



**HAL**  
open science

## Estimate of quantile-oriented sensitivity indices

Thomas Browne, Jean-Claude Fort, Bertrand Iooss, Loïc Le Gratiet

► **To cite this version:**

Thomas Browne, Jean-Claude Fort, Bertrand Iooss, Loïc Le Gratiet. Estimate of quantile-oriented sensitivity indices. 2017. hal-01450891

**HAL Id: hal-01450891**

**<https://hal.science/hal-01450891v1>**

Preprint submitted on 31 Jan 2017

**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.

# Estimate of quantile-oriented sensitivity indices

Thomas BROWNE<sup>a,b</sup>, Jean-Claude FORT<sup>b,c</sup>, Bertrand IOOSS<sup>a,c</sup>, Loïc LE GRATIET<sup>a</sup>

<sup>a</sup>EDF Lab Chatou, Département Management des Risques Industriels, France; <sup>b</sup>Université Paris-Descartes, Laboratoire MAP5, France; <sup>c</sup>Institut de mathématiques de Toulouse, France

## ARTICLE HISTORY

Compiled January 31, 2017

## ABSTRACT

In the context of black-box numerical codes, it is relevant to use sensitivity analysis in order to assess the influence of each random input  $X$  over the output  $Y$ . Goal-oriented sensitivity analysis states that one must first focus on a certain probability feature  $\theta(Y)$  from the distribution of  $Y$  (such as its mean, quantile, or a probability of failure etc...), which would be chosen regarding a relevant strategy. The wish is to evaluate the impact of each input over  $\theta(Y)$ . In order to get supplementary information about sensitivity, we set that  $\theta(Y)$  is the  $\alpha$ -level quantile of  $Y$ , where  $\alpha \in ]0, 1[$ . Throughout some examples, it has been pointed out that in some cases quantile-oriented sensitivity indices can detect some influence that Sobol indices would not. Mainly, the influence over each level of quantile displays how an input distribution entirely propagates through the output. We establish further results for the quantile-oriented indices properties in order to justify their relevancy. The main contribution of this paper comes when a statistical estimator for this index is introduced.

## KEYWORDS

Sensitivity analysis; goal-oriented sensitivity analysis; output quantiles; kernel-based estimators

## 1. Introduction

In computer experiments, sensitivity analysis (SA) aims to quantify the influence of each random input over the studied output. In many industrial contexts such as safety in nuclear industry [8], geophysics and oil reservoir, soil pollution (see [10] of interest. Besides, if the study points out that several inputs do not affect  $Y$ , the user can consider neglecting the corresponding input distributions. This is particularly advantageous when the number of inputs is high and/or each computation of the code is heavy. On the other hand, if one input happens to be highly influential, the user could wonder how to reduce its variability [3]. In this case, a significant decrease of the uncertainty on  $Y$  would result. for agricultural examples), complex black-box models are highly used. The model that is considered writes  $Y = g(X, Z)$ , where  $g$  is a deterministic function,  $Y$  is the output and  $X$  the random input whose influence over  $Y$  must be quantified.  $Z$  regroups all the other inputs, parameters of the function. Carrying out a SA study provides useful information about the model. First, it enables the user to understand how each input impacts the output. Many definitions for the influence can be given. For instance, the goal of global sensitivity analysis (GSA) is to assess how much the variability of an input propagates through the output distribution. This is the case for the works in [11, 14]. More recently, new indices have been set in [1, 2], respectively based on importance measures and independence measures. In this paper, we focus on a more precise interpretation of the influence: the goal-oriented sensitivity analysis (GOSA) introduced

in [12]. Rather than focusing on the impact of  $X$  over  $Y$ 's whole distribution, we only consider a one-dimension feature from the distribution of  $Y$ . GOSA states: if the user sets one feature of interest,  $\theta(Y)$ , from the distribution of  $Y$ , which can be for instance its expectation, any level of quantile or probability of exceedance, it is relevant to quantify how an input distribution affects the value of  $\theta(Y)$ .

In this study, we understand the influence of the input,  $X$ , over  $\theta(Y)$  as the amount of variability of the conditional feature,  $\theta(Y | X)$ . Indeed, we have: if, when we set  $X$  to several different values,  $\theta(Y | X)$  varies a lot, then  $X$  is highly responsible for the value of  $\theta(Y)$ . In this case, we state that  $X$  is influential over  $\theta(Y)$ . SA indices with respect to a contrast, introduced in [6], precisely quantify the variability of  $\theta(Y | X)$  regarding a relevant distance. In this paper, we focus on the quantile-oriented versions of these indices. The main contribution of this paper is to provide an estimator for these indices.

The structure of the paper is as follows. In Section 2, we recall the definition of SA indices with respect to a contrast for the only case of the quantile-oriented SA. We also prove some of their properties that confirms their relevancy. In Section 3, we introduce an estimator for these indices by justifying their construction step by step. To this effect, we enumerate and prove consistency results for some kernel-based estimators on which we rely. In Section 4, we raise the question of the choice for the bandwidth. By setting different quantities related to the error of the estimator, we propose a method to set an efficient bandwidth in practice. This method is approved through numerical applications to a toy example. Finally, the estimator is computed in Section 5 for an industrial case. The considered data is taken from runs of a numerical code, CIVA, which simulates ultrasonic non-destructive examinations over a split in the inner wall of a pipe. We provide several interpretations of the values that we computed, regarding the respective influence of the inputs over the output quantiles. A conclusion synthesizes the work.

## 2. Quantile-Oriented Sensitivity Indices

We study the scalar output  $Y = g(X, Z)$ , with a real random variable input of interest,  $X$ ,  $Z$  which denotes the other inputs and  $g$  a deterministic function. Given a level of quantile  $\alpha \in ]0, 1[$ , let us recall the quantile-oriented sensitivity indices introduced in [6]:

$$S_{c_\alpha}^X(Y) = \min_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta)] - \mathbb{E}_X \left[ \min_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X] \right], \quad (1)$$

with  $c_\alpha$  the quantile-oriented simple contrast:

$$\forall y, \theta \in \mathbb{R} \quad c_\alpha(y, \theta) = (y - \theta)(\mathbf{1}_{y \leq \theta} - \alpha).$$

Indeed, let us remember that quantiles and conditional quantiles can be defined with the mean contrast:

$$q^\alpha(Y) = \arg \min_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta)] \quad \text{and} \quad q^\alpha(Y | X) = \arg \min_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X], \quad (2)$$

where  $q^\alpha(Y)$  is the  $\alpha$ -level of quantile of  $Y$ , and  $q^\alpha(Y | X)$  is the  $\alpha$ -level of the conditional quantile. In order to ensure the uniqueness for  $q^\alpha(Y)$  and  $q^\alpha(Y | X)$ , let us set that the density of  $Y$  is strictly positive on its support, as well as its conditional density with respect to  $X$  a.s. The purpose of this index (1) is precisely to quantify the variability of the conditional quantile. This is suggested in the following properties.

This index is well defined as long as  $Y \in L^1(\Omega)$  in the sense that, as for the first term,

$$\mathbb{E}[c_\alpha(Y, q^\alpha(Y))] \leq \mathbb{E}[|Y - q^\alpha(Y)|] \leq \|Y\|_{L^1(\Omega)} + |q^\alpha(Y)|, \quad (3)$$

and we easily prove that  $|q^\alpha(Y)| < +\infty$  as

$$\|Y\|_{L^1(\Omega)} \geq \mathbb{E}\left[|Y| \cdot \mathbf{1}_{q^\alpha(Y) \leq Y}\right] \geq (1 - \alpha) q^\alpha(Y).$$

Thus one can conclude that  $\mathbb{E}[c_\alpha(Y, q^\alpha(Y))] < +\infty$ . As for the second term, it is clear that

$$\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X] \leq \mathbb{E}[c_\alpha(Y, q^\alpha(Y)) | X] \quad a.s. \quad (4)$$

We get after integrating:  $\mathbb{E}[c_\alpha(Y, q^\alpha(Y | X))] \leq \mathbb{E}[c_\alpha(Y, q^\alpha(Y))]$ .

The previous inequality also yields the positiveness of the index. In practice, we normalize it as we divide it by the first term  $\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta)]$ . From now on, we only consider the normalized expression of the index:

$$S_{c_\alpha}^X(Y) = \frac{\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta)] - \mathbb{E}_X \left[ \min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X] \right]}{\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta)]} \quad (5)$$

It is more meaningful for the user as one can prove the two following properties.

**Property 1.** *The sensitivity index  $S_{c_\alpha}^X(Y)$  verifies:*

$$0 \leq S_{c_\alpha}^X(Y) \leq 1.$$

**Property 2.** *The sensitivity index  $S_{c_\alpha}^X(Y)$  verifies:*

- $S_{c_\alpha}^X(Y) = 0$  if and only if  $q^\alpha(Y | X) = q^\alpha(Y)$  a.s.
- $S_{c_\alpha}^X(Y) = 1$  if and only if there exists a real function  $h$  such that  $Y = h(X)$  a.s.

The first statement of Property 2 justifies the expected meaning of the index regarding the sensitivity analysis. Indeed,  $q^\alpha(Y | X) = q^\alpha(Y)$  a.s. reveals that setting  $X$  to any of its possible realizations does not influence the value of the output  $\alpha$ -quantile. The second statement is more global as it states that, as soon as one sets  $X$  to any of its possible realizations, the output  $Y$  becomes a constant. This means that in the case where  $S_{c_\alpha}^X(Y) = 1$ , the whole variability of  $Y$  is induced by the variability of  $X$ .

**Proof.** • One has:  $S_{c_\alpha}^X(Y) = 0$  if and only if  $\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta)] = \mathbb{E} \left[ \min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X] \right]$ .

With the equation (4), the previous equivalence brings:

$$\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X] = \mathbb{E}[c_\alpha(Y, q^\alpha(Y)) | X] \quad \text{almost surely (a.s.).}$$

The uniqueness of the minimum for  $\mathbb{E}[c_\alpha(Y, \theta) | X]$  proves that  $S_{c_\alpha}^X(Y) = 0$  if and only if  $q^\alpha(Y | X) = q^\alpha(Y)$  a.s.

