ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Neumeyer, Natalie; Selk, Leonie

Article — Published Version

Testing for changes in the error distribution in functional linear models

Statistical Papers

Suggested Citation: Neumeyer, Natalie; Selk, Leonie (2025) : Testing for changes in the error distribution in functional linear models, Statistical Papers, ISSN 1613-9798, Springer Berlin Heidelberg, Berlin/Heidelberg, Vol. 66, Iss. 2, https://doi.org/10.1007/s00362-024-01656-9

This Version is available at:
https://hdl.handle.net/10419/319258

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

http://creativecommons.org/licenses/by/4.0/

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Statistical Papers (2025) 66:33
https://doi.org/10.1007/s00362-024-01656-9

REGULAR ARTICLE

Testing for changes in the error distribution in functional linear models

Natalie Neumeyer[1] · **Leonie Selk**[1]

Received: 3 April 2024 / Revised: 6 November 2024 / Published online: 15 January 2025
© The Author(s) 2025

Abstract
We consider linear models with scalar responses and covariates from a separable Hilbert space. The aim is to detect change points in the error distribution, based on sequential residual empirical distribution functions. Expansions for those estimated functions are more challenging in models with infinite-dimensional covariates than in regression models with scalar or vector-valued covariates due to a slower rate of convergence of the parameter estimators. Yet the suggested change point test is asymptotically distribution-free and consistent for one-change point alternatives. In the latter case we also show consistency of a change point estimator.

Keywords Change-points · Functional data analysis · Regularized function estimators · Regression · Residual processes

Mathematics Subject Classification Primary 62R10; Secondary 62G10 · 62G30

1 Introduction

We consider a functional linear model $Y = \alpha + \langle X, \beta\rangle + \varepsilon$ with scalar response Y and covariates X from a separable Hilbert space, e.g. $L^2([0, 1])$. Structural changes in the distribution can appear, even when the parameters α and β do not change. For this reason we focus on detecting changes in the error distribution. If the errors were observable one could use the classical test (and change point estimators) based on the difference of the sequential empirical distribution functions of the first $\lfloor nt \rfloor$ and the last $n - \lfloor nt \rfloor$ error terms from a sample of n observations, see Csörgö et al. (1997), Picard (1985), Carlstein (1988), Dümbgen (1991), Hariz et al. (2005) and Hariz et al. (2007). In a regression model those tests have to be based on estimated residuals $\hat{\varepsilon} = Y - \hat{\alpha} - \langle X, \hat{\beta}\rangle$. Similar tests have been considered by Bai (1994) in

✉ Natalie Neumeyer
natalie.neumeyer@uni-hamburg.de

Leonie Selk
leonie.selk@uni-hamburg.de

[1] Fachbereich Mathematik, Universität Hamburg, Hamburg, Germany

the context of ARMA-models, by Koul (1996) in the context of nonlinear time series, by Ling (1998) for nonstationary autoregressive models, and by Neumeyer and Van Keilegom (2009) and Selk and Neumeyer (2013) for nonparametric independent and time series regression models. Typically the asymptotic distribution is derived using asymptotic expansions of residual-based empirical distribution functions. For models with functional covariates those expansions can be problematic because inner products $\langle X, \hat{\beta} - \beta \rangle$ appear and those can have a slow rate of convergence [see Cardot et al. (2007), Shang and Cheng (2015), Yeon et al. (2023)]. However, we show that under very simple non-restrictive assumptions those terms cancel for the suggested change point test statistic and thus the asymptotic distribution is the same as based on true (unobserved) errors.

Change point testing and estimation for functional data, and for the parameter in functional linear models have been considered in the literature, but not for the error distribution. Tests for changes in the functional mean and in the parameter function of autoregressive models are considered in chapters 6 and 14 in Horváth and Kokoszka (2012). Berkes et al. (2009) propose a CUSUM testing procedure to detect a change in the mean of functional observations. They apply projections on principal components of the data to estimate the mean. Aue et al. (2009) extend this result and introduce an estimator for the change point in this model and derive its limit distribution. Aston and Kirch (2012) consider the same type of model with epidemic changes and dependent data. Aue et al. (2018) consider how to detect and date structural breaks in the mean of functional observations without the application of dimension reduction techniques (as functional principal component analysis). Aue et al. (2014) propose a monitoring procedure to detect structural changes in functional linear models with functional response, allowing for dependence in the data, including functional autoregressive processes. They test for a change in the regression operator, which is the analogue to our β, based on functional principal component analysis. A linear regression model with scalar response is considered in Horváth et al. (2024) who propose a tests for the detection of multiple change points in the regression parameter. The regressors in their model can be functional and can include lagged values of the response.

The paper is organized as follows. In Sect. 2 we define the test statistic and present model assumptions to obtain the asymptotically distribution-free hypothesis test. In Sect. 3 we discuss the assumptions on the parameter estimators and some examples. In Sect. 4 consistency of the test as well as of a change point estimator is considered in the context of one change point. Finite sample properties are shown in Sect. 5. Section 6 concludes the paper, in particular with an outlook on goodness-of-fit testing. The proofs are given in the appendix.

2 Model, test statistic and main result under the null

Let $\mathcal{H}$ be a separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$, corresponding norm $\| \cdot \|$ and Borel-sigma field. Let (X_i, Y_i), $i = 1, \ldots, n$, be an independent sample of $(\mathcal{H} \times \mathbb{R})$-valued random variables defined on the same probability space with probability

measure $\mathbb{P}$. The data are modeled as functional linear model

$$Y_i = \alpha + \langle X_i, \beta \rangle + \varepsilon_i, \quad i = 1, \ldots, n,$$

with scalar response Y_i and $\mathcal{H}$-valued covariate X_i, and with parameters $\alpha \in \mathbb{R}, \beta in H$. The covariates $X_1, \ldots, X_n$ are assumed to be iid with $E\|X_i\| < \infty$, and the errors $\varepsilon_1, \ldots, \varepsilon_n$ are independent, centered, and independent of the covariates. Our aim is to test for change-points in the error distribution. In this section we consider the test statistic under the null hypothesis, where the errors are identically distributed.

Let $\hat{\alpha}$ and $\hat{\beta}$ denote estimators for the parameters $\alpha \in \mathbb{R}$ and $\beta \in \mathcal{H}$. We build residuals $\hat{\varepsilon}_i = Y_i - \hat{\alpha} - \langle X_i, \hat{\beta} \rangle, i = 1 \ldots, n$. The test statistic

$$T_n = \sup_{t \in [0,1]} \sup_{z \in \mathbb{R}} |\hat{G}_n(t, z)|$$

based on the process

$$\hat{G}_n(t, z) = \frac{\lfloor nt \rfloor \, (n - \lfloor nt \rfloor)}{n^{3/2}} \left(\hat{F}_{\lfloor nt \rfloor}(z) - \tilde{F}_{\lfloor nt \rfloor}(z) \right),$$

compares for each $k = 1, \ldots, n-1$ the empirical distribution functions

$$\hat{F}_k(z) = \frac{1}{k} \sum_{i=1}^{k} I\{\hat{\varepsilon}_i \le z\}, \qquad \tilde{F}_k(z) = \frac{1}{n-k} \sum_{i=k+1}^{n} I\{\hat{\varepsilon}_i \le z\}$$

of the first k and last $n-k$ residuals, respectively. Note that one can write

$$\hat{G}_n(t, z) = \frac{\lfloor nt \rfloor}{n^{1/2}} \left(\hat{F}_{\lfloor nt \rfloor}(z) - \hat{F}_n(z) \right).$$

For the asymptotic distribution of the test statistic under the null hypothesis we assume the following conditions. Let P denote the distribution of (X_1, ε_1).

