Goal-Oriented Shapley Effects with Special Attention to the Quantile-Oriented Case

Kevin Elie-Dit-Cosaque, Véronique Maume-Deschamps

HAL Id: hal-03137402
https://hal.science/hal-03137402v1

Submitted on 10 Feb 2021 (v1), last revised 4 Feb 2023 (v2)

On quantile oriented sensitivity analysis

Kévin Elie-Dit-Cosaque [*1, 2] and Véronique Maume-Deschamps [†1]

[1]Institut Camille Jordan, Université Claude Bernard Lyon 1, Lyon, FRANCE
[2]Actuarial Modelling, SCOR, Paris, FRANCE

February 10, 2021

Abstract

We propose to study quantile oriented sensitivity indices (QOSA indices) and quantile oriented Shapley effects (QOSE). Some theoretical properties of QOSA indices will be given and several calculations of QOSA indices and QOSE will allow to better understand the behaviour and the interest of these indices.

Keywords: Quantile Oriented Sensisitivy Analysis, Shapley effects.

1 Introduction

Sensitivity Analysis (SA) is defined by Saltelli et al. (2004) as "the study of how the uncertainty in the output of a model can be apportioned to different sources of uncertainty in the model input". Various tools exist today to perform a SA (see e.g. Iooss and Lemaître (2015) for a review of SA methods).

We are especially interested in *Global Sensitivity Analysis methods* - GSA - which allow to study the effects of simultaneous variation of the inputs on the model output in their entire domain. For a detailed description of sensitivity analysis methods, the interested reader can refer to the various survey papers dedicated to this topic (Saltelli et al., 2004, 2008; Faivre et al., 2016; Borgonovo and Plischke, 2016; Borgonovo et al., 2017). Variance-based methods are common tools in the analysis of complex physical phenomenons. Most of them rest on an ANalysis Of VAriance (ANOVA) of the model output and give information on the sensitivity around the mean (as it is variance based). In this paper, we are interested in Quantile Oriented indices, in order to obtain informations on the sensitivity around quantiles. Much less work has been done on Quantile Oriented Sensitivity Analysis (QOSA). We shall focus on indices defined firstly in Fort et al. (2016).

For completeness, Section 2 shortly recalls variance-based methods dealing both with indepedent and dependent inputs. Section 3 introduces QOSA indices which allows to quantify the sensitivity over a quantile. Note that in Kucherenko et al. (2019), other quantile oriented indices have been defined. Nevertheless, QOSA indices have more natural interpretations as will be seen in Section 3. Some properties of QOSA indices are also proposed within this section. Several calculations of QOSA indices are done in Section 4 and a preliminary work is carried out

---

[*]edckev@gmail.com
[†]veronique.maume@univ-lyon1.fr

in order to understand the impact of the statistical dependence between the inputs over these indices. Facing some interpretation issues, we finally propose in Section 5 new indices based on Shapley values (Shapley, 1953) which seem to be a promising alternative. Section 6 presents some further perspective of research.

Consider a model $Y = \eta(\mathbf{X})$ with d random inputs denoted by $\mathbf{X} = (X_1, X_2, \dots, X_d)$. Let $\mathbf{X}_{\mathcal{J}}$ indicate the vector of inputs corresponding to the index set $\mathcal{J} \subseteq \mathcal{D}$ where $\mathcal{D} = \{1, 2, \dots, d\}$. GSA aims at quantifying the impact of the inputs $X_1, X_2, \dots, X_d$ on the output Y.

2 Variance-based sensitivity indices

We shall briefly recall the framework of Sobol indices and Shapley effects.

2.1 Sobol indices

Sobol' sensitivity indices stem from the works of Fisher and Mackenzie (1923) and Hoeffding (1948) on the U-statistics taken up by various authors over time such as Efron and Stein (1981). Those ultimately lead to a functional ANOVA expansion of the model output η:

$$\eta(\mathbf{X}) = \eta_0 + \sum_{i=1}^{d} \eta_i(X_i) + \sum_{1 \leqslant i < j \leqslant d} \eta_{ij}(X_i, X_j) + \cdots + \eta_{1,\dots,d}(\mathbf{X}) \ . \tag{2.1}$$

This decomposition is unique under orthogonality constraints (see Sobol (1993)), that are verified when dealing with independent inputs, which lead to the decomposition of the global variance as

$$\mathrm{Var}(Y) = \sum_{i=1}^{d} V_i + \sum_{1 \leqslant i < j \leqslant d} V_{ij} + \cdots + V_{1,\dots,d} \ , \tag{2.2}$$

where

$$V_i = \mathrm{Var}\left(\eta_i\left(X_i\right)\right) = \mathrm{Var}\left(\mathbb{E}\left[Y \mid X_i\right]\right) \ ,$$
$$V_{ij} = \mathrm{Var}\left(\eta_{ij}\left(X_i, X_j\right)\right) = \mathrm{Var}\left(\mathbb{E}\left[Y \mid X_i, X_j\right]\right) - V_i - V_j \ ,$$
$$V_{ijk} = \mathrm{Var}\left(\eta_{ijk}\left(X_i, X_j, X_k\right)\right) = \mathrm{Var}\left(\mathbb{E}\left[Y \mid X_i, X_j, X_k\right]\right) - V_{ij} - V_{ik} - V_{jk} - V_i - V_j - V_k \ ,$$
$$\vdots$$

$$V_{1,\dots,d} = \mathrm{Var}\left(\eta_{1,\dots,d}(\mathbf{X})\right) = \mathrm{Var}\left(Y\right) - \sum_{i=1}^{d} V_i - \sum_{1 \leqslant i < j \leqslant d} V_{ij} - \dots - \sum_{1 \leqslant i_1 < \dots < i_{d-1} \leqslant d} V_{i_1 \dots i_{d-1}} \ .$$

