



**HAL**  
open science

# Estimation in a change-point non linear quantile model

Gabriela Ciuperca

► **To cite this version:**

Gabriela Ciuperca. Estimation in a change-point non linear quantile model. Communications in Statistics - Theory and Methods, 2017, 46 (12), pp.6017-6034. 10.1080/03610926.2015.1116576 . hal-04447054

**HAL Id: hal-04447054**

**<https://hal.science/hal-04447054>**

Submitted on 8 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Estimation in a change-point nonlinear quantile model

GABRIELA CIUPERCA \*

Université Lyon 1, UMR 5208, Institut Camille Jordan, France

March 21, 2022

**ABSTRACT.** *This paper considers a nonlinear quantile model with change-points. The quantile estimation method, which as a particular case includes median model, is more robust with respect to other traditional methods when model errors contain outliers. Under relatively weak assumptions, the convergence rate and asymptotic distribution of change-point and of regression parameter estimators are obtained. Numerical study by Monte Carlo simulations shows the performance of the proposed method for nonlinear model with change-points.*

**Keywords:** Multiple change-points; Quantile regression; Asymptotic behaviour.

**Mathematics Subject Classification:** Primary 62F10, 62F12 ; Secondary 62J02.

## 1 Introduction

Classically, for linear or nonlinear models, the errors are supposed with mean zero and bounded variance. In this case, model parameters are estimated generally by least squares (LS) method. If these conditions are not satisfied or if model contains outliers,

---

\*Université Lyon 1, Institut Camille Jordan, Bat. Braconnier, 43, blvd du 11 novembre 1918, F - 69622 Villeurbanne Cedex, France, E-mail: *Gabriela.Ciuperca@univ-lyon1.fr*,

then the LS estimators of the model parameters can have a large error. A very interesting and robust alternative method was proposed by Koenker and Bassett (1978) by the introduction of the quantile method. A particular case of this method is that of least absolute deviation (LAD). For a complete overview on quantile method, we refer the reader to book of Koenker (2005). Properties of a nonlinear quantile model are studied also in the papers Chen et al. (2013), Choi et al. (2005), Oberhofer and Haupt (2014).

On the other hand, in applications, it is possible that we have not one but several models, the localization where model changes being unknown. We obtain which is called as a change-point model. The purpose of this paper is to study the properties of this type of model estimated by quantile method, when between two consecutive change-points the model is nonlinear. For this study we need to know the asymptotic behaviour of the objective function.

To our knowledge, most previous studies of change-point models have focused on linear models. On this subject, we can mention the following papers: Bai (1998) for LAD method, Bai and Perron (1998) for LS method, Koul and Qian (2002) for maximum likelihood method, Koul et al. (2003) for M-estimation method. For quantile method, Oka and Qu (2011) estimate the change-points location and the coefficient parameters of each phase, Furno (2012) realize a Lagrange multiplier test for detecting the structural breaks. For change-point nonlinear model, because of difficulties caused by the nonlinearity, literature is less rich: Boldea and Hall (2013) use LS method to estimate and test the number of breaks. In Ciuperca (2009), the M-estimation method is used to estimate a multiphase nonlinear model with random design and changes in the model due to some (unknown) values in design. A general criterion is proposed in Ciuperca (2011a) to determine the change-point number. If changes in the model occur in time, the LAD estimation method was studied by Ciuperca (2011b).

Present paper generalizes Ciuperca (2011b), considering a method, for estimating and for choosing the change-point number criterion, based on the quantile framework. This is because, often in practice, especially in the case of change-point models, the quantile index  $\tau$  of errors is not  $1/2$ .

We note the important fact that, in a multiple change-point model, the change-point estimation could affect the estimator properties. Moreover, it is difficult to study, theoretically but also numerically, a change-point model since it depends of two parameter types: the regression parameters and the change-points.

The plan of this paper is as follows. In Section 2 we first introduce some notations and assumptions. Next, we study the asymptotic behaviour of the objective function. In Section 3, we define and study the quantile estimator in a nonlinear model with change-points. Convergence rate and asymptotic distributions of the estimators are obtained. Finally, in Section 4, simulation results illustrate the performance of the quantile method for change-point nonlinear model. In Appendix Bernstein's inequality is recalled.

## 2 Quantile regression without change

In this section we study the asymptotic behaviour of the quantile process.

Let us consider the following regression model

$$Y_i = g(\mathbf{X}_i, \boldsymbol{\phi}) + \varepsilon_i, \quad i = 1, \dots, n, \quad (1)$$

where the regression function  $g : \Upsilon \times \Gamma \rightarrow \mathbb{R}$ , with  $\boldsymbol{\phi} \in \Gamma \subseteq \mathbb{R}^p$ ,  $\mathbf{x} \in \Upsilon$ ,  $\Upsilon \subseteq \mathbb{R}^d$  is known up to the parameter  $\boldsymbol{\phi}$ . We suppose that the set  $\Gamma$  is compact.

For a fixed quantile index  $\tau \in (0, 1)$ , the  $\tau$ th conditional quantile regression of  $Y$ , given  $\mathbf{x}$ , is  $g(\mathbf{x}, \boldsymbol{\phi}) + F^{-1}(\tau)$ , with  $F^{-1}(\tau)$  the  $\tau$ th quantile ( $F^{-1}$  is the inverse of the distribution function  $F$ ) of error  $\varepsilon$ . We suppose that  $F(0) = \tau$ .

In the case when the model contains intercept, noted  $\phi_0$ , the regression function has the form  $g(\mathbf{x}, \boldsymbol{\phi}) = \phi_0 + g_0(\mathbf{x}, \boldsymbol{\phi}_1)$ . Then, the following parameter vector is considered  $\boldsymbol{\phi}(\tau) = (\phi_0(\tau), \boldsymbol{\phi}_1)^t$ , with  $\phi_0(\tau) = \phi_0 + F^{-1}(\tau)$ . Thus, when model contains an intercept  $\phi_0$ , we estimate first  $\phi_0(\tau)$  from which we then have the estimation of  $\phi_0$ . For linear models, in literature, the presence or not of intercept intervenes in the result proofs (see for example Oka and Qu (2011)).

The estimator of  $\boldsymbol{\phi}(\tau)$  by Koenker and Bassett technique (see Koenker and Bassett

(1978)) is called *quantile regression*. We suppose  $\tau$  fixed, therefore, for simplicity reasons, we will note  $\phi(\tau)$  by  $\phi$ . Contrary to the papers where linear models are studied (see for example Oka and Qu (2011)) when  $\phi_0 \neq 0$ , in this paper we shall consider simultaneously the two cases presented above.

## 2.1 Assumptions and notations

In this subsection we give assumptions and notations needed in this paper.

For simplicity, we suppose that the regressors  $\mathbf{X}_i$  are non random, although the results will, typically hold for random  $\mathbf{X}_i$ 's independent of the  $\varepsilon_i$ 's and if  $\mathbf{X}_i$  independent of  $\mathbf{X}_j$  for  $i \neq j$ .

For the model (1), we consider  $\phi^0$  the true value (unknown) of  $\phi$ , with  $\phi^0$  an inner point of the compact  $\Gamma$ .

For a fixed quantile index  $\tau \in (0, 1)$ , consider the check function  $\rho : \mathbb{R} \rightarrow \mathbb{R}$  given by

$$\rho_\tau(u) = u[\tau - \mathbb{1}_{u \leq 0}]$$

and random variable

$$D_i(\tau) \equiv \tau - \mathbb{1}_{\varepsilon_i \leq 0}.$$

Since  $F(0) = \tau$  we have  $\mathbb{E}[D_i(\tau)] = 0$ .

The quantile estimator of parameter  $\phi$  is defined by

$$\hat{\phi}_n^{(\tau)} \equiv \arg \min_{\phi \in \Gamma} \sum_{i=1}^n \rho_\tau(Y_i - g(\mathbf{X}_i, \phi)). \quad (2)$$

Its consistency, with  $n^{-1/2}$  convergence rate and asymptotic normality of estimator  $\hat{\phi}_n^{(\tau)}$  have been proved in previous papers (see for example Koenker (2005)).

For reading convenience, all throughout the paper, let us consider the following notation, for all  $\phi \in \Gamma$  and  $i = 1, \dots, n$ ,

$$h_i(\phi) \equiv g(\mathbf{X}_i, \phi) - g(\mathbf{X}_i, \phi^0).$$

In order to study the quantile model, let us consider the following two random processes:

$$\begin{cases} W_n(\tau; \boldsymbol{\phi}, \boldsymbol{\phi}^0) \equiv - \sum_{i=1}^n D_i(\tau) h_i(\boldsymbol{\phi}), \\ Z_n(\boldsymbol{\phi}, \boldsymbol{\phi}^0) \equiv \sum_{i=1}^n \int_0^{h_i(\boldsymbol{\phi})} (\mathbb{1}_{\varepsilon_i \leq s} - \mathbb{1}_{\varepsilon_i \leq 0}) ds. \end{cases} \quad (3)$$

Obviously  $\mathbb{E}[W_n(\tau; \boldsymbol{\phi}, \boldsymbol{\phi}^0)] = 0$ .

For each sample  $i \in \{1, \dots, n\}$ , in order to study always the quantile model, consider the following difference

$$G_i^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0) \equiv \rho_\tau(\varepsilon_i - g(\mathbf{X}_i, \boldsymbol{\phi}) + g(\mathbf{X}_i, \boldsymbol{\phi}^0)) - \rho_\tau(\varepsilon_i),$$

from which us define the random process

$$\mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0) \equiv \sum_{i=1}^n G_i^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0),$$

which is in fact the objective function for finding the quantile estimator  $\hat{\boldsymbol{\phi}}_n^{(\tau)}$  of model (1).

Using identity of Knight (1998), for any real nonzero number  $r$ , we have that  $\rho_\tau(r - s) - \rho_\tau(r) = s[\mathbb{1}_{r < 0} - \tau] + \int_0^s [\mathbb{1}_{r \leq t} - \mathbb{1}_{r \leq 0}] dt$ . Then, the process  $\mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)$  can also write as

$$\mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0) = W_n(\tau; \boldsymbol{\phi}, \boldsymbol{\phi}^0) + Z_n(\boldsymbol{\phi}, \boldsymbol{\phi}^0). \quad (4)$$

In the case of a nonlinear model, function  $\mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)$  is not convex in  $\boldsymbol{\phi}$  which means that the study of estimator  $\hat{\boldsymbol{\phi}}_n^{(\tau)}$  and of function  $\mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)$  will be different than for a linear model, based on the convexity.