- One has:  $S_{c_\alpha}^X(Y) = 1$  if and only if  $\mathbb{E}[c_\alpha(Y, q^\alpha(Y | X))] = 0$ , which, given the positiveness of the function  $c_\alpha$ , writes  $c_\alpha(Y, q^\alpha(Y | X)) = 0$  a.s. Therefore  $S_{c_\alpha}^X(Y) = 1$  is equivalent to  $Y = q^\alpha(Y | X)$  a.s. We have that  $Y$  is a constant as soon as  $X$  is set:  $X$  contains all the variability observed in the output.

□

As for the exact interpretation of the index, one can rewrite  $S_{c_\alpha}^X(Y)$  as:

$$S_{c_\alpha}^X(Y) = 1 - \frac{\mathbb{E}[c_\alpha(Y, q^\alpha(Y | X))]}{\mathbb{E}[c_\alpha(Y, q^\alpha(Y))]}.$$

In other words,  $S_{c_\alpha}^X(Y)$  compares the mean distance between  $Y$  and its conditional quantile to the mean distance between  $Y$  and its quantile, where the considered distance is the contrast,  $c_\alpha$ . For instance, if one computes  $S_{c_\alpha}^X(Y) = 0.8$ , then  $\mathbb{E}[c_\alpha(Y, q^\alpha(Y | X))]$  is 5 times smaller than  $\mathbb{E}[c_\alpha(Y, q^\alpha(Y))]$ . Since  $q^\alpha(Y | X)$  is statistically “closer” to  $Y$ , it must vary significantly. Therefore, the conditional quantile,  $q^\alpha(Y | X)$ , has a significant variability when  $S_{c_\alpha}^X(Y)$  is close to one. On the other hand,  $S_{c_\alpha}^X(Y)$  being nearly zero would induce that  $q^\alpha(Y | X)$  is statistically as distant (regarding  $c_\alpha$ ) to  $Y$  as  $q^\alpha(Y)$ :  $q^\alpha(Y | X)$  has not much variability and is statistically close to  $q^\alpha(Y)$ .

### 3. Estimate of the Quantile-Oriented Index

Let us recall that we focus on the quantile-oriented *SA* index,  $S_{c_\alpha}^X(Y)$  (introduced in [6]), such that:

$$S_{c_\alpha}^X(Y) = \frac{\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta)] - \mathbb{E}_X \left[ \min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X] \right]}{\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta)]}$$

Given a budget of computations  $n \in \mathbb{N}$ , our goal is to provide an estimator for  $S_{c_\alpha}^X(Y)$  from a iid  $n$ -sample  $((X^1, Z^1, Y^1), \dots, (X^n, Z^n, Y^n))$  such that:

$$\forall j \in \{1, \dots, n\} \quad Y^j = g(X^j, Z^j).$$

We propose respective estimators for the two terms of the index. Let us focus on the first term as it is more natural to estimate:

$$\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta)].$$

We propose a classical empirical estimator:

$$\min_{\theta \in \mathbb{R}} \frac{1}{n} \sum_{j=1}^n c_\alpha(Y^j, \theta).$$

As, for any  $j \in \{1, \dots, n\}$ ,  $\left(\theta \mapsto \frac{1}{n} \sum_{j=1}^n c_\alpha(Y^j, \theta)\right)$  is a convex and continuous piecewise linear function, its minimum is reached on the subset  $\{Y^1, \dots, Y^n\}$ . Then finding the minimum only requires to evaluate this function on these  $n$  points and pick the lowest value. It is important to mention that this minimization is fast to proceed as no run of the code is performed. Let us write  $\hat{q}^\alpha(Y)$ , the classical empirical estimator for  $q^\alpha(Y)$ , defined as follows:

$$\hat{q}_n^\alpha(Y) := Y^{(i_0)} \quad \text{with} \quad i_0 := \lfloor n \cdot \alpha \rfloor + 1, \tag{6}$$

where  $\{Y^{(1)} \leq \dots \leq Y^{(n)}\}$  is the order statistic of  $\{Y^1, \dots, Y^n\}$ . The expression (6) holds:

$$\hat{q}_n^\alpha(Y) = \arg \min_{\theta \in \mathbb{R}} \frac{1}{n} \sum_{j=1}^n c_\alpha(Y^j, \theta),$$

Proposition 1 provides the consistency of the index first term estimate.

**Proposition 1.** *If  $Y \in L^1(\Omega)$ , then*

$$\min_{\theta \in \mathbb{R}} \frac{1}{n} \sum_{j=1}^n c_\alpha(Y^j, \theta) \xrightarrow[n \rightarrow +\infty]{a.s.} \mathbb{E}[c_\alpha(Y, q^\alpha(Y))].$$

**Proof.** Let us first show that:

$$\frac{1}{n} \sum_{j=1}^n c_\alpha(Y^j, \hat{q}_n^\alpha(Y)) - \frac{1}{n} \sum_{j=1}^n c_\alpha(Y^j, q^\alpha(Y)) \xrightarrow[n \rightarrow +\infty]{a.s.} 0 \quad a.s., \quad (7)$$

with  $\hat{q}_n^\alpha(Y)$  introduced in (6). One easily gets:

$$\forall y, \theta, \theta' \in \mathbb{R} \quad |c_\alpha(y, \theta) - c_\alpha(y, \theta')| \leq \max(\alpha, 1 - \alpha)|\theta - \theta'|$$

Therefore:

$$\left| \frac{1}{n} \sum_{j=1}^n [c_\alpha(Y^j, \hat{q}_n^\alpha(Y)) - c_\alpha(Y^j, q^\alpha(Y))] \right| \leq \max(\alpha, 1 - \alpha) |\hat{q}_n^\alpha(Y) - q^\alpha(Y)| \quad a.s..$$

A classical result states that  $\hat{q}_n^\alpha(Y) \xrightarrow[n \rightarrow +\infty]{a.s.} q^\alpha(Y)$  (see [15] for further details). This proves (7). As for the second step, since  $(c_\alpha(Y^1, q^\alpha(Y)), \dots, c_\alpha(Y^n, q^\alpha(Y)))$  is a *iid*  $n$ -sample and  $\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta)] \in L^1(\Omega)$  (see (3)), the law of large numbers yields:

$$\frac{1}{n} \sum_{j=1}^n c_\alpha(Y^j, q^\alpha(Y)) \xrightarrow[n \rightarrow +\infty]{a.s.} \mathbb{E}[c_\alpha(Y, q^\alpha(Y))]. \quad (8)$$

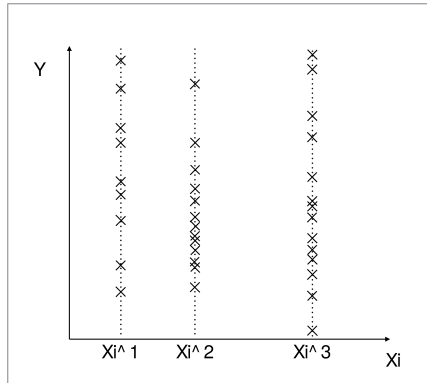
(7) and (8) together lead to the result of the lemma.  $\square$

We now focus on the second term:

$$\mathbb{E} \left[ \min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X] \right], \quad (9)$$

which requires much more development to estimate as it contains a double expectation, including a conditional expectation, and a minimization problem. In the following we show the successive steps on which we relied to build an estimator for this term. First of all, approximating  $\mathbb{E} \left[ \min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X] \right]$  suggests to be able to estimate  $\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X = x]$ , for  $x$  any “likely” realization of  $X$ . We write  $f$  the *pdf* of  $X$  and for any  $x \in \mathbb{R}$ , we call a “likely” (or possible) realization of  $X$  any real number such that  $f(x) \neq 0$ . One could get back to the needed estimator by calculating the average of the estimators of  $\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X = x]$  with a Monte-Carlo approach. In order to estimate  $\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X = x]$ , one needs to focus on the conditional expectation.

One harmful way to proceed would be to sample  $X$ ,  $(X^1, \dots, X^n)$  and successively set  $X$  to  $X^k$ , for  $k \in \{1, \dots, n\}$ , and compute in each case  $(Y | X = X^k)$  a number of times  $n' \in \mathbb{N}$  and get the  $n'$ -sample  $((Y | X = X^k)^1, \dots, (Y | X = X^k)^{n'})$ . Even though it could be seen like a classical approach, both  $n$  and  $n'$  would need to be high in this case. As  $n \times n'$  calls to  $g$  would be needed to estimate  $\mathbb{E} \left[ \min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X] \right]$ , it may be too high regarding a reasonable time budget. We illustrate in Figure 1 this not desirable method. One additional problem relies on the fact



**Figure 1.** High-time consuming approach for the quantile-oriented estimate of  $S_{c_\alpha}^X(Y)$ .

that for any  $k \in \{1, \dots, n\}$ , the  $n'$ -sample  $((Y | X = X^k)^1, \dots, (Y | X = X^k)^{n'})$  has relevant information to estimate only  $\mathbb{E}[c_\alpha(Y, \theta) | X = X^k]$ . A usual method to overcome this double-loop issue is to use a kernel-based estimator. Through the kernel-approach, we justify step by step the construction of the future estimator for the second term.