The so-called Sobol indices given in Sobol (1993) are derived from (2.2):

$$S_i = \frac{V_i}{\mathrm{Var}(Y)}, \qquad S_{ij} = \frac{V_{ij}}{\mathrm{Var}(Y)}, \qquad \dots \tag{2.3}$$

where the first-order index S_i (also called main effect) measures the part of variance of the model output that stems from the variability in X_i, the second-order index S_{ij} measures the part of variance of the model output due to the interaction between X_i and X_j and so on for higher interaction orders.

Using the S_i, S_{ij} and higher order indices given above, one can build a picture of the importance of each variable in the output variance. However, when the number of variables is large, this requires the evaluation of $2^d - 1$ indices, which can be too computationally demanding

2

and whose interpretation becomes difficult. For this reason, another popular variance based coefficient called Total-order index by Homma and Saltelli (1996) is used. It measures the contribution to the output variance of X_i, including its main effect as well as all its interaction effects, of any order, with any other input variables. This index is defined by

$$
\begin{aligned}
ST_i &= S_i + \sum_{1 \leqslant i < j \leqslant d} S_{ij} + \cdots + S_{1,\ldots,d} \\
&= 1 - \frac{\operatorname{Var}_{\mathbf{X}_{-i}}\left(\mathbb{E}_{X_i}\left[Y \mid \mathbf{X}_{-i}\right]\right)}{\operatorname{Var}(Y)} = \frac{\mathbb{E}_{\mathbf{X}_{-i}}\left[\operatorname{Var}_{X_i}\left(Y \mid \mathbf{X}_{-i}\right)\right]}{\operatorname{Var}(Y)} \;,
\end{aligned}
\tag{2.4}
$$

where the notation $\mathbf{X}_{-i}$ indicates the set of all variables except X_i. Note that the following property can easily be deduced :

$$
0 \leqslant S_i \leqslant ST_i \leqslant 1 \;.
$$

Hence, the closer the index ST_i is to 1, the more influential the variable is. It should be noted that the case of equality $S_i = ST_i$ occurs if we have a purely additive model.

Sobol indices are well-defined for independent inputs. Indeed, the functional decomposition of the output variance given in (2.2) is unique in this context. This makes possible to clearly identify the contribution of each input or group of inputs to the variance output.

However, in many applications, it is common for inputs to have a statistical dependence structure imposed by a probabilistic dependence structure. In such a case, the uniqueness of the functional ANOVA decomposition is no longer guaranteed. The classical Sobol indices can still be calculated but their interpretation becomes difficult. Indeed, as mentioned in Song et al. (2016), the sum of first-order effects may exceed the total variance of the output or the sum of the total effects may be lower than the total variance of the output.

Several works have been carried out to overcome this limitation and extend Sobol indices to the case of stochastic dependence such as Caniou (2012); Kucherenko et al. (2012); Mara and Tarantola (2012); Mara et al. (2015). However, none of these works has given an univocal definition of the functional ANOVA decomposition for dependent inputs as the one provided by Sobol (1993) when inputs are independent. A new variable importance measure has been defined in Chastaing et al. (2012) through a generalization of ANOVA when inputs are dependent (Stone, 1994). But this measure comes with two conceptual problems: it requires some restrictive conditions on the joint probability distribution of the inputs as underlined in Owen and Prieur (2017) and its interpretation remains difficult because it can be negative.

New indices called Shapley effects have been proposed in Owen (2014). They present good properties in the presence of dependence: they are non negative, they sum to the total output variance and they are easy to interpret as highlighted by Song et al. (2016); Iooss and Prieur (2019).

2.2 Shapley effects

Shapley values have been introduced in game theory by Shapley (1953). The motivation, in the context of cooperative game theory, was to define an attribution method to allocate fairly the value created by a team effort to its individual members. Turning now to variance-based sensitivity analysis, it appears that the idea of assigning a portion of the output variance to each input variable has some similarities. These were highlighted and brought to the SA community, in the context of variance-based sensitivity analysis, by Owen (2014).

Formally, in Song et al. (2016) a d-player game with the set of players $\mathcal{D} = \{1, 2, \ldots, d\}$ is defined as a real-valued function that maps a subset of $\mathcal{D}$ to its corresponding cost, i.e.,

$c : 2^{\mathcal{D}} \mapsto \mathbb{R}$ with $c(\emptyset) = 0$. Hence, $c(\mathcal{J})$ represents the cost that arises when the players in the subset $\mathcal{J}$ of $\mathcal{D}$ participate in the game. Let $v^i = v^i(c)$, $i = 1, \ldots, d$, be the Shapley value for each player that will be defined below. According to Winter et al. (2002), an attribution method should have the next four compelling properties:

- **Efficiency:** $\sum\limits_{i=1}^{d} v^i = c(\mathcal{D})$. The sum of the values attributed to the players must be equal to what the coalition of all the players can obtain.