All throughout the paper,  $C$  denotes a positive generic constant which may take different values in different formula or even in different parts of the same formula. All vectors are column and  $\mathbf{v}^t$  denotes the transposed of  $\mathbf{v}$ . All vectors and matrices are in bold. Concerning the used norms, for a  $m$ -vector  $\mathbf{v} = (v_1, \dots, v_m)$ , let us denote by  $\|\mathbf{v}\|_1 = \sum_{j=1}^m |v_j|$  and  $\|\mathbf{v}\|_2 = (\sum_{j=1}^m v_j^2)^{1/2}$ . For a matrix  $\mathbf{M} = (a_{ij})_{\substack{1 \leq i \leq m_1 \\ 1 \leq j \leq m_2}}$ , we denote by  $\|\mathbf{M}\|_1 = \max_{j=1, \dots, m_2} (\sum_{i=1}^{m_1} |a_{ij}|)$ , the subordinate norm to the vector norm  $\|\cdot\|_1$ .

We now state the assumptions on the errors and on the regression function.

The errors  $(\varepsilon_i)_{1 \leq i \leq n}$  are supposed independent identically distributed (i.i.d.) random variables. We denote by  $f$  the density and by  $F$  the distribution function of  $\varepsilon$ .

**(A1)** There exists two constants  $c_0 > 0$  and  $\delta_0 > 0$  such that for all  $|x| \leq \delta_0$ , we have

$$\min(F(|x|) - F(0), F(0) - F(-|x|)) \geq c_0|x|.$$

Contrary to the classic assumptions for a nonlinear regression model, we do not impose the condition that the mean of errors  $\varepsilon_i$  is zero or that their variance is bounded.

The regression function  $g(\mathbf{x}, \phi)$  is supposed twice differentiable in  $\phi$  and continuous on  $\Upsilon$ . In the following, for  $\mathbf{x} \in \Upsilon$  and  $\phi \in \Gamma$  we use notation  $\dot{\mathbf{g}}(\mathbf{x}, \phi) \equiv \partial g(\mathbf{x}, \phi) / \partial \phi$  and  $\ddot{\mathbf{g}}(\mathbf{x}, \phi) \equiv \partial^2 g(\mathbf{x}, \phi) / \partial \phi^2$ . Moreover, for the function  $g$ , following assumptions are considered:

**(A2)** For all  $\mathbf{x} \in \Upsilon$ , function  $\dot{\mathbf{g}}(\mathbf{x}, \phi)$  is bounded in every  $\eta$ -neighbourhood of  $\phi^0$ , when  $\eta \rightarrow 0$ .

**(A3)** There exists  $c_1 > 0$  such that  $n^{-1} \sum_{i=1}^n \sup_{\phi \in \Gamma} \|\dot{\mathbf{g}}(\mathbf{X}_i, \phi)\|_2 \leq c_1 < \infty$ .

**(A4)** There exist two positive constants  $c_2, c_3 > 0$  and natural  $n_0$  such that for all  $\phi_1, \phi_2 \in \Gamma$  and  $n \geq n_0$ :  $c_2 \|\phi_1 - \phi_2\|_2 \leq (n^{-1} \sum_{i=1}^n [g(\mathbf{X}_i, \phi_1) - g(\mathbf{X}_i, \phi_2)]^2)^{1/2} \leq c_3 \|\phi_1 - \phi_2\|_2$ . Moreover, we have  $n^{-1} \sum_{i=1}^n \dot{\mathbf{g}}(\mathbf{X}_i, \phi^0) \dot{\mathbf{g}}^t(\mathbf{X}_i, \phi^0)$  converges, as  $n \rightarrow \infty$ , to a positive definite matrix. Furthermore,  $\max_{1 \leq i \leq n} n^{-1/2} \|\ddot{\mathbf{g}}(\mathbf{X}_i, \phi^0)\|_2 \rightarrow 0$ .

**(A5)** For all  $\mathbf{x} \in \Upsilon$ ,  $\phi \in \mathcal{V}_\eta(\phi^0)$ , with  $\eta \rightarrow 0$ , we have that  $\|\ddot{\mathbf{g}}(\mathbf{x}, \phi)\|_1$  is bounded.

For certain results, stronger assumptions are necessary:

**(A6)** For all  $\phi \in \Gamma$ ,  $\mathbf{x} \in \Upsilon$ , we have that  $\|\dot{\mathbf{g}}(\mathbf{x}, \phi)\|_2$  is bounded.

**(A7)** For all  $\phi \in \Gamma$ ,  $\mathbf{x} \in \Upsilon$ , we have that  $\|\ddot{\mathbf{g}}(\mathbf{x}, \phi)\|_1$  is bounded.

We wish to emphasise the fact that, with respect to the particular case  $\tau = 1/2$ , considered in Ciuperca (2011b), we consider here multidimensional regressors  $\mathbf{X}_i$ .

Assumption (A2) means that for every  $\eta \rightarrow 0$ , the function  $\|\dot{\mathbf{g}}(\mathbf{x}, \phi)\|_2$  is bounded for all  $\mathbf{x} \in \Upsilon$  and for all  $\phi \in \mathcal{V}_\eta$ , with

$$\mathcal{V}_\eta \equiv \{\phi \in \Gamma; \|\phi - \phi^0\|_2 \leq \eta\}.$$

Assumptions (A1), (A4) are needed that the objective function has a unique minimum at  $\phi^0$  and for convergence and asymptotic normality of the quantile estimator (see Koenker (2005), page 124). Obviously that, assumption (A7) implies (A5). We have also that (A6) implies (A2), (A3) and third condition of (A4).

Note that assumption (A3) is the same as in paper Ciuperca (2011b), for a median nonlinear model and (A1) is supposition (C4) of Oberhofer and Haupt (2014)'s paper, for a nonlinear quantile regression with weakly dependent errors. As noted in the last paper, assumption (A1) is stronger than the usual assumption in literature:  $f(x)$  exists in a neighbourhood of  $x = 0$  and  $f(0) \geq c_0 > 0$ .

## 2.2 Asymptotic behaviour of objective function $\mathcal{G}_n^{(\tau)}(\phi, \phi^0)$

In order to study the main part of this paper devoted to a change-point model estimated by quantile framework, in this subsection the asymptotic behaviour of process  $\mathcal{G}_n^{(\tau)}(\phi, \phi^0)$  is studied.

Recall that under assumptions (A1) and (A4), the quantile estimator  $\hat{\phi}_n^{(\tau)}$ , given by (2), is weakly  $n^{-1/2}$ -consistent (see Koenker (2005) or Oberhofer (1982)).

**Remark 2.1** *By elementary calculations we show that, for all  $\phi \in \Gamma$ :*

- if  $h_i(\phi) \geq 0$  we have

$$\int_0^{h_i(\phi)} \mathbb{1}_{\varepsilon_i < t} dt = h_i(\phi) \mathbb{1}_{\varepsilon_i < 0} + [h_i(\phi) - \varepsilon_i] \mathbb{1}_{0 \leq \varepsilon_i \leq h_i(\phi)} + 0 \cdot \mathbb{1}_{\varepsilon_i > h_i(\phi)},$$

- if  $h_i(\phi) < 0$  we have

$$\int_0^{h_i(\phi)} \mathbb{1}_{\varepsilon_i < t} dt = 0 \cdot \mathbb{1}_{\varepsilon_i > 0} - \varepsilon_i \mathbb{1}_{h_i(\phi) \leq \varepsilon_i \leq 0} + h_i(\phi) \mathbb{1}_{\varepsilon_i < 0}.$$

*These imply that*

- if  $h_i(\phi) \geq 0$  we have

$$\int_0^{h_i(\phi)} [\mathbb{1}_{\varepsilon_i < t} - \mathbb{1}_{\varepsilon_i < 0}] dt = [h_i(\phi) - \varepsilon_i] \mathbb{1}_{0 \leq \varepsilon_i \leq h_i(\phi)},$$

- if  $h_i(\boldsymbol{\phi}) < 0$  we have

$$\int_0^{h_i(\boldsymbol{\phi})} [\mathbb{1}_{\varepsilon_i < t} - \mathbb{1}_{\varepsilon_i < 0}] dt = [-\varepsilon_i - h_i(\boldsymbol{\phi})] \mathbb{1}_{h_i(\boldsymbol{\phi}) \leq \varepsilon_i \leq 0}.$$

Thus, for  $Z_n(\boldsymbol{\phi}, \boldsymbol{\phi}^0)$  defined by (3), we have with probability one that,  $Z_n(\boldsymbol{\phi}, \boldsymbol{\phi}^0) \geq 0$  for all  $\boldsymbol{\phi} \in \Gamma$ .

A consequence of this remark, taking into account that  $\mathbb{E}[W_n(\tau; \boldsymbol{\phi}, \boldsymbol{\phi}^0)] = 0$  for all  $\boldsymbol{\phi} \in \Gamma$ , is that

$$\mathbb{E}[\mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)] \geq 0, \quad \text{for all } \boldsymbol{\phi} \in \Gamma. \quad (5)$$

We will now study the asymptotic behaviour of the objective function  $\mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)$ . In this purpose, for a bounded deterministic sequence  $c_n$ , let us consider the following parameter set:

$$\Omega_{c_n} \equiv \{\boldsymbol{\phi} \in \Gamma; \|\boldsymbol{\phi} - \boldsymbol{\phi}^0\|_2 \leq c_n\}.$$

Emphasize that for the following Proposition, claim (i), the sequence  $(c_n)$  is a constant  $c$ . Thus, we denote the set  $\Omega_{c_n}$  by  $\Omega_c$ . The proof idea is the same as in Bai (1998), Lemma 4, only now the nonlinearity of  $g(\mathbf{x}, \boldsymbol{\phi})$  and the quantile regression intervene significantly.

