_i)) = \gamma^\top B\left(\alpha^\top X_i\right) + \delta^\top B_1(W_i), \text{ for } i = 1, \dots, n. \quad (11)$$

The mean function estimator $\hat{\mu}$ is given by the evaluation of the parameters $\hat{\theta} = (\hat{\alpha}^\top, \hat{\gamma}^\top, \hat{\delta}^\top)^\top$ and inverting Equation (11). In fact, $\hat{\theta} = (\hat{\alpha}^\top, \hat{\gamma}^\top, \hat{\delta}^\top)^\top$ is determined by maximizing the following quasi-likelihood

$$\hat{\theta} = (\hat{\alpha}^\top, \hat{\gamma}^\top, \hat{\delta}^\top)^\top = \arg \max_{\theta = (\alpha^\top, \gamma^\top, \delta^\top)^\top \in \mathbb{R}^d \times \mathbb{R}^{N_n} \times \mathbb{R}^{N'}} l(\theta), \quad (12)$$

where

$$l(\theta) := l(\alpha, \gamma, \delta) = \frac{1}{n} \sum_{i=1}^n Q\left(g^{-1}\left\{\gamma^\top B\left(\alpha^\top X_i\right) + \delta^\top B_1(W_i)\right\}, Y_i\right) = \frac{1}{n} \sum_{i=1}^n Q\left(g^{-1}\{m_i\}, Y_i\right),$$

with $m_i = \gamma^\top B(U_i) + \delta^\top B_1(W_i)$, where $U_i = \alpha^\top X_i$, $\alpha_0, \gamma_0, \delta_0$, and $\eta_0(\cdot)$ denote the true values, respectively, of α, γ, δ , and $\eta(\cdot)$. So, the spline estimator of $\eta_0(\cdot)$ is $\hat{\eta}(\cdot) = \hat{\gamma}^\top B(\cdot)$.

Let $R(\tau) = \begin{pmatrix} J(\tau) & 0 \\ 0 & I_{N' \times N'} \end{pmatrix}$ be the Jacobian matrix of $(\alpha^\top(\tau), \delta^\top)^\top$ and

$$(\tilde{\alpha}, \tilde{\delta}) = \arg \max_{\|\alpha\|_d=1, \delta \in \mathbb{R}^{N'}} \frac{1}{n} \sum_{i=1}^n Q\left(g^{-1}\left\{\tilde{\eta}\left(\alpha^\top X_i\right) + \delta^\top B_1(W_i)\right\}, Y_i\right). \quad (13)$$

Then,

$$(\tilde{\tau}, \tilde{\delta}) = \arg \max_{\tau \in \mathbb{R}^{d-1}, \delta \in \mathbb{R}^{N'}} \tilde{l}(\tau, \delta), \quad (14)$$

where $\tilde{l}(\tau, \delta) = \frac{1}{n} \sum_{i=1}^n Q\left(g^{-1}\left\{\tilde{\eta}\left(\alpha^\top(\tau) X_i\right) + \delta^\top B_1(W_i)\right\}, Y_i\right)$ and $\eta(\cdot)$ was replaced by $\tilde{\eta}(\cdot)$.

We define $\tilde{\theta}_\tau = (\tilde{\tau}^\top, \tilde{\gamma}^\top, \tilde{\delta}^\top)^\top$ such that

$$(\tilde{\tau}^\top, \tilde{\gamma}^\top, \tilde{\delta}^\top)^\top = \arg \max_{\tau \in \mathbb{R}^{d-1}, \gamma \in \mathbb{R}^{N_n}, \delta \in \mathbb{R}^{N'}} \frac{1}{n} \sum_{i=1}^n Q\left(g^{-1}\left\{\gamma^\top B\left(\alpha^\top(\tau) X_i\right) + \delta^\top B_1(W_i)\right\}, Y_i\right). \quad (15)$$

Then, $l(\theta_\tau)$ becomes

$$l(\theta_\tau) = \frac{1}{n} \sum_{i=1}^n Q\left(g^{-1}\left\{\gamma^\top B\left(\alpha^\top(\tau) X_i\right) + \delta^\top B_1(W_i)\right\}, Y_i\right) = \frac{1}{n} \sum_{i=1}^n Q\left(g^{-1}\{m_i\}, Y_i\right).$$

The score vector is given by

$$S(\theta_\tau) = \frac{\partial l}{\partial \theta_\tau}(\theta)|_{\theta=\theta_\tau} = \frac{1}{n} \sum_{i=1}^n q_1(m_i, Y_i) \xi_i(\tau, \gamma, \delta), \quad (16)$$

$$\text{where } \xi_i(\tau, \gamma, \delta) = \begin{pmatrix} \gamma^\top B'(\alpha^\top(\tau) X_i) J^\top(\tau) X_i \\ B(\alpha^\top(\tau) X_i) \\ B_1(W_i) \end{pmatrix}.$$

Then, the Hessian matrix of the quasi-likelihood function is

$$H(\theta_\tau) = \frac{1}{n} \sum_{i=1}^n \rho_2(m_i) \xi_i(\tau, \gamma, \delta) \xi_i^\top(\tau, \gamma, \delta).$$

We have $\tilde{\theta}_\tau = (\tilde{\tau}^\top, \tilde{\gamma}^\top, \tilde{\delta}^\top)^\top = \arg \max_{\theta_\tau = (\tau, \gamma, \delta) \in \mathbb{R}^{d-1} \times \mathbb{R}^N \times \mathbb{R}^{N'}}$ $l(\theta_\tau)$. Then, the Fisher scoring update equations become

$$\begin{aligned} \theta_\tau^{(k+1)} &= \theta_\tau^{(k)} + [H(\theta_\tau^{(k)})]^{-1} S(\theta_\tau^{(k)}) \\ &= \theta_\tau^{(k)} + \left[\sum_{i=1}^n \rho_2(m_i^{(k)}) \xi_i(\tau^{(k)}, \gamma^{(k)}, \delta^{(k)}) \xi_i^\top(\tau^{(k)}, \gamma^{(k)}, \delta^{(k)}) \right]^{-1} \\ &\quad \times \left[\sum_{i=1}^n (Y_i - \mu_i^{(k)}) \rho_1(m_i^{(k)}) \xi_i(\tau^{(k)}, \gamma^{(k)}, \delta^{(k)}) \right], \end{aligned} \quad (17)$$

where for $1 \leq i \leq n$,

$$\begin{aligned} m_i^{(k)} &= \gamma^{(k)\top} B(\alpha^{(k)\top}(\tau^{(k)}) X_i) + \delta^{(k)\top} B_1(W_i), \\ \mu_i^{(k)} &= g^{-1}(m_i^{(k)}), \\ \hat{\eta}(t) &= \hat{\gamma}^\top B(t) \approx \gamma^{(k)\top} B(t) = \sum_{j=1}^{N_n} \gamma_j^{(k)} B_j(t), \quad \hat{R}(\cdot) = \sum_{j=1}^{N'} \delta_j^{(k)} B_{1,j}(\cdot), \\ \hat{m}_i &= \hat{\gamma}^\top B(\alpha^\top(\hat{\tau}) X_i) + \delta^{(k)\top} B_1(W_i) \approx \sum_{j=1}^{N_n} \gamma_j^{(k)} B_j(\alpha^\top(\tau^k) X_i) + \sum_{j=1}^{N'} \delta_j^{(k)} B_{1,j}(W_i). \end{aligned}$$

Then, $\hat{\mu}_i = g^{-1}(\hat{m}_i)$ and $\hat{\alpha} = \alpha(\tau^{(k)})$ is the estimator of the single-index coefficient vector of the CVGPLSFIM model. The statistic $\hat{\beta}$ is the estimated functional single-index, and $\hat{R}$ is the estimated nonlinear regression operator R obtained by the CVGPLSFIM model.

3. Main Asymptotic Properties

In this section, we establish the asymptotic properties of the estimators of (i) the nonparametric components, (ii) the parametric components, (iii) the unique index, (iv) the nonlinear regression operator, and (v) the convergence of the estimators of the univariate components. These properties are established under a set of specific assumptions.