(a.1) $|\hat{\alpha} - \alpha| = o_{\mathbb{P}}(1), \|\hat{\beta} - \beta\| = o_{\mathbb{P}}(1)$

(a.2) Let $\varepsilon_1, \ldots, \varepsilon_n$ be independent and identically distributed with cdf F that is Hölder-continuous of order $\gamma \in (0, 1]$ with Hölder-constant c.

(a.3) $\mathbb{P}\big(\hat{\beta} - \beta \in \mathcal{B}\big) \to 1$ as $n \to \infty$ for a class $\mathcal{B} \subset \mathcal{H}$ such that the function class

$$\mathcal{F} = \{(x, e) \mapsto I\{e \le v + \langle x, b \rangle\} \mid v \in \mathbb{R}, b \in \mathcal{B}\}$$

is P-Donsker.

Remark 2.1 The assumptions are very mild and in particular less restrictive than typical assumptions for asymptotic distribution of residual-based empirical processes, even for finite-dimensional covariates. In assumption (a.1) only consistency is needed, no rates of convergence. Typically in the literature about residual-based procedures a bounded error density is assumed, see e.g. Akritas and Van Keilegom (2001). Then

(a.2) is fulfilled for $\gamma = 1$, but (a.2) is less restrictive in the cases $\gamma \in (0, 1)$. Suitable conditions for the general assumption (a.3) are discussed in Sect. 3. One possibility for $\mathcal{H} = L^2([0, 1])$ is to assume smoothness of β which is a typical assumption. If $\gamma \in (\frac{1}{2}, 1]$ in assumption (a.2), and β is in a Sobolev-space with third derivatives, (a.2) holds for the estimator $\hat{\beta}$ from Yuan and Cai (2010). This estimator can also be applied for smaller γ in (a.2) if higher smoothness of β is assumed.

Define the process G_n as $\hat{G}_n$, but based on the true errors instead of residuals, i.e.

$$G_n(t, z) = \frac{\lfloor nt \rfloor}{n^{1/2}} \left(F_{\lfloor nt \rfloor}(z) - F_n(z) \right)$$

with

$$F_{\lfloor nt \rfloor}(z) = \frac{1}{\lfloor nt \rfloor} \sum_{i=1}^{\lfloor nt \rfloor} I\{\varepsilon_i \leq z\}. \tag{2.1}$$

Further let G be a completely tucked Brownian sheet, i.e. a centered Gaussian process on $[0, 1]^2$ with covariance structure

$$\mathrm{Cov}(G(s, u), G(t, v)) = (s \wedge t - st)(u \wedge v - uv).$$

Theorem 2.2 *Under the assumptions* (a.1)–(a.3)*,*

$$\sup_{t \in [0,1]} \sup_{z \in \mathbb{R}} |\hat{G}_n(t, z) - G_n(t, z)| = o_{\mathbb{P}}(1), \tag{2.2}$$

and thus the process $(\hat{G}_n(t, z))_{t \in [0,1], z \in \mathbb{R}}$ *converges weakly to* $(G(t, F(z)))_{t \in [0,1], z \in \mathbb{R}}$.

The proof of (2.2) in the theorem is given in the appendix. The weak convergence of G_n is a classical result, see Bickel and Wichura (1971), Shorack and Wellner (1986). With the continuous mapping theorem one obtains the asymptotic distribution of the test statistic T_n under the null hypothesis of no change-point, which is the distribution of $T = \sup_{t,u \in [0,1]} |G(t, u)|$ because F is continuous. The test statistic is asymptotically distribution-free with the same limit distribution as for corresponding changepoint tests based on iid observations (not residuals). Let $\bar{\alpha} \in (0, 1)$ and q be the $(1 - \bar{\alpha})$-quantile of T. Then the test that rejects the null hypothesis if $T_n > q$ has asymptotic level $\bar{\alpha}$. Consistency is considered in Sect. 4.

Remark 2.3 The choice of T_n as a Kolmogorov–Smirnov type test statistic is not mandatory. In principle, any continuous functional of the process $\hat{G}_n$ can be considered. The most common ones, besides T_n, are of Cramér-von–Mises type, e. g. $T_{n,2} = \sup_{t \in [0,1]} \int_{\mathbb{R}} |\hat{G}_n(t, z)|^2 dF(z)$ or $T_{n,3} = \int_0^1 \int_{\mathbb{R}} |\hat{G}_n(t, z)|^2 dF(z) dt$. The asymptotic distribution of these test statistics under the null hypothesis also follows from Theorem 2.2 and with

$$\int_{\mathbb{R}} |\hat{G}_n(t, z)|^2 dF(z) \to \int_{\mathbb{R}} |G(t, F(z))|^2 dF(z) = \int_0^1 |G(t, x)|^2 dx,$$

and thus these test statistics are asymptotically distribution-free as well. However, $T_{n,2}$ and $T_{n,3}$ contain the unknown quantity F and must therefore be modified in order to be applied. This can be done by replacing the integral with the sample mean: $\tilde{T}_{n,2} = \sup_{t\in[0,1]} \frac{1}{n}\sum_{i=1}^{n} |\hat{G}_n(t,\hat{\varepsilon}_i)|^2$ and $\tilde{T}_{n,3} = \int_0^1 \frac{1}{n}\sum_{i=1}^{n} |\hat{G}_n(t,\hat{\varepsilon}_i)|^2 dt$.

3 Discussion of assumptions and examples

To show validity of the Donsker-class assumption (a.3) there are sufficient conditions on covering numbers or bracketing numbers. We discuss some specific conditions on the class $\mathcal{B}$, examples for Hilbert spaces $\mathcal{H}$, and estimators for the parameter function β that fulfill the conditions.

3.1 VC-class condition

Assumption (a.3) can be derived from a VC-function class condition formulated as follows. Assume that $\mathbb{P}\big(\hat{\beta} - \beta \in \mathcal{B}\big) \to 1$ as $n \to \infty$ for a class $\mathcal{B} \subset \mathcal{H}$ such that the class of maps

$$\{\mathcal{H} \to \mathbb{R}, x \mapsto \langle x, b\rangle + v \mid b \in \mathcal{B}, v \in \mathbb{R}\} \tag{3.1}$$

is a VC-subgraph class. By definition then

$$\{\{(x,e) \in \mathcal{H}\times\mathbb{R} \mid e \le \langle x, b\rangle + v\} \mid b \in \mathcal{B}, v \in \mathbb{R}\}$$

is a VC-class of sets. The class $\mathcal{F}$ from (a.3) is the class of the corresponding indicator functions and (a.3) is fulfilled by Theorems 8.19 and 9.2 in Kosorok (2008).