**Proposition 2.1** *Let us consider a deterministic positive sequence  $(a_n)$  such that  $a_n \rightarrow \infty$ , as  $n \rightarrow \infty$ .*

(i) *Under assumption (A6), if sequence  $(a_n)$  satisfies in addition the conditions  $n^{-1}a_n = O(1)$ ,  $n^{-1}a_n^2/\log n \rightarrow \infty$  as  $n \rightarrow \infty$ , and the parameters belong to the set  $\Omega_c = \{\boldsymbol{\phi} \in \Gamma; \|\boldsymbol{\phi} - \boldsymbol{\phi}^0\|_2 \leq c\}$ , we have that, for all  $\epsilon > 0$ , there exists a constant  $C > 0$  and a natural number  $n_\epsilon \in \mathbb{N}$  such that for all  $n \geq n_\epsilon$ ,*

$$\mathbb{P} \left[ \sup_{\boldsymbol{\phi} \in \Omega_c} \left| \frac{1}{a_n} \left[ \mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0) - \mathbb{E}[\mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)] \right] \right| > \epsilon \right] \leq \exp(-C\epsilon^2 n^{-1} a_n^2).$$

(ii) *Under assumptions (A2), (A3), if we have furthermore an another sequence  $(c_n)$  such that  $c_n \rightarrow 0$ ,  $n^{-1}c_n^{-1}a_n = O(1)$ ,  $a_n^2/(nc_n^2 \log n) \rightarrow \infty$  as  $n \rightarrow \infty$  and the parameters*

belong to the set  $\Omega_{c_n}$ , we have that, for all  $\epsilon > 0$ , there exists a constant  $C > 0$  and a natural number  $n_\epsilon \in \mathbb{N}$  such that for all  $n \geq n_\epsilon$ ,

$$\mathbb{P} \left[ \sup_{\phi \in \Omega_{c_n}} \left| \frac{1}{a_n} [\mathcal{G}_n^{(\tau)}(\phi, \phi^0) - \mathbb{E}[\mathcal{G}_n^{(\tau)}(\phi, \phi^0)]] \right| > \epsilon \right] \leq \exp(-C\epsilon^2 n^{-1} c_n^{-2} a_n^2).$$

**Proof.** (i) We decompose  $\Omega_c$  in subsets, such that the diameter of each subset is less than  $cn^{-1/2}$ . Thus, we can write  $\Omega_c = \cup_{j=1}^{C_p n^{p/2}} \omega_j^n$ , where  $\omega_j^n \equiv \{\phi \in \Omega_c; \|\phi - \phi_j\|_2 \leq cn^{-1/2}\}$ , with  $\phi_j \in \Omega_c$ . This decomposition depends on the dimension of the parameter set  $\Gamma$ , then positive constant  $C_p (< \infty)$  depends on  $p$ . Diameter  $cn^{-1/2}$  of each cell  $\omega_j^n$  was taken as that, for two points belonging to the same cell, the difference which occurs in the left hand side of following relation (6) converges to 0 as  $n \rightarrow \infty$ . Then, in order to study the behaviour of  $\mathcal{G}_n^{(\tau)}(\phi, \phi^0) - \mathbb{E}[\mathcal{G}_n^{(\tau)}(\phi, \phi^0)]$  for all  $\phi \in \Gamma$ , we take one representative point, noted  $\phi_j$ , of  $\omega_j^n$ , for each  $j = 1, \dots, C_p n^{p/2}$ . Then, in order to study the probability of the left hand side of claim (i) we just study the probability of the left hand side of (7). Since  $n^{-1}a_n^2/\log n \rightarrow \infty$ , then we have that  $n^{-1/2}a_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Thus, for each  $\phi_1, \phi_2 \in \omega_j^n$ , using assumption  $n^{-1} \sum_{i=1}^n \sup_{\phi \in \Gamma} \|\dot{\mathbf{g}}(\mathbf{X}_i, \phi)\|_2 \leq c_1 < \infty$  of (A6), we have

$$\begin{aligned} & a_n^{-1} |\mathcal{G}_n^{(\tau)}(\phi_1, \phi^0) - \mathbb{E}[\mathcal{G}_n^{(\tau)}(\phi_1, \phi^0)] - \mathcal{G}_n^{(\tau)}(\phi_2, \phi^0) + \mathbb{E}[\mathcal{G}_n^{(\tau)}(\phi_2, \phi^0)] \\ & \leq C n a_n^{-1} \|\phi_1 - \phi_2\|_2 \leq C n^{1/2} a_n^{-1} \rightarrow 0. \end{aligned} \quad (6)$$

For  $\phi_j \in \omega_j^n$ , for any  $j = 1, \dots, C_p n^{p/2}$ , we have the following obvious inequality

$$\mathbb{P} \left[ \sup_j a_n^{-1} |\mathcal{G}_n^{(\tau)}(\phi_j, \phi^0) - \mathbb{E}[\mathcal{G}_n^{(\tau)}(\phi_j, \phi^0)]| > \epsilon \right] \leq \sum_{j=1}^{C_p n^{p/2}} \mathbb{P} \left[ |\mathcal{G}_n^{(\tau)}(\phi_j, \phi^0) - \mathbb{E}[\mathcal{G}_n^{(\tau)}(\phi_j, \phi^0)]| > \epsilon a_n \right]. \quad (7)$$

On the other hand, taking into account assumption (A6), we obtain that there exists a constant  $C > 0$  such that  $|\mathcal{G}_i^{(\tau)}(\phi_j, \phi^0) - \mathbb{E}[\mathcal{G}_i^{(\tau)}(\phi_j, \phi^0)]| \leq C \|\phi_j - \phi^0\|_2 \leq Cc$ , with probability 1.

Moreover  $\text{Var} [\mathcal{G}_n^{(\tau)}(\phi_j, \phi^0)] \leq Cnc^2$ , with  $C > 0$ .

Then, since  $n^{-1}a_n = O(1)$ ,  $a_n^2/(n \log n) \rightarrow \infty$  as  $n \rightarrow \infty$ , we can apply Bernstein's inequality (24) (see Appendix), for  $\beta = Cc$ ,  $V = Cnc^2$ ,  $s = 1/2$ ,  $z = a_n \epsilon$  and we obtain

that

$$\begin{aligned} \sum_{j=1}^{C_p n^{p/2}} \mathbb{P}[|\mathcal{G}_n^{(\tau)}(\phi_j, \phi^0) - \mathbb{E}[\mathcal{G}_n^{(\tau)}(\phi_j, \phi^0)]| > \epsilon a_n] &\leq 2C_p n^{p/2} \exp(-C\epsilon^2 n^{-1} a_n^2) \\ &= 2C_p \exp(-C\epsilon^2 n^{-1} a_n^2 + p/2 \log n) \rightarrow 0, \end{aligned}$$

with  $C > 0$ .

The claim follows combining the last relation together with (6) and (7).

(ii) In this case, we decompose the set  $\Omega_{c_n} = \cup_{j=1}^{C_p n^{p/2}} \omega_j^n$ , with the subsets  $\omega_j^n \equiv \{\phi \in \Omega_{c_n}; \|\phi - \phi_j\|_2 \leq c_n n^{-1/2}\}$ , with  $\phi_j \in \Omega_{c_n}$ . Since  $a_n^2/(nc_n^2 \log n) \rightarrow \infty$  we have that  $n^{1/2}c_n/a_n \rightarrow 0$  as  $n \rightarrow \infty$ . Then, for  $\phi_1, \phi_2 \in \omega_j^n$  we have by similar arguments as in (6):  $a_n^{-1}|\mathcal{G}_n^{(\tau)}(\phi_1, \phi^0) - \mathbb{E}[\mathcal{G}_n^{(\tau)}(\phi_1, \phi^0)] - \mathcal{G}_n^{(\tau)}(\phi_2, \phi^0) + \mathbb{E}[\mathcal{G}_n^{(\tau)}(\phi_2, \phi^0)]| \leq Cn^{1/2}c_n a_n^{-1} \rightarrow 0$ , as  $n \rightarrow \infty$ . On the other hand, for  $\phi_j \in \omega_j^n$ , for any  $j = 1, \dots, C_p n^{p/2}$ , using assumption (A2), (A3) we have:  $|\mathcal{G}_i^{(\tau)}(\phi_j, \phi^0) - \mathbb{E}[\mathcal{G}_i^{(\tau)}(\phi_j, \phi^0)]| \leq Cc_n$ , with probability 1. The end of proof is similar to that of (i), using  $\text{Var}[\mathcal{G}_n^{(\tau)}(\phi_j, \phi^0)] \leq Cnc_n^2$  and applying Bernstein's inequality (24) for  $\beta = Cc_n$ ,  $V = Cnc_n^2$ ,  $s = 1/2$ ,  $z = a_n\epsilon$ .  $\blacksquare$

**Remark 2.2** An example of sequence  $(a_n)$  which satisfies the conditions of Proposition 2.1(i) is  $a_n = n/(\log n)^s$ , with  $s \geq 0$ .

Applying Proposition 2.1 and Borel-Cantelli lemma we obtain that for all  $\epsilon > 0$ , we have

$$\limsup_{n \rightarrow \infty} \left( \sup_{\phi \in \Omega_{c_n}} \frac{1}{a_n} \left| \mathcal{G}_n^{(\tau)}(\phi, \phi^0) - \mathbb{E}[\mathcal{G}_n^{(\tau)}(\phi, \phi^0)] \right| \right) \leq \epsilon, \quad a.s. \quad (8)$$

which is the equivalent of Lemma A.2 of Oka and Qu (2011).

The following Proposition proves a general result that the infimum of objective function  $\mathcal{G}_n^{(\tau)}(\phi, \phi^0)$  is strictly positive for  $\phi$  outside a ball centred at  $\phi^0$  with radius  $w_n$ , when  $w_n \rightarrow 0$ .

**Proposition 2.2** *Suppose that assumptions (A1), (A4)-(A6) are satisfied, that density  $f$  of  $\varepsilon$  is differentiable in a neighbourhood of 0 and that  $f'(x)$  is bounded in this neighbourhood.*

*Let  $(w_n)$  be a monotone deterministic sequence converging to 0, such that  $nw_n^2/\log n \rightarrow \infty$ , as  $n \rightarrow \infty$ .*

*Then, there exists  $\epsilon_1 > 0$  such that,*

$$\liminf_{n \rightarrow \infty} \left( \inf_{\|\phi - \phi^0\|_2 \geq w_n} \frac{1}{nw_n^2} \mathcal{G}_n^{(\tau)}(\phi, \phi^0) \right) > \epsilon_1,$$

*with probability 1.*

**Proof.** Let us first consider a monotone positive sequence  $(v_n)$  such that  $v_n \rightarrow 0$ , as  $n \rightarrow \infty$  and  $v_n \geq w_n$  for  $n$  large enough.