- **Symmetry:** if $c(\mathcal{J} \cup i) = c(\mathcal{J} \cup j)$ for all $\mathcal{J} \subseteq \mathcal{D} \setminus \{i, j\}$, then $v^i = v^j$. The contribution of two players should be the same if they contribute equally to all possible coalitions.

- **Dummy:** if $c(\mathcal{J} \cup i) = c(\mathcal{J})$ for all $\mathcal{J} \subseteq \mathcal{D} \setminus \{i\}$, then $v^i = 0$. A player who does not change the predicted value, no matter to which coalition of players it is added, should have a contribution value of 0.

- **Additivity:** if the i-th player has a contribution v^i (resp. v'^i) in the coalitional game described by the gain function c (resp. c'). Then, the contribution of the i-th player in the new coalitional game described by the gain function $c + c'$ is $v^i + v'^i$ for $i \in \mathcal{D}$.

Shapley (1953) showed that the unique valuation v^i that satisfies these properties is given by

$$v^i = \sum_{\mathcal{J} \subseteq \mathcal{D} \setminus \{i\}} \frac{(d - |\mathcal{J}| - 1)! |\mathcal{J}|!}{d!} \left(c\left(\mathcal{J} \cup \{i\} \right) - c\left(\mathcal{J} \right) \right) , \tag{2.5}$$

defined for the player i with respect to the cost function $c(\cdot)$ and where $|\mathcal{J}|$ indicates the size of $\mathcal{J}$. In other words, v^i is the incremental cost of including player i in the set $\mathcal{J}$ averaged over all sets $\mathcal{J} \subseteq \mathcal{D} \setminus \{i\}$.

In the framework of global sensitivity analysis, we may consider the set of inputs of $\eta(\cdot)$ as the set of players $\mathcal{D}$. We then need to define a $c(\cdot)$ cost function such that for $\mathcal{J} \subseteq \mathcal{D}$, $c(\mathcal{J})$ measures the part of variance of Y caused by the uncertainty of the inputs in $\mathcal{J}$. To this aim, we want a cost function that verifies $c(\emptyset) = 0$ and $c(\mathcal{D}) = \mathrm{Var}(Y)$.

Owen (2014) proposed the cost function $\tilde{c}(\mathcal{J}) = \mathrm{Var}\left(\mathbb{E}\left[Y | \mathbf{X}_{\mathcal{J}} \right] \right)$ for considering the Shapley value in the framework of variance-based sensitivity indices. Song et al. (2016) showed in their Theorem 1 that the Shapley values defined using cost functions $\tilde{c}(\mathcal{J})$ and $c(\mathcal{J}) = \mathbb{E}\left[\mathrm{Var}\left(Y | \mathbf{X}_{-\mathcal{J}} \right) \right]$ are equal. They used the term *Shapley effects* to describe variance based Shapley values as new importance measures in SA. Note that in Owen (2014); Song et al. (2016), the cost function is not normalized by the variance of Y, whereas, in this paper, we consider its normalized version to quantify relative importance of each input with respect to the output variance. We denote hereafter the *Shapley effect* by Sh^i and a generic *Shapley value* by v^i.
The purpose of Sobol indices defined in Subsection 2.1 is to decompose $\mathrm{Var}(Y)$ and allocate it to *each subset* $\mathcal{J} \subseteq \mathcal{D}$ whereas Shapley effects decompose $\mathrm{Var}(Y)$ and allocate it to *each input* X_i. This difference allows to consider any variables regardless of their dependence with other inputs. Shapley effects rely on an equitable allocation of part of the output variance to each input. In case of dependent inputs, several test-cases where the Shapley effects can be analytically computed have been investigated in Owen and Prieur (2017); Iooss and Prieur (2019); Benoumechiara and Elie-Dit-Cosaque (2019) in order to understand the effect of the dependence between inputs on the variance based Shapley values. These studies highlighted several properties of these indices. Benoumechiara and Elie-Dit-Cosaque (2019) have also compared the Shapley effects with the strategy proposed in Mara et al. (2015) based on the estimation of four sensitivity indices per input. In this last case, the Shapley effects can be a good alternative to

the existing extensions of classical Sobol indices. Indeed, Shapley effects allow an apportionment of the interaction and dependence contributions between the input involved, making them condensed and easy-to-interpret indices. At last, we can mention that Rabitti and Borgonovo (2019) have recently introduced Shapley-Owen interaction effects which are a generalization of Shapley effects in order to study the synergistic/antagonistic nature of interactions among inputs.