Example 3.1 We consider the Hilbert space $\mathcal{H} = L^2([0,1])$ with inner product $\langle g, h\rangle = \int_0^1 g(t)h(t)\,dt$ and norm $\|g\| = (\int_0^1 g^2(t)\,dt)^{1/2}$. For the parameter function β we assume sparsity as in Lee and Park (2012). Let $(\phi_j)_{j\in\mathbb{N}}$ be a basis of $\mathcal{H}$ and assume $\beta = \sum_{j\in J}\beta_j\phi_j$ for some finite, but unknown index set J. Lee and Park (2012) consider the estimator $\hat{\beta} = \sum_{j=1}^{k}\hat{\beta}_j\phi_j$ with

$$(\hat{\beta}_1,\ldots,\hat{\beta}_k) = \arg\min_{b_1,\ldots,b_k\in\mathbb{R}} \left(\frac{1}{n}\sum_{i=1}^{n}\Big(Y_i - \overline{Y}_n - \sum_{j=1}^{k} b_j\langle X_i - \overline{X}_n, \phi_j\rangle\Big)^2 + \sum_{j=1}^{k}\hat{w}_j|b_j|\right)^2,$$

where k is a chosen dimension-cut-off, $\hat{w}_j$ are suitable weights based on initial estimators, and $\overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$, $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Further, $\hat{\alpha} = \overline{Y}_n - \langle\hat{\beta}, \overline{X}_n\rangle$. Under suitable assumptions, in particular $E\|X\|^2 < \infty$, and k is larger than the largest index in J, Lee and Park (2012) show in their Theorem 2 that $\mathbb{P}(\hat{\beta}_j = 0 \text{ for } j \notin J) \to 1$ for

$n \to \infty$. Thus we can set

$$\mathcal{B} = \left\{ \sum_{j \in J} b_j \phi_j \;\middle|\; b_j \in \mathbb{R} \, \forall j \in J \right\}$$

and obtain $\mathbb{P}(\hat{\beta} - \beta \in \mathcal{B}) \to 1$ for $n \to \infty$. Further, the class of maps in (3.1), i.e.

$$\left\{ \mathcal{H} \to \mathbb{R}, x \mapsto \sum_{j \in J} b_j \langle x, \phi_j \rangle + v \;\middle|\; b_j \in \mathbb{R} \, \forall j \in J, v \in \mathbb{R} \right\}$$

is a finite dimensional vector space and thus a VC-class, see Lemma 2.6.15 in van der Vaart and Wellner (1996). Then as discussed above validity of (a.3) follows. Furthermore, from Theorem 2 in Lee and Park (2012) it also follows that our assumption (a.1) is fulfilled, and thus under assumption (a.2) the assertion of Theorem 2.2 holds.

3.2 Bracketing number condition

In this subsection we assume that $\mathcal{H}$ is a separable Hilbert space of real-valued functions (or vectors with real components) and the inner product is increasing in the sense that from $h \leq g$ (pointwise for functions; componentwise for vectors) it follows that $\langle h, x \rangle \leq \langle g, x \rangle$ for all $x \in \mathcal{H}$ with $x \geq 0$. Then assumption (a.3) can be replaced by the condition in the next lemma.

Lemma 3.2 *Assume* (a.1), (a.2) *and* $\mathbb{P}\big(\hat{\beta} - \beta \in \mathcal{B}\big) \to 1$ *as* $n \to \infty$ *for a function class* $\mathcal{B} \subset \mathcal{H}$ *such that the bracketing number fulfills* $\log N_{[\,]}(\mathcal{B}, \epsilon, \|\cdot\|) \leq K/\epsilon^{1/k}$ *for some* $k > 1/\gamma$. *Here* γ *is the Hölder-order from assumption* (a.2). *Then assumption* (a.3) *holds.*

The proof is given in the appendix.

Example 3.3 We consider the Hilbert space $\mathcal{H} = L^2([0,1])$ with inner product $\langle g, h \rangle = \int_0^1 g(t)h(t)\,dt$ and norm $\|g\| = (\int_0^1 g^2(t)\,dt)^{1/2}$. We assume $\beta \in \mathcal{W}_2^m([0,1])$ for some $m > 2$ and the Sobolev-space

$$\mathcal{W}_2^m([0,1]) = \big\{ b : [0,1] \to \mathbb{R} \mid b^{(j)} \text{ is absolutely continuous for } j = 0, \ldots, m-1, \text{ and } \|b^{(m)}\| < \infty \big\},$$

where $b^{(0)} = b$, and $b^{(j)}$ denotes the j-th derivative of b, $j \geq 1$. We consider the regularized estimators in Yuan and Cai (2010), i.e.

$$\big(\hat{\alpha}, \hat{\beta}\big) = \arg \min_{a \in \mathbb{R}, b \in \mathcal{W}_2^m([0,1])} \left\{ \frac{1}{n} \sum_{i=1}^n \Big(Y_i - \big(a + \langle X_i, b \rangle\big) \Big)^2 + \lambda_n \big\| b^{(m)} \big\|^2 \right\}$$

for a suitable positive sequence λ_n converging to zero. Convergence rates of $\hat{\beta}$ and its derivatives can be found in Corollaries 10 and 11 in Yuan and Cai (2010). Under suitable assumptions one obtains $\left\|\hat{\beta}^{(j)} - \beta^{(j)}\right\| = o_{\mathbb{P}}(1)$ for $j = 0, 1, 2$, and thus $\mathbb{P}(\hat{\beta} - \beta \in \mathcal{B}) \to 1$ for the function class

$$\mathcal{B} = \left\{b \in \mathcal{W}_2^2([0,1]) : \|b\| + \|b^{(2)}\| \leq 1\right\}.$$

By Corollary 4.3.38 in Giné and Nickl (2021) and Lemma 9.21 in Kosorok (2008) the class $\mathcal{B}$ fulfills the bracketing number condition in Lemma 3.2 for $k = 2$. Thus the assumptions (a.1)–(a.3) are fulfilled if F is Hölder-continuous of order $\gamma \in (\frac{1}{2}, 1]$. Less restrictive assumptions on F, i.e. $\gamma \leq \frac{1}{2}$, require for this concept higher smoothness of β.

4 Fixed one-change point alternative: consistency of the test and change point estimator

In this section we consider fixed alternatives with one change point at index $k_n^* = \lfloor n\vartheta^* \rfloor$ with $\vartheta^* \in (0, 1)$. We write the functional linear model as in Sect. 2 under the following assumption.

(a.2)' Assume $\varepsilon_1, \ldots, \varepsilon_{k_n^*}$ are iid with cdf F_1, and $\varepsilon_{k_n^*+1}, \ldots, \varepsilon_n$ are iid with cdf $F_2 \neq F_1$. Let F_1 and F_2 be Hölder-continuous of order $\gamma_1, \gamma_2 \in (0, 1]$ with Hölder-constant c_1, c_2, respectively.

Let further P_1 denote the distribution of (X_1, ε_1) (before the change) and P_2 denote the distribution of (X_n, ε_n) (after the change). For the empirical distribution functions $\hat{F}_k$ and $\tilde{F}_k$ as in Sect. 2 we obtain the following asymptotic result.