3.1. Assumptions

Let φ , φ_1 , and φ_2 be measurable functions on $[a, b]$. We define the empirical inner product and its corresponding norm as follows:

$$\langle \varphi_1, \varphi_2 \rangle_n = \frac{1}{n} \sum_{i=1}^n \varphi_1(U_i) \varphi_2(U_i) \text{ and } \|\varphi\|_n^2 = \frac{1}{n} \sum_{i=1}^n \varphi^2(U_i), \text{ where } U_i = \alpha^\top X_i.$$

If φ , φ_1 , and φ_2 are L^2 -integrable, we define the theoretical inner product and its corresponding norm as follows:

$$\langle \varphi_1, \varphi_2 \rangle = \mathbb{E}[\varphi_1(U) \varphi_2(U)] \text{ and } \|\varphi\|_2^2 = \mathbb{E}[\varphi^2(U)] = \int_a^b \varphi^2(u) f(u) du.$$

Let $\varepsilon = Y - g^{-1}(m_0(T))$, where $T = (X^\top, \bar{W}^\top)^\top$ and $\bar{W} = B_1(W)$. We assume that (C1) The single-index link function $\eta_0(\cdot) \in \mathcal{H}(p)$, where $\mathcal{H}(p)$ is defined as above.

- (C2) For all $m \in \mathbb{R}$ and for all y in the range of the response variable Y , we have for $k = 1, 2$ that

$$q_2(m, y) < 0 \text{ and } c_q < |q_2^k(m, y)| < C_q,$$

for some positive constants c_q and C_q .

- (C3) The ν -th order partial derivative of the joint density function of X satisfies the Lipschitz condition of order κ ($\kappa \in (0, 1]$).

The marginal density function of $\alpha^\top X$ is continuous and bounded away from zero and supported within $[a, b]$.

- (C4) For any vector τ , there exists positive constants c_τ and C_τ such that

$$c_\tau I_{t \times t} \leq \mathbb{E} \left[\begin{pmatrix} 1 \\ T \end{pmatrix} \begin{pmatrix} 1 \\ T \end{pmatrix}^\top \middle| \alpha^\top(\tau)X = \alpha^\top(\tau)x \right] \leq C_\tau I_{t \times t},$$

where $t = 1 + d + N_n + N'$ and $T = (X^\top, \bar{W}^\top)^\top$.

- (C5) The number N_n of knots satisfies $n^{\frac{1}{2(p+1)}} \ll N_n \ll n^{\frac{1}{8}}$ ($p > 3$).

- (C6) The fourth-order moment of the random variable Z is finite, i.e., $\mathbb{E}\|Z(\cdot)\|^4 \leq C$, where C denotes a generic positive constant.

- (C7) The covariance function $K(t, s) = \text{Cov}(Z(t), Z(s))$ is positive definite.

- (C8) For some finite positive constants C_ρ , C_ρ^* , and M_0 ,

$$|\rho_1(m_0)| \leq C_\rho \text{ and } |\rho_1(m) - \rho_1(m_0)| \leq C_\rho^* |m - m_0| \text{ for all } |m - m_0| \leq M_0.$$

- (C9) For some finite positive constants C_g , C_g^* , and M_1 , the link function g in the model (3) satisfies:

$$\left| \frac{d}{dm} g(m) \right|_{m=m_0} \leq C_g \text{ and, for all } |m - m_0| \leq M_1,$$

$$\left| \frac{d}{dm} g^{-1}(m) - \frac{d}{dm} g^{-1}(m) \right|_{m=m_0} \leq C_g^* |m - m_0|.$$

- (C10) There exists a positive constant C_0 such that $\mathbb{E}(\varepsilon^2 | U_{\tau,0}) \leq C_0$.

- (C11) We assume that all the random variables $\langle \beta, Z \rangle$ for all $\beta \in \mathcal{H}$ have values on a set $\mathcal{C}$, where $\mathcal{C}$ is a compact subset of $\mathbb{R}$.

- (C12) The nonlinear regression operator $R(\cdot) \in \mathcal{H}(p)$.

Comments on the Assumptions: The smoothness condition in (C1) describes that the single-index function $\eta_0(\cdot)$ can be approximated by functions in the B-spline space with a normalized basis. On the other hand, condition (C2) ensures the uniqueness of the solution, where condition (C3) is a smoothness condition on the joint and marginal density functions of $\alpha^\top X$ and X . Condition (C5) allows us to obtain the rate of growth of the dimension of the spline spaces relative to the sample size. Conditions (C6) and (C7) are required for the covariate function Z , whereas conditions (C4) and (C8)–(C10) are technical hypotheses that will be needed for the results's proofs. Conditions (C11) and (C12) are smoothness conditions of the nonlinear operator regression R .

3.2. The Consistency Study

3.2.1. Technical Lemmas

In this subsection, we present the needed lemmas to prove Theorems 2 and 3, for which the proofs take into account the behaviors of all the components involved in the model (3).

Lemma 1. Under assumptions (C1)–(C4) and (C6)–(C8), we have

$$\sqrt{n} \begin{pmatrix} \tilde{\tau} - \tau_0 \\ \tilde{\delta} - \delta_0 \end{pmatrix} \xrightarrow{D} \mathcal{N}(0, A^{-1} \Sigma_1 A^{-1}),$$

where Σ_1 and A will be defined below in the Appendix for more details; the symbol $\xrightarrow{D}$ denotes the convergence in distribution and $\bar{W} = B_1(W)$, $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^\top & A_{22} \end{pmatrix}$ with

$$\begin{aligned} A_{11} &= \mathbb{E} \left[\rho_2(m_0(T)) \{ \eta'_0(U_{\tau,0}) \}^2 J^\top(\tau_0) X \bar{W}^\top J(\tau_0) \right], \\ A_{22} &= \mathbb{E} \left[\rho_2(m_0(T)) \bar{W} \bar{W}^\top \right], \\ A_{12} &= \mathbb{E} \left[q_1^2(m_0(T), Y) \cdot \left(\eta'_0(U_{\tau,0}) J^\top(\tau_0) X \right) \left(\eta'_0(U_{\tau,0}) J^\top(\tau_0) X \right)^\top \right] \end{aligned}$$

By applying the δ -method, we obtain the following lemma.

Lemma 2. Under assumptions (C1)–(C4) and (C6)–(C8), we have

$$\sqrt{n} \begin{pmatrix} \alpha(\tilde{\tau}) - \alpha(\tau_0) \\ \tilde{\delta} - \delta_0 \end{pmatrix} \xrightarrow{D} \mathcal{N} \left(0, R(\tau_0) A^{-1} \Sigma_1 A^{-1} R^\top(\tau_0) \right),$$

where

$$R(\tau) = \begin{pmatrix} J(\tau) & 0 \\ 0 & I_{N' \times N'} \end{pmatrix},$$

and

$$\Sigma_1 = \mathbb{E} \left[q_1^2(m_0(T)) \left(\eta'_0(U_{\tau,0}) J^\top(\tau_0) X \right) \left(\eta'_0(U_{\tau,0}) J^\top(\tau_0) X \right)^\top \right].$$

Furthermore, $\alpha(\tilde{\tau}) - \alpha(\tau_0) = o_{\mathbb{P}} \left(\frac{1}{\sqrt{n}} \right)$ and $\tilde{\delta} - \delta_0 = o_{\mathbb{P}} \left(\frac{1}{\sqrt{n}} \right)$.

Lemma 3. Under assumptions (C1)–(C5), we have

$$\|\hat{\theta} - \tilde{\theta}\| = O_{\mathbb{P}} \left\{ \sqrt{N_n} \left(h^p + \frac{1}{\sqrt{nh}} \right) \right\}. \quad (18)$$

where N_n is the number of B-splines basis functions of order r .

The proofs of the previous results are supported by the following lemmas.

3.2.2. Convergence of the Estimated Univariate Components

So, for the nonlinear regression operator R behavior, we have the following theorem:

Theorem 1. Under assumptions (C1)–(C8) and (C11)–(C12), we have

$$\|\hat{R} - R\|_2 = O_{\mathbb{P}} \left\{ \sqrt{N_n} \left(\frac{1}{\sqrt{nh}} + h^p \right) \right\}.$$

The proof of Theorem 1 will be given in the Supplementary Materials.

3.2.3. Estimation of the Systematic Component Function

Theorem 2. Under assumptions (C1)–(C7), we have

$$\|\hat{\eta} - \eta_0\|_2 = O_{\mathbb{P}} \left\{ \sqrt{N_n} \left(\frac{1}{\sqrt{nh}} + h^p \right) \right\}.$$

and

$$\|\hat{\eta} - \eta_0\|_n = O_{\mathbb{P}} \left\{ \sqrt{N_n} \left(\frac{1}{\sqrt{nh}} + h^p \right) \right\}.$$

The proof of Theorem 2 will be given in the Supplementary Materials. This proof takes into account the components of our model, which makes it different from the results obtained by Alahiane et al. [28,29].