3 QOSA indices

In the previous section, we have reviewed variance-based measures. Even if these indices are extremely popular and informative importance measures, they only study the impact of the input variables around the expectation of the output distribution as they use the variance as a distance measure. However, as highlighted by Borgonovo (2006), in some cases, this one poorly represents the variability/uncertainty of the output distribution. Different approaches have been developed to overcome this issue including, for example, moment independent importance measures proposed by Borgonovo (2007); Borgonovo et al. (2011) that quantify the influence of an input over the whole distribution of the output.

Another approach presented in Fort et al. (2016) is to define indices that quantify the impact of inputs $\mathbf{X} = (X_1, \ldots, X_d)$ on a feature of interest of the output distribution depending on the problem (mean, quantiles and so on). They refer to this method as *Goal-Oriented Sensitivity Analysis* (GOSA). These indices rely on contrast functions and are members of a wider family containing sensitivity indices based on dissimilarity measures (Da Veiga, 2015). We shall consider the specific case of *Quantile Oriented Sensitivity Analysis* (QOSA).

3.1 First-order QOSA indices

Let us focus on QOSA indices measuring the impact of the inputs over the α-quantile of the output distribution. Given a level of quantile $\alpha \in]0, 1[$, let us recall the expression of the first-order QOSA index:

$$
\begin{aligned}
S_i^\alpha &= \frac{\min\limits_{\theta \in \mathbb{R}} \mathbb{E}\left[\psi_\alpha\left(Y, \theta\right)\right] - \mathbb{E}\left[\min\limits_{\theta \in \mathbb{R}} \mathbb{E}\left[\psi_\alpha\left(Y, \theta\right)\middle| X_i\right]\right]}{\min\limits_{\theta \in \mathbb{R}} \mathbb{E}\left[\psi_\alpha\left(Y, \theta\right)\right]} \\
&= \frac{\mathbb{E}\left[\psi_\alpha\left(Y, q^\alpha\left(Y\right)\right)\right] - \mathbb{E}\left[\psi_\alpha\left(Y, q^\alpha\left(Y\middle| X_i\right)\right)\right]}{\mathbb{E}\left[\psi_\alpha\left(Y, q^\alpha\left(Y\right)\right)\right]} \,,
\end{aligned}
\tag{3.1}
$$

where $\psi_\alpha : (y, \theta) \mapsto (y - \theta)\left(\alpha - \mathbb{1}_{\{y \leqslant \theta\}}\right)$ is the contrast function associated to the α-quantile and the q's are the quantiles

$$
q^\alpha\left(Y\right) = \underset{\theta \in \mathbb{R}}{\arg\min}\, \mathbb{E}\left[\psi_\alpha\left(Y, \theta\right)\right] \quad \text{and} \quad q^\alpha\left(Y\middle| X_i = x_i\right) = \underset{\theta \in \mathbb{R}}{\arg\min}\, \mathbb{E}\left[\psi_\alpha\left(Y, \theta\right)\middle| X_i = x_i\right] \,.
$$

Remark that replacing ψ_α in the above equation by $(y, \theta) \mapsto (y - \theta)^2$ leads to the definition of first-order Sobol indices. In order to interpret QOSA indices, one has to consider $\psi_\alpha(Y, \theta)$ as a dispersion measure of Y which is minimized for $\theta = q^\alpha(Y)$. So that QOSA indices compare the dispersion of Y around its quantile with its conditional counterpart.

The first-order QOSA indices have been defined in Fort et al. (2016), studied and estimated in Browne et al. (2017); Maume-Deschamps and Niang (2018); Elie-Dit-Cosaque and Maume-Deschamps (2021). They may be rewritten as follows

$$
S_i^\alpha = \frac{\mathbb{E}\left[Y\mathbb{1}_{\{Y \leqslant q^\alpha(Y|X_i)\}}\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right]}{\alpha\mathbb{E}\left[Y\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right]} = 1 - \frac{\alpha\mathbb{E}\left[Y\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y \leqslant q^\alpha(Y|X_i)\}}\right]}{\alpha\mathbb{E}\left[Y\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right]} \,.
$$

Let us mention that Kucherenko et al. (2019) have recently proposed new indices to assess the impact of inputs over the α-quantile of the output distribution. Instead of considering the expression of the first-order Sobol index based on a contrast function as in (3.1), they consider the expression of Sobol' indices with numerator $\mathrm{Var}\left(\mathbb{E}\left[Y\,|\,X_i\right]\right) = \mathbb{E}\left[\left(\mathbb{E}\left[Y\,|\,X_i\right] - \mathbb{E}\left[Y\right]\right)^2\right]$ and simply replace the expectations by α-quantiles to define the following indices

$$\bar{q}_{i,1}^{\alpha} = \mathbb{E}\left[|q^{\alpha}\left(Y\right) - q^{\alpha}\left(Y\,|\,X_i\right)|\right] \quad\text{and}\quad \bar{q}_{i,2}^{\alpha} = \mathbb{E}\left[\left(q^{\alpha}\left(Y\right) - q^{\alpha}\left(Y\,|\,X_i\right)\right)^2\right] .$$