Lemma 4.1 *Under assumptions* (a.1) *and* (a.2)' *and if* (a.3) *is valid for* $P = P_1$ *and* $P = P_2$, *it holds that*

$$\sup_{z\in\mathbb{R}} |\hat{F}_{k_n^*}(z) - F_1(z)| = o_{\mathbb{P}}(1) \quad \text{and} \quad \sup_{z\in\mathbb{R}} |\tilde{F}_{k_n^*}(z) - F_2(z)| = o_{\mathbb{P}}(1).$$

The proof is given in the appendix. Now note that

$$\frac{T_n}{n^{1/2}} \geq \frac{k_n^*(n - k_n^*)}{n^2} \sup_{z\in\mathbb{R}} \left|\hat{F}_{k_n^*}(z) - \tilde{F}_{k_n^*}(z)\right|,$$

and by Lemma 4.1 the right hand side converges in probability to the positive constant

$$\vartheta^*(1 - \vartheta^*) \sup_{z\in\mathbb{R}} |F_1(z) - F_2(z)|.$$

From this it follows that tests that reject the null hypothesis of no change-point if $T_n > q$ for some $q > 0$ (see Sect. 2) are consistent.

The estimator for the change point ϑ^* is based on the process $\hat{G}_n$ and is defined as

$$\hat{\vartheta}_n = \min \left\{ t : \sup_{z \in \mathbb{R}} |\hat{G}_n(t, z)| = \sup_{t' \in [0,1]} \sup_{z \in \mathbb{R}} |\hat{G}_n(t', z)| \right\}.$$

Lemma 4.2 *Under assumptions* (a.1), (a.2)' *and if* (a.3) *holds for* $P = P_1$ *and* $P = P_2$, *the change point estimator is consistent, i. e.*

$$|\hat{\vartheta}_n - \vartheta^*| = o_{\mathbb{P}}(1).$$

The proof is given in the appendix.

5 Finite sample properties

We consider the Hilbert space $\mathcal{H} = L^2([0, 1])$. For $i = 1, \ldots, n$ the functional observations $X_i(t)$, $t \in [0, 1]$, are generated according to

$$X_i(t) = \frac{1}{2} \sum_{l=1}^{5} \Big(B_{i,l} \sin\big(t(5 - B_{i,l})2\pi\big) - M_{i,l} - E[B_{i,l} \sin\big((5 - B_{i,l})2\pi\big) - M_{i,l}] \Big),$$

where $B_{i,l} \sim \mathcal{U}[0, 5]$ and $M_{i,l} \sim \mathcal{U}[0, 2\pi]$ for $l = 1, \ldots, 5$, $i = 1, \ldots, n$. $\mathcal{U}$ stands for the (continuous) uniform distribution. The functional linear model is built as

$$Y_i = \int X_i(t) \gamma_{3,\frac{1}{3}}(t) dt + \varepsilon_i,$$

where the coefficient function $\gamma_{a,b}(t) = b^a / \Gamma(a) t^{a-1} e^{-bt} I\{t > 0\}$ is the density of the Gamma distribution. Furthermore, we assume that each X_i is observed on a dense, equidistant grid of 300 evaluation points.

The parameter estimators are the regularized estimators described in Example 3.3 with $m = 3$ and a data-driven tuning parameter λ_n chosen by generalized cross-validation as described in Yuan and Cai (2010).

We model three similar types of change points, such that

$$\varepsilon_1, \ldots, \varepsilon_{\lfloor \frac{n}{2} \rfloor} \sim \mathcal{N}(0, 1), \ \varepsilon_{\lfloor \frac{n}{2} \rfloor + 1}, \ldots, \varepsilon_n \sim \tilde{F}_{1,\delta} \text{ (respectively } \tilde{F}_{2,\delta}, \ \tilde{F}_{3,\delta}),$$

where $\tilde{F}_{1,\delta}$, $\tilde{F}_{2,\delta}$, $\tilde{F}_{3,\delta}$ have in common that the mean remains zero and the variance remains one. In particular

- $\tilde{F}_{1,\delta}$ is the distribution function of a random variable that is $\mathcal{N}(-2\delta, 1)$ distributed with probability 0.5 and $\mathcal{N}(2\delta, 1)$ distributed with probability 0.5.
- $\tilde{F}_{2,\delta}$ is the distribution function of a random variable that is $\mathcal{N}(0, (1 - \delta)^2)$ distributed with probability 0.5 and $\mathcal{N}(0, 2 - (1 - \delta)^2)$ distributed with probability 0.5.

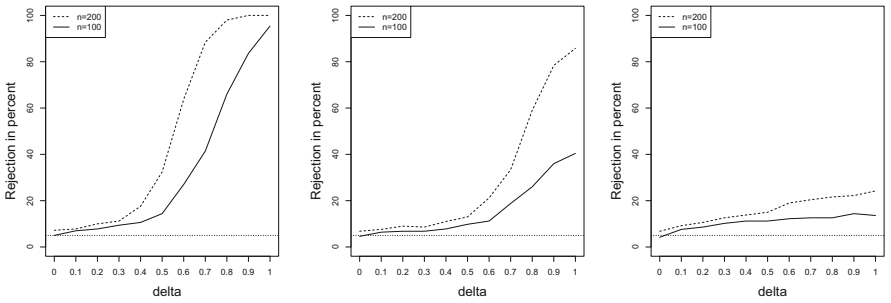

Fig. 1 Rejection probabilities with a change in the error distribution from $\mathcal{N}(0, 1)$ to $\tilde{F}_{1,\delta}$ (left), to $\tilde{F}_{2,\delta}$ (middle) and to $\tilde{F}_{3,\delta}$ (right). The dotted line marks the 5% level

- $\tilde{F}_{3,\delta}$ is the "skew-normal"-distribution

$$SN\left(-\sqrt{\frac{2\pi\left((10\delta)^2+(10\delta)^4\right)}{\pi^2+\left(2\pi^2-2\pi\right)\cdot(10\delta)^2+\left(\pi^2-2\pi\right)\cdot(10\delta)^4}},\sqrt{\frac{\pi\left(1+(10\delta)^2\right)}{\pi+(\pi-2)\cdot(10\delta)^2}},10\delta\right).$$

A random variable Z is distributed $SN(\lambda_1, \lambda_2, \lambda_3)$ if $Z = \lambda_1 + \lambda_2 \cdot Z_0$ and Z_0 has the density $2\phi(x)\Phi(\lambda_3 x)$, where ϕ is the density and Φ is the distribution function of the standard normal distribution [see Azzalini and Capitanio (1999)]. The expected value of such a random variable Z is calculated as $\lambda_1+\lambda_2\sqrt{\frac{2}{\pi}}\cdot\frac{\lambda_3}{\sqrt{1+\lambda_3{}^2}}$ and the variance as $\lambda_2{}^2\left(1-\frac{2}{\pi}\cdot\frac{\lambda_3{}^2}{1+\lambda_3{}^2}\right)$. This results in the parameters for the distribution after the change point, such that the the expected value of the errors remains 0 and the variance remains 1.

So $\delta = 0$ represents the null hypothesis of no change point, and the difference between the distribution before and after the change point grows with δ in all three cases.

In Fig. 1 the rejection probabilities for 500 repetitions, level 5% [critical value tabled in Picard (1985)] and sample sizes $n \in \{100, 200\}$ are shown. In all three cases it can be seen that the level is approximated well and the power increases for increasing parameter δ as well as for increasing sample size n. In the case of a change in skewness, the increase with δ is not as pronounced as in the other two cases. This is because the distributions for different values of δ become more similar as δ increases. The same types of changes (from $\mathcal{N}(0, 1)$ to $\tilde{F}_{1,\delta}$ and to $\tilde{F}_{2,\delta}$) were also simulated in Selk and Neumeyer (2013) for a real-valued nonparametric autoregression model with lag 1. The results are comparable with an even higher power in the paper at hand.