We now consider parameter  $\phi$  such that  $\|\phi - \phi^0\|_2 = v_n$ . For a  $p$ -vector  $\mathbf{u}$  in an open set of  $\mathbb{R}^p$ , with  $\|\mathbf{u}\|_2 = 1$ , we have using (4),

$$\mathbb{E}[\mathcal{G}_n^{(\tau)}(\phi^0 + v_n \mathbf{u}, \phi^0)] = \mathbb{E}[Z_n(\phi^0 + v_n \mathbf{u}, \phi^0)] = \sum_{i=1}^n \int_0^{h_i(\phi^0 + v_n \mathbf{u})} [F(s) - F(0)] ds.$$

We make the Taylor's expansion for the distribution function  $F$  and we obtain

$$\mathbb{E}[\mathcal{G}_n^{(\tau)}(\phi^0 + v_n \mathbf{u}, \phi^0)] = \sum_{i=1}^n \int_0^{h_i(\phi^0 + v_n \mathbf{u})} s f(0) ds + \sum_{i=1}^n \int_0^{h_i(\phi^0 + v_n \mathbf{u})} \frac{s^2}{2} f'(\zeta_i s) ds \equiv R_1 + R_2, \quad (9)$$

with  $0 < \zeta_i < 1$ .

The Taylor's expansion up to order 1 of  $g$  at  $\phi = \phi^0$ , the assumptions (A5), (A6) imply that  $R_1$  is equal to

$$v_n^2 \frac{f(0)}{2} \sum_{i=1}^n [\mathbf{u}^t \dot{\mathbf{g}}(\mathbf{X}_i, \phi^0) + v_n \mathbf{u}^t \ddot{\mathbf{g}}(\mathbf{X}_i, \tilde{\phi}_i) \mathbf{u}]^2 = v_n^2 \frac{f(0)}{2} \sum_{i=1}^n [\mathbf{u}^t \dot{\mathbf{g}}(\mathbf{X}_i, \phi^0)]^2 (1 + o(1)), \quad (10)$$

with  $\tilde{\phi}_i = \phi^0 + b_i v_n \mathbf{u}$ ,  $b_i \in [0, 1]$ ,  $i = 1, \dots, n$ .

For  $R_2$ , by similar arguments and since  $f'$  is bounded in the neighbourhood of 0, we have

$$|R_2| \leq C \sum_{i=1}^n \int_0^{h_i(\phi^0 + v_n \mathbf{u})} \frac{s^2}{2} ds = C v_n^3 \sum_{i=1}^n [\mathbf{u}^t \dot{\mathbf{g}}(\mathbf{X}_i, \phi^0)]^3 (1 + o(1)) = o(R_1). \quad (11)$$

Then, taking into account (9), (10) and (11) we obtain:

$$\mathbb{E}[\mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}^0 + v_n \mathbf{u}, \boldsymbol{\phi}^0)] = w_n^2 \frac{f(0)}{2} \sum_{i=1}^n [\mathbf{u}^t \dot{\mathbf{g}}(\mathbf{X}_i, \boldsymbol{\phi}^0)]^2 (1 + o(1)),$$

with  $o(1)$  uniformly in  $\mathbf{u}$ . Since  $\|\mathbf{u}\|_2 = 1$ , we have then that

$$\mathbb{E}[\mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}^0 + v_n \mathbf{u}, \boldsymbol{\phi}^0)] \geq n w_n^2 \frac{f(0)}{2} \lambda_{\min, n} (1 + o(1)),$$

where  $\lambda_{\min, n}$  is the smallest eigenvalue of matrix  $n^{-1} \sum_{i=1}^n \dot{\mathbf{g}}(\mathbf{X}_i, \boldsymbol{\phi}^0) \dot{\mathbf{g}}^t(\mathbf{X}_i, \boldsymbol{\phi}^0)$ .

On the other hand, assumption (A4) implies that there exists a  $\lambda > 0$  such that  $\lambda_{\min, n} \rightarrow \lambda$ , as  $n \rightarrow \infty$ . Then, since  $n v_n^2 \rightarrow \infty$ , we have that for all large enough  $n$ :

$$\frac{1}{n v_n^2} \mathbb{E}[\mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}^0 + v_n \mathbf{u}, \boldsymbol{\phi}^0)] \geq \lambda \frac{f(0)}{2} (1 + o(1)). \quad (12)$$

Under assumptions (A6), we take  $\epsilon = \lambda f(0)/4$ ,  $\Omega_n \equiv \{\boldsymbol{\phi} \in \Gamma; \|\boldsymbol{\phi} - \boldsymbol{\phi}^0\|_2 = v_n\}$  and  $a_n = n v_n^2$  for Proposition 2.1(ii). Then, relation (8) follows. Moreover, since  $(n v_n^2)^{-1} \text{Var}[\mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)] < \infty$ ,  $n v_n^2 / \log n \rightarrow \infty$ , together with relation (12), we have that

$$\liminf_{n \rightarrow \infty} \left( \inf_{\|\boldsymbol{\phi} - \boldsymbol{\phi}^0\|_2 = v_n} (n v_n^2)^{-1} \mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0) \right) > \eta, \quad (13)$$

with  $\eta = \lambda f(0)/4$ .

Let us now consider a monotone positive sequence  $(v_n)$  such that  $v_n$  (or no sub-sequence) does not converge to 0. Then, there exists  $\delta > 0$  such that  $v_n \geq \delta$  for any  $n$  large enough. By simple algebraic computations we obtain that:

$$\mathbb{E}[G_i^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)] = \begin{cases} \int_0^{-h_i(\boldsymbol{\phi})} [|h_i(\boldsymbol{\phi})| - x] dF(x), & \text{if } h_i(\boldsymbol{\phi}) < 0, \\ \int_{-h_i(\boldsymbol{\phi})}^0 [|h_i(\boldsymbol{\phi})| + x] dF(x), & \text{if } h_i(\boldsymbol{\phi}) \geq 0, \end{cases}$$

Then

$$\begin{aligned} \mathbb{E}[G_i^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)] &\geq \mathbb{1}_{h_i(\boldsymbol{\phi}) \geq 0} \int_{-\frac{h_i(\boldsymbol{\phi})}{2}}^0 (|h_i(\boldsymbol{\phi})| + x) dF(x) + \mathbb{1}_{h_i(\boldsymbol{\phi}) < 0} \int_0^{-\frac{h_i(\boldsymbol{\phi})}{2}} (|h_i(\boldsymbol{\phi})| - x) dF(x) \\ &\geq \mathbb{1}_{h_i(\boldsymbol{\phi}) \geq 0} \frac{|h_i(\boldsymbol{\phi})|}{2} \int_{-\frac{h_i(\boldsymbol{\phi})}{2}}^0 dF(x) + \mathbb{1}_{h_i(\boldsymbol{\phi}) < 0} \frac{|h_i(\boldsymbol{\phi})|}{2} \int_0^{-\frac{h_i(\boldsymbol{\phi})}{2}} dF(x) \\ &= \frac{|h_i(\boldsymbol{\phi})|}{2} \left( \mathbb{1}_{h_i(\boldsymbol{\phi}) \geq 0} [F(0) - F(-\frac{h_i(\boldsymbol{\phi})}{2})] + \mathbb{1}_{h_i(\boldsymbol{\phi}) < 0} [F(-\frac{h_i(\boldsymbol{\phi})}{2}) - F(0)] \right). \end{aligned}$$

Taking into account the assumptions (A1) and (A4), for  $\|\phi - \phi^0\|_2 = v_n \geq \delta$ , we have that there exists a constant  $c > 0$  such that:

$$\mathbb{E}[n^{-1}\mathcal{G}_n^{(\tau)}(\phi, \phi^0)] > c. \quad (14)$$

Under assumption (A6), taking  $\epsilon = c/2$ ,  $\Omega_n \equiv \{\phi \in \Gamma; \|\phi - \phi^0\|_2 = v_n\}$  and  $a_n = nv_n^2$ , for Proposition 2.1, we obtain the relation (8). Then, together with relation (14) we have that

$$\liminf_{n \rightarrow \infty} \left( \inf_{\|\phi - \phi^0\|_2 = v_n} (nv_n^2)^{-1} \mathcal{G}_n^{(\tau)}(\phi, \phi^0) \right) > \eta, \quad (15)$$

with  $\eta = \epsilon = c/2$ .

Since relations (13) and (15) are valid for any positive sequence  $(v_n)$ , such that  $v_n \geq w_n$ , and since  $(w_n)$  is monotonic, the Proposition follows if we consider  $\epsilon_1 = \eta$ . ■

The following two lemma will be needed in the next section, where model contains change-points. The change-points are the observations where model changes. In the next section, we will estimate simultaneously these change-points but also the model parameters between two change-points. The following Lemma will be used to find the convergence rate of the change-points estimators.

**Lemma 2.1** *For  $1 \leq l < k \leq n$  such that  $k - l \rightarrow \infty$ , as  $n \rightarrow \infty$ , under assumptions (A2), (A7), we have, for all  $\alpha > 1/2$ ,*

$$\sup_{1 \leq l < k \leq n} \left| \inf_{\phi} \sum_{i=l}^k G_i^{(\tau)}(\phi, \phi^0) \right| = O_{\mathbb{P}}(n^{-\alpha}).$$

**Proof.** Since for all  $r_1, r_2 \in \mathbb{R}$  we have that  $|\rho_{\tau}(r_1) - \rho_{\tau}(r_2)| < |r_1 - r_2|$ , then

$$|\rho(\varepsilon_i - h_i(\phi_1)) - \rho(\varepsilon_i - h_i(\phi_2))| \leq |h_i(\phi_1) - h_i(\phi_2)|.$$

For  $\|\phi_1 - \phi_2\|_2 \leq Cn^{-1/2}$ , using the Taylor expansion up to order 1 for each of the two functions  $h_i(\phi_1)$ ,  $h_i(\phi_2)$  in respect to  $\phi$ , around  $\phi^0$ , and using assumptions (A2), (A7), we obtain that  $\sum_{i=1}^n \{\rho(\varepsilon_i - h_i(\phi_1)) - \rho(\varepsilon_i - h_i(\phi_2)) - \mathbb{E}[\rho(\varepsilon_i - h_i(\phi_1)) - \rho(\varepsilon_i - h_i(\phi_2))]\} \leq$

$O_{\mathcal{P}}(n^{1/2})$ .