They also provide the normalized versions as follows

$$Q_{i,1}^{\alpha} = \frac{\bar{q}_{i,1}^{\alpha}}{\sum\limits_{j=1}^{d} \bar{q}_{j,1}^{\alpha}} \quad\text{and}\quad Q_{i,2}^{\alpha} = \frac{\bar{q}_{i,2}^{\alpha}}{\sum\limits_{j=1}^{d} \bar{q}_{j,2}^{\alpha}} .$$

These measures thereby quantify the mean distance between quantiles $q^{\alpha}\left(Y\right)$ and $q^{\alpha}\left(Y\,|\,X_i\right)$ rather than the mean distance between average contrast functions like in the first-order QOSA index given in (3.1). The indices $\bar{q}_{i,1}^{\alpha}$ and $Q_{i,1}^{\alpha}$ will be called below *absolute value indices*; $\bar{q}_{i,2}^{\alpha}$ and $Q_{i,2}^{\alpha}$ will be called below *squared indices*. The example below shows that their interpretation as sensitivity indices is questionnable, it is why we shall focus on QOSA indices.

Consider the simple model also studied in Fort et al. (2016): $Y = X_1 - X_2$ where X_1 and X_2 are two independent exponential random variables with expectation 1. So that Y follows a Laplace distribution. QOSA indices have closed form formulas that may be found in Fort et al. (2016). The indices proposed in Kucherenko et al. (2019) may also be computed. Indeed, with

$$\gamma_1 = \begin{cases} -\log\left(2\alpha\left(1-\alpha\right)\right) \text{ if } \alpha \geq \frac{1}{2} \\ \log(2) \text{ if } \alpha < \frac{1}{2} \end{cases},$$

and

$$\gamma_2 = \begin{cases} \log(2) \text{ if } \alpha \geq \frac{1}{2} \\ -\log\left(2\alpha\left(1-\alpha\right)\right) \text{ if } \alpha < \frac{1}{2} \end{cases},$$

we have

$$\bar{q}_{1,1}^{\alpha} = \gamma_1 + 2e^{-\gamma_1} - 1 \quad\text{and}\quad \bar{q}_{2,1}^{\alpha} = \gamma_2 + 2e^{-\gamma_2} - 1 ;$$
$$\bar{q}_{1,2}^{\alpha} = \gamma_1^2 - 2\gamma_1 + 2 \quad\text{and}\quad \bar{q}_{2,2}^{\alpha} = \gamma_2^2 - 2\gamma_2 + 2 .$$

Below we show the behaviour of QOSA, $\bar{q}_{i,1}^{\alpha}$ and $\bar{q}_{i,2}^{\alpha}$ indices. As expected, QOSA indices show that X_1 has more influence on quantiles of level higher that $\frac{1}{2}$ and X_2 is more influent for α lower than $\frac{1}{2}$. The interpretation of $\bar{q}_{i,j}^{\alpha}$ is not so clear since $\bar{q}_{1,j}^{\alpha}$ are constant for α less than $\frac{1}{2}$ and $\bar{q}_{2,j}^{\alpha}$ are constant for α greater than $\frac{1}{2}$. Moreover, the $\bar{q}_{i,2}^{\alpha}$'s are not monotonic.
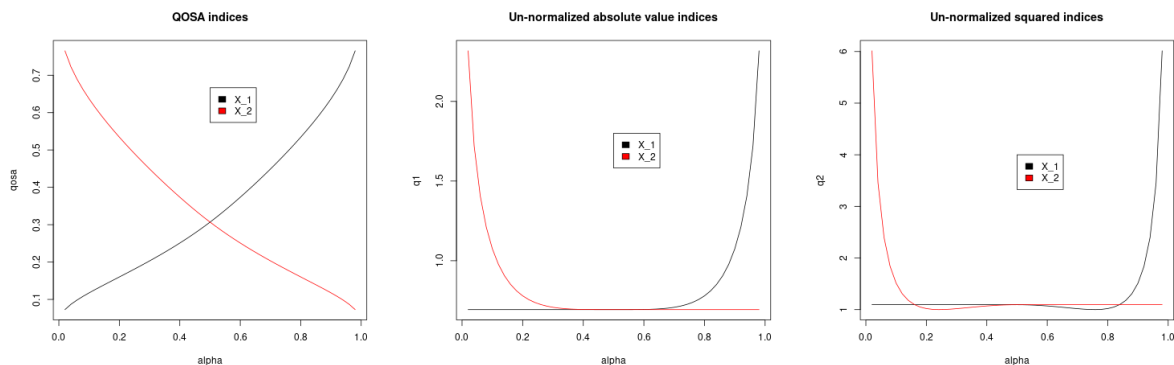
The normalized indices below show that the interpretation for $Q_{i,j}^\alpha$ remains questionnable while the normalized QOSA indices keep the interpretation of the un-normalized ones.
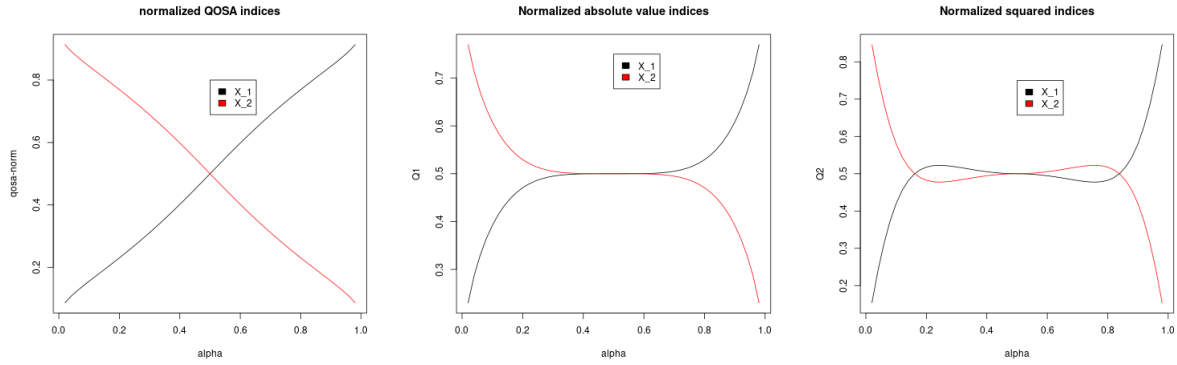
3.2 Some elementary properties of QOSA indices