3.2.4. Estimation of the Parametric Components

The next theorem shows that the maximum quasi-likelihood estimator is root- n -consistent and is asymptotically normal, although the convergence rate of the nonparametric component $\hat{\eta}$ is slower than root- n . Before enouncing the theorem, let us denote

$$Y(u_{\tau,0}) = \frac{\mathbb{E}[X\rho_2(m_0(T))|U_{\tau,0} = u_{\tau,0}]}{\mathbb{E}[\rho_2(m_0(T))|U_{\tau,0} = u_{\tau,0}]}, \quad \Gamma(u_{\tau,0}) = \frac{\mathbb{E}[W\rho_2(m_0(T))|U_{\tau,0} = u_{\tau,0}]}{\mathbb{E}[\rho_2(m_0(T))|U_{\tau,0} = u_{\tau,0}]},$$

$$\Phi(x) = \Phi(U_{\tau,0}, x) = x - Y(u_{\tau,0}) \quad \text{and} \quad \Psi(w) = \Psi(U_{\tau,0}, w) = w - \Gamma(u_{\tau,0}).$$

Theorem 3. Under assumptions (C1)–(C10), the constrained quasi-likelihood estimators $\hat{\alpha}$ and $\hat{\delta}$ with $\|\hat{\alpha}\|_d = 1$ are jointly asymptotically normally distributed, i.e.,

$$\sqrt{n} \begin{pmatrix} \hat{\alpha} - \alpha_0 \\ \hat{\delta} - \delta_0 \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}(0, R(\tau_0)D^{-1}R^{\top}(\tau_0)),$$

where $\xrightarrow{\mathcal{D}}$ denotes the convergence in distribution,

$$D = \mathbb{E} \left[\rho_2(m_0(T)) \begin{pmatrix} \eta'_0(U_{\tau,0})J^{\top}(\tau_0)\Phi(X) \\ \Psi(W) \end{pmatrix} \begin{pmatrix} \eta'_0(U_{\tau,0})J^{\top}(\tau_0)\Phi(X) \\ \Psi(W) \end{pmatrix}^{\top} \right],$$

and

$$\alpha(\hat{\tau}) - \alpha(\tau_0) = O_{\mathbb{P}} \left(\frac{1}{\sqrt{n}} \right), \quad \hat{\delta} - \delta_0 = O_{\mathbb{P}} \left(\frac{1}{\sqrt{n}} \right).$$

where $O_{\mathbb{P}}$ denotes the Bachmann–Landau notation “in probability”.

Proof of Theorem 3. The proof of Theorem 3 is given in the Supplementary Materials. This proof takes into account all the components involved in the model (3), which makes it different from the results obtained by Alahiane et al. [28,29]. $\square$

4. A Simulation Study

We aim to show the performance of various estimators of the parameters τ , γ , δ , the nonparametric function η , the functional index β , and the nonlinear regression operator R of the model (3) through numerical simulations under both the Gaussian and the logistic cases. The conditional density of Y given $X = x, Z = z$ is described by Equation (1).

We believe that the model is given by the following equation:

$$g(\mu(X_i, Z_i)) = \sin \left\{ \frac{\pi(\alpha^{\top} X_i - A)}{B - A} \right\} + R(\langle \beta, Z_i \rangle) + \varepsilon_i, \quad \text{for } i = 1, \dots, n. \quad (19)$$

The responses Y_i are simulated according to the Equation (19); X_i are taken uniformly over the interval $[-0.5, 0.5]$, whereas the errors $\varepsilon_i \sim \mathcal{N}(0, 0.025)$. Moreover, we take the following coefficients:

$$\alpha = \frac{1}{\sqrt{3}}(1, 1, 1)^{\top}, \quad A = \frac{\sqrt{3}}{2} - \frac{1.645}{\sqrt{12}} \quad \text{and} \quad B = \frac{\sqrt{3}}{2} + \frac{1.645}{\sqrt{12}}.$$

The functional real variable $Z_i(\cdot)$ is taken as $Z(t) = a \cos(2\pi t) + b \sin(4\pi t) + 2c(t - 0.25)(t - 0.5)$, $t \in [-1, 1]$, where $a \sim \mathcal{U}(0, 1)$, $b \sim \mathcal{U}(0, 1)$ and $c \sim \mathcal{U}(0, 1)$. A set of 1000 independent curves is generated according to the following model: $Z_i(t) = a_i \cos(2\pi t) + b_i \sin(4\pi t) + 2c_i(t - 0.25)(t - 0.5)$, $t \in [-1, 1]$, where a_i , b_i , and c_i are uniformly distributed on $[0, 1]$, respectively. Curves are discretized over a very fine mesh of one thousand equispaced points; the set of curves is stored in the matrix $\mathcal{Z} = [Z_i(t_j)]$, $i = 1, \dots, 1000$, $j = 1, \dots, 1000$. A random selection of 30 of these functional data is plotted in Figure 1.

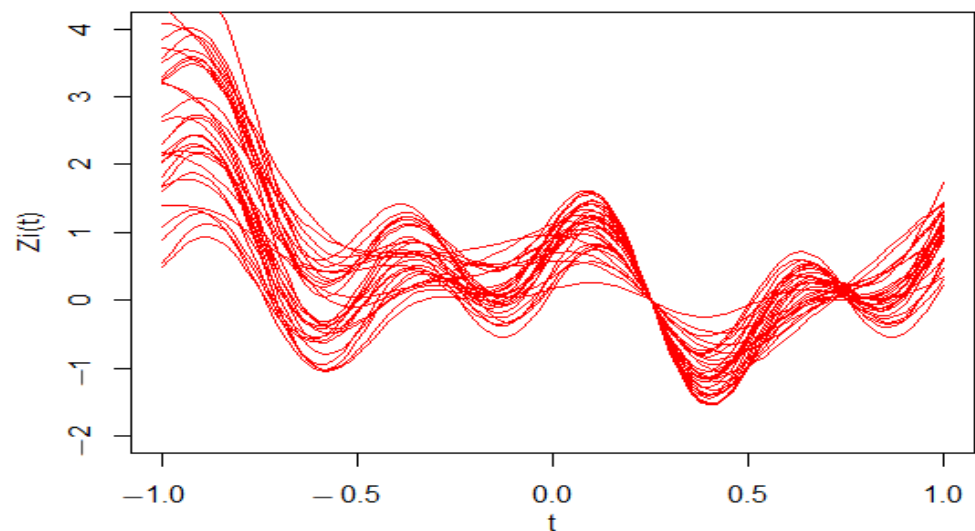


Figure 1. A random selection of 30 simulated curves of Z .

We consider the functional index $\beta(t) = \frac{1}{\sqrt{2}} \left[\sin\left(\frac{3}{2}\pi t\right) + \sin\left(\frac{1}{2}\pi t\right) \right]$ and the regression operator $R(u) = \left\{ \frac{1}{\sqrt{2}} \int_{-1}^1 \left[\sin\left(\frac{3}{2}\pi t\right) + \sin\left(\frac{1}{2}\pi t\right) \right] Z(t) dt \right\}^3$.

We set directions, and we use the first to the fourth eigenfunctions of the variance operator of Z ; note that the cumulative explained variances of the corresponding linear principal components are 71.3%, 87.5%, 98.8%, and 99.6%.

In the first step of our algorithm, we base our simulation experiments on samples of 300 couples (Z_i, Y_i) that we randomly extracted; we use the first 200 couples as the training set and the remaining 100 couples as the test set. We do not use all the columns of $\mathcal{Z}$, only a selection. At each step, we use cubic splines, and the number of knots is set to six.

Concerning the nonparametric functional estimator, the semimetric used is the standard distance L_2 between the curves.

In the second step, the nodes are selected according to the formula

$$C n^{\frac{1}{2r}} \log(n), \text{ where } C \in [0.3, 1] \text{ (see [11])}.$$

We choose $C = 0.6$, and we perform replications of 3000 samples of sizes $n = 500$.

Through the plug-in process (the second step), we estimate the model parameters (8) using the GPLFSIM algorithm as described previously.

$$g(\mu(X_i, Z_i)) = \eta(\alpha^\top X_i) + R(\langle \hat{\beta}, Z_i \rangle) \text{ for } i = 1, \dots, n. \quad (20)$$

Then, the computed bias, the standard deviation (SD), and the mean squared error (MSE) with respect to the parameter τ , the parameter γ , and the parameter δ are summarized in Tables 1–7.