Since  $\mathbb{E}[G_i^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)] \geq 0$  for all  $\boldsymbol{\phi}$ , we have that  $G_i^{(\tau)}(\boldsymbol{\phi}^0, \boldsymbol{\phi}^0) = 0 \geq \inf_{\boldsymbol{\phi}} \sum_{i=l}^k G_i^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0) \geq \inf_{\boldsymbol{\phi}} \sum_{i=l}^k (G_i^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0) - \mathbb{E}[G_i^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)])$ .

The rest of proof is similar to that of Lemma 2.3 of Ciuperca (2014).  $\blacksquare$

By the following Lemma we prove that the objective function  $\mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)$  given by (4) varies little when a small portion of observations is ignored.

**Lemma 2.2** *Under assumptions (A2), (A5), for all parameter  $\boldsymbol{\phi}$  such that  $\|\boldsymbol{\phi} - \boldsymbol{\phi}^0\|_2 \leq n^{-1/2}$  and for  $M \in \mathbb{N}$  arbitrary, we have, as  $n \rightarrow \infty$ ,*

$$\sup_{m \in [n-M, n]} \sup_{\boldsymbol{\phi}, \|\boldsymbol{\phi} - \boldsymbol{\phi}^0\|_2 \leq n^{-1/2}} |\mathcal{G}_m^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0) - \mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)| = o_{\mathcal{P}}(1).$$

**Proof.** Let us consider a natural number  $m$  such that  $m \in [n-M, n]$ . Since  $|\rho_{\tau}(\varepsilon - r) - \rho_{\tau}(\varepsilon)| \leq |r|$ , we have that  $|\mathcal{G}_m^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0) - \mathcal{G}_n^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)| \leq \sum_{i=m+1}^n |g(\mathbf{X}_i, \boldsymbol{\phi}) - g(\mathbf{X}_i, \boldsymbol{\phi}^0)| \leq \sum_{i=m+1}^n \|\boldsymbol{\phi} - \boldsymbol{\phi}^0\|_1 \|\dot{\mathbf{g}}(\mathbf{X}_i, \boldsymbol{\phi}^0) + \ddot{\mathbf{g}}(\mathbf{X}_i, \tilde{\boldsymbol{\phi}})(\boldsymbol{\phi} - \boldsymbol{\phi}^0)^t\|_1$ , with  $\tilde{\boldsymbol{\phi}} = \boldsymbol{\phi}^0 + a(\boldsymbol{\phi} - \boldsymbol{\phi}^0)$ ,  $a \in [0, 1]$ . Since  $\ddot{\mathbf{g}}$  is bounded in a neighbourhood of  $\boldsymbol{\phi}^0$  by assumption (A5), applying also the Markov inequality and assumption (A2), the last sum is smaller than  $n^{-1}O_{\mathcal{P}}(1) = O_{\mathcal{P}}(n^{-1})$ .  $\blacksquare$

**Lemma 2.3** *Under the same assumptions as in Proposition 2.2, we have, for all  $\delta \in (0, 1)$ :*

$$\sup_{[n\delta] \leq m \leq n} \left| \inf_{\boldsymbol{\phi}, \|\boldsymbol{\phi} - \boldsymbol{\phi}^0\|_2 \leq n^{-1/2}} \mathcal{G}_m^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0) \right| = O_{\mathcal{P}}(1).$$

**Proof.**

Given the convergence rate of the quantile estimator, we have

$$\sup_{[n\delta] \leq m \leq n} \left| \arg \min_{\boldsymbol{\phi} \in \Gamma} \mathcal{G}_m^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0) \right| = O_{\mathcal{P}}(n^{-1/2}).$$

Then, it is sufficient to prove that

$$\sup_{1 \leq m \leq n} \sup_{\|\mathbf{u}\|_2 \leq M} |\mathcal{G}_m^{(\tau)}(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0)| = O_{\mathcal{P}}(1). \quad (16)$$

On the other hand, by relation (3), for any  $\mathbf{u} \in \mathbb{R}^p$  such that  $\|\mathbf{u}\|_2 \leq M$ , we have  $\mathcal{G}_m^{(\tau)}(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0) = W_m(\tau; \boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0) + Z_m(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0)$ , with, by definition:  $Z_m(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0) = \sum_{i=1}^m \int_0^{h_i(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u})} [\mathbb{1}_{\varepsilon_i \leq s} - \mathbb{1}_{\varepsilon_i \leq 0}] ds$  and  $W_m(\tau; \boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0) = -\sum_{i=1}^m D_i(\tau) h_i(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u})$ .

We first study  $W_m$ . Recall that  $\mathbb{E}[W_m(\tau; \boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0)] = 0$ . On the other hand, taking the Taylor expansion up to order 1 of  $g$ , we have:

$$W_m(\tau; \boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0) = -\sum_{i=1}^m D_i \mathbf{u}^t [n^{-1/2} \dot{\mathbf{g}}(\mathbf{X}_i, \boldsymbol{\phi}^0) + n^{-1} \ddot{\mathbf{g}}(\mathbf{X}_i, \tilde{\boldsymbol{\phi}}_i) \mathbf{u}]$$

with  $\tilde{\boldsymbol{\phi}}_i = \boldsymbol{\phi}^0 + b_i n^{-1/2} \mathbf{u}$ ,  $b_i \in [0, 1]$ ,  $i = 1, \dots, m$ .

By the central limit theorem, using assumptions (A4) and (A6), we have:

$$n^{-1/2} \sum_{i=1}^m D_i \mathbf{u}^t \dot{\mathbf{g}}(\mathbf{X}_i, \boldsymbol{\phi}^0) = O_{\mathbb{P}}(1), \quad (17)$$

uniformly in  $m$  and  $\mathbf{u}$ . Since  $\|\ddot{\mathbf{g}}(\mathbf{X}_i, \tilde{\boldsymbol{\phi}}_i)\|_1$  is bounded by assumption (A5), we obtain:

$$n^{-1} \sum_{i=1}^m \mathbf{u}^t \ddot{\mathbf{g}}(\mathbf{X}_i, \tilde{\boldsymbol{\phi}}_i) \mathbf{u} = O_{\mathbb{P}}(1), \quad (18)$$

uniformly in  $m$  and  $\mathbf{u}$ . Thus, by (17) and (18), we have

$$W_m(\tau; \boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0) = O_{\mathbb{P}}(1), \quad (19)$$

uniformly in  $m$  and  $\mathbf{u}$ .

We study now  $Z_m(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0)$ :

$$\begin{aligned} \mathbb{E}[Z_m(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0)] &= \sum_{i=1}^m \int_0^{h_i(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u})} [F(s) - F(0)] ds \\ &= \sum_{i=1}^m \int_0^{h_i(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u})} s f(0) ds + \sum_{i=1}^m \int_0^{h_i(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u})} \frac{s^2}{2} f'(\zeta_i s) ds \end{aligned}$$

with  $0 < \zeta_i < 1$ .

Using the same arguments as in the proof of Proposition 2.2, we obtain:

$$\mathbb{E}[Z_m(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0)] = \frac{f(0)}{2n} \sum_{i=1}^m [\mathbf{u}^t \dot{\mathbf{g}}(\mathbf{X}_i, \boldsymbol{\phi}^0)]^2 (1 + o(1)) = O(1).$$

On the other hand, using similar arguments to those for  $W_m(\tau; \boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0)$ , we have that

$$\text{Var}[Z_m(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0)] \leq \sum_{i=1}^m h_i^2(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0) = O(1),$$

uniformly in  $m$  and  $\mathbf{u}$ .

Then, the Bienaymé-Tchebychev inequality, we obtain that  $Z_m(\boldsymbol{\phi}^0 + n^{-1/2}\mathbf{u}, \boldsymbol{\phi}^0) = O_{\mathcal{P}}(1)$ . Taking into account (16) and (19), the Lemma follows.  $\blacksquare$

### 3 Quantile regression with multiple change-points

This section considers that the nonlinear model changes to unknown observations. More specifically, the regression parameters change to unknown times. First, we define the quantile estimators of the model parameters. If the number of changes is known, we give the convergence rate and the limiting distribution of the all estimators. Next, we give a consistent criterion for estimating the change-point number.

Consider a model with  $K$  change-points, i.e. a model which changes to observations  $l_1, \dots, l_K$ , with  $1 < l_1 < \dots, l_K < n$ ,

$$Y_i = \sum_{r=0}^K g(\mathbf{X}_i, \boldsymbol{\phi}_{r+1}) \mathbb{1}_{l_r \leq i < l_{r+1}} + \varepsilon_i, \quad (20)$$

$i = 1, \dots, n$ , with  $l_0 = 1$  and  $l_{K+1} = n$ .

We assume that numbers of changes  $K$  is known.

Concerning the change-point location, we suppose that each segment contains a significant proportion of samples:

**(A8)**  $l_{r+1} - l_r \geq n^a$ ,  $a > 1/2$ , for all  $r = 0, \dots, K$ , with  $l_0 = 1$  and  $l_{K+1} = n$ .

This condition is necessary in order to apply Lemma 2.1, therefore constant "a" must be strictly greater than 1/2.

For fixed  $K$ , the parameters of model (20) are the regression parameters  $\boldsymbol{\theta}_1 \equiv (\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_{K+1}) \in \Gamma^{K+1}$  and the change-points  $\boldsymbol{\theta}_2 \equiv (l_1, \dots, l_K) \in \mathbb{N}^K$ . The true values of the parameters are  $\boldsymbol{\theta}_1^0 \equiv (\boldsymbol{\phi}_1^0, \dots, \boldsymbol{\phi}_{K+1}^0)$  for the regression parameters and  $\boldsymbol{\theta}_2^0 \equiv (l_1^0, \dots, l_K^0)$  for the change-points. Obviously  $\boldsymbol{\phi}_{r+1}^0 \neq \boldsymbol{\phi}_r^0$ , for all  $r = 1, \dots, K$ .