Let  $K$  be a positive and 2-order kernel, i.e.:

$$\int uK(u)du = 0 \quad \text{and} \quad 0 \neq \int u^2K(u)du < +\infty.$$

In the following, we assume the kernel  $K$  to verify the condition  $(K)$ :

$$(K) \equiv \begin{cases} K \text{ has a compact support } \Delta_K \subset \mathbb{R} \\ K \in L^2(\Delta_K) \\ \forall u \in \Delta_K \quad K(-u) = K(u) \geq 0 \\ K \text{ is a second-order kernel.} \end{cases}$$

We introduce the following estimator for the function  $(C_x : \theta \mapsto \mathbb{E}[c_\alpha(Y, \theta) | X = x])$ , for any likely  $x \in \mathbb{R}$ :

$$\forall \theta \in \mathbb{R} \quad \hat{C}_x(\theta) := \frac{1}{nf(x)} \sum_{j=1}^n c_\alpha(Y^j, \theta) K_{h_n}(X^j - x), \quad (10)$$

where  $K_{h_n}$  is set as follows:

$$K_{h_n}(x) := \frac{1}{h_n} K\left(\frac{x}{h_n}\right).$$

$(h_n)$  is the bandwidth sequence:  $\forall n \in \mathbb{N} \quad h_n > 0$ . We provide the sketch of proof for the pointwise consistency of the estimator (10) below. To this effect, we set the condition  $(J)$  for the joint density

$((x, y) \mapsto f(x, y))$  of  $(X, Y)$ :

$$(J) \equiv \begin{cases} f \text{ is } C^2 \text{ on } \mathbb{R}^2, \\ \forall k = 1, 2 \int y^2 |\partial_x^k f(x, y)| dx dy < \infty, \\ \exists (y \mapsto C(y)) \in C^0, \exists \delta > 0 \text{ s.t. } C(y) \underset{\pm\infty}{\approx} y^{-3-\delta} \\ \text{and for any } y \in \mathbb{R}, (x \mapsto \partial_x^\nu f(x, y)) \text{ is } C(y)\text{-Lipschitz.} \end{cases}$$

**Proposition 2.** *If the joint density  $((x, y) \mapsto f(x, y))$  of  $(X, Y)$  verifies (J),  $K$  verifies (K) and  $x \in \mathbb{R}$  s.t.  $f(x) \neq 0$ , then we have:*

$$\forall \theta \in \mathbb{R} \quad \frac{1}{nf(x)} \left[ \sum_{j=1}^n c_\alpha(Y^j, \theta) K_{h_n}(x - X^j) \right] \xrightarrow[n \rightarrow +\infty]{L^2} \mathbb{E}[c_\alpha(Y, \theta) | X = x],$$

with  $K$  a 2-order kernel,  $(h_n)_{n \in \mathbb{N}}$  the bandwidth sequence such that  $h_n \xrightarrow[n \rightarrow +\infty]{} 0$  and  $n \times h_n \xrightarrow[n \rightarrow +\infty]{} +\infty$ .

As the previous proposition states the pointwise consistency of the estimation for the function  $C_x$ , for  $x$  s.t.  $f(x) \neq 0$ , we point out that the second term of the SA index (9) also writes  $\mathbb{E}_X \left[ \min_{\theta \in \mathbb{R}} C_X(\theta) \right]$ . As we aim to estimate the minimum, for each  $x$ , of the function  $C_x$ , we rely on existing consistency results from quantile regression. In this case, the minimizer, and not the minimum, is efficiently estimated from the minimizer of  $C_x$  [5].

**Proposition 3.** *If the joint density  $((x, y) \mapsto f(x, y))$  of  $(X, Y)$  verifies (J),  $K$  verifies (K) and  $x \in \mathbb{R}$  s.t.  $f(x) \neq 0$ , then we have:*

$$\arg \min_{\theta} \sum_{j=1}^n c_\alpha(Y^j, \theta) K_{h_n}(X^j - x) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \arg \min_{\theta} \mathbb{E}[c_\alpha(Y, \theta) | X = x].$$

One can see that the last estimation is based on the minimization of a certain function. It is not a real issue as for any  $j \in \{1, \dots, n\}$ ,  $\theta \mapsto \sum_{j=1}^n c_\alpha(Y^j, \theta) K_{h_n}(x - X^j)$  is a convex piecewise linear function. Therefore the minimum of such a function is finite, unique and is reached on one element of  $\{Y^1, \dots, Y^n\}$ . At this stage, solving this minimization problem only consists in evaluating the low time-consuming function  $(\theta \mapsto \sum_{j=1}^n c_\alpha(Y^j, \theta) K_{h_n}(X^j - x))$  for  $\theta = Y^1, \dots, Y^n$  and picking the lowest value. Proposition 3 states that the minimizer of  $\hat{C}_x$  is a consistent estimator of the minimizer of  $C_x$ . In order to bring relevant information to our case, we prove a similar consistency result regarding the minimum of  $C_x$ .

**Proposition 4.** *If the joint density  $((x, y) \mapsto f(x, y))$  of  $(X, Y)$  verifies (J),  $K$  verifies (K) and  $x \in \mathbb{R}$  s.t.  $f(x) \neq 0$ , then we have:*

$$\min_{\theta} \frac{1}{nf(x)} \sum_{j=1}^n c_\alpha(Y^j, \theta) K_{h_n}(X^j - x) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X = x].$$

As the second term of the estimator (9) is the expectation of the random variable  $\min_{\theta \in \mathbb{R}} C_X(\theta)$ , we set its estimator as the empirical mean of the estimate  $\min_{\theta \in \mathbb{R}} \hat{C}_X(\theta)$ . In order to do



so, we generate an iid  $m$ -sample of  $X$ ,  $(X^{l_1}, \dots, X^{l_m})$ , with  $m \in \mathbb{N}$ , independent from  $((X^1, Z^1, Y^1), \dots, (X^n, Z^n, Y^n))$ . We set the estimator for the second term as follows:

$$\frac{1}{m} \sum_{k=1}^m \min_{\theta} \frac{1}{nf(X^{l_k})} \sum_{j=1}^n c_{\alpha}(Y^j, \theta) K_{h_n}(X^j - X^{l_k}), \quad (11)$$

which also writes  $\frac{1}{m} \sum_{k=1}^m \min_{\theta \in \mathbb{R}} \hat{C}_{X^{l_k}}(\theta)$ , or even:

$$\frac{1}{m} \sum_{k=1}^m \frac{1}{nf(X^{l_k})} \sum_{j=1}^n c_{\alpha}(Y^j, \hat{q}_n^{\alpha}(Y | X^{l_k})) K_{h_n}(X^j - X^{l_k}).$$

It is important to point out that generating  $(X^{l_1}, \dots, X^{l_m})$  is completely affordable as we know its distribution  $f$  and it does not require any run of the code. At the end, we set the expression for the estimator of the index,  $S_{c_{\alpha}}^X(Y)$ :

$$\begin{aligned} \hat{S}_{c_{\alpha}}^X(Y) := & \min_{\theta \in \mathbb{R}} \frac{1}{n} \sum_{j=1}^n c_{\alpha}(Y^j, \theta) \\ & + \frac{1}{m} \sum_{k=1}^m \min_{\theta} \frac{1}{nf(X^{l_k})} \sum_{j=1}^n c_{\alpha}(Y^j, \theta) K_{h_n}(X^j - X^{l_k}). \end{aligned} \quad (12)$$

So far, no consistency result has been proved for  $\hat{S}_{c_{\alpha}}^X(Y)$ . In this paper, we focus on its efficiency in practice as we try it for two different application cases.

**Remark.** In Proposition 4, the minimizer of the asymptotic quantity,  $\min_{\theta \in \mathbb{R}} \mathbb{E}[c_{\alpha}(Y, \theta) | X = x]$ , is the conditional quantile  $q^{\alpha}(Y | X = x)$ . To this effect, we introduce the estimator for the conditional quantile,  $\hat{q}_n^{\alpha}(Y | X = x)$ , as follows:

$$\hat{q}_n^{\alpha}(Y | X = x) := \arg \min_{\theta \in \mathbb{R}} \sum_{j=1}^n c_{\alpha}(Y^j, \theta) K_{h_n}(X^j - x).$$

## 4. Optimal Bandwidth Determination

For any Kernel-based estimator, arguments have to be provided in order to choose a relevant bandwidth. Proposition 4 only suggests the following asymptotic property:  $n \cdot h_n \rightarrow +\infty$  while  $h_n \rightarrow 0$ . Yet in practice  $n$  is set and the user must choose  $h(n)$  as the minimizer of a quantity which denotes the error of  $\hat{S}_{c_{\alpha}}^X(Y)$ . This section focuses on mean-squared errors (MSE). The goal is to find a MSE for an estimator that is possible to express, theoretically, as well as to estimate.

### 4.1. Theoretical Optimal Bandwidth

As  $\hat{S}_{c_{\alpha}}^X(Y)$  is a kernel-based estimator, its efficiency strongly relies on the determination of a relevant bandwidth sequence  $(\tilde{h}_n)$ . A classical way to proceed is to set  $\tilde{h}_n$  as the bandwidth which minimizes the MSE of the estimator. However no expressions for the MSE of  $\hat{S}_{c_{\alpha}}^X(Y)$  are available. An other idea is to minimize the error of each term for the second term in the estimator

(9) e.g. for any  $x$  s.t.  $f(x) \neq 0$ :

$$\frac{1}{nf(x)} \sum_{j=1}^n c_\alpha(Y, \hat{q}_n^\alpha(Y | x)) K_{h^{(k)}}(X^j - x)$$

which is the estimator for (10)

$$\min_{\theta \in \mathbb{R}} \hat{C}_x(\theta) = \min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X = x].$$

Roughly speaking, if one can minimize the error of all the terms, the error of the mean of these very terms must be low. The MSE of the point-wise estimators  $\min_{\theta \in \mathbb{R}} \hat{C}_x(\theta)$  is given by:

$$\begin{aligned} \text{MSE} \left( \min_{\theta \in \mathbb{R}} \hat{C}_x(\theta) \right) &= \mathbb{E} \left[ \left( \min_{\theta \in \mathbb{R}} C_x(\theta) - \min_{\theta \in \mathbb{R}} \hat{C}_x(\theta) \right)^2 \right] \\ &= \frac{h^2}{f(x)} \mu_2(K) \int c_\alpha(y, q^\alpha(Y | x)) \partial_x f(x, y) dy + o(h^2) \\ &\quad + \frac{\text{Var}(c_\alpha(y, q^\alpha(Y | x))) R(K)}{f(x)nh} + o(nh), \end{aligned}$$

where  $R(K)$  and  $\mu_2(K)$  are inherent constants of the kernel,  $K$ :

$$\mu_2(K) = \int u^2 K(u) du, \text{ and } R(K) = \int K^2. \quad (13)$$

Since  $\text{MSE} \left( \min_{\theta \in \mathbb{R}} \hat{C}_x(\theta) \right)$  relies on unknown parameters, its minimum in  $h$  cannot be determined. Still regarding the same idea, we decide to minimize the error over the minimizer, and not the minimum, i.e.  $\arg \min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X^k] = q^\alpha(Y | x)$ . We assume that if the error over the minimizer is low, so is the error over the minimum. This is partly due to the fact that we observe that the convex function  $(\theta \mapsto \mathbb{E}[c_\alpha(Y, \theta) | x])$  is flat around its minimum. Therefore a small variation around its minimizer implies a small variation around its minimum. The asymptotic expression for the MSE of  $\hat{q}_k^\alpha(Y | x)$  is given in [5], with  $x \in \Delta_x$ :

$$\mathbb{E} \left[ (q^\alpha(Y | x) - \hat{q}_n^\alpha(Y | x))^2 \right] = \beta(x)^2 h_n^4 + \frac{v^2(x)}{nh_n} + o(h_n^2) + o\left(\frac{1}{nh_n}\right),$$

with:

$$\beta(x) = \frac{\mu_2(K)}{2} \frac{d^2 q^\alpha(Y | x)}{dx^2}, \quad v^2(x) = \frac{R(K)\alpha(1-\alpha)}{f(x)f(q^\alpha(Y | x) | x)},$$

where  $f(\cdot | x)$  is the pdf of the conditional distribution  $(Y | X = x)$ . Hence, for the given  $n$ -sample, and  $x \in \Delta_x$ , the optimal bandwidth  $h_n^*(x) = \arg \min_{h>0} \text{MSE}(x)$  writes:

$$h_n^*(x)^5 = \frac{R(K)\alpha(1-\alpha)}{n\mu_2(K)^2 \left( \frac{d^2 q^\alpha(x)}{dx^2} \right)^2 f(x) f(q^\alpha(x) | x)^2}. \quad (14)$$