In addition, we model a more distinct change, that is

$$\varepsilon_1, \ldots, \varepsilon_{\lfloor \frac{n}{2} \rfloor} \sim \mathcal{N}(0, 0.5^2),\ \varepsilon_{\lfloor \frac{n}{2} \rfloor+1}, \ldots, \varepsilon_n \sim \mathcal{N}(0, (0.5+\delta)^2).$$

As expected for a change in the variance the power grows faster with increasing δ than in the other three cases, especially for small δ. The results are shown in Fig. 2. This kind of change point was also simulated in Neumeyer and Van Keilegom (2009)

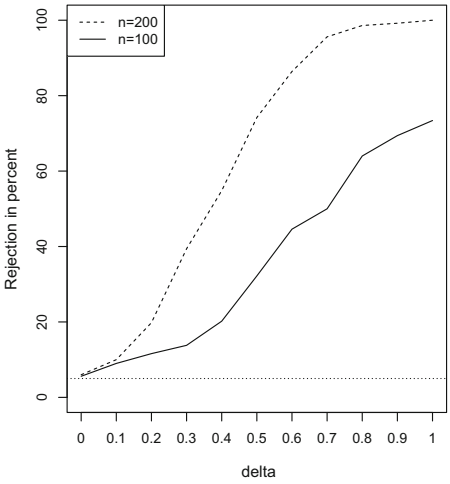

Fig. 2 Rejection probabilities with a change in the error distribution from $\mathcal{N}(0, 0.5^2)$ to $\mathcal{N}(0, (0.5+\delta)^2)$. The dotted line marks the 5% level

for a nonparametric regression model with one-dimensional regressor. The results are similar.

Next to the Kolmogorov–Smirnov type test with test statistic T_n, we also applied Cramér-von-Mises type tests with test statistics $T_{n,2}$ and $T_{n,3}$. The results are very similar and are not presented here for the sake of brevity.

6 Concluding remarks

To detect structural changes in functional linear models, we considered the classical test by Bickel and Wichura (1971) for a change in the distribution, but based on estimated errors. We gave simple assumptions under which the asymptotic distribution of the test statistic under the null is the same as for iid data. The test as well as the corresponding change point estimators are consistent in one-change point models. The same test can be considered in more complex regression models with functional covariates, e.g. a quadratic model as in Boente and Parada (2023) or nonparametric models, see Ferraty and Vieu (2006). We only consider independent data, but testing for change points in the innovation distribution in times series models that include functional covariates is a very interesting topic. However, the proofs for asymptotic distributions will be more complicated. In future work we are planning to consider a time series model $Y_t = m(X_t) + \varepsilon_t$, where Y_t and ε_t are real-valued and X_t contains a functional part, but can also contain past values $Y_{t-1}, \ldots, Y_{t-p}$. We presume that the proofs as in Selk and Neumeyer (2013) (for nonparametric autoregression time series with independent errors) and of Sect. 4.2 in Neumeyer and Omelka (2025) (for linear models with finite-dimensional covariates and beta-mixing errors) can be combined with the proofs in the paper at hand to consider change-point tests for the innovation distribution under the assumption that (X_t, ε_t), $t \in \mathbb{Z}$, is a strictly stationary beta-mixing time series.

In the proof of Theorem 2.2 we derive an expansion for the sequential residual-based empirical distribution function,

$$\hat{F}_{\lfloor nt \rfloor}(z) = F_{\lfloor nt \rfloor}(z) + R_n(z) + o_{\mathbb{P}}(n^{-1/2})$$

uniformly in $t \in [0, 1]$, $z \in \mathbb{R}$, where $F_{\lfloor nt \rfloor}$ is defined in (2.1), and the term

$$R_n(z) = E_X[F(z + \hat{\alpha} - \alpha + \langle X, \hat{\beta} - \beta \rangle)] - F(z)$$

appears from estimating the parameters [see (A.1) in the appendix]. Here E_X denotes the expectation with respect to X, which has the same distribution as X_i, but is independent of $\hat{\alpha}, \hat{\beta}$. For change-point testing the remainder term R_n cancels when considering the test statistic T_n. For other testing procedures, e. g. goodness-of-fit tests for the error distribution, this typically nonnegligible term is of relevance, see Koul (2002) and Neumeyer et al. (2006) for linear models and Akritas and Van Keilegom (2001) for nonparametric regression. Under more restrictive assumptions one can further expand the remainder term as follows. Assume that F is twice differentiable with density $F' = f$ and bounded f' and further $|\hat{\alpha} - \alpha| + \|\hat{\beta} - \beta\| = o_{\mathbb{P}}(n^{-1/4})$, and $E\|X\|^2 < \infty$. Then by Taylor's expansion one obtains

$$\hat{F}_{\lfloor nt \rfloor}(z) = F_{\lfloor nt \rfloor}(z) + f(z)\left(\hat{\alpha} - \alpha + \langle E[X], \hat{\beta} - \beta \rangle\right) + o_{\mathbb{P}}(n^{-1/2}).$$

In models with intercept α, where the estimator for α is chosen as $\hat{\alpha} = \overline{Y}_n - \langle \overline{X}_n, \hat{\beta} \rangle$ the remainder term is

$$R_n(z) = f(z)\left(\overline{\varepsilon}_n + \langle \overline{X}_n - E[X], \hat{\beta} - \beta \rangle\right) + o_{\mathbb{P}}(n^{-1/2}).$$

(Here, $\overline{X}_n = n^{-1}\sum_{i=1}^n X_i$ and analogous for $\overline{Y}_n$ and $\overline{\varepsilon}_n$.) By Cauchy–Schwarz-inequality and the central limit theorem for $n^{1/2}(\overline{X}_n - E[X])$ one obtains that the dominating part of the remainder term is $f(z)\overline{\varepsilon}_n$. This is the same as in homoscedastic finite-dimensional linear models with intercept and nonparametric regression models. Note that α and β are identifiable if the kernel of the covariance operator of the covariate X is $\{0\}$. But often functional linear models without intercept are considered in the literature. So in our model assume $\alpha = \hat{\alpha} = 0$. Then the remainder term is

$$R_n(z) = f(z)\langle E[X], \hat{\beta} - \beta \rangle + o_{\mathbb{P}}(n^{-1/2}),$$

and $\langle x, \hat{\beta} - \beta \rangle$ for fixed $x \in \mathcal{H}$ typically has a slower rate than $n^{-1/2}$, see Cardot et al. (2007), Shang and Cheng (2015), Yeon et al. (2023). If one assumes $E[X] = 0$, then this problematic term cancels [similar as e.g. for centered ARMA-processes, see Bai (1994)], but otherwise $f(z)\langle E[X], \hat{\beta} - \beta \rangle$ will dominate the asymptotic distribution of the process $(\hat{F}_{\lfloor nt \rfloor}(z) - F(z))_{t \in [0,1], z \in \mathbb{R}}$. For our change-point test this dominating term vanishes. The same holds when estimating the conditional copula of the response in multidimensional functional linear models, given the covariate, see Theorem 5 in Neumeyer and Omelka (2025). But e.g. for goodness-of-fit testing the remainder term

would be of relevance. We consider goodness-of-fit tests for the error distribution in the different cases explained above in future work. With the derived expansion for residual empirical distribution functions one can also develop other tests for the error distribution as e.g. for symmetry, or equality of error distributions in different models, see e.g. Pierce and Kopecky (1979), Neumeyer et al. (2005), Pardo Fernandez (2007), among many others, in the cases of regression models with finite-dimensional covariates.