Table 1. Bias, SD, and MSE according to the parameter τ for CVGPLFSIM with the identity link function and $n = 500$ in both the Gaussian and the logistic cases.

Sample Size	Gaussian Case ¹		Logistic Case ²	
	τ_1	τ_2	τ_1	τ_2
Bias	0.0013	−0.0012	0.0027	−0.0061
SD	0.0012	0.0034	0.0031	0.0201
MSE	3.13×10^{-6}	1.3×10^{-5}	4.45×10^{-5}	4.14×10^{-4}

Note: This table summarizes the bias, SD, and MSE of τ with sample size $n = 500$. ¹ τ for the Gaussian case. ² τ for the logistic case.

Table 2. Bias, SD, and MSE evolutions with respect to the parameter γ variation for CVPGLFSIM with the identity link function and $n = 500$.

	γ_1	γ_2	γ_3	γ_4	γ_5
Bias	−0.0026	0.0142	−0.0231	0.0242	−0.0037
SD	0.0102	0.0165	0.0141	0.0140	0.0071
MSE	1.1080×10^{-4}	4.7389×10^{-4}	7.3242×10^{-4}	7.8164×10^{-4}	6.4100×10^{-5}

Table 3. Bias, SD, and MSE evolutions with respect to the parameter γ variation for CVGPLFSIM with the identity link function and $n = 500$.

	γ_6	γ_7	γ_8	γ_9	γ_{10}
Bias	0.0032	0.0023	−0.0021	0.0012	−0.0045
SD	0.0033	0.0041	0.0014	0.0027	0.0042
MSE	2.113×10^{-5}	2.21×10^{-5}	6.37×10^{-6}	8.73×10^{-6}	3.789×10^{-5}

Table 4. Bias, SD, and MSE evolutions with respect to the parameter δ variation for CVGPLFSIM with the identity link function and $n = 500$.

	δ_1	δ_2	δ_3	δ_4	δ_5
Bias	0.0009	0.0037	−0.0045	0.0082	−0.0035
SD	0.0036	0.0012	0.0071	0.0038	0.0091
MSE	1.377×10^{-5}	1.513×10^{-5}	7.066×10^{-5}	8.168×10^{-5}	2.450×10^{-5}

Table 5. Bias, SD, and MSE evolutions with respect to the parameter γ variation for CVGPLFSIM with the logistic link function and $n = 500$.

	γ_1	γ_2	γ_3	γ_4	γ_5
Bias	−0.0032	0.0156	−0.0233	0.0341	−0.0079
SD	0.0107	0.0126	0.0306	0.0321	0.0042
MSE	1.2473×10^{-4}	4.0212×10^{-4}	1.4792×10^{-3}	2.1932×10^{-3}	8.005×10^{-5}

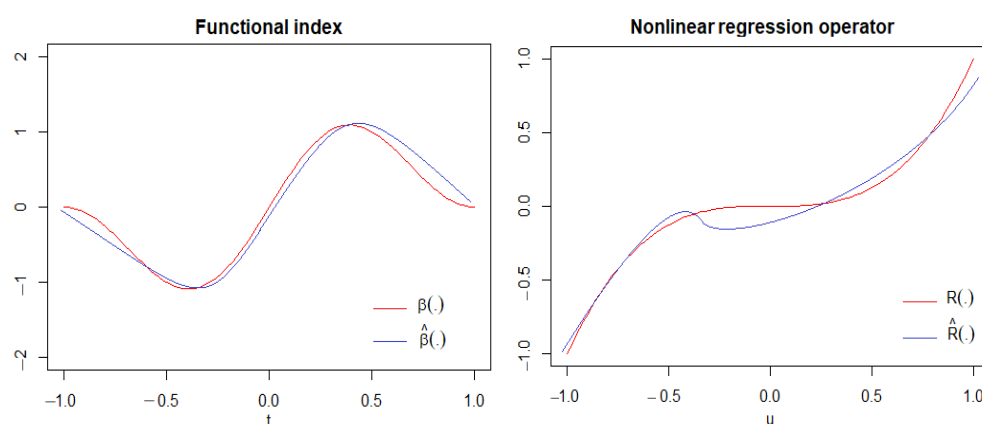
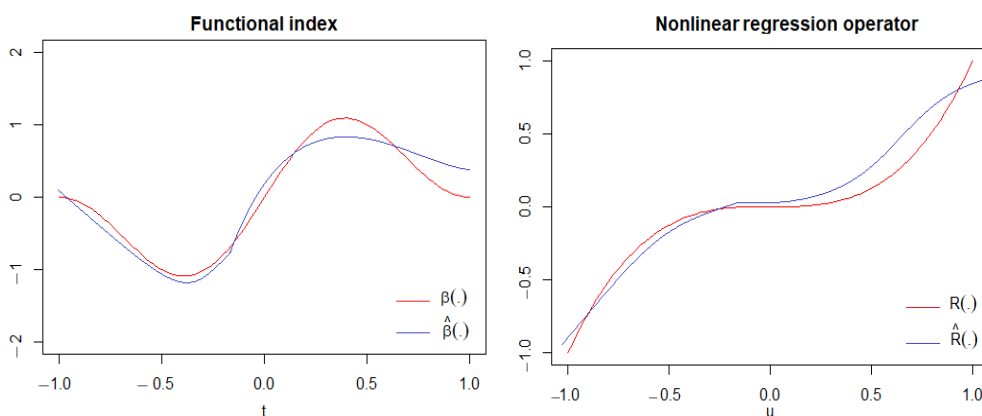
Table 6. Bias, SD, and MSE evolutions with respect to the parameter γ variation for CVGPLFSIM with the logistic link function and $n = 500$.

	γ_6	γ_7	γ_8	γ_9	γ_{10}
Bias	0.0063	0.0022	0.0327	0.0005	−0.0045
SD	0.0051	0.0062	0.0027	0.0091	0.0021
MSE	6.57×10^{-5}	4.32×10^{-5}	1.076×10^{-3}	8.306×10^{-5}	2.466×10^{-3}

Table 7. Bias, SD, and MSE evolutions with respect to the parameter δ variation for CVGPLFSIM with the logistic link function and $n = 500$.

	δ_1	δ_2	δ_3	δ_4	δ_5
Bias	0.0036	0.0038	−0.0029	0.0145	−0.0169
SD	0.0054	0.0037	0.0122	0.0231	0.0132
MSE $\times 10^{-4}$	4.212×10^{-1}	2.813×10^{-1}	1.5725	7.4386	4.5985

We present below in Figures 2 and 3 the functional index and nonlinear regression operator obtained by the GPLFSIM algorithm in both cases: the Gaussian and the logistic cases. We chose the two distributions (Gaussian and logistic) because of their popularity and because of the use of activation functions (sigmoid functions) in neural networks, especially when the input data are high-dimensional.

**Figure 2.** The Gaussian case: estimators of $\hat{\beta}(\cdot)$ (the left plot) and estimators of the nonlinear regression operator $\hat{R}(u)$ (the right plot), where u stands for $\langle \hat{\beta}, Z \rangle$.**Figure 3.** The logistic case: estimators of $\hat{\beta}(\cdot)$ (the left plot) and estimators of the nonlinear regression operator $\hat{R}(u)$ (the right plot), where u stands for $\langle \hat{\beta}, Z \rangle$.

In order to compute the bias, SD, and MSE, we recorded 3000 replications of the CVGPLFSIM algorithm in the Gaussian case and in the logistic case with $n = 500$ as follows (see Tables 1–7).

It is obvious to remark that the quality of the estimators is illustrated through these simulations, as the method works quite well. The bias, SD, and MSE are generally reasonably low. The parametric and the nonparametric components, the single-index α , the functional index $\beta(\cdot)$, and the nonlinear regression operator R of Y over X, Z are calculated by the procedure described above. The two tables therefore indicate the consistency of $\hat{\alpha}$ and $\hat{\delta}$ as the bias, SD, and MSE decrease as the sample size increases.

We have developed our algorithm in both cases, the identity link function and the logistic link function. The simulations show that the CVPGPLSFIM algorithm works well in both cases. We present below in Figure 4 the single index estimated by the model in both cases, the Gaussian and logistic cases.