We define the quantile estimators of parameters  $\boldsymbol{\theta}_1$  and  $\boldsymbol{\theta}_2$  by

$$(\hat{\boldsymbol{\theta}}_{1n}^{(\tau)}, \hat{\boldsymbol{\theta}}_{2n}^{(\tau)}) \equiv \arg \min_{(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)} \sum_{r=1}^{K+1} \sum_{i=l_{r-1}+1}^{l_r} \rho_\tau(Y_i - g(\mathbf{X}_i, \boldsymbol{\phi}_r)). \quad (21)$$

See Ciuperca (2011b) for a discussion on the construction of the estimators in a change-point model.

In order to prove the convergence rate of the change-point quantile estimator  $\hat{\boldsymbol{\theta}}_{2n}^{(\tau)}$ , we first prove that if in a phase we take in the place of the true regression parameter those of the nearby phase, then the value of the objective function is different from that calculated for the true value.

**Lemma 3.1** *Under assumption (A6), we have for every  $r = 1, \dots, K$ , when  $l_r < l_r^0$  such that  $l_r^0 - l_r \rightarrow \infty$ , that there exists  $\eta > 0$ ,  $C > 0$  such that*

$$\mathbb{P} \left[ \left| \sum_{i=l_r+1}^{l_r^0} (\rho_\tau(\varepsilon_i - g(\mathbf{X}_i, \boldsymbol{\phi}_{r+1}^0) + g(\mathbf{X}_i, \boldsymbol{\phi}_r^0)) - \rho(\varepsilon_i)) \right| \geq \eta(l_r^0 - l_r) \right] \geq 1 - \exp(-C(l_r^0 - l_r)).$$

**Proof.** Since  $\boldsymbol{\phi}_{r+1}^0 \neq \boldsymbol{\phi}_r^0$ , there exists  $\delta$ , and  $\epsilon_0 > 0$  such that  $|g(\mathbf{X}_i, \boldsymbol{\phi}_{r+1}^0) - g(\mathbf{X}_i, \boldsymbol{\phi}_r^0)| \geq \delta$  for  $(l_r^0 - l_r)\epsilon_0$  observations. Then

$$\mathbb{E} \left[ \sum_{i=l_r+1}^{l_r^0} G_i^{(\tau)}(\boldsymbol{\phi}_{r+1}^0, \boldsymbol{\phi}_r^0) \right] = (l_r^0 - l_r)\epsilon_0 \int_{-\delta}^0 (x + \delta) dF(x).$$

Applying Proposition 2.1(i) for  $c = \|\boldsymbol{\phi}_{r+1}^0 - \boldsymbol{\phi}_r^0\|_2$  and  $a_n = l_r^0 - l_r$ , we have that for all  $\epsilon > 0$ , the following inequality

$$\mathbb{P} \left[ \left| \sum_{i=l_r+1}^{l_r^0} G_i^{(\tau)}(\boldsymbol{\phi}_{r+1}^0, \boldsymbol{\phi}_r^0) - \mathbb{E} \left[ \sum_{i=l_r+1}^{l_r^0} G_i^{(\tau)}(\boldsymbol{\phi}_{r+1}^0, \boldsymbol{\phi}_r^0) \right] \right| \geq \epsilon(l_r^0 - l_r) \right] \leq \exp(-C\epsilon^2(l_r^0 - l_r)).$$

Then, Lemma follows considering  $\eta = \epsilon = 2^{-1}\epsilon_0 \int_{-\delta}^0 (x + \delta)dF(x)$ . ■

**Remark 3.1** *Using Lemma 2.3 and Proposition 2.1, by similar technique to one use in the paper of Ciuperca (2011b), for Lemmas 7 and 8, and in the paper Ciuperca (2014), for Lemmas 3 and 4, we obtain their equivalent. That is, if data come from two different models, the quantile estimator is close to the parameter of the model from where most of the data came.*

Following result shows that the distance between the change-point quantile estimator and the true value is finished.

**Theorem 3.1** *Under assumptions (A1), (A4), (A6)-(A8) and if density function  $f$  of  $\epsilon$  satisfies conditions of Proposition 2.2, then we have  $\|\hat{\boldsymbol{\theta}}_{2n}^{(\tau)} - \boldsymbol{\theta}_2^0\|_2 = O_{\mathcal{P}}(1)$ .*

**Proof.** The proof is similarly to that of Theorem 3.1 of Ciuperca (2014), using relation (8),  $\mathbb{E}[G_i^{(\tau)}(\boldsymbol{\phi}, \boldsymbol{\phi}^0)] \geq 0$  by (5) and Lemma 2.1, Proposition 2.2, Lemma 3.1, Remark 3.1. We omit all details. ■

With this result we can now give the asymptotic distributions, first for the change-point estimator and then for the regression parameter estimator. This result is the generalization of that obtained in Ciuperca (2011b) for LAD method ( $\tau = 1/2$ ), where the proof was based on norm  $L_1$  of objective function. The asymptotic distribution of the change-point quantile estimator depends on regression function  $g$ , on the true regression parameters to the left and right of the estimated break point and of quantile index  $\tau$  of  $\epsilon$ . The asymptotic distribution of regression parameter quantile estimator is Gaussian, with covariance matrix dependent of  $\tau$ . Theorem 3.2 is a standard result in quantile model without change-point (Koenker (2005)) and in a change-point model estimated by other methods. See i.e. Boldea and Hall (2013) for LS method.

We consider by convention that  $l_0^0 = \hat{l}_0^{(\tau)} = 1$  and  $l_{K+1}^0 = \hat{l}_{K+1}^{(\tau)} = n$ .

**Theorem 3.2** *Under the same conditions of Theorem 3.1, we have the following asymptotic laws of the change-point quantile estimators:*

(i) for each  $r = 1, \dots, K$ ,

$$(\hat{l}_r^{(\tau)} - l_r^0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \arg \min_{j \in \mathbb{Z}} Z_{r,j}^{(\tau)},$$

where:

- if  $j = 1, 2, \dots$ ,

$$Z_{r,j}^{(\tau)} \equiv \sum_{i=l_r^0+1}^{l_r^0+j} [\rho_\tau(\varepsilon_i - g(\mathbf{X}_i, \phi_r^0) + g(\mathbf{X}_i, \phi_{r+1}^0)) - \rho_\tau(\varepsilon_i)].$$

- if  $j = -1, -2, \dots$ ,

$$Z_{r,j}^{(\tau)} \equiv \sum_{i=l_r^0+j}^{l_r^0} [\rho_\tau(\varepsilon_i - g(\mathbf{X}_i, \phi_{r+1}^0) + g(\mathbf{X}_i, \phi_r^0)) - \rho_\tau(\varepsilon_i)].$$

(ii) for each  $r = 1, \dots, K + 1$ ,

$$(\hat{l}_r^{(\tau)} - \hat{l}_{r-1}^{(\tau)})^{1/2} (\hat{\phi}_r^{(\tau)} - \phi_r^0) \Sigma_r^{1/2} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(\mathbf{0}, \frac{\tau(1-\tau)}{f^2(0)} \mathbf{I}_p),$$

with

$$\Sigma_r \equiv (l_r^0 - l_{r-1}^0)^{-1} \sum_{i=l_{r-1}^0+1}^{l_r^0} \dot{g}(\mathbf{X}_i, \phi_r^0) \dot{g}^t(\mathbf{X}_i, \phi_r^0)$$

and  $\mathbf{I}_p$  the identity matrix of order  $p$ .

**Proof.** Let us consider the set of change-point vectors

$$\Theta_2 \equiv \{\boldsymbol{\theta}_2 = (l_1, \dots, l_K); l_j = l_j^0 + m, |m| \leq C_2, \forall j = 1, \dots, K\}$$

and the set of regression parameter vectors

$$\Theta_1 \equiv \{\boldsymbol{\theta}_1 = (\phi_1, \dots, \phi_K); (l_j^0 - l_{j-1}^0)^{1/2} \|\phi_j - \phi_j^0\|_2 \leq C_3, \forall j = 1, \dots, K + 1\},$$

with  $C_2, C_3 > 0$  finite constants. Let be the sum

$$S_n(\tau, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \equiv \sum_{j=0}^K \sum_{i=l_j+1}^{l_{j+1}} \rho_\tau(Y_i - g(\mathbf{X}_i, \phi_{j+1})).$$

Consider the following identity

$$\inf_{\boldsymbol{\theta}_2} \inf_{\boldsymbol{\theta}_1} S_n(\tau, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \inf_{\boldsymbol{\theta}_2} \inf_{\boldsymbol{\theta}_1} (S_n(\tau, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2^0) + S_n(\tau, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) - S_n(\tau, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2^0)). \quad (22)$$

By Lemma 2.2 we have that  $[S_n(\tau, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) - S_n(\tau, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2^0)] - [S_n(\tau, \boldsymbol{\theta}_1^0, \boldsymbol{\theta}_2) - S_n(\tau, \boldsymbol{\theta}_1^0, \boldsymbol{\theta}_2^0)] = o_{\mathcal{P}}(1)$  uniformly in  $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$  belonging in  $\Theta_1 \times \Theta_2$ .

Without loss of generality, we suppose that  $\hat{l}_r^{(\tau)} \leq l_r^0$ .