A Proofs

For ease of notation let (X, Y, ε) be some generic random variable with the same distribution as $(X_1, Y_1, \varepsilon_1)$ under the null, but independent from the sample (X_i, Y_i), $i = 1, \ldots, n$. Let P denote the distribution of (X, ε). Further let E_X denote the expectation with respect to X, which in the context below is the conditional expectation given (X_i, Y_i), $i = 1, \ldots, n$.

The proofs of Theorem 2.2 and Lemma 3.2 are similar as a part of the proof of Theorem 5 in Neumeyer and Omelka (2025), but under less restrictive assumptions.

A.1 Proof of Theorem 2.2

From the Donsker-property in assumption (a.3) and Corollary 9.31 in Kosorok (2008) it follows that

$$\{(x, e) \mapsto I\{e \le z + a + \langle x, b\rangle\} - I\{e \le z\} \mid z \in \mathbb{R}, a \in \mathbb{R}, b \in \mathcal{B}\}$$

is also P-Donsker. From Theorem 2.12.1 in van der Vaart and Wellner (1996) it follows that also the centered sequential process

$$H_n(t, z, a, b) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \big(I\{\varepsilon_i \le z + a + \langle X_i, b\rangle\} - I\{\varepsilon_i \le z\} - E[F(z + a + \langle X, b\rangle] + F(z) \big),$$

indexed in $t \in [0, 1]$, $z \in \mathbb{R}$, $a \in \mathbb{R}$, $b \in \mathcal{B}$, converges weakly to a centered Gaussian process. Thus the process H_n is asymptotically stochastic equicontinuous with respect to the semi-metric

$$\rho((t_1, z_1, a_1, b_1), (t_2, z_2, a_2, b_2)) = |t_1 - t_2| + \mathrm{Var}(I\{\varepsilon \le z_1 + a_1 + \langle X, b_1\rangle\} - I\{\varepsilon \le z_2 + a_2 + \langle X, b_2\rangle\}),$$

see van der Vaart and Wellner (1996), problem 2, p. 93, and Sect 2.12. In particular we need

$$\rho((t, z, a, b), (t, z, 0, 0)) = \mathrm{Var}\left(I\{\varepsilon \le z + a + \langle X, b\rangle\} - I\{\varepsilon \le z\}\right)$$

$$\begin{aligned} &\leq |E\left[F(z+a+\langle X,b\rangle)-F(z)\right]| \\ &\leq cE\left[|a+\langle X,b\rangle|^{\gamma}\right] \;\leq\; c\big(|a|+\|b\|E\|X\|\big)^{\gamma} \end{aligned}$$

by Cauchy–Schwarz and Jensen's inequality. Now setting $a=\hat{\alpha}-\alpha$ and $b=\hat{\beta}-\beta$ we obtain convergence to zero in probability by assumption (a.1). Thus from asymptotic stochastic equicontinuity of the process H_n, and $H_n(t,z,0,0)=0$ we obtain that

$$\sup_{t\in[0,1],z\in\mathbb{R}} |H_n(t,z,\hat{\alpha}-\alpha,\hat{\beta}-\beta)| = o_{\mathbb{P}}(1),$$

which means that

$$\begin{aligned} \frac{\lfloor nt\rfloor}{\sqrt{n}}\hat{F}_{\lfloor nt\rfloor}(z) = \frac{\lfloor nt\rfloor}{\sqrt{n}}F_{\lfloor nt\rfloor}(z) + \frac{\lfloor nt\rfloor}{\sqrt{n}}\Big(E_X[F(z+\hat{\alpha}-\alpha+\langle X,\hat{\beta}-\beta\rangle)]-F(z)\Big) \\ +o_{\mathbb{P}}(1) \end{aligned} \tag{A.1}$$

uniformly in $t\in[0,1]$, $z\in\mathbb{R}$, where $F_{\lfloor nt\rfloor}$ was defined in (2.1) and is based on the true errors. In particular for $t=1$ we have

$$\hat{F}_n(z) = F_n(z) + \Big(E_X[F(z+\hat{\alpha}-\alpha+\langle X,\hat{\beta}-\beta\rangle)]-F(z)\Big) + o_{\mathbb{P}}(n^{-1/2})$$

uniformly in $z\in\mathbb{R}$. From those expansions we obtain

$$\begin{aligned} \hat{G}_n(t,z) &= \frac{\lfloor nt\rfloor}{\sqrt{n}}\hat{F}_{\lfloor nt\rfloor}(z) - \frac{\lfloor nt\rfloor}{\sqrt{n}}\hat{F}_n(z) \\ &= \frac{\lfloor nt\rfloor}{\sqrt{n}}F_{\lfloor nt\rfloor}(z) - \frac{\lfloor nt\rfloor}{\sqrt{n}}F_n(z) + o_{\mathbb{P}}(1) \\ &= G_n(t,z) + o_{\mathbb{P}}(1) \end{aligned} \tag{A.2}$$

uniformly in $t\in[0,1]$, $z\in\mathbb{R}$. □

A.2 Proof of Lemma 3.2

Let $\epsilon>0$ and let

$$\big[b_i^L, b_i^U\big], \quad i=1,\ldots,N(\epsilon)=O\big(\exp(\epsilon^{-2/(k\gamma)})\big)$$

be brackets for $\mathcal{B}$ of $\|\cdot\|$-length $\epsilon^{2/\gamma}$ [see assumption (a.3)]. Now for $b\in[b_i^L,b_i^U]$ the indicator function $I\{e\leq v+\langle x,b\rangle\}$ is contained in the bracket

$$\begin{aligned} \big[\, &I\{e\leq v+\langle xI\{x\geq 0\},b_i^L\rangle+\langle xI\{x<0\},b_i^U\rangle\},\; I\{e\leq v+\langle xI\{x\geq 0\},b_i^U\rangle \\ &+\langle xI\{x<0\},b_i^L\rangle\}\big] \end{aligned}$$

for each $v \in \mathbb{R}$. Further the above bracket has $L^2(P)$-length

$$\begin{aligned}
&\Big(E\big[I\{\varepsilon \le v + \langle XI\{X \ge 0\}, b_i^U\rangle + \langle XI\{X < 0\}, b_i^L\rangle\} \\
&\quad - I\{\varepsilon \le v + \langle XI\{X \ge 0\}, b_i^L\rangle + \langle XI\{X < 0\}, b_i^U\rangle\}\big]^2\Big)^{1/2} \\
&\le \Big(E\big[F\big(v + \langle X, b_i^U\rangle\big) - F\big(v + \langle X, b_i^L\rangle\big)\big]\Big)^{1/2} \\
&\le \Big(E\big[c|\langle X, b_i^U - b_i^L\rangle|^\gamma\big]\Big)^{1/2} \\
&\le c^{1/2}(E\|X\|)^{\gamma/2}\|b_i^U - b_i^L\|^{\gamma/2} \\
&= O(\epsilon),
\end{aligned}$$

by assumption (a.2), Cauchy–Schwarz and Jensen's inequality. Similar to the proof of Lemma 1 in Akritas and Van Keilegom (2001) one obtains an upper bound $O(\epsilon^{-2}\exp(\epsilon^{-2/k}))$ for the $L^2(P)$-bracketing number of the class $\mathcal{F}$. Thus $\mathcal{F}$ is P-Donsker by the bracketing integral condition in Theorem 19.5 of van der Vaart (1998). □