As  $\frac{d^2 q^\alpha(x)}{dx^2}$  and  $f(q^\alpha(x) | x)$  are unknown, the authors in [16] propose to express  $h_n^*(x)$  in terms of  $h_{mean}(x)$ , the optimal bandwidth for the estimation of  $m(x)$  (i.e. which minimizes its asymptotic MSE), with  $m$  the conditional mean function defined as follows:

$$m : x \mapsto \mathbb{E}[Y | X = x].$$

The relation is approximated as follows:

$$h_n^*(x) = h_{mean}(x) \left[ \frac{\alpha(1-\alpha)}{\phi(\Phi^{-1}(\alpha))} \right]^{1/5} \quad (15)$$

As for  $h_{mean}(x)$ , we have from [4]:

$$h_{mean}(x)^5 = \frac{R(K)\sigma^2(x)}{n\mu_2(K)^2 m''(x)^2 f(x)}, \quad (16)$$

where  $\sigma^2(x) := \text{Var}(Y | X = x)$ . Nevertheless, the problem of unknown parameters still persists in (16) as  $\sigma^2(x)$  and  $m''(x)$  need to be estimated.

#### 4.2. Computation of an Optimal Bandwidth in Practice

The goal is to get an estimate of  $h_{mean}(x)$ , which mainly relies on a good approximation of the second-order derivative of  $m$ . Based on the development in [13], we suggest to perform a least squares quadratic fit (in the quoted article, the authors preferred a fourth-order fit, as they needed to estimate the fourth-order derivative of  $m$ ). This writes:

$$\forall j \in \{1, \dots, n\} \quad \left( \hat{m}(X^j), \hat{m}'(X^j), \hat{m}''(X^j) \right) = \arg \min_{(a,b,c) \in \mathbb{R}^3} \sum_{k=1}^n \left( Y^k - a - b(X^k - X^j) - c(X^k - X^j)^2 \right)^2$$

However, this is inadequate for regression functions having many oscillations (see [7] for further details). To this effect, the authors in [7] suggest to partition the range of the  $X$  data into  $N \in \mathbb{N}$  blocks and perform a quadratic fit over each block. In the following, the  $N$  blocks,  $B_1, \dots, B_N$ , are equally sized, with  $n$  divisible by  $N$ . We write  $t = n/N$ . For each  $i \in \{1, \dots, N\}$ ,  $B_i$  is the  $i$ -th sub-sample of the ordered data,  $X$ , i.e.

$$B_i := \left\{ X^{((i-1)t+1)}, \dots, X^{(it)} \right\}.$$

We perform the block-wise quadratic fit over the  $i$ -th block for its components, as follows, for any  $j \in \{(i-1)t+1, \dots, it\}$ :

$$\left( \hat{m}(X^{(j)}), \hat{m}'(X^{(j)}), \hat{m}''(X^{(j)}) \right) = \arg \min_{(a,b,c) \in \mathbb{R}^3} \sum_{k=(i-1)t+1}^{it} \left( Y^{(k)} - a - b(X^{(k)} - X^{(j)}) - c(X^{(k)} - X^{(j)})^2 \right)^2.$$

We define the blocked quadratic estimator for  $\sigma^2(\cdot)$  of the  $i$ -th block below:

$$\hat{\sigma}_{B_i}^2 := \frac{1}{t-3} \sum_{k=(i-1)t+1}^{it} \left( Y^{((i-1)t+s)} - \hat{m}(X^{((i-1)t+s)}) \right)^2,$$

which means that we assume  $\sigma^2(\cdot)$  to be constant over each block. The latter is as relevant as long as the blocked fit is efficient. The next issue is to set the best number of blocks,  $N$ , in order to get both satisfying blocked fit and a good estimation of the second derivative of the function  $m$ . To this effect, as suggested in [13], we use Mallows's  $CP$ , introduced in [9], with  $N \in \{1, \dots, N_{max}\}$ :

$$CP(N) = \frac{RSS(N)}{RSS(N_{max})}(n - 3N_{max}) - (n - 6N), \quad (17)$$

with  $N_{max}$  to be set in the following and  $RSS$  the classical residual sum of squares, defined as follows:

$$RSS(N) = \sum_{j=1}^n \left( Y^j - \hat{m}(X^{(j)}) \right)^2,$$

where  $\hat{m}(\cdot)$  corresponds to the  $N$  block-division of the initial  $X$  data. We introduce the data-driven best number of blocks,  $\hat{N}$ , as follows:

$$\hat{N} := \arg \min_{N \in \{1, \dots, N_{max}\}} CP(N), \quad (18)$$

as it guarantees a relevant fit. One can see that the expression in (17) includes a penalty on the number of blocks,  $N$ . Indeed, a high number of blocks would ensure a good fit, but the blocked estimation for  $m''$  would rely only on a few elements as the blocks would get small. This may lead to a very high variance-estimator for  $m''$ . As soon as one has determined  $\hat{N}$ , we can set a logical estimator for  $h_{mean}(X^j)$  from (16), with  $X^k$  a component of  $B_i$ :

$$\hat{h}_{mean}(x) := \left[ \frac{R(K)}{\mu_2(K)^2} \frac{\hat{\sigma}_{B_i}^2}{\hat{m}''(X^{(j)})nf(x)} \right]^{1/5} \quad (19)$$

One gets to the optimal bandwidth estimation for the conditional quantile with (15):

$$\hat{h}_n(x) := \hat{h}_{mean}(x) \left[ \frac{\alpha(1-\alpha)}{\phi(\Phi^{-1}(\alpha))} \right]^{1/5}. \quad (20)$$

We plug the expression (20) into the definition of the estimator,  $\hat{S}_{c_\alpha}^X(Y)$  (12).

**Remark.** This is a variant to the method displayed in [16]. Indeed they suggest to use the estimate provided [13] for  $h_{mean}(X^j)$ . However, the problem that they consider in the latter is to determine the best bandwidth regarding the integrated MSE for  $h_{mean}(X)$ . Therefore, this estimator cannot be plugged into the expression in (15) as it requires to be pointwise. While the authors in [13] defined  $\hat{h}_{mean}(X^k)$  from integrated parameters, e.g. from estimates for  $\int_{\Delta_x} m''(x)f(x)dx$  and  $\int_{\Delta_x} \sigma^2(x)dx$ , we defined  $\hat{h}_{mean}(X^k)$  from the pointwise (resp. blockwise) parameters,  $\hat{m}''(X^{(j)})$  (resp.  $\hat{\sigma}_{B_i}^2$ ).

## 5. Numerical Results

### 5.1. Toy Example

In order to test the efficiency of the optimal bandwidth estimation, we apply the method that we introduce to the following toy example:

$$Y = X - Z,$$

with  $X, Z$  iid  $\sim \text{Exp}(1)$ . With  $n = 10^3$ , we generate a  $n$ -sample  $((X^1, Z^1, Y^1), \dots, (X^n, Z^n, Y^n))$ .

#### 5.1.1. Estimates for the conditional quantile and the minimum of the mean contrast

Successively, we use the estimated bandwidth introduced in (20) to estimate conditional quantiles (i.e.  $q^\alpha(Y | X = x)$ , for  $x > 0$ ) and minima of the quantile-oriented mean contrast,  $c_\alpha$ ,  $\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X = x]$ , for  $x > 0$ . We assess the efficiency of these estimators for only  $\alpha = 0.5$ .

#### Conditional Quantile Estimate

For three different values of  $x$ ,  $x^1 = 0.5$ ,  $x^2 = 1$  and  $x^3 = 2.4$ , we computed the estimator for the conditional quantile, i.e. for  $i = 1, 2, 3$ :

$$\hat{q}_n^\alpha(Y | X = x^i) = \arg \min_{\theta \in \{Y^1, \dots, Y^n\}} \left[ \frac{1}{nf(x^i)} \sum_{j=1}^n c_\alpha(Y^j, \theta) K_{\hat{h}_n(x^i)}(X^j - x^i) \right],$$

with  $\hat{h}_n(x^i)$  introduced in (20). The values that we seek to estimate are:

$$q^\alpha(Y | X = x^1) \simeq -0.19, \quad q^\alpha(Y | X = x^2) \simeq 0.31, \quad \text{and} \quad q^\alpha(Y | X = x^3) \simeq 1.71$$

In order to display the efficiency of this estimator, we repeated the experiments  $n_{iter} = 10^3$  times and computed the following values, for  $i = 1, 2, 3$ :

$$\text{Mean\_error}(i) := \sqrt{\frac{1}{n_{iter}} \sum_{j=1}^{n_{iter}} (\hat{q}_n^\alpha(Y | X = x^i) - q^\alpha(Y | X = x^i))^2},$$