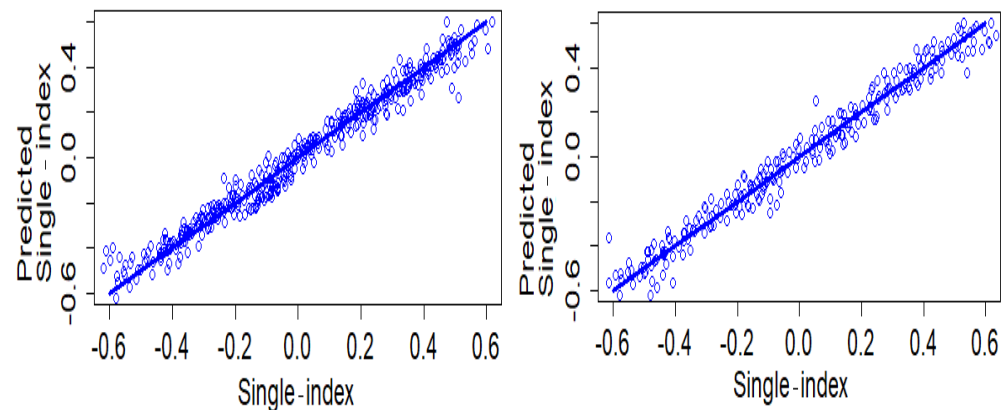


Figure 4. On the **left** plot: single-index α versus predicted single-index $\hat{\alpha}$: Gaussian case; on the **right** plot: single-index α versus predicted single-index $\hat{\alpha}$: logistic case.

We observe that the single-index estimated by our model fits well with the single-index.

We present below in Figure 5 the systematic component η estimated by the model in both cases: the Gaussian and the logistic cases.

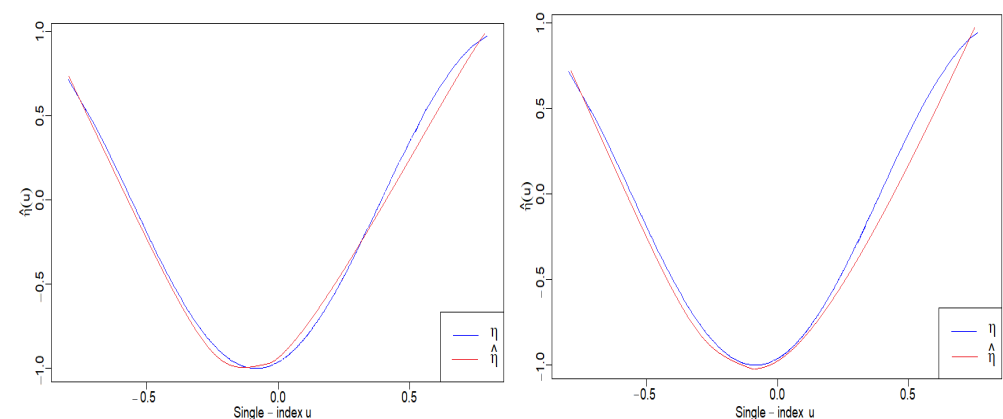


Figure 5. The function η versus its estimator $\hat{\eta}$ for the Gaussian case (**left** plot), and the function η versus its estimator $\hat{\eta}$ for the logistic case (**right** plot).

We consider the root of the averaged squared error criterion (see [35]) in both the Gaussian and the logistic cases:

$$\text{RASE}_1 = \left(\frac{1}{n} \sum_{i=1}^n (\hat{\eta}(u_i) - \eta(u_i))^2 \right)^{\frac{1}{2}} \quad \text{and} \quad \text{RASE}_2 = \left(\frac{1}{n} \sum_{i=1}^n (\hat{R}(u_i) - R(u_i))^2 \right)^{\frac{1}{2}}$$

Tables 8 and 9 summarize the samples means, medians, and variances of the RASE_1 and RASE_2 with different sample sizes in both the Gaussian and logistic cases.

Table 8. The RASE criterion for the function η for both cases and for $n = 500$ and $n = 1000$.

Sample Size	Gaussian Case			Logistic Case		
	Mean	Median	Variance	Mean	Median	Variance
$n = 500$	0.033	0.035	0.012	0.033	0.035	0.012
$n = 1000$	0.031	0.032	0.011	0.031	0.032	0.011

We conclude that as the sample size n increases from 500 to 1000, the sample mean, median and variance of $RASE_1$ decreases.

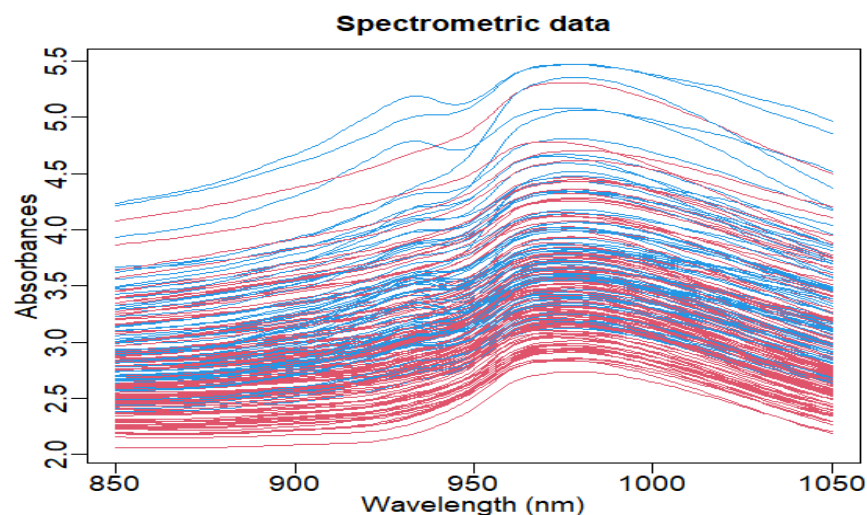
Table 9. The $RASE_2$ criterion with the nonlinear regression operator R in the both cases and for $n = 500$ and $n = 1000$.

Sample Size	Gaussian Case			Logistic Case		
	Mean	Median	Variance	Mean	Median	Variance
$n = 500$	0.035	0.037	0.006	0.046	0.039	0.031
$n = 1000$	0.031	0.032	0.004	0.042	0.033	0.025

We conclude that when the sample size n increases from 500 to 1000, the sample mean, the median, and the variance of $RASE_2$ decreases.

5. Application for Tecator Data

In this section, we employ the CVGPLFSIM model to analyze the Tecator data, which is a well-known dataset in the field of FDA. The dataset can be obtained from the following link: <http://lib.stat.cmu.edu/datasets/tecator> (accessed on 1 March 2024). It consists of 215 finely chopped meat samples, each associated with its corresponding fat content (Y_i for $i = 1, \dots, 215$), near-infrared absorbance spectra (Z_i for $i = 1, \dots, 215$) measured at 100 wavelengths ranging from 850 to 1050 nm, as well as the protein content $X_{1,i}$ and moisture content $X_{2,i}$ of the meat samples. For more comprehensive information and insights, we refer readers to Ferraty et al. [13]. We aim to forecast the fat content of the finely chopped meat samples. Figure 6 shows a sample of the absorbance curves.

**Figure 6.** Sample of 100 absorbance curves Z .

In order to evaluate the effectiveness of the model (3), we employ a random splitting of the sample into two subsets: a training subset, denoted as I_1 , consisting of 160 observations, and a test subset, denoted as I_2 , consisting of 55 observations. The purpose of the training subset is to estimate the model parameters, while the test subset is used to assess the

accuracy of the predictors. We utilize the mean square error of prediction (MSEP), as defined in Aneiros et al. [14] and given by

$$\text{MSEP} = \frac{1}{55} \sum_{i \in I_2} (Y_i - \hat{Y}_i)^2 / \text{var}_{I_2}(Y_i).$$

where $\hat{Y}_i$ represents the predicted value based on the training subset and $\text{var}_{I_2}(Y_i)$ denotes the variance of the response variables in the test subset. This indicator allows us to assess the accuracy of our predictions with respect to the variability in the test dataset.

The performance comparison of the CVGPLFSIM model with other models is presented in Tables 10 and 11. Based on the obtained results, we can infer that the CVGPLFSIM model demonstrates competitiveness and effectiveness in analyzing the given dataset.

Tables 10 and 11 show the performance of the CVGPLFSIM model by comparing it with other models. The CVGPLFSIM model is a competitive one for such data.

Table 10. MSEP for different models in the Gaussian case.