By the definition of  $S_n$ , we have that

$$S_n(\tau, \boldsymbol{\theta}_1^0, \boldsymbol{\theta}_2) - S_n(\tau, \boldsymbol{\theta}_1^0, \boldsymbol{\theta}_2^0) = \sum_{r=1}^K \sum_{i=l_{r+1}^0}^{l_r^0} [\rho_\tau(Y_i - g(\mathbf{X}_i, \phi_{r+1}^0)) - \rho_\tau(Y_i - g(\mathbf{X}_i, \phi_r^0))].$$

Then, relation (22) becomes

$$\inf_{\boldsymbol{\theta}_1 \in \Theta_1} S_n(\tau, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2^0) + \inf_{\boldsymbol{\theta}_2 \in \Theta_2} \sum_{r=1}^K \sum_{i=l_{r+1}^0}^{l_r^0} [\rho_\tau(Y_i - g(\mathbf{X}_i, \phi_{r+1}^0)) - \rho_\tau(Y_i - g(\mathbf{X}_i, \phi_r^0))] (1 + o_{\mathcal{P}}(1)).$$

Theorem results taking into account that every term of this last relation depends on different parameters, together with convergence rate of the estimators (by Theorem 3.1 for the change-point estimator and  $(l_r^0 - l_{r-1}^0)^{-1/2}$  for the regression parameter quantile estimator) and limit law of quantile estimator for a nonlinear model (see for example Koenker (2005)). ■

**Remark 3.2** *In the case presented here, parameters  $\phi_r, \phi_{r+1}$  from a segment to the other are fixed. In the paper of Oka and Qu (2011) for linear model, it is supposed that the difference between two consecutive parameter tends to zero as  $n \rightarrow \infty$ . Then, the limit law of the change-points estimators is totally different, it is the maximizer of a Wiener process with drift.*

**Remark 3.3** *In order to determine the number of change points, we can use a similar criterion to that proposed in the paper of Ciuperca (2014) for a linear quantile model. Under conditions that  $\mathbb{E}[\rho_\tau(\varepsilon)] > 0$  and  $\mathbb{E}[\rho_\tau^2(\varepsilon)] < \infty$ , we propose the following consistent estimator of the change-point number  $K$*

$$\hat{K}_n^{(\tau)} \equiv \arg \min_K \left( n \log \left( n^{-1} S_n(\tau, \hat{\boldsymbol{\theta}}_{1n}^{(\tau)}(K), \hat{\boldsymbol{\theta}}_{2n}^{(\tau)}(K)) \right) + P(K, p) B_n \right), \quad (23)$$

where the function  $S_n$  is defined in the proof of Theorem 3.2,  $(\hat{\boldsymbol{\theta}}_{1n}^{(\tau)}(K), \hat{\boldsymbol{\theta}}_{2n}^{(\tau)}(K))$  is the quantile estimators of  $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$  for a fixed  $K$ ,  $(B_n)$  is a deterministic sequence converging to infinity such that  $B_n n^{-a} \rightarrow 0$ ,  $B_n n^{-1/2} \rightarrow \infty$ , as  $n \rightarrow \infty$  and the penalty function  $P(K, p)$  is such that  $P(K_1, p) \leq P(K_2, p)$  for all number change-points  $K_1 \leq K_2$ . Recall that the constant  $a$  is that of the supposition (A8) and  $p$  is parameter number of the regression function  $g$ .

The proof of the consistency of the criterion is similar to that in Ciuperca (2014). We do not give the details.

## 4 Simulation study for change-point nonlinear models

To evaluate the performance of the quantile method in a change-point nonlinear model, Monte Carlo simulations are realized. We compare the performance of the least squares (LS) and quantile estimation methods. We use *quantreg*, *VGAM* packages in *R* to run the simulations.

For each model, 100 Monte Carlo samples of size  $n$  are generated for regressor  $X$  and error  $\varepsilon$ .

Throughout this section, we generate the design  $X \sim \mathcal{N}(1, 1)$  and the regression function  $g(x, \boldsymbol{\phi})$  is growth function  $b_1 - \exp(-b_2 x)$ , or more exactly the mono-molecular model (see Seber and Wild (2003)), with  $\boldsymbol{\phi} = (b_1, b_2)$ . The same regression function has been considered in Ciuperca (2011a) using the M-method that has the least squares method as a special case. For the errors  $\varepsilon$ , three distributions were considered: standard Normal  $\mathcal{N}(0, 1)$ , Laplace  $\mathcal{L}(0, 1)$ , and Cauchy  $\mathcal{C}(0, 1)$ .

The quantile estimations of the regression parameters and of the change-points, for a fixed number  $K$  of change-points, are calculated using relation (21). The corresponding LS estimations are obtained by minimizing in  $\boldsymbol{\theta}_1$  and  $\boldsymbol{\theta}_2$  following sum (see Boldea and

Hall (2013):

$$\sum_{r=1}^K \sum_{i=l_{r-1}+1}^{l_r} [Y_i - g(\mathbf{X}_i, \phi_r)]^2.$$

#### 4.1 Known change-point number

First, in Tables 1, 2, 3, the change-point number is known and it is equal to two (model with three phases). In Tables 1 and 2 the number of observations is  $n=100$ , with the particular case of epidemic model in Tables 2, when model is the same in the first and the third phase ( $\phi_1^0 = \phi_3^0 = (0.5, 1)$ ).

Since the asymptotic distribution of the change-point quantile estimators can not be symmetric (see Theorem 3.2), the median of change-point estimations are given. Asymptotic distributions of regression parameter estimators by LS and quantile methods in a change-point nonlinear model are Gaussian (see Theorem 3.2 and corresponding result of Boldea and Hall (2013) for LS method). Then, the mean and standard-deviation(*sd*) of corresponding estimations are reported.

In all situations (see Tables 1, 2, 3), the median of the change-point estimations are very close to the true values. When the errors are Gaussian, very good results are obtained by the two estimation methods. For Laplace errors, the results deteriorate slightly, while for Cauchy errors, the quantile method gives very satisfactory results, while by LS method, the obtained estimates are biased and with a wide variation, when  $n = 100$  or when  $n$  is greater (Tables 1 and 3).

#### 4.2 Unknown change-point number

In view of these results, in order to study the selection criterion of the change-point number, we will consider only Normal and Cauchy distributions for errors. We simulate a model with one change-point in  $l_1^0 = 20$  for  $n = 100$  observations. The estimation  $\hat{K}_n^{(\tau)}$  of the change-point number associated to quantile method is calculated using Remark 3.3. For criterion (23), for the penalty we consider  $P(K, p) = Kp$  and deterministic sequence  $B_n = n^{5/8}$ . The estimation of the change-point number associated to LS method, is the

Table 1: Model with two change-points  $l_1^0 = 20$ ,  $l_2^0 = 85$ ,  $n = 100$ .  $\phi_1^0 = (0.5, 1)$ ,  $\phi_2^0 = (1, -0.5)$ ,  $\phi_3^0 = (2.5, 1)$ . Estimation by LS and quantile methods.

Estimation method	$\varepsilon$ law	median( $\hat{l}_1^{(\tau)}$ )	median( $\hat{l}_2^{(\tau)}$ )	mean( $\hat{\phi}_1^{(\tau)}$ ) <i>sd</i> ( $\hat{\phi}_1^{(\tau)}$ )	mean( $\hat{\phi}_2^{(\tau)}$ ) <i>sd</i> ( $\hat{\phi}_2^{(\tau)}$ )	mean( $\hat{\phi}_3^{(\tau)}$ ) <i>sd</i> ( $\hat{\phi}_3^{(\tau)}$ )
LS	$\varepsilon \sim \mathcal{N}$	19	84	(0.52, 1.06) <i>(0.17, 0.64)</i>	(0.98, -0.5) <i>(0.09, 0.02)</i>	(2.52, 1.07) <i>(0.15, 0.5)</i>
	$\varepsilon \sim \mathcal{L}$	19	84	(0.58, 1.28) <i>(0.42, 1.56)</i>	(0.98, -0.5) <i>(0.22, 0.05)</i>	(2.65, 1.13) <i>(0.46, 1.1)</i>
	$\varepsilon \sim \mathcal{C}$	22	85	(2.51, 1.26) <i>(18.7, 2.2)</i>	(2.34, -0.24) <i>(12.7, 0.96)</i>	(7.7, 1.75) <i>(42, 3.4)</i>
quantile	$\varepsilon \sim \mathcal{N}$	19	84	(0.52, 1.1) <i>(0.16, 0.78)</i>	(0.99, -0.5) <i>(0.09, 0.02)</i>	(2.53, 1.1) <i>(0.18, 0.8)</i>
	$\varepsilon \sim \mathcal{L}$	19	84	(0.57, 1.17) <i>(0.37, 1.28)</i>	(1, -0.5) <i>(0.13, 0.04)</i>	(2.6, 1.45) <i>(0.32, 3.2)</i>
	$\varepsilon \sim \mathcal{C}$	20	84	(0.58, 1.2) <i>(0.55, 1.1)</i>	(0.98, -0.48) <i>(0.29, 0.23)</i>	(2.7, 1.75) <i>(0.6, 3.1)</i>

minimizer in  $K$  of

$$n \log \left( n^{-1} \min_{(\theta_1, \theta_2)} \left\{ \sum_{r=1}^K \sum_{i=l_{r-1}+1}^{l_r} [Y_i - g(\mathbf{X}_i, \phi_r)]^2 \right\} \right) + P(K, p) B_n.$$

Two cases are considered for the true regression parameters: the parameters of the two (true) phases are far (Figure 1),  $\phi_1^0 = (0.5, 1)$ ,  $\phi_2^0 = (10, 2.5)$  and the parameters are closely (Figure 2),  $\phi_1^0 = (0.5, 1)$ ,  $\phi_2^0 = (1, -0.5)$ . In the case of Gaussian errors, the criterion associated to the LS method is slightly better when the parameters are far. The criteria associated to the two methods (LS and quantile) give the same good results if the parameters are closely. In the case of Cauchy errors, the quantile criterion selects well the change-point number when the parameters are far, while when the parameters are closely, the two criteria rather prefer a model without change-points (Tables 4).

### 4.3 Conclusion

These simulations allow us to conclude that for a nonlinear model with change-points, when the errors are Gaussian, the quantile method, proposed in this paper, gives similar

Table 2: Model with two change-points  $l_1^0 = 20$ ,  $l_2^0 = 85$ ,  $n = 100$ .  $\phi_1^0 = \phi_3^0 = (0.5, 1)$ ,  $\phi_2^0 = (1, -0.5)$ . Estimation by LS and quantile methods.

Estimation method	$\varepsilon$ law	median( $\hat{l}_1^{(\tau)}$ )	median( $\hat{l}_2^{(\tau)}$ )	mean( $\hat{\phi}_1^{(\tau)}$ ) $sd(\hat{\phi}_1^{(\tau)})$	mean( $\hat{\phi}_2^{(\tau)}$ ) $sd(\hat{\phi}_2^{(\tau)})$	mean( $\hat{\phi}_3^{(\tau)}$ ) $sd(\hat{\phi}_3^{(\tau)})$
LS	$\varepsilon \sim \mathcal{N}$	19	84	(0.5, 1.02) (0.12, 0.31)	(1, -0.5) (0.07, 0.02)	(0.5, 1.2) (0.14, 1.7)
	$\varepsilon \sim \mathcal{L}$	19	84	(0.58, 1.03) (0.35, 0.62)	(1.03, -0.51) (0.24, 0.05)	(0.5, 2.01) (0.39, 6.5)
	$\varepsilon \sim \mathcal{C}$	21	85	(2.28, 1.75) (11.8, 3)	(-0.07, -0.07) (17, 1.2)	(1.3, 1.8) (7.9, 6.4)
quantile	$\varepsilon \sim \mathcal{N}$	19	84	(0.52, 1.05) (0.15, 0.52)	(1, -0.5) (0.09, 0.02)	(0.5, 1.02) (0.17, 0.4)
	$\varepsilon \sim \mathcal{L}$	19	84	(0.56, 1.06) (0.3, 0.8)	(1.03, -0.5) (0.15, 0.04)	(0.51, 1.8) (0.3, 5.6)
	$\varepsilon \sim \mathcal{C}$	19	84	(0.64, 1.45) (0.56, 1.6)	(0.93, -0.49) (0.2, 0.14)	(0.7, 1.3) (0.56, 1.54)

results to those obtained by least squares method. On the other hand, for heavy-tailed errors, the performance of the quantile method is better than LS method, whether in estimation or in selection criterion.