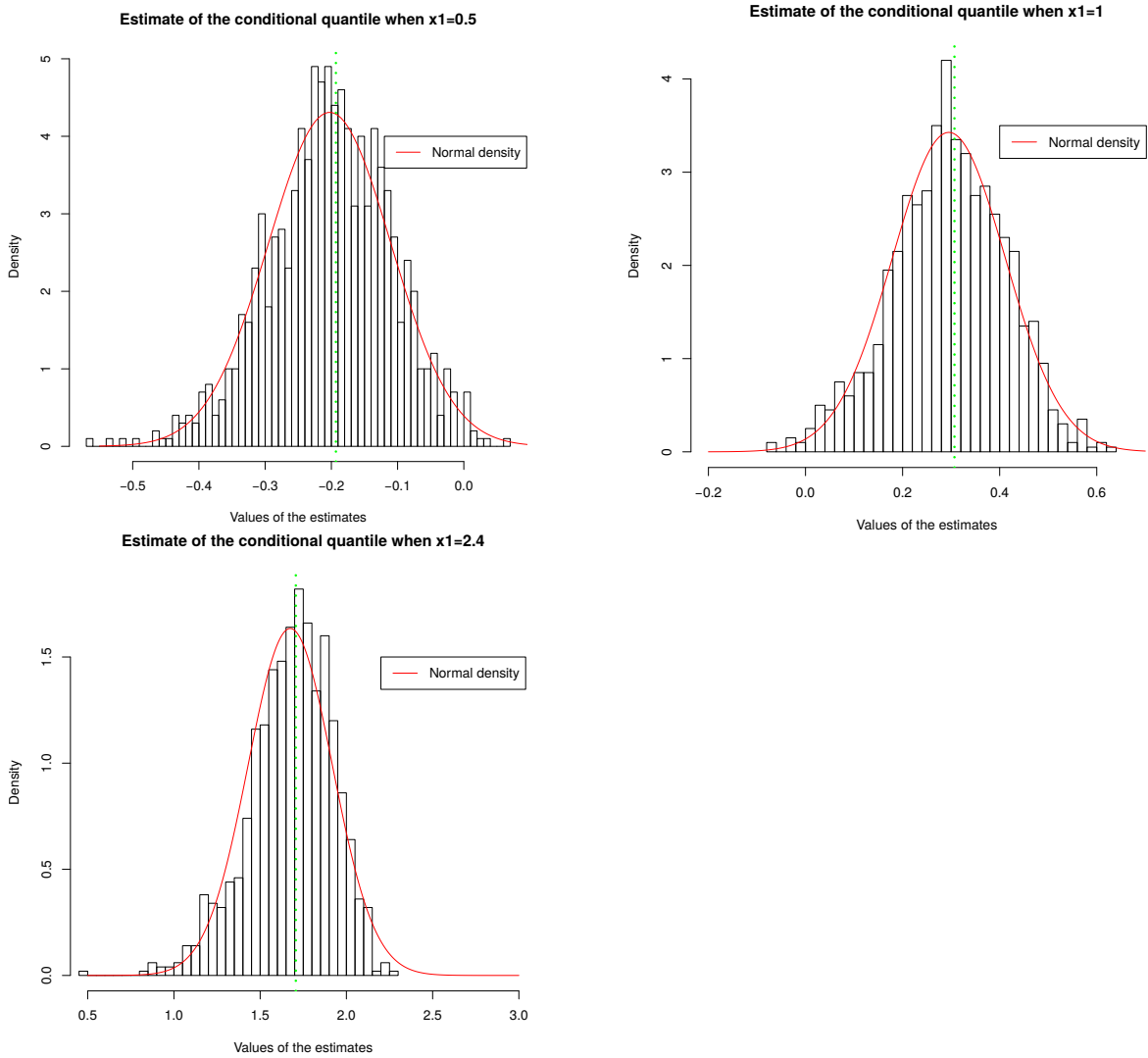
$$\text{Mean\_bias}(i) := \frac{1}{n_{iter}} \sum_{j=1}^{n_{iter}} (\hat{q}_n^\alpha(Y | X = x^i) - q^\alpha(Y | X = x^i)),$$

$$\text{Mean\_variance}(i) := \frac{1}{n_{iter}} \sum_{j=1}^{n_{iter}} (\hat{q}_n^\alpha(Y | X = x^i) - q^\alpha(Y | X = x^i) - \text{Mean\_bias}(i))^2.$$

The numerical results for the three types of error are listed in Table 1 and displayed in the histograms of Figure 2. As for Figure 2, we add each time the Gaussian pdf whose expectation is  $\text{Mean\_bias}(i)$  and variance is  $\text{Mean\_variance}(i)$ . We can see that the estimation of the conditional quantile is accurate as long as the conditioning value,  $x^i$ , is likely enough. On the other hand, the error is quite high when conditioning by  $x^3$ , with  $q^{0.90}(X) < x^3$ . In Figure 2, we can observe the asymptotic normality of the estimation of  $q^\alpha(Y | X = x^i)$ . But once again, it seems less accurate when  $x$  is not likely.

$x$	$x^1=0.5$	$x^2=1$	$x^3=2.4$
$Mean\_error(i)$	0.03	0.15	0.92
$Mean\_bias(i)$	$-3.3e^{-3}$	$-5.4.e^{-3}$	$-1.5e^{-2}$
$Mean\_variance(i)$	$4.2e^{-4}$	$7.2e^{-3}$	$3.3e^{-2}$

**Table 1.**  $n_{iter}$  repetitions of  $n$ -sample in order to generate  $n_{iter}$  estimates of  $q^\alpha(Y | X = x^i)$ , with  $i = 1, 2, 3$  for  $\alpha = 0.5$ .



**Figure 2.** Distribution of the repetitions of the estimator,  $\hat{q}_n^\alpha(Y | x^i)$ , with  $i = 1, 2, 3$  for  $\alpha = 0.5$  and  $n = 10^3$ . The real value is in green dotted line.

$x$	$x^1=0.5$	$x^2=1$	$x^3=2.4$
$Mean\_error(i)$	$3.7e^{-4}$	$4.8e^{-4}$	$9.8e^{-4}$
$Mean\_bias(i)$	$-1.3e^{-3}$	$-1.2e^{-3}$	$-1.5e^{-3}$
$Mean\_variance(i)$	$2.1e^{-3}$	$3.7e^{-3}$	$1.5e^{-2}$

**Table 2.** Errors of the  $n_{iter}$  estimates of  $\widehat{\min}_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X = x^i]$ , with  $i = 1, 2, 3$  for  $\alpha = 0.5$ .

### Minimum of the quantile-oriented contrast

For three different values of  $x$ ,  $x^1 = 0.5$ ,  $x^2 = 1$  and  $x^3 = 2.4$ , we computed the estimator for the minimum of the quantile-oriented contrast, i.e. for  $i = 1, 2, 3$ :

$$\widehat{\min}_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X = x^i]_n := \min_{\theta \in \{Y^1, \dots, Y^n\}} \left[ \frac{1}{nf(x^i)} \sum_{j=1}^n c_\alpha(Y^j, \theta) K_{\hat{h}_n(x^i)}(X^j - x^i) \right].$$

In [6], the exact expression of  $\mathbb{E}[c_\alpha(Y, \theta) | X = x]$ , for any  $x \in \mathbb{R}$  and  $\alpha \in ]0, 1[$ , is computed. We have:  $\mathbb{E}[c_\alpha(Y, \theta) | X = x] = -\alpha \log(\alpha)$ , if  $1/2 \leq \alpha$ . The values that we seek to estimate are all equal in this case:  $\min_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X = x^i] \simeq 0.35$ , for  $\alpha = 0.5$ . In order to display the efficiency of this estimator, we repeated the whole experiments  $n_{iter} = 10^3$  times and computed the following values, for  $i = 1, 2, 3$ :

$$Mean\_error(i) := \frac{1}{n_{iter}} \sum \left| \widehat{\min}_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X = x^i]_n - \min_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X = x^i] \right|,$$

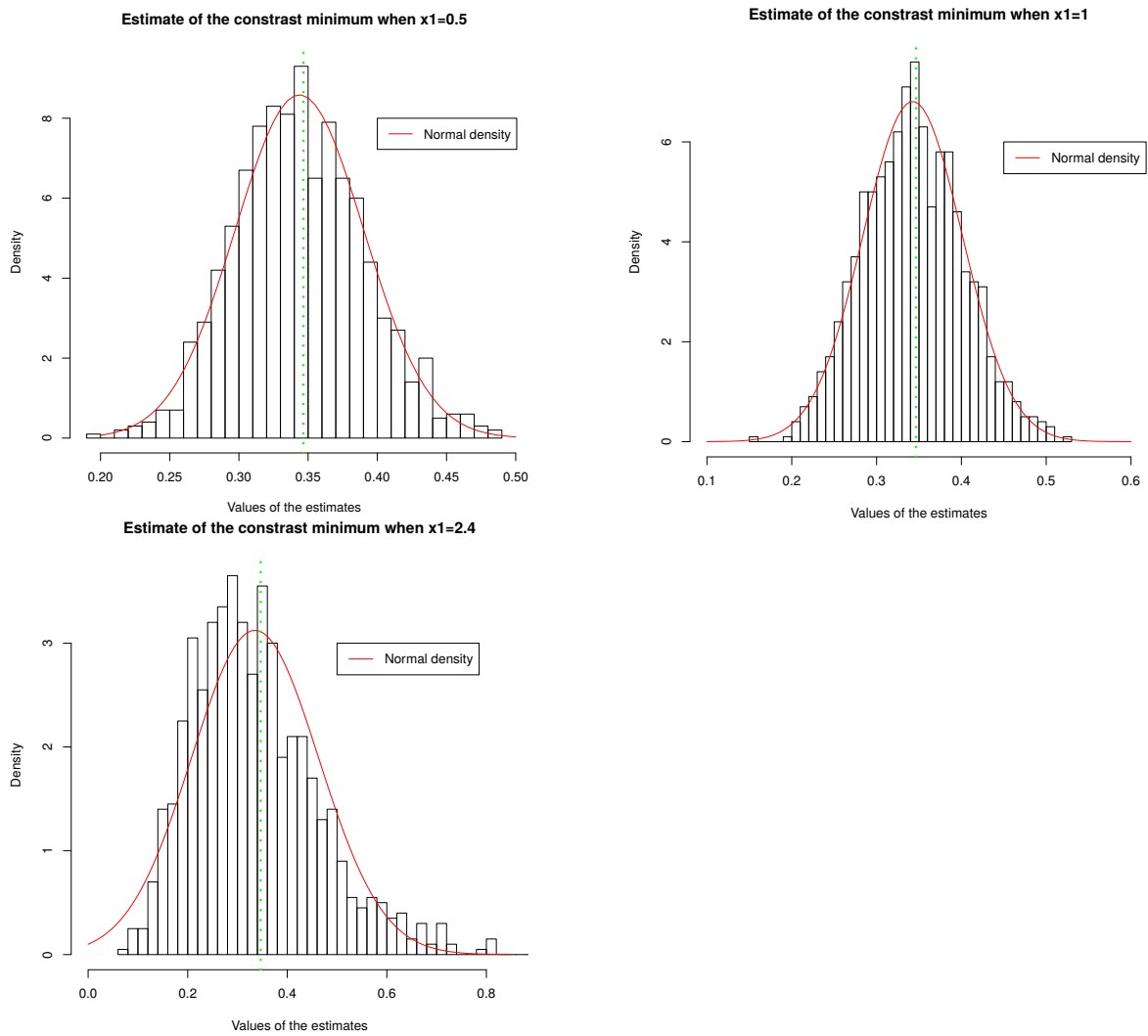
$$Mean\_bias(i) := \frac{1}{n_{iter}} \sum \left( \widehat{\min}_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X = x^i]_n - \min_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X = x^i] \right),$$

$$Mean\_variance(i) := \frac{1}{n_{iter}} \sum \left( \widehat{\min}_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X = x^i]_n - \min_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X = x^i] - Mean\_bias(i) \right)^2.$$