Let us turn to some elementary properties of QOSA indices.

The following lemma is useful, it is closely related to the proof of sub-additivity of TVaR in risk theory (see Marceau (2013) e.g.).

Lemma 3.1.
Consider any event E such that $\mathbb{P}(E) = \alpha$. Then, for any random variable X, we have

$$\mathbb{E}\left[X\mathbb{1}_{\{X \leqslant q^\alpha(X)\}}\right] \leqslant \mathbb{E}\left[X\mathbb{1}_E\right] \ ,$$

with $q^\alpha(X)$ the α-quantile of X.

Proof of Lemma 3.1.

We have:

$$
\begin{aligned}
\mathbb{E}\left[X\mathbb{1}_{\{X \leqslant q^\alpha(X)\}}\right] - \mathbb{E}\left[X\mathbb{1}_E\right] &= \mathbb{E}\left[X\left(\mathbb{1}_{\{X \leqslant q^\alpha(X)\}} - \mathbb{1}_E\right)\right] \\
&= \mathbb{E}\left[(X - q^\alpha(X))\left(\mathbb{1}_{\{X \leqslant q^\alpha(X)\}} - \mathbb{1}_E\right)\right] \leqslant 0 \ .
\end{aligned}
$$

■

As a consequence, since $\mathbb{P}\left(Y \leqslant q^\alpha\left(Y \mid X_i\right) \mid X_i\right) = \alpha$ and $\mathbb{P}\left(Y \leqslant q^\alpha\left(Y\right)\right) = \alpha$, we get that

$$\alpha\mathbb{E}\left[Y\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y \leqslant q^\alpha(Y|X_i)\}}\right] = \mathbb{E}\left[\left(Y - q^\alpha\left(Y \mid X_i\right)\right)\left(\alpha - \mathbb{1}_{\{Y \leqslant q^\alpha(Y|X_i)\}}\right)\right] \geqslant 0 \tag{3.2}$$

$$\alpha\mathbb{E}\left[Y\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right] = \mathbb{E}\left[\left(Y - q^\alpha\left(Y\right)\right)\left(\alpha - \mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right)\right] \geqslant 0 \tag{3.3}$$

$$\mathbb{E}\left[Y\mathbb{1}_{\{Y \leqslant q^\alpha(Y|X_i)\}}\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right] \geqslant 0 \ . \tag{3.4}$$

This implies that $0 \leqslant S_i^\alpha \leqslant 1$ as already noticed in Fort et al. (2016). The index S_i^α also has the three following interesting properties.

Proposition 3.2.

1. *S_i^α is invariant with respect to translations of the output Y.*

2. *S_i^α is invariant by homothety with strictly positive ratio of the output Y.*

3. *A homothety with strictly negative ratio of the output Y gives the index $S_i^{1-\alpha}$ associated to the $1-\alpha$ level.*

Proof of Proposition 3.2.

Let us consider any model $Y = \eta(\mathbf{X})$. We will denote by $S_i'^\alpha$ the QOSA indices related to a r.v. Y'.

1. Let $Y' = Y + k$, $k \in \mathbb{R}$. Then, we have $q^\alpha(Y') = q^\alpha(Y) + k$ and $q^\alpha(Y'|X_i) = q^\alpha(Y|X_i) + k$. It is easy to check that $S_i'^\alpha = S_i^\alpha$.

2. Let $Y' = k \times Y$, $k > 0$. Then we have, $q^\alpha(Y') = k \times q^\alpha(Y)$ and $q^\alpha(Y'|X_i) = k \times q^\alpha(Y|X_i)$. We can easily show that $S_i'^\alpha = S_i^\alpha$.

3. Let $Y' = k \times Y$, $k < 0$. Then we have, $q^\alpha(Y') = k \times q^{1-\alpha}(Y)$ and $q^\alpha(Y'|X_i) = k \times q^{1-\alpha}(Y|X_i)$. It leads to $S_i'^\alpha = S_i^{1-\alpha}$.