## A Bernstein's Inequality

**Bernstein's Inequality** (see for example Pollard (1984)).

Let  $Z_i$  be a sequence of independent random variables with mean zero and  $|Z_i| \leq \beta$  for some  $\beta > 0$ . Let also  $V \geq \sum_{i=1}^n \mathbb{E}[Z_i^2]$ . Then for all  $0 < s < 1$  and  $0 \leq z \leq V/(s\beta)$ , we have

$$\mathbb{P} \left[ \left| \sum_{i=1}^n Z_i \right| > z \right] \leq 2 \exp(-z^2 s(1-s)/V). \quad (24)$$

## References

Bai, J., Estimation of multiple-regime regressions with least absolute deviation. *Journal of Statistical Planning Inference*, **74**, 103-134, (1998).

Table 3: Model with two change-points  $l_1^0 = 100$ ,  $l_2^0 = 200$ ,  $n = 300$ .  $\phi_1^0 = (0.5, 1)$ ,  $\phi_2^0 = (1, -0.5)$ ,  $\phi_3^0 = (2.5, 1)$ . Estimation by LS and quantile methods.

Estimation method	$\varepsilon$ law	median( $\hat{l}_1^{(\tau)}$ )	median( $\hat{l}_2^{(\tau)}$ )	mean( $\hat{\phi}_1^{(\tau)}$ ) $sd(\hat{\phi}_1^{(\tau)})$	mean( $\hat{\phi}_2^{(\tau)}$ ) $sd(\hat{\phi}_2^{(\tau)})$	mean( $\hat{\phi}_3^{(\tau)}$ ) $sd(\hat{\phi}_3^{(\tau)})$
LS	$\varepsilon \sim \mathcal{N}$	99	199	(0.5, 1) (0.05, 0.07)	(1, -0.5) (0.06, 0.01)	(2.5, 1) (0.05, 0.08)
	$\varepsilon \sim \mathcal{L}$	99	199	(0.48, 1.02) (0.16, 0.19)	(1, -0.5) (0.17, 0.05)	(2.5, 1) (0.13, 0.16)
	$\varepsilon \sim \mathcal{C}$	100	200	(3.05, 1.02) (16.2, 1.17)	(1.99, -0.32) (10.6, 0.77)	(3.2, 1.1) (4.9, 1.06)
quantile	$\varepsilon \sim \mathcal{N}$	99	199	(0.5, 1) (0.05, 0.08)	(1, -0.5) (0.07, 0.02)	(2.5, 1) (0.05, 0.08)
	$\varepsilon \sim \mathcal{L}$	99	199	(0.51, 1) (0.09, 0.16)	(0.99, -0.5) (0.11, 0.03)	(2.5, 1.01) (0.1, 0.1)
	$\varepsilon \sim \mathcal{C}$	99	199	(0.54, 1.06) (0.14, 0.75)	(0.98, -0.5) (0.16, 0.04)	(2.5, 0.98) (0.1, 0.4)

Bai, J., Perron P., Estimating and testing linear models with multiple structural changes, *Econometrica* **66**(1), 47-78, (1998).

Boldea, O., Hall, A.R., Estimation and inference in unstable nonlinear least squares models. *Journal of Econometrics*, **172**(1), 158-167, (2013).

Chen L.A., Tran L.T., Lin L.C., Symmetric regression quantile and its application to robust estimation for the nonlinear regression model, *Journal of Statistical Planning and Inference*, **126**(2), 423-440, (2004).

Choi S.H., Kim K.J., Lee M.S., Robust test based on nonlinear regression quantile estimators, *Communications of the Korean Mathematical Society*, **20**(1), 145-159, (2005).

Ciuperca G., The M-estimation in a multi-phase random nonlinear model. *Statistics and Probability Letters*, **75**(5), 573-580, (2009).

Table 4: Results on the choice of the change-point number by criteria associated to LS and quantile methods. The true change-point number is 1 in  $l_1^0 = 20$  for  $n = 100$  observations. 100 Monte Carlo replications.

True parameters	LS method						Quantile method					
	$\varepsilon \sim \mathcal{N}$			$\varepsilon \sim \mathcal{C}$			$\varepsilon \sim \mathcal{N}$			$\varepsilon \sim \mathcal{C}$		
	$\hat{K}_n^{(\tau)} = 0$	$= 1$	$= 2$	$= 0$	$= 1$	$= 2$	$= 0$	$= 1$	$= 2$	$= 0$	$= 1$	$= 2$
$\phi_1^0 = (0.5, 1), \phi_2^0 = (10, 2.5)$	0	100	0	56	43	1	2	92	6	13	86	1
$\phi_1^0 = (0.5, 1), \phi_2^0 = (1, -0.5)$	4	96	0	91	9	0	5	95	0	92	8	0

Ciuperca, G., A general criterion to determinate the number of change-points. *Statistics and Probability Letters*, **81**, no 8, 1267-1275, (2011).

Ciuperca G., Estimating nonlinear regression with and without change-points by the LAD-method. *Annals of the Institute of Statistical Mathematics*, **63**(4), 717-743, (2011).

Ciuperca G., Adaptive model selection in a high-dimension multiphase quantile regression. *arXiv:1309:1262*, (2014).

Furno M., Tests for structural break in quantile regressions. *AStA Advances in Statistical Analysis*, **96**(4), 493-515, (2012).

Knight K., Limiting distributions for  $L_1$  regression estimators under general conditions, *Annals of Statistics*, **26**(2), 755-770, (1998).

Koenker R., Quantiles regression, Econometric Society Monographs, No **38**, Cambridge University Press, (2005).

Koenker R., Bassett G., Regression Quantiles, *Econometrica*, **46**, 33-50, (1978).

Koul, H.L., Qian, L., Asymptotics of maximum likelihood estimator in a two-phase linear regression model. *Journal of Statistical Planning and Inference* **108**, 99-119, (2002).

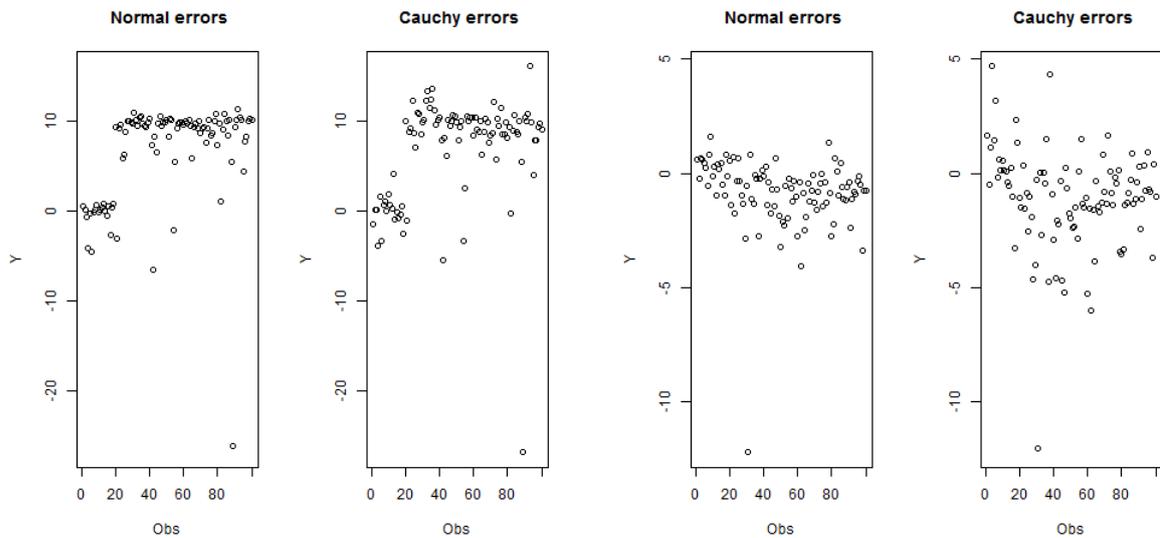


Figure 1: *Model with normal and Cauchy errors, with one change-point in  $l^0 = 20$ , far parameters for the two phases.*

Figure 2: *Model with normal and Cauchy errors, with one change-point in  $l^0 = 20$ , close parameters for the two phases.*

Koul, H.L., Qian, L., Surgailis, D., Asymptotics of M-estimators in two-phase linear regression models. *Stochastic Processes and their Applications* **103**, 123-154, (2003).

Oberhofer W., The consistency of nonlinear regression minimizing the L1-norm. *Annals of Statistics*, **10**(1), 316-319, (1982).

Oberhofer W., Haupt H., Asymptotic theory for nonlinear quantile regression under weak dependence. Working Paper. <https://www.researchgate.net/publication/29858831-Asymptotic-theory-for-nonlinear-quantile-regression-under-weak-dependence>, (2014).

Oka T., Qu Z., Estimating structural changes in regression quantiles. *Journal of Econometrics*, **162**, 248-267, (2011).

Pollard D., *Convergence of stochastic processes*, Springer, New York, (1984).

Qu Z., Testing for structural change in regression quantiles. *Journal of Econometrics*, **146**, 170-184, (2008).

Seber, G.A.F., Wild, C.J., *Nonlinear regression*. Wiley Series in Probability and Statistics, New Jersey, (2003).

van der Vaart A.W., Wellner J.A.,. *Weak convergence and empirical processes. With applications to statistics*, Springer Series in Statistics, New York, (1996).