The numerical results for the three types of error are listed in Table 2 and in the histograms of Figure 3. As for Figure 2, we add each time the Gaussian *pdf* whose expectation is  $Mean\_bias(i)$  and variance is  $Mean\_variance(i)$ . The error seems even lower than on  $q^\alpha(Y | x^i)$ , which justifies our choice to minimize the error of the minimizer in order to get a good estimation of the minimum. Mainly, the estimation does not seem to be harmed by the choice of  $x$ , as the different types of error do not really vary whether we choose  $x^1, x^2$  or  $x^3$ . We observe once again in Figure 3 the asymptotic normality for the estimator of  $\min_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X = x^i]$ . Unlike the errors in Table 2, it seems more affected by the choice of  $x$ .

### 5.1.2. Application of the selected bandwidth to the quantile-oriented SA estimator

Given the previous development regarding the determination of an efficient bandwidth for the estimation of  $\min_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X = x]$ , our goal is to integrate it into the expression of the estimator  $\hat{S}_{c_\alpha}^X(Y)$ . To do so, let us recall that one can express the second term of the estimator with



**Figure 3.** Distribution of the repetitions of the estimators  $\widehat{\min}_{\theta \in \mathbb{R}} \mathbb{E} [c_\alpha(Y, \theta) | X_1 = x^i]_n$ , with  $i = 1, 2, 3$  for  $\alpha = 0.5$  and  $n = 10^3$ . The real value is in green dotted line.



$\hat{C}_x$ , introduced in (10):

$$\frac{1}{m} \sum_{k=1}^m \min_{\theta \in \mathbb{R}} \hat{C}_{X^{tk}}(\theta),$$

where  $\min_{\theta \in \mathbb{R}} \hat{C}_{X^{tk}}$  is the estimator for  $\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) \mid X = x]$ . Therefore, for each  $k \in \{1, \dots, m\}$ , we substitute  $h_n$  by  $\hat{h}_n(X^{tk})$ , introduced in (20), into the expression of  $\hat{C}_{X^{tk}}$ . We finally get:

$$\begin{aligned} \hat{S}_{c_\alpha}^X(Y) &:= \min_{\theta \in \mathbb{R}} \frac{1}{n} \sum_{j=1}^n c_\alpha(Y^j, \theta) \\ &- \frac{1}{m} \sum_{k=1}^m \min_{\theta} \frac{1}{nf(X^{tk})} \sum_{j=1}^n c_\alpha(Y^j, \theta) K_{\hat{h}_n(X^{tk})}(X^j - X^{tk}). \end{aligned} \quad (21)$$

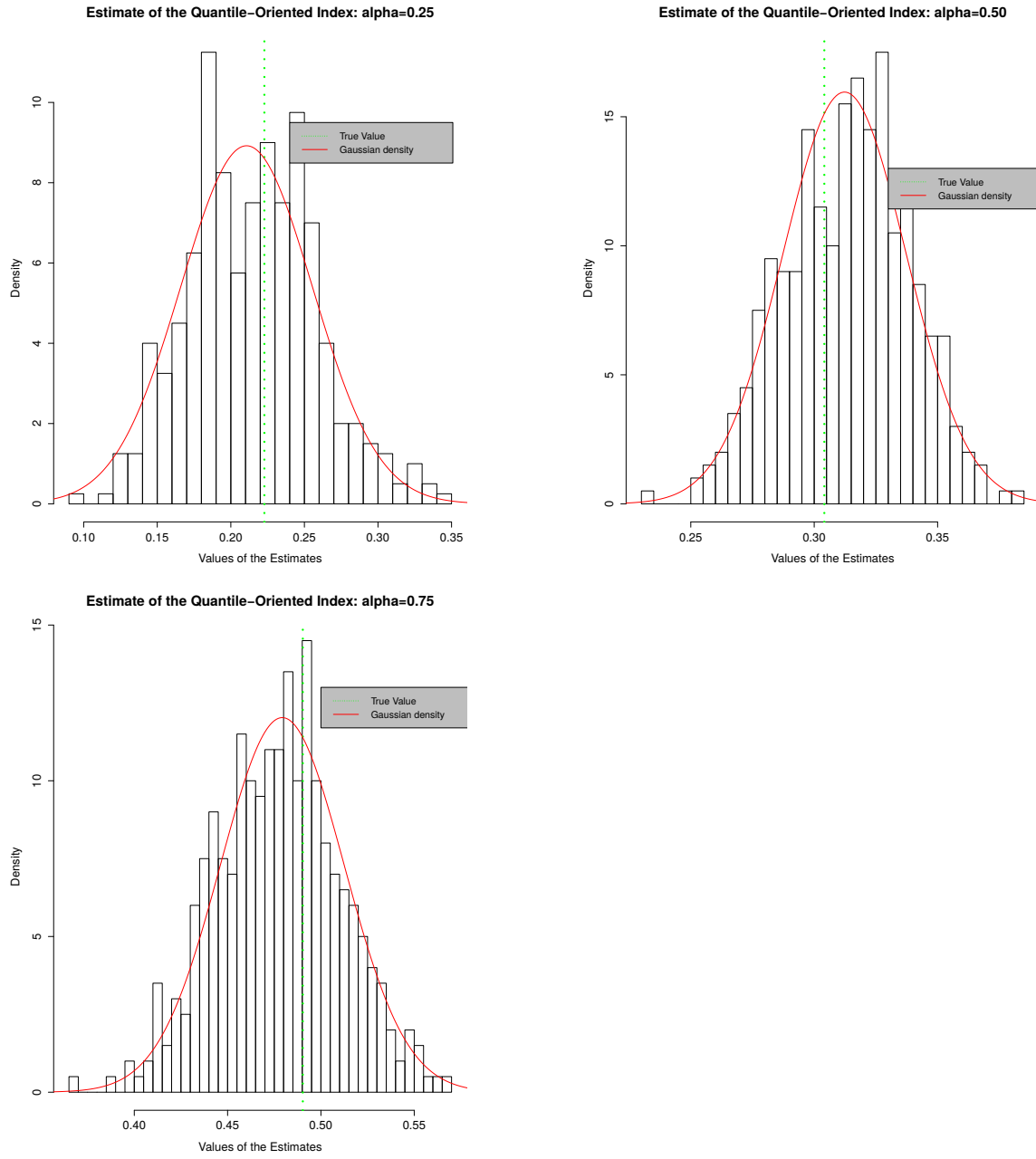
Let us recall the toy example:  $Y = X - Z$ , with  $X, Z$  iid  $\sim \text{Exp}(1)$ . With  $n = 10^3$ , we generate a  $n$ -sample  $((X^1, Z^1, Y^1), \dots, (X^n, Z^n, Y^n))$ . The aim of the following paragraph is to assess the estimator we defined in (21). The development provided for this ‘‘toy example’’ in [6] gives the exact expression for the quantile-oriented SA analysis,  $S_{c_\alpha}^X(Y)$ :

$$\begin{aligned} S_{c_\alpha}^X(Y) &= \frac{(1 - \alpha)(1 - \log(2(1 - \alpha))) + \alpha \log(\alpha)}{(1 - \alpha)(1 - \log(2(1 - \alpha)))} \quad \text{if } 1/2 \leq \alpha \\ S_{c_\alpha}^X(Y) &= \frac{\alpha(1 - \log(2\alpha)) + \alpha \log(\alpha)}{(1 - \alpha)(1 - \log(2(1 - \alpha)))} \quad \text{if } 1/2 > \alpha \end{aligned}$$

The value that we seek to estimate is:  $S_{c_\alpha}^X(Y) = 0.22, 0.30$  and  $0.49$  for  $\alpha = 0.25, 0.5$  and  $0.75$ . It is important to mention that nothing has been said so far about the value of  $m$ , the size of the sample generated from  $X'$ . It turns out that its value is independent from the size of sample,  $n$ . We consider  $n$  to be set prior to the SA study (at least in our case). As we have already stated, the value  $n$  matters more as it requires more runs of the code when it grows. Increasing  $m$  just impacts the number of the minimization problems needed for the estimator. Let us recall that we have  $n$  needed evaluations of a low-time consuming function, and this for each element  $X^{tk}$  of the  $m$ -sample  $(X^1, \dots, X^m)$ . In the following, we set  $m = 10^3$  as we seem to have the following convergence knowing the  $n$ -sample  $((X^1, Y^1), \dots, (X^n, Y^n))$ :

$$\frac{1}{m} \sum_{k=1}^m \min_{\theta \in \mathbb{R}} \hat{C}_{X^{tk}}(\theta) \xrightarrow{m \rightarrow \infty} \mathbb{E}_{X'} \left[ \min_{\theta \in \mathbb{R}} \hat{C}_{X'}(\theta) \mid \left( (X^1, Y^1), \dots, (X^n, Y^n) \right) \right],$$

with  $\hat{C}_{X'}$ , introduced in (10). Through this, we mean that according to our numerical applications, increasing  $m$  would not improve the quality of the estimator. Indeed, it is certain that increasing  $m$  infinitely will not ensure any consistency for the estimator as long as  $n$  is too low. Therefore, the quality of our estimator mainly relies on the pointwise quality of the estimator  $\min_{\theta \in \mathbb{R}} \hat{C}_{X'}(\theta)$ , which depends on the value of  $n$  (see Proposition 4). We repeat the experiments  $n_{iter} = 10^3$  times in order to observe the estimators’ empirical respective distributions. Classical errors that we observe are listed in Table 3. We display the histograms of the results from these repetitions for  $\alpha = 0.5$  in Figure 4. We add to the histogram the pdf’s of Gaussian distribution with similar bias and variance than on the sample. It seems that the samples of  $\hat{S}_{c_\alpha}^X(Y)$  are distributed according to a Gaussian law.