$\blacksquare$

Now, we are going to investigate the sum $\mathcal{S}$ of the first-order QOSA indices:

$$\mathcal{S} = \sum_{i=1}^{d} S_i^\alpha = \frac{\sum_{i=1}^{d} \mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y|X_i)\}}\right] - d\mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right]}{\alpha \mathbb{E}[Y] - \mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right]} .$$

We see that $\mathcal{S} \leqslant 1$ if and only if

$$\sum_{i=1}^{d} \mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y|X_i)\}}\right] - d\mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right] \leqslant \left(\alpha \mathbb{E}[Y] - \mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right]\right) . \tag{3.5}$$

Or equivalently:

$$\alpha \mathbb{E}[Y] + (d-1)\mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right] - \sum_{i=1}^{d} \mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y|X_i)\}}\right] \geqslant 0 .$$

As proved in the following proposition, $\mathcal{S}$ is smaller than 1 in the case of an additive model with independent inputs. Unfortunately, this result is not true in the general case as showed with a counterexample in Subsection 4.1.

Proposition 3.3.

Let $\mathbf{X} = (X_1, \ldots, X_d)$ with independent X_i's. Let $Y = m_0 + \sum_{i=1}^{d} m_i(X_i)$ be an additive model. Then, the sum of the first-order QOSA indices $\mathcal{S}$ satisfies $\mathcal{S} \leqslant 1$.

Proof of Proposition 3.3.

Given a random variable X, we denote by $q^\alpha(X)$ its α-quantile. For any $i = 1, \ldots, d$, let $Xs_{(-i)} = \sum_{\substack{1 \leqslant j \leqslant d \\ j \neq i}} m_j(X_j)$, thanks to the independence of the X_i's, we have

$$q^\alpha(Y|X_i) = m_0 + m_i(X_i) + q^\alpha\left(Xs_{(-i)}\right), \text{ and } \{Y \leqslant q^\alpha(Y|X_i)\} = \{Xs_{(-i)} \leqslant q^\alpha\left(Xs_{(-i)}\right)\} .$$

We have

$$g(\alpha) := \alpha \mathbb{E}[Y] + (d-1)\mathbb{E}\left[Y\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right] - \sum_{i=1}^{d} \mathbb{E}\left[Y\mathbb{1}_{\{Y \leqslant q^\alpha(Y|X_i)\}}\right]$$

$$= \alpha \mathbb{E}[Y] + (d-1)\mathbb{E}\left[Y\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right] - \sum_{i=1}^{d}\left(\alpha \times m_0 + \mathbb{E}\left[m_i(X_i)\mathbb{1}_{\{Xs_{-(i)} \leqslant q^\alpha(Xs_{(-i)})\}}\right]\right.$$

$$\left. + \mathbb{E}\left[Xs_{(-i)}\mathbb{1}_{\{Xs_{-(i)} \leqslant q^\alpha(Xs_{(-i)})\}}\right]\right) .$$

The independence of the X_i's implies that $\mathbb{E}\left[m_i(X_i)\mathbb{1}_{\{Xs_{(-i)} \leqslant q^\alpha(Xs_{(-i)})\}}\right] = \alpha \mathbb{E}[m_i(X_i)]$ and thus,

$$g(\alpha) = (d-1)\mathbb{E}\left[\left(\sum_{j=1}^{d} m_j(X_j)\right)\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right] - \sum_{i=1}^{d} \mathbb{E}\left[Xs_{(-i)}\mathbb{1}_{\{Xs_{(-i)} \leqslant q^\alpha(Xs_{(-i)})\}}\right] .$$

Now, we use Lemma 3.1 which gives:

$$(d-1)\mathbb{E}\left[\left(\sum_{j=1}^{d} m_j(X_j)\right)\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right] = \sum_{i=1}^{d-1}\left(\mathbb{E}\left[m_i(X_i)\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right] + \mathbb{E}\left[Xs_{(-i)}\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right]\right)$$

$$\geqslant \sum_{i=1}^{d-1}\left(\mathbb{E}\left[m_i(X_i)\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right]\right.$$

$$\left. + \mathbb{E}\left[Xs_{(-i)}\mathbb{1}_{\{Xs_{(-i)} \leqslant q^\alpha(Xs_{(-i)})\}}\right]\right) .$$

As a consequence,

$$g(\alpha) \geqslant \sum_{i=1}^{d-1} \mathbb{E}\left[m_i(X_i)\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right] - \mathbb{E}\left[Xs_{(-d)}\mathbb{1}_{\{Xs_{(-d)} \leqslant q^\alpha(Xs_{(-d)})\}}\right]$$

$$= \mathbb{E}\left[Xs_{(-d)}\mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right] - \mathbb{E}\left[Xs_{(-d)}\mathbb{1}_{\{Xs_{(-d)} \leqslant q^\alpha(Xs_{(-d)})\}}\right]$$

$$\geqslant 0 \text{ using again Lemma 3.1.}$$

∎

First-order QOSA indices capture only the main effect of the i-th input. Following Kala (2019), we consider below higher order QOSA indices.