A.3 Proof of Lemma 4.1

To show the assertion for $\hat{F}_{k_n^*}$ we use the arguments as in the proof of Theorem 2.2 for the process H_n, but based on the iid sample $(X_1, Y_1), \ldots, (X_{k_n^*}, Y_{k_n^*})$ before the change. Then as in the proof of Theorem 2.2 asymptotic stochastic equicontinuity of the process $H_{k_n^*}$ holds and thus

$$\sup_{t\in[0,1], z\in\mathbb{R}} |H_{k_n^*}(t, z, \hat{\alpha} - \alpha, \hat{\beta} - \beta)| = o_{\mathbb{P}}(1).$$

Here, $\hat{\alpha}$ and $\hat{\beta}$ depend on the whole sample and assumption (a.1) is used. Thus as in Eq. (A.1) we obtain

$$\hat{F}_{k_n^*}(z) = F_{k_n^*}(z) + E_{X_1}[F_1(z + \hat{\alpha} - \alpha + \langle X_1, \hat{\beta} - \beta\rangle)] - F_1(z) + o_{\mathbb{P}}\Big(\frac{1}{\sqrt{n}}\Big)$$

uniformly in $z \in \mathbb{R}$. By the classical Glivenko–Cantelli result $F_{k_n^*}$ converges uniformly almost surely to F_1. By assumptions (a.1), (a.2)' and $E\|X_1\| < \infty$ the remainder term is $o_{\mathbb{P}}(1)$ and the assertion for $\hat{F}_{k_n^*}$ follows. The assertion for $\tilde{F}_{k_n^*}$ can be shown analogously. □

A.4 Proof of Lemma 4.2

First note that

$$\hat{\vartheta}_n \in \arg\max_{t\in[0,1]} \left(\sup_{z\in\mathbb{R}} |\hat{G}_n(t,z)| \right) = \arg\max_{t\in[0,1]} \left(\sup_{z\in\mathbb{R}} \left| \frac{\hat{G}_n(t,z)}{n^{1/2}} \right| \right).$$

Further it holds

$$\begin{aligned}
\frac{\hat{G}_n(t,z)}{n^{1/2}} &= \frac{\lfloor nt \rfloor (n - \lfloor nt \rfloor)}{n^2} \Big(\frac{1}{\lfloor nt \rfloor} \sum_{i=1}^{\lfloor nt \rfloor} I\{\hat{\varepsilon}_i \le z\} - \frac{1}{n - \lfloor nt \rfloor} \sum_{i=\lfloor nt \rfloor + 1}^{n} I\{\hat{\varepsilon}_i \le z\} \Big) \\
&= \frac{\lfloor nt \rfloor (n - \lfloor nt \rfloor)}{n^2} \Big(\frac{1}{\lfloor nt \rfloor} \sum_{i=1}^{\lfloor nt \rfloor \wedge \lfloor n\vartheta^* \rfloor} I\{\hat{\varepsilon}_i \le z\} + I\{t > \vartheta^*\} \frac{1}{\lfloor nt \rfloor} \sum_{i=\lfloor n\vartheta^* \rfloor + 1}^{\lfloor nt \rfloor} I\{\hat{\varepsilon}_i \le z\} \\
&\quad - \frac{1}{n - \lfloor nt \rfloor} \sum_{i=\lfloor nt \rfloor \vee \lfloor n\vartheta^* \rfloor + 1}^{n} I\{\hat{\varepsilon}_i \le z\} - I\{t < \vartheta^*\} \frac{1}{n - \lfloor nt \rfloor} \sum_{i=\lfloor nt \rfloor + 1}^{\lfloor n\vartheta^* \rfloor} I\{\hat{\varepsilon}_i \le z\} \Big) \\
&= \frac{\lfloor nt \rfloor (n - \lfloor nt \rfloor)}{n^2} \Big(\frac{\lfloor nt \rfloor \wedge \lfloor n\vartheta^* \rfloor}{\lfloor nt \rfloor} F_1(z) + I\{t > \vartheta^*\} \frac{\lfloor nt \rfloor - \lfloor n\vartheta^* \rfloor}{\lfloor nt \rfloor} F_2(z) \\
&\quad - \frac{n - \lfloor nt \rfloor \vee \lfloor n\vartheta^* \rfloor}{n - \lfloor nt \rfloor} F_2(z) - I\{t < \vartheta^*\} \frac{\lfloor n\vartheta^* \rfloor - \lfloor nt \rfloor}{n - \lfloor nt \rfloor} F_1(z) \Big) + o_{\mathbb{P}}(1),
\end{aligned}$$

since we have

$$\begin{aligned}
&\sup_{t\in[0,\vartheta^*]} \sup_{z\in\mathbb{R}} \frac{\lfloor nt \rfloor}{n} \left| \frac{1}{\lfloor nt \rfloor} \sum_{i=1}^{\lfloor nt \rfloor} I\{\hat{\varepsilon}_i \le z\} - F_1(z) \right| \\
&\le \underbrace{\sup_{t\in[0,\vartheta^*]} \sup_{z\in\mathbb{R}} \frac{\lfloor nt \rfloor}{\lfloor n\vartheta^* \rfloor} \left| \frac{1}{\lfloor nt \rfloor} \sum_{i=1}^{\lfloor nt \rfloor} I\{\hat{\varepsilon}_i \le z\} - \frac{1}{\lfloor n\vartheta^* \rfloor} \sum_{i=1}^{\lfloor n\vartheta^* \rfloor} I\{\hat{\varepsilon}_i \le z\} \right|}_{= \frac{1}{\lfloor n\vartheta^* \rfloor^{1/2}} \tilde{G}_{\lfloor n\vartheta^* \rfloor}(t,z)} \qquad \text{(A.3)} \\
&\quad + \frac{\lfloor nt \rfloor}{\lfloor n\vartheta^* \rfloor} \left| \frac{1}{\lfloor n\vartheta^* \rfloor} \sum_{i=1}^{\lfloor n\vartheta^* \rfloor} I\{\hat{\varepsilon}_i \le z\} - F_1(z) \right| \qquad \text{(A.4)} \\
&= o_{\mathbb{P}}(1).
\end{aligned}$$

Here we have used Lemma 4.1 for the term (A.4). Further $\tilde{G}_{\lfloor n\vartheta^* \rfloor}$ is defined as $\hat{G}_n$ based on the iid-sample $(X_1, Y_1), \ldots, (X_{k_n^*}, Y_{k_n^*})$, but where the residuals are built with $\hat{\alpha}$, $\hat{\beta}$ based on the whole sample. With the same argument as in the proof of Theorem 2.2 it holds that

$$\tilde{G}_{\lfloor n\vartheta^* \rfloor}(t,z) = G_{\lfloor n\vartheta^* \rfloor}(t,z) + o_{\mathbb{P}}(1)$$

uniformly in $t \in [0,1]$, $z \in \mathbb{R}$, see (A.2), and thus the term (A.3) is $o_{\mathbb{P}}(1)$.