Functional Models		MSEP
M-1 (CVGPLFSIM)	$g(\mu_i(X_i, Z_i)) = \eta(\alpha_1 X_{1,i} + \alpha_2 X_{2,i}) + R(\langle \beta, Z_i \rangle) + \varepsilon_i$	0.056
M-2 (GNP-FPLSIM)	$g(\mu_i(X_i, Z_i)) = \eta(\alpha_1 X_{1,i} + \alpha_2 X_{2,i}) + R(Z_i) + \varepsilon_i$	0.082
M-3 (FPLSIM)	$Y_i = \eta(\alpha_1 X_{1,i} + \alpha_2 X_{2,i}) + R(Z_i) + \varepsilon_i$	0.102
M-4 (SIM)	$Y_i = \eta(\alpha_1 X_{1,i} + \alpha_2 X_{2,i}) + \varepsilon_i$	1.102

Table 11. MSEP for different models in the logistic case.

Functional Models		MSEP
M-1 (CVGPLFSIM)	$g(\mu_i(X_i, Z_i)) = \eta(\alpha_1 X_{1,i} + \alpha_2 X_{2,i}) + R(\langle \beta, Z_i \rangle) + \varepsilon_i$	0.045
M-2 (GNP-FPLSIM)	$g(\mu_i(X_i, Z_i)) = \eta(\alpha_1 X_{1,i} + \alpha_2 X_{2,i}) + R(Z_i) + \varepsilon_i$	0.067
M-3 (FPLSIM)	$Y_i = \eta(\alpha_1 X_{1,i} + \alpha_2 X_{2,i}) + R(Z_i) + \varepsilon_i$	0.102
M-4 (SIM)	$Y_i = \eta(\alpha_1 X_{1,i} + \alpha_2 X_{2,i}) + \varepsilon_i$	1.102

Moreover, Figure 7 shows the nonparametric estimator of the function η , in both the Gaussian and the logistic cases.

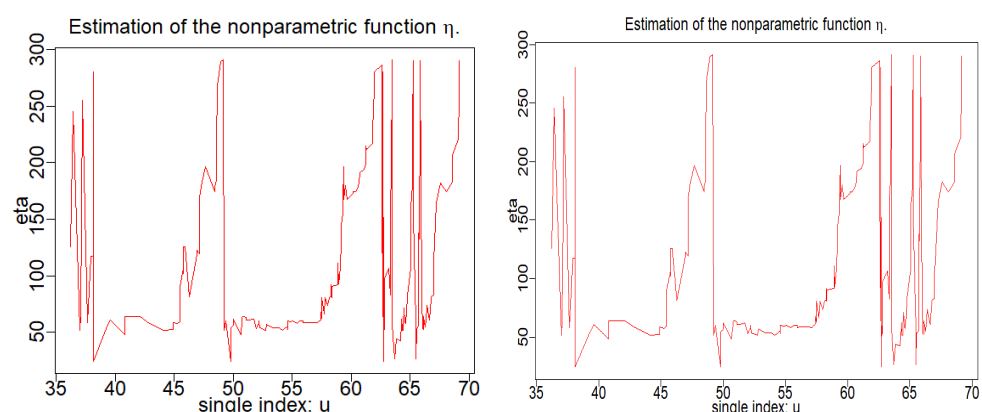


Figure 7. The estimator $\hat{\eta}(\cdot)$ in the Gaussian case (left plot), and the estimator $\hat{\eta}(\cdot)$ in the logistic case (right plot).

Figures 8 and 9 show the estimator functional index $\hat{\beta}(\cdot)$ and the estimator of the nonlinear regression operator $\hat{R}$.

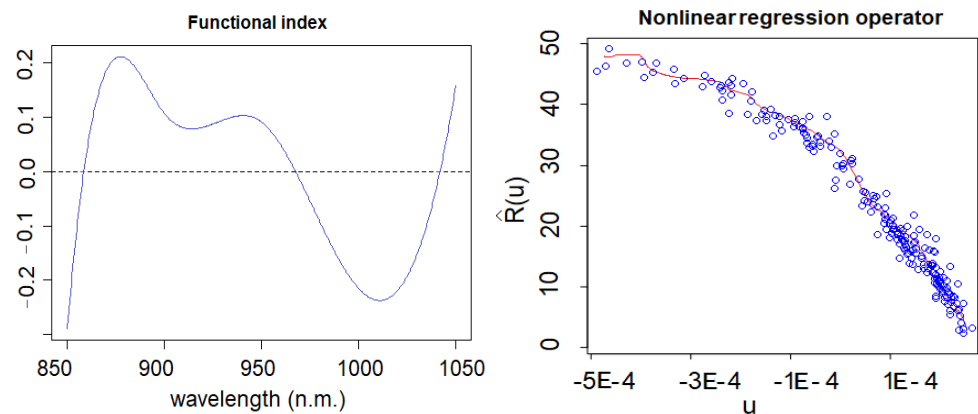


Figure 8. The Gaussian case: the estimator $\hat{\beta}$ (left plot) and the estimator $\hat{R}$ (right plot).

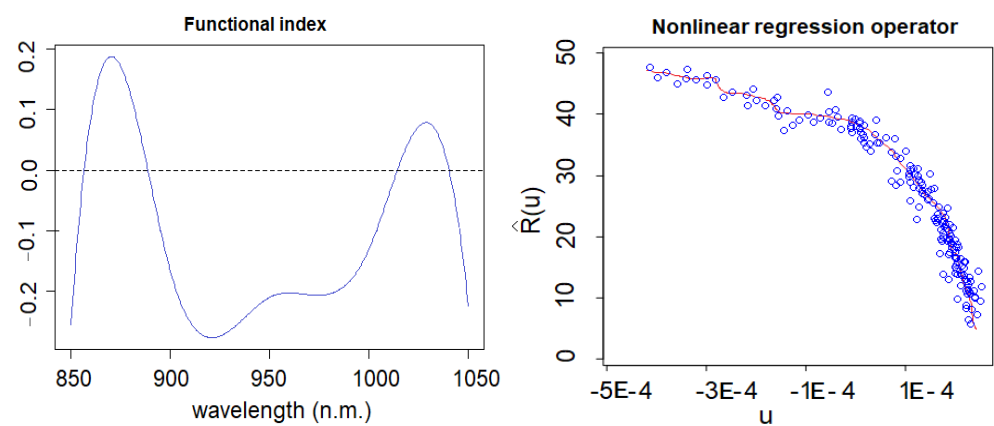


Figure 9. The logistic case: the estimator $\hat{\beta}$ (left plot) and the estimator $\hat{R}$ (right plot).

Figure 10 illustrates the difference between the fat content and its estimation from the model for both the Gaussian and the logistic cases.

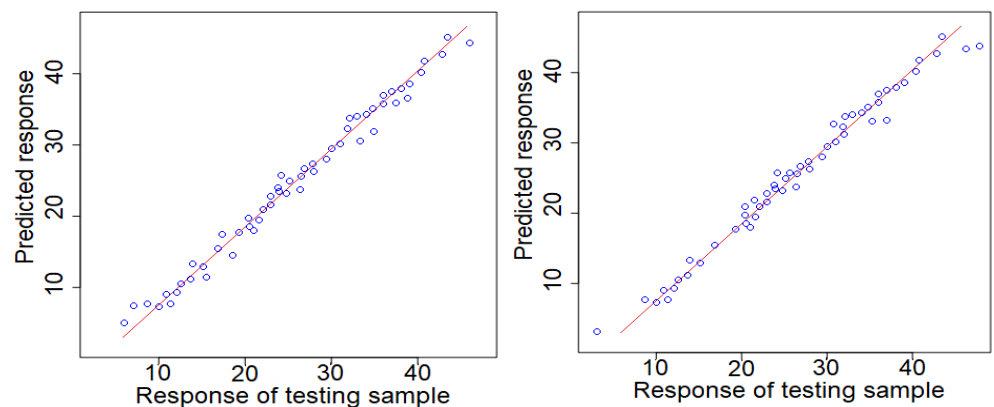


Figure 10. The content of fatness and its estimation: the Gaussian case (left plot) and the logistic case (right plot).

We can see that our model fits well the content of fatness “215 pieces of meat”.

Possible extensions to our manuscript include hypothesis testing and confidence bands using the bootstrap method, the extension of the activation function (sigmoid functions) in the context of neural networks (deep learning) when the input data are high-dimensional, and functional projection pursuit regression for GFPLSIM models $g(\mu(X, Z)) = \eta(\alpha^\top X) + r(Z)$, where $r(Z) = \sum_{i=1}^m \langle \beta_i, Z \rangle$ -type models take advantage of the information required from the various revealing directions.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1. In the supplementary material, we present the appendix and proofs of the various results obtained in this paper. References [7,10,14,34,36,37] are cited in Supplementary Materials.