**Figure 4.** Distribution of the repetitions of the estimator  $\hat{S}_{c_\alpha}^X(Y)$  for  $\alpha = 0.25, 0.5$  and  $0.75$ , with  $n = 10^3$ . The real value is in green dotted line.

$\hat{S}_{c_\alpha}^{X_1}(Y)$	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$
<i>Mean_error</i>	$6.1e^{-2}$	$2.9e^{-2}$	$5.5e^{-2}$
<i>Mean_bias</i>	$1.2e^{-2}$	$8.2e^{-3}$	$1.1e^{-2}$
<i>Mean_variance</i>	$2.0e^{-3}$	$7.5e^{-4}$	$1.1e^{-3}$

**Table 3.** Errors of the  $n_{iter}$  estimates of  $S_{c_\alpha}^X(Y)$ .

$\tilde{S}_{c_\alpha}^{X^1}(Y)$	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$
<i>Mean_error</i>	$2.0e^{-1}$	$1.3e^{-1}$	$2.6e^{-1}$
<i>Mean_bias</i>	$1.2e^{-1}$	$8.5e^{-2}$	$1.4e^{-1}$
<i>Mean_variance</i>	$8.0e^{-1}$	$3.5e^{-1}$	$7.2e^{-1}$

**Table 4.**  $n_{iter}$  repetitions of  $\tilde{S}_{c_\alpha}^X(Y)$ , the empirical estimator of  $S_{c_\alpha}^X(Y)$

**Remark.** In order to prove the accuracy of this kernel-based approach for the estimate of  $S_{c_\alpha}^X(Y)$ , we compare it with a classical empirical estimator (we mention it in Section 3 as a non-desirable method, see Figure 1). With the same number of simulations  $n = 10^3$ , we allocate  $n_1$  simulations for the estimation of  $\min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta)]$  and  $n_2$  for  $\mathbb{E}_X \left[ \min_{\theta \in \mathbb{R}} \mathbb{E}[c_\alpha(Y, \theta) | X] \right]$ , with  $n = n_1 + n_2$ . Let us generate the iid samples  $((X^1, Z^1), \dots, (X^{n_1}, Z^{n_1}))$  of  $(X, Z)$ ,  $(X'_1, \dots, X'_N)$  of  $X$  and  $(Z'_{k'}{}^j)_{k' \in \{1, \dots, N\}, j \in \{1, \dots, n_2/N\}}$ , with  $n_2$  divisible by  $N$ . We introduce the expression for the classical empirical estimator:

$$\tilde{S}_{c_\alpha}^X(Y) := \frac{\frac{1}{n_1} \min_{\theta \in \mathbb{R}} \sum_{k=1}^{n_1} c_\alpha(Y^k, \theta) - \frac{1}{N} \sum_{k'=1}^N \frac{N}{n_2} \min_{\theta \in \mathbb{R}} \sum_{j=1}^{n_2/N} c_\alpha(Y'_{k'}{}^j, \theta)}{\frac{1}{n_1} \min_{\theta \in \mathbb{R}} \sum_{k=1}^{n_1} c_\alpha(Y^k, \theta)}, \quad (22)$$

with:

$$\forall k \in \{1, \dots, n_1\} \quad Y^k := g(X^k, Z^k),$$

and:

$$\forall k' \in \{1, \dots, N\}, \forall j \in \{1, \dots, n_2/N\} \quad Y'_{k'}{}^j := g(X'_{k'}, Z'_{k'}{}^j).$$

We set  $n_1 = 100$ ,  $n_2 = 900$  and  $N = 30$ . The whole experiments are repeated  $n_{iter} = 10^3$  times and the results are displayed in Table 4. Compared to Table 3, the errors are very high in this case. As expected, the empirical approach for  $S_{c_\alpha}^X(Y)$  estimate seems highly inaccurate.

## 5.2. Application to Non-Destructive Examination

### 5.2.1. Presentation

In industries, it is common to carry out Non-Destructive Examinations (NDE) in order to ensure the integrity of an important structure. The goal is to detect any defect, split or flaw, that could severely damage the system. It consists in sending an ultrasonic wave through the structure to study and measure its amplitude after reflection. The main idea is that the amplitude of the wave increases with the size of the defect, if there is any. To this effect, engineers set a threshold  $t_s > 0$  such that one concludes that there is a defect as soon as the measured amplitude is greater than  $t_s$ . As for the mathematical framework, we have:

- $Y > 0$ , the amplitude of the signal after examination,
- $a > 0$ , the size of defect,
- $X \in \chi \subset \mathbb{R}^d$ , with  $d = 6$ , the influential random parameters of the examination.  $X_1, \dots, X_d$  are independent.

$Y$  is a deterministic function of  $a$  and  $X$ . We write  $g$  the function so that  $Y = g(a, X)$ . It is important to note out that, for any  $x \in \mathbb{R}^d$ , the function  $(a \rightarrow f(a, x))$  is increasing. Therefore,

$X_1$	Thickness of the structure
$X_2$	Transverse wave speed
$X_3$	Attenuation coefficient
$X_4$	Rate of refraction
$X_5$	Angle control
$X_6$	Tilt

**Table 5.** List of the random inputs in the defect detection control.

we understand that the higher the signal is, the more likely it is to have a defect in the inspected structure. To this effect, engineers set a threshold  $t_s$ , so that:  $Y > t_s$  means that a defect has been detected. In order to assess the capability for the NDE's to detect any defect by repeating the experiments a high number of times, engineers use simulators that model the detection process.

We work with the numerical code CIVA from the company CEA which simulates ultrasonic controls. We display in Table 5 the different meanings of each input. For  $N = 30$  different sizes of defect  $(a_1, \dots, a_N)$ ,  $M = 20$  simulations are performed with  $M$  independent realizations of the random vector  $X$ :

$$\begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix}, \begin{pmatrix} X^{1,1} & \dots & X^{1,M} \\ \vdots & & \vdots \\ X^{N,1} & \dots & X^{N,M} \end{pmatrix}$$

As the signal  $Y$  is function of  $a$  and  $X$ , we write:

$$\forall k \in \{1, \dots, N\}, j \in \{1, \dots, M\} \quad Y^{k,j} = Y(a_k, X_1^{k,j}, \dots, X_d^{k,j}).$$

In order to have a first idea of the inputs' respective influence, we display in Figure 5 the scatter plots for the first four inputs,  $(a, X_1, X_2, X_3)$ , with respect to the output,  $Y$ . The idea is to observe how some aspects of  $Y$ 's distribution can be affected by different values for the considered input. We can see that  $a$  seems far more influential. One could even assume a linear relationship with  $Y$ . As for  $X_1$ , its only influence that can be assume over  $Y$  regards the higher quantiles,  $q^\alpha(Y)$ , for  $\alpha$  close to 1. Indeed, we observe that mainly the upper tail of  $Y$ 's distribution seems to be affected by the different values of  $X_1$ . As for  $X_2$  and  $X_3$ , no significant modification is to be pointed out. We did not represent the scatter plots for the last 3 inputs as they do not significantly differ from  $X_2$  and  $X_3$ .

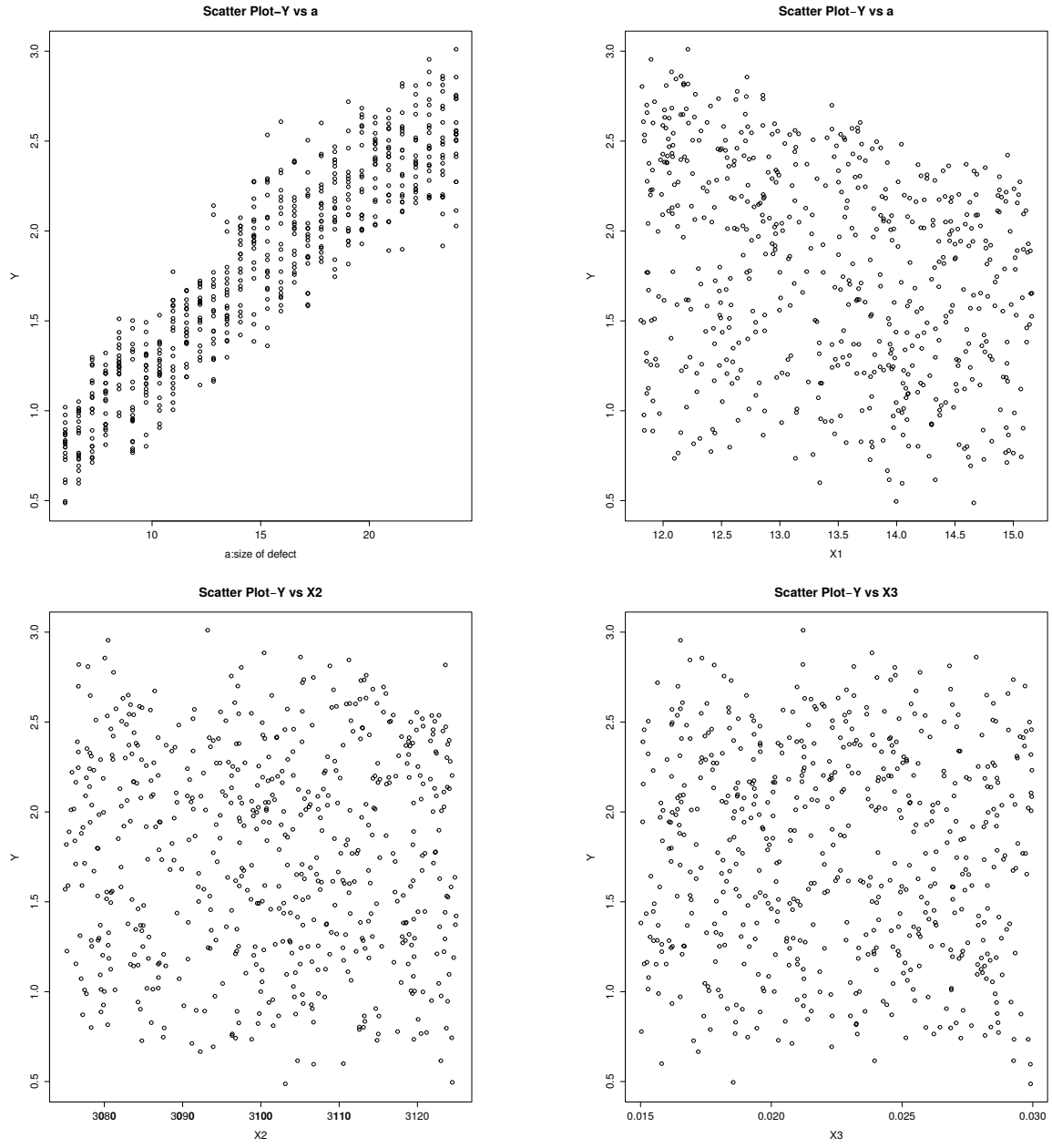
### 5.2.2. Numerical Results: Quantile-Oriented SA