Analogously one can show that $\sup_{t\in[\vartheta^*,1]} \sup_{z\in\mathbb{R}} \frac{n-\lfloor nt\rfloor}{n} \left| \frac{1}{n-\lfloor nt\rfloor} \sum_{i=\lfloor nt\rfloor+1}^{n} I\{\hat{\varepsilon}_i \le z\} - F_2(z)\right| = o_{\mathbb{P}}(1)$.

Thus, it holds uniformly in $t \in [0, 1]$

$$\begin{aligned}\frac{\hat{G}_n(t,z)}{n^{1/2}} &= I\{t > \vartheta^*\}\Big(\frac{\lfloor n\vartheta^*\rfloor(n-\lfloor nt\rfloor)}{n^2}(F_1(z)-F_2(z))\Big)\\ &\quad + I\{t \le \vartheta^*\}\Big(\frac{\lfloor nt\rfloor(n-\lfloor n\vartheta^*\rfloor)}{n^2}(F_1(z)-F_2(z))\Big) + o_{\mathbb{P}}(1)\\ &= \Big(I\{t > \vartheta^*\}\vartheta^*(1-t) + I\{t \le \vartheta^*\}t(1-\vartheta^*)\Big)(F_1(z)-F_2(z)) + o_{\mathbb{P}}(1).\end{aligned}$$

The assertion then follows by Theorem 2.12 in Kosorok (2008) as ϑ^* is well-separated maximum of $t \mapsto I\{t > \vartheta^*\}\vartheta^*(1-t) + I\{t \le \vartheta^*\}t(1-\vartheta^*)$. □

Acknowledgements The authors are grateful to the Editors and Guest Editors for the organization of the Special Issue "Goodness-of-Fit, Change-Point, and Related Problems", and to the referees, the Associate Editor and the Guest Editor Simos Meintanis for their constructive comments and interesting ideas to expand the topic.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Akritas MG, Van Keilegom I (2001) Non-parametric estimation of the residual distribution. Scand J Stat 28(3):549–567

Aston JAD, Kirch C (2012) Detecting and estimating changes in dependent functional data. J Multivar Anal 109:204–220

Aue A, Gabrys R, Horváth L, Kokoszka P (2009) Estimation of a change-point in the mean function of functional data. J Multivar Anal 100(10):2254–2269

Aue A, Hörmann S, Horváth L, Hušková M (2014) Dependent functional linear models with applications to monitoring structural change. Stat Sin 24:1043–1073

Aue A, Rice G, Sönmez O (2018) Detecting and dating structural breaks in functional data without dimension reduction. J R Stat Soc Ser B Stat Methodol 80(3):509–529

Azzalini A, Capitanio A (1999) Statistical applications of the multivariate skew-normal distribution. J R Stat Soc Ser B Stat Methodol 61:579–602

Bai J (1994) Weak convergence of the sequential empirical processes of residuals in ARMA models. Ann Stat 22:2051–2061

Berkes I, Gabrys R, Horváth L, Kokoszka P (2009) Detecting changes in the mean of functional observations. J R Stat Soc Ser B Stat Methodol 71(5):927–946

Bickel PJ, Wichura MJ (1971) Convergence criteria for multiparameter stochastic processes and some applications. Ann Math Stat 42:1656–1670

Boente G, Parada D (2023) Robust estimation for functional quadratic regression models. Comput Stat Data Anal 187:107798

Cardot H, Mas A, Sarda P (2007) Clt in functional linear regression models. Probab Theory Relat Fields 138:325–361
Carlstein E (1988) Nonparametric change-point estimation. Ann Stat 16(1):188–197
Csörgö M, Horváth L, Szyszkowicz B (1997) Integral tests for suprema of kiefer processes with application. Stat Decis 15:365–377
Dümbgen L (1991) The asymptotic behavior of some nonparametric change-point estimators. Ann Stat 19(3):1471–1495
Ferraty F, Vieu P (2006) Nonparametric functional data analysis, Springer series in statistics. Springer, New York
Giné E, Nickl R (2021) Mathematical foundations of infinite-dimensional statistical models. Cambridge series in statistical and probabilistic mathematics. Cambridge University Press
Hariz SB, Wylie JJ, Zhang Q (2005) Nonparametric change-point estimation for dependent sequences. CR Math 341(10):627–630
Hariz SB, Wylie JJ, Zhang Q (2007) Optimal rate of convergence for nonparametric change-point estimators for nonstationary sequences. Ann Stat 35(4):1802–1826
Horváth L, Kokoszka P (2012) Inference for functional data with applications, Springer series in statistics. Springer, New York
Horvath et al (2024) Variable selection based testing for parameter changes in regression with autoregressive dependence. J Bus Econ Stat 42(4):1331–1343
Kosorok MR (2008) Introduction to empirical processes and semiparametric inference. Springer, New York
Koul HL (1996) Asymptotics of some estimators and sequential residual empiricals in nonlinear time series. Ann Stat 24:380–404
Koul HL (2002) Weighted empirical processes in dynamic nonlinear models, Lecture notes in statistics. Springer, New York
Lee ER, Park BU (2012) Sparse estimation in functional linear regression. J Multivar Anal 105(1):1–17
Ling S (1998) Weak convergence of the sequential empirical processes of residuals in nonstationary autoregressive models. Ann Stat 26:741–754
Neumeyer N, Van Keilegom I (2009) Change-point tests for the error distribution in non-parametric regression. Scand J Stat 36(3):518–541
Neumeyer N, Dette H, Nagel E-R (2005) A note on testing symmetry of the error distribution in linear regression models. J Nonparametr Stat 17(6):697–715
Neumeyer N, Dette H, Nagel E-R (2006) Bootstrap tests for the error distribution in linear and nonparametric regression models. Austr New Zealand J Stat 48(2):129–156
Neumeyer N, Omelka M (2025) Generalized Hadamard differentiability of the copula mapping and its applications. Bernoulli (to appear)
Pardo Fernandez JC (2007) Comparison of error distributions in nonparametric regression. Stat Probab Lett 77:350–356
Picard D (1985) Testing and estimating change-points in time series. Adv Appl Probab 17:841–867
Pierce DA, Kopecky KJ (1979) Testing goodness of fit for the distribution of errors in regression models. Biometrika 66(1):1–5
Selk L, Neumeyer N (2013) Testing for a change of the innovation distribution in nonparametric autoregression: the sequential empirical process approach. Scand J Stat 40(4):770–788
Shang Z, Cheng G (2015) Nonparametric inference in generalized functional linear models. Ann Stat 43(4):1742–1773
Shorack GR, Wellner JA (1986) Empirical processes with applications to statistics. Wiley
van der Vaart A, Wellner JA (1996) Weak convergence and empirical processes. Springer, New York
van der Vaart A (1998) Asymptotic statistic. Cambridge series in statistical and probabilistic mathematics. Cambridge University Press
Yeon H, Dai X, Nordman DJ (2023) Bootstrap inference in functional linear regression models with scalar response. Bernoulli 29(4):2599–2626
Yuan M, Cai T (2010) A reproducing kernel Hilbert space approach to functional linear regression. Ann Stat 38(6):3412–3444

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.