Author Contributions: Conceptualization, M.R., M.A., I.O., A.A. and L.H.; Methodology, M.R., M.A., I.O., A.A. and L.H.; Software, M.R., M.A., I.O., A.A. and L.H.; Validation, M.R., M.A., I.O., A.A. and L.H.; Formal analysis, M.R., M.A., I.O., A.A. and L.H.; Investigation, M.R., M.A., I.O., A.A. and L.H.; Resources, M.R., M.A., I.O., A.A. and L.H.; Data curation, M.R., M.A., I.O., A.A. and L.H.; Writing—original draft, M.R., M.A., I.O., A.A. and L.H.; Writing—review & editing, M.R., M.A., I.O., A.A. and L.H.; Visualization, M.R., M.A., I.O., A.A. and L.H.; Supervision, M.R., M.A. and I.O.; Project administration, M.R., M.A. and I.O.; Funding acquisition, M.A. and I.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: As in Ferraty et al. [13], Aneiros et al. [14], and Alahiane et al. [28,29], to value our theoretical results, we consider the public spectrometric data widely used in FDA and that are available at the website <http://lib.stat.cmu.edu/datasets/tecator>, accessed on 1 March 2024.

Acknowledgments: The authors thank the editor and the reviewers for their helpful and constructive comments.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

Appendix A

We will present technical lemmas that will be used for proving Theorems 2 and 3. In what follows, for all probability measures Q , we define

$$L^2(Q) = \left\{ f \text{ such that } Qf^2 = \int f^2 dQ < \infty \right\}.$$

Let $\mathcal{F}$ be a subclass of $L^2(Q)$. For all $f \in \mathcal{F}$, $\|f\| = \left(\int f^2 dQ \right)^{\frac{1}{2}}$.

Denote $N(\delta, \mathcal{F}, L^2(Q))$ as the δ -covering number of $\mathcal{F}$, i.e., the smallest value of N for which there exist functions $f_1, f_2, \dots, f_N$ (which are not necessarily in $\mathcal{F}$), such that for each $f \in \mathcal{F}$, there exists $j \in \{1, \dots, N\}$, $\|f - f_j\| < \delta$ or $\mathcal{F} \subset \bigcup_{j=1}^N B(f_j, \delta)$.

For two functions l and u , a bracketing $[l, u]$ is the set of functions f such that $l \leq f \leq u$, $[l, u] = \{f \text{ such that } l \leq f \leq u\}$.

The δ -covering number with bracketing $N_{[\cdot]}(\delta, \mathcal{F}, L^2(Q))$ is defined as the smallest value of N , necessary to cover the whole $\mathcal{F}$, for which there exist pairs of functions $\left\{ [f_j^L, f_j^U] \text{ for } j = 1, \dots, N \right\}$ with $\|f_j^U - f_j^L\| \leq \delta$, such that for each $f \in \mathcal{F}$, there is a $j \in \{1, \dots, N\}$ such that $f_j^L \leq f \leq f_j^U$ (f_j^U and f_j^L are not necessarily belonging to $\mathcal{F}$).

The δ -entropy with bracketing is $\log N_{[\cdot]}(\delta, \mathcal{F}, L^2(Q))$. The uniform entropy integral $J_{[\cdot]}(\delta, \mathcal{F}, L^2(Q))$ is defined as

$$J_{[\cdot]}(\delta, \mathcal{F}, L^2(Q)) = \int_0^\delta \left\{ 1 + \log N_{[\cdot]}(\kappa, \mathcal{F}, L^2(Q)) \right\}^{\frac{1}{2}} d\kappa.$$

Let Q_n be the empirical measure of Q , i.e., $Q_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}(\cdot)$ such that

$$Q_n f = \mathbb{E}^{Q_n}[f] = \int f dQ_n = \frac{1}{n} \sum_{i=1}^n \int f \delta_{X_i} = \frac{1}{n} \sum_{i=1}^n f(X_i).$$

Denote $G_n = \sqrt{n}(Q_n - Q)$ as the standardized empirical process indexed by $\mathcal{F}$ and $\|G_n\|_{\mathcal{F}} = \sum_{f \in \mathcal{F}} |G_n f|$ for any measurable class of functions $\mathcal{F}$.

For all $f \in \mathcal{F}$, we have $Qf = \mathbb{E}^Q[f(X)] = \int f dQ$, and

$$G_n f = \sqrt{n}(Q_n f - Qf) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (f(X_i) - \mathbb{E}[f(X)]).$$

Lemma A1 (Lemma 3.4.2. in Van Der Van et al. [38]). Let $M_0 > 0$ and $\mathcal{F}$ be uniformly bounded class of measurable functions such that

$$\begin{cases} \text{for all } f \in \mathcal{F}, \|f\|_{\infty} < M_0, \\ Qf^2 < \delta^2. \end{cases}$$

Then

$$\mathbb{E}^Q[\|G_n\|_{\mathcal{F}}] \leq c_0 J_{[]}(\delta, \mathcal{F}, L^2(Q)) \left\{ 1 + \frac{J_{[]}(\delta, \mathcal{F}, L^2(Q))}{\delta^2 \sqrt{n}} M_0 \right\},$$

where c_0 is a finite constant not dependent on n .

Lemma A2 (Lemma A.1. in Huang [39]). For any $\lambda > 0$, let $\Theta_n = \{\eta(\alpha_0^\top x) \text{ such that } \|\delta - \delta_0\| \leq \lambda, \eta \in S_n, \|\eta - \eta_0\|_2 \leq \lambda\}$. Then, for any $\epsilon \leq \lambda$

$$\log N_{[]}(\lambda, \Theta_n, L^2(\mathbb{P})) \leq C N_n \log\left(\frac{\lambda}{\epsilon}\right).$$

Lemma A3 (Lemma A.2. in Wang et al. [40] and Lemma A.4. in Xue et al. [41]). Under assumptions (C1)–(C5), we have

$$A_n = \sup_{\eta_1, \eta_2 \in S_n} \left| \frac{\langle \eta_1, \eta_2 \rangle_n - \langle \eta_1, \eta_2 \rangle}{\|\eta_1\|_2 \|\eta_2\|_2} \right| = O_{a.co.} \left\{ \sqrt{\frac{\log n}{nh}} \right\},$$

where $O_{a.co.}$ denotes the “O” Lanadau symbol for the almost-complete convergence.

$$\text{Let } D_{n,\theta} = \begin{pmatrix} \gamma^\top B'(\alpha^\top(\tau) X_i) J^\top(\tau) & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & B(\alpha^\top(\tau) X_i) \end{pmatrix}; \text{ we denote } T_i = \begin{pmatrix} X_i^\top & \bar{W}_i^\top \end{pmatrix},$$

$$W_{n,\theta} = \frac{1}{n} \sum_{i=1}^n D_{i,\theta} \begin{pmatrix} T_i \\ 1 \end{pmatrix} \begin{pmatrix} T_i \\ 1 \end{pmatrix}^\top D_{i,\theta}^\top, \text{ and } W_\theta = \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[D_{i,\theta} \begin{pmatrix} T_i \\ 1 \end{pmatrix} \begin{pmatrix} T_i \\ 1 \end{pmatrix}^\top D_{i,\theta}^\top \right].$$

Then, we have the following lemma.

Lemma A4 (Lemma A.3 in the Supplementary Material of Wang et al. [40]). Under assumptions (C1)–(C8), there exists $C > 0$ such that

$$\sup_{\theta} \|W_\theta^{-1}\|_2 \leq C \sqrt{N_n} \text{ a.co. and } \sup_{\theta} \|W_{n,\theta}^{-1}\|_2 \leq C \sqrt{N_n} \text{ a.co.,}$$

$$\text{where } \|M\|_2 = \sup_{x \neq 0} \frac{\|Mx\|}{\|x\|} = \sup_{\|x\|=1} \|Mx\|.$$

In what follows, we will give lemmas that allow us to prove Theorem 3. The lemmas and theorem proofs will be developed in the Supplementary Materials. The proofs of these

lemmas are different from those of Alahiane et al. [28,29] as they take into account the components of our new model.