We perform the numerical results for the estimation of  $S_{c_\alpha}^{X_i}(Y)$ ,  $i = 1, \dots, d$  out of the data set which we were provided. As mentioned in the previous numerical results,  $m = 10^3$  still seems to ensure the following convergence:

$$\frac{1}{m} \sum_{k=1}^m \min_{\theta \in \mathbb{R}} \hat{C}_{X^{rk}}(\theta) \xrightarrow{m \rightarrow \infty} \mathbb{E}_{X'} \left[ \min_{\theta \in \mathbb{R}} \hat{C}_{X'}(\theta) \mid \left( (X^1, Y^1), \dots, (X^n, Y^n) \right) \right],$$

with  $\hat{C}_{X'}$  introduced in (10). We computed estimates introduced in (21) for  $S_{c_\alpha}^{X_i}(Y)$ ,  $i = 1, \dots, d$ , as well as  $S_{c_\alpha}^a(Y)$ , for different levels of quantile,  $\alpha \in ]0, 1[$ . We listed the results in Table 6. The values of the estimates underline three facts:

- For each input but  $X_1$ , it seems that there is equal influence over all the quantiles of the signal,  $q^\alpha(Y)$ ,  $\alpha \in ]0, 1[$ . This means that the uncertainty propagation of each input is well



**Figure 5.** Scatter Plots of the first four inputs,  $Y$  vs  $a$ ,  $Y$  vs  $X_i$ ,  $i = 1, \dots, 3$ .

$\alpha$	$a$	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
0.1	0.92	0.10	0.15	0.12	0.14	0.12	0.13
0.25	0.89	0.11	0.13	0.11	0.13	0.10	0.14
0.5	0.90	0.13	0.13	0.10	0.13	0.10	0.14
0.75	0.91	0.16	0.13	0.10	0.14	0.09	0.13
0.9	0.87	0.20	0.14	0.11	0.16	0.11	0.14

**Table 6.** List of the values of the estimates for  $S_{c_\alpha}^-(Y)$ , for all the inputs, at different levels of quantiles,  $\alpha \in ]0, 1[$ .

$\alpha$	$a$	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
0.1	[0.89;0.93]	[0.06;0.12]	[0.06;0.17]	[0.06;0.16]	[0.07;0.16]	[0.09;0.16]	[0.09;0.15]
0.25	[0.88;0.91]	[0.08;0.13]	[0.07;0.16]	[0.09;0.18]	[0.06;0.15]	[0.09;0.15]	[0.08;0.16]
0.5	[0.91;0.92]	[0.12;0.15]	[0.10;0.17]	[0.07;0.12]	[0.07;0.15]	[0.08;0.15]	[0.09;0.14]
0.75	[0.91;0.93]	[0.13;0.19]	[0.08;0.18]	[0.07;0.15]	[0.06;0.15]	[0.07;0.17]	[0.10;0.15]
0.9	[0.86;0.91]	[0.18;0.26]	[0.06;0.16]	[0.07;0.16]	[0.05;0.15]	[0.07;0.14]	[0.08;0.15]

**Table 7.** List of the 90%-Confidence Intervals for the estimates of  $S_{c_\alpha}^\cdot(Y)$  after bootstrap.

spread throughout the whole output distribution (vertical reading).

- $X_1$ 's influence seems to be more important for higher levels of quantiles (moving from 0.11 for  $\alpha = 0.1$  to 0.20 for  $\alpha = 0.9$ ). This confirms the observation made from the corresponding scatter plot.
- Overall,  $a$  is far more influential over the output distribution than every marginal of  $X$ , which is not surprising at all as it is the main the purpose of this detection process. On the other hand, together the marginals of  $X$  seem to be equally influential.

In order to verify the robustness of the estimator for the given data, we perform a bootstrap procedure. It consists in generating  $n$ -samples,  $n_{iter}$  times, from the initial  $n$ -sample by sampling with replacements. Then, we apply the estimator to all the samples to get  $(S_{c_\alpha}^\cdot(Y)_{BS}^1, \dots, S_{c_\alpha}^\cdot(Y)_{BS}^{n_{iter}})$  and compare their values. A classical method suggests to set  $1 - \beta$ -confidence intervals over  $S_{c_\alpha}^\cdot(Y)$ ,  $CI^\beta(S_{c_\alpha}^\cdot(Y))$ , with  $\beta \in ]0, 0.5[$ , as follows:

$$CI^\beta(S_{c_\alpha}^\cdot(Y)) := \left[ \hat{q}_n^{\beta/2}(S_{c_\alpha}^\cdot(Y)_{BS}) ; \hat{q}_n^{1-\beta/2}(S_{c_\alpha}^\cdot(Y)_{BS}) \right], \quad (23)$$

where  $\hat{q}_n$  is the classical empirical estimator for the quantiles, introduced in (6). The corresponding 90%-confidence intervals are displayed in Table 7. One can conclude that once again,  $a$  has a different behaviour. Indeed, the confidence-intervals for  $a$  are pretty thin. This suggests that the corresponding values are reliable. As for the other inputs, one can see that the initial sample would require to have  $n$  higher to ensure reliable estimators.

## 6. Conclusion

In this paper, we introduced estimators for the quantile-oriented SA indices with respect to a contrast. Several side consistency results are proved in order to justify its efficiency. Yet, the proof for the estimators' consistency remains a major wish for the future. The computations of the estimators for the "toy example" suggest that these estimators might have an asymptotic Gaussian distribution. This would be a significant step forward for their use. Through the applications, we could see that they provide relevant information that other common methods, such as Sobol indices, could not detect. It is important to remember that these indices do not have an analogue to the variance decomposition offered by Sobol indices through the Hoeffding theorem. Thus, it is for example impossible in practice to set a random input to a unique value after proving that it is barely influential over the quantile. These indices provide only local information. To this effect, it should be interesting to set a general methodology when a SA study is needed: for instance, how and in what order to combine the different indices that we know to solve a precise problem.

## **7. Acknowledgments**

Part of this work has been backed by French National Research Agency (ANR) through the Chorus project (no. ANR-13-MONU-0005-08). We are grateful to CEA for providing us with the numerical data. We also thank Thierry Klein, Nabil Rachdi, Vincent Feuillard and Sébastien Da Veiga for helpful discussions.

## **8. Disclosure statement**

The authors declare that there is no conflict of interest.

## References

- [1] E. Borgonovo. A new uncertainty importance measure. *Reliability Engineering & System Safety*, 92(6):771–784, 2007.
- [2] S. Da Veiga. Global sensitivity analysis with dependence measures. *Journal of Statistical Computation and Simulation*, 85(7):1283–1305, 2015.
- [3] E. de Rocquigny, N. Devictor, and S. Tarantola, editors. *Uncertainty in industrial practice*. Wiley, 2008.
- [4] J. Fan. Local Linear Regression Smoothing and their Minimax Efficiencies. *The Annals of Statistics*, 21:196–216, 1993.
- [5] J. Fan, T.-C. Hu, and Y. K. Truong. Robust Non-parametric Function Estimation. *Foundation of the Scandinavian Journal of Statistics*, 21:433–446, 1994.
- [6] J.-C. Fort, T. Klein, and N. Rachdi. New Sensitivity Analysis Subordinated to a Contrast. *Communication in Statistics-Theory and Methods*, 45(15):4349–4364, 2016.
- [7] W. Härdle and J. S. Marron. Fast and simple scatterplot smoothing. *Computational Statistics & Data Analysis*, 20(1):1–17, 1995.
- [8] L. Le Gratiet, B. Iooss, G. Blatman, T. Browne, S. Cordeiro, and B. Goursaud. *Journal of Nondestructive Evaluation*, 36(1):8, 2017.
- [9] C. L. Mallows. Some comments on Cp. *Technometrics*, 15(4):661–675, 1973.
- [10] H. Monod, C. Naud, and D. Makowski. Uncertainty and sensitivity analysis for crop models. In D. Wallach, D. Makowski, and J. W. Jones, editors, *Working with Dynamic Crop Models: Evaluation, Analysis, Parameterization, and Applications*, pages 55–99. Emsevier, 2006.
- [11] M. D. Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161–174, 1991.
- [12] N. Rachdi. *Statistical Learning and Computer Experiments*. PhD thesis, Université de Toulouse, France, 2011.
- [13] D. Ruppert, S. J. Sheather, and M. P. Wand. An Effective Bandwidth Selector for Local Squares Regression. *Journal of the American Statistical Association*, 90(432):1257–1270, 1995.
- [14] I. M. Sobol. Sensitivity estimates for nonlinear mathematical models. *MMCE*, 1:407–414, 1993.
- [15] A. W. Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.
- [16] K. Yu and M. Jones. Local linear quantile regression. *Journal of the American statistical Association*, 93(441):228–237, 1998.