Lemma A5. Under conditions (C1)–(C8), we have

$$\frac{1}{n} \sum_{i=1}^n \rho_2(m_{0i}) \{\hat{\eta}(U_{\tau,0i}) - \eta_0(U_{\tau,0i})\} \eta'_0(U_{\tau,0i}) J^\top(\tau_0) \Phi(X_i) = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right), \quad (\text{A1})$$

$$\frac{1}{n} \sum_{i=1}^n \rho_2(m_{0i}) \eta'_0(U_{\tau,0i}) \Phi(X_i) Y^\top(U_{\tau,0i}) J(\tau_0) (\hat{\tau} - \tau_0) = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right), \quad (\text{A2})$$

$$\frac{1}{n} \sum_{i=1}^n \rho_2(m_{0i}) \eta'_0(U_{\tau,0i}) \Phi(X_i) \Gamma^\top(U_{\tau,0i}) J(\tau_0) (\hat{\delta} - \delta_0) = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right). \quad (\text{A3})$$

Lemma A6. Under conditions (C1)–(C8), we have

$$\frac{1}{n} \sum_{i=1}^n \rho_2(m_{oi}) \{\hat{\eta}(U_{\tau,oi}) - \eta_0(U_{\tau,oi})\} \Psi(T_i) = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right), \quad (\text{A4})$$

$$\frac{1}{n} \sum_{i=1}^n \rho_2(m_{oi}) \eta'_0(U_{\tau,oi}) \Psi(T_i) Y^\top(U_{\tau,oi}) J(\tau_0) (\hat{\tau} - \tau_0) = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right), \quad (\text{A5})$$

$$\frac{1}{n} \sum_{i=1}^n \rho_2(m_{oi}) \Psi(U_{\tau,oi}, Z_i) \Gamma^\top(U_{\tau,oi}) (\hat{\delta} - \delta_0) = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right). \quad (\text{A6})$$

References

- McCullagh, P.; Nelder, J. *Generalized Linear Models*; Routledge: London, UK, 1989; Volume 37.
- Nelder, J.; Wedderburn, R. Generalized linear models. *J. R. Stat. Soc. Ser. (Stat. Soc.)* **1972**, *135*, 370–384.
- Hastie, T.; Tibshirani, R. *Generalized Additive Models*; Chapman & Hall/CRC: New York, NY, USA, 1990; Volume 43.
- Wood, S. *Generalized Additive Models. An Introduction with R*; CRC/Taylor & Francis: New York, NY, USA, 2017.
- Hardle, W.; Hall, P.; Ichimura, H. Optimal smoothing in single index models. *Ann. Stat.* **1993**, *21*, 157–178.
- Hristache, M.; Juditsky, A.; Spokoiny, V. Direct estimation of the index coefficient in a single-index model. *Ann. Stat.* **2001**, *29*, 595–623.
- Ait-Saïdi, A.; Ferraty, F.; Kassa, R.; Vieu, P. Cross-validated estimations in the single-functional index model. *Stat. J. Theor. Appl. Stat.* **2008**, *42*, 475–494.
- Liang, H.; Wang, N. Partially linear single index measurement error models. *Stat. Sin.* **2005**, *15*, 99–116.
- Chen, J.; Li, D.; Liang, H.; Wang, S. Semiparametric GEE analysis in partially linear single-index models for longitudinal data. *Ann. Stat.* **2015**, *43*, 1682–1715.
- Carroll, R.; Fan, J.; Gijbels, I.; Wand, M. Generalized partially linear single-index models. *J. Am. Stat. Assoc.* **1997**, *92*, 477–489.
- Wang, L.; Cao, G. Efficient estimation for generalized partially linear single index models. *Bernoulli* **2018**, *24*, 1101–1127.
- Ramsay, J.; Silverman, B. *Functional Data Analysis*; Springer: New York, NY, USA, 2005.
- Ferraty, F.; Vieu, P. *Nonparametric Functional Data Analysis: Theory and Practice*; Springer: New York, NY, USA, 2006; Volume 76.
- Aneiros-Pérez, G.; Vieu, P. Semi-functional partial linear regression. *Stat. Probab. Lett.* **2006**, *76*, 1102–1110.
- Aneiros, P.; Goia, V.; Vieu, P. Partial linear modelling with multi-functional covariates. *Comput. Stat.* **2015**, *30*, 647–671.
- Horváth, L.; Kokoszka, P. *Inference for Functional Data with Applications*; Springer Science and Business Media: New York, NY, USA, 2012; Volume 200.
- Kokoszka, P.; Reimherr, M. *Introduction to Functional Data Analysis*; CRC Press: Boca Raton, FL, USA, 2021.
- Schumaker, L. *Spline Functions: Basic Theory*; Cambridge University Press: Cambridge, UK, 2007.
- Cao, R.; Du, J.; Zhou, J.; Xie, T. FPCA-based estimation for generalized functional partially linear models. *Stat. Pap.* **2020**, *61*, 2715–2735.
- Li, C.; Lu, M. A lack-of-fit test for generalized linear models via single-index techniques. *Comput. Stat.* **2018**, *33*, 731–756.
- Ould Said, E.; Ouassou, I.; M, R. *Functional Statistics and Applications: Selected Papers from MICPS-2013 (Contributions to Statistics) Hardcover*; Springer: New York, NY, USA, 2015.
- Laksaci, A.; Kaid, Z.; Alahiane, M.; Ouassou, I.; Rachdi, M. Non parametric estimations of the conditional density and mode when the regressor and the response are curves. *Commun. -Stat.-Theory Methods* **2023**, *52*, 4659–4674.
- Ouassou, I.; Rachdi, M. Regression operator estimation by delta sequences method for functional data and its applications. *ASTA Adv. Stat. Anal.* **2010**, *96*, 451–465.
- Ouassou, I.; Rachdi, M. Stein type estimation of the regression operator for functional data. *Adv. Appl. Stat. Sci.* **2010**, *1*, 233–250.

25. Yu, P.; Du, J.; Zhang, Z. Single index partially functional linear regression model. *Stat. Pap.* **2020**, *61*, 1107–1123.
26. Yu, Y.; Ruppert, D. Penalized spline estimation for partially linear single index models. *J. Am. Stat. Assoc.* **2002**, *97*, 1042–1054.
27. Rachdi, M.; Alahiane, M.; Ouassou, I.; Vieu, P. Generalized functional partially linear single index models. In *Functional and High Dimensional Statistics and Related Fields*; Springer: New York, NY, USA, 2020; pp. 221–228.
28. Alahiane, M.; Ouassou, I.; Rachdi, M.; Vieu, P. Partially Linear Generalized single Index Models for Functional Data PLGSIMF. *Stats Funct. Data Anal. FDA* **2021**, *4*, 793–813.
29. Alahiane, M.; Ouassou, I.; Rachdi, M.; Vieu, P. High-Dimensional Statistics: Non-Parametric Generalized Functional Partially Linear Single Index Model. *Math. Adv. Stat. Theory Methodol. Appl. Data Anal.* **2022**, *10*, 2704.
30. Friedman, J.; Stuetzle, W. Projection pursuit regression. *J. Am. Stat. Assoc.* **1981**, *76*, 817–823.
31. Hall, P. On projection pursuit regression. *Ann. Stat.* **1989**, *17*, 573–588.
32. Huber, P. Projection pursuit. *Ann. Stat.* **1985**, *13*, 435–475.
33. Ferraty, F.; Goia, A.; Salinelli, E.; Vieu, P. Functional projection pursuit regression. *Test* **2013**, *22*, 293–320.
34. De Boor, C. *A Practical Guide to Splines*; Springer: New York, NY, USA, 2001; Volume 27.
35. Lai, P.; Tian, Y.; Lian, H. Estimation and variable selection for generalised partially linear single-index models. *J. Nonparametric Stat.* **2014**, *26*, 171–185.
36. Pollard, D. Asymptotics for least absolute deviation regression estimators. *Econom. Theory* **1991**, *7*, 186–199.
37. Stone, C. The dimensionality reduction principle for generalized additive models. *Ann. Stat.* **1986**, *14*, 590–606.
38. Van Der Vaart, A.W.; Wellner, J. *Weak Convergence*; Springer: New York, NY, USA, 1996.
39. Huang, J. Efficient estimation of the partly linear additive Cox model. *Ann. Stat.* **1999**, *27*, 1536–1563.
40. Wang, L.; Yang, L. Spline estimation of single index models. *Stat. Sin.* **2009**, *19*, 765–783.
41. Xue, L.; Yang, L. Additive coefficient modeling via polynomial spline. *Stat. Sin.* **2006**, *16*, 1423–1446.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.