3.3 Higher order QOSA indices

Kala (2019) introduced second-order QOSA indices in order to assess the impact of the interaction effect of two inputs on the α-quantile. He proposes to measure the joint effect of the pair (X_i, X_j) by:

$$S_{ij}^\alpha = \frac{\min_{\theta \in \mathbb{R}} \mathbb{E}\left[\psi_\alpha(Y, \theta)\right] - \mathbb{E}\left[\min_{\theta \in \mathbb{R}} \mathbb{E}\left[\psi_\alpha(Y, \theta)\mid X_i, X_j\right]\right]}{\min_{\theta \in \mathbb{R}} \mathbb{E}\left[\psi_\alpha(Y, \theta)\right]} - S_i^\alpha - S_j^\alpha .$$

Higher-order QOSA indices can be expressed analogously. Hence, one obtain a variance-like decomposition for quantiles in the case of independent inputs:

$$\sum_{i=1}^{d} S_i^\alpha + \sum_{1 \leqslant i < j \leqslant d} S_{ij}^\alpha + \cdots + S_{1,\ldots,d}^\alpha = 1 . \tag{3.6}$$

9

Although this decomposition appears similar to that of Sobol indices, it is quite different since it does not stem from an univocal decomposition.

The indices $S_i^\alpha, S_{ij}^\alpha$ and higher order allow to assess thoroughly the impact of each input over the α-quantile of the output distribution. However, as for Sobol indices, in the case of a large number of inputs, it would require the evaluation of $2^d - 1$ indices, which could be computationally demanding. Therefore, it is suitable to introduce the so-called total QOSA index as suggested by Kala (2019) that measures the contribution of an input, including its main effect as well as its interactions effects, of any order, with other input variables:

$$ST_i^\alpha = \frac{\mathbb{E}\left[\min_{\theta \in \mathbb{R}} \mathbb{E}\left[\psi_\alpha\left(Y, \theta\right) | \mathbf{X}_{-i}\right]\right]}{\min_{\theta \in \mathbb{R}} \mathbb{E}\left[\psi_\alpha\left(Y, \theta\right)\right]} = \frac{\mathbb{E}\left[\psi_\alpha\left(Y, q^\alpha\left(Y | \mathbf{X}_{-i}\right)\right)\right]}{\mathbb{E}\left[\psi_\alpha\left(Y, q^\alpha\left(Y\right)\right)\right]} \ .$$

The total QOSA index may be rewritten as follows

$$ST_i^\alpha = \frac{\alpha \mathbb{E}\left[Y\right] - \mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y | \mathbf{X}_{-i})\}}\right]}{\alpha \mathbb{E}\left[Y\right] - \mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right]} \ .$$

As for the first-order QOSA index, the total one has the three following interesting properties.

Proposition 3.4.

1. ST_i^α is invariant with respect to translations of the output Y.

2. ST_i^α is invariant by homothety with strictly positive ratio of the output Y.

3. A homothety with strictly negative ratio of the output Y gives the index $ST_i^{1-\alpha}$ associated to the $1-\alpha$ level.

Proof of Proposition 3.4.

Just adapting the steps of the proof of the Proposition 3.2 for the total QOSA index. ∎

Proposition 3.5 below shows that the total QOSA index is greater than or equal to the first-order one for any α-level in the case of an additive model with $\mathbf{X}$ that has independent marginals. This is a major difference from variance-based methods, specifically Sobol indices. Indeed, it is well-known that for a purely additive model with independent inputs, we have for the Sobol indices $ST_i = S_i$, $\forall i \in \mathcal{D}$, which is not the case for the QOSA indices. It therefore appears that the total QOSA index captures some interaction between the inputs when using an additive model. The origin of this phenomenom is not yet understood at this stage and requires further analysis.
Besides, it should be noted that Proposition 3.5 is not verified in the general non additive case as emphasized with a counterexample in Subsection 4.3.

Proposition 3.5.

Let $\mathbf{X} = (X_1, \ldots, X_d)$ with the X_i's independent. Let $Y = m_0 + \sum_{i=1}^{d} m_i\left(X_i\right)$ be an additive model with m_i, $i = 1, \ldots, d$, the one-dimensional nonparametric functions operating on each element of the vector $\mathbf{X}$. Then,

$$\forall \ \alpha \in \ ]0, 1[\ , \quad S_i^\alpha \leqslant ST_i^\alpha \ .$$

10

Proof of Proposition 3.5.

We have:

$$ST_i^\alpha - S_i^\alpha = \frac{\left(\alpha\mathbb{E}\left[Y\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y\leqslant q^\alpha(Y|\mathbf{X}_{-i})\}}\right]\right) - \left(\mathbb{E}\left[Y\mathbb{1}_{\{Y\leqslant q^\alpha(Y|X_i)\}}\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y\leqslant q^\alpha(Y)\}}\right]\right)}{\alpha\mathbb{E}\left[Y\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y\leqslant q^\alpha(Y)\}}\right]} \; .$$

As the denominator is non negative according to Equation (3.3), we just have to show that the numerator is also non negative.

Let

$$g(\alpha) = \alpha\mathbb{E}\left[Y\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y\leqslant q^\alpha(Y|\mathbf{X}_{-i})\}}\right] + \mathbb{E}\left[Y\mathbb{1}_{\{Y\leqslant q^\alpha(Y)\}}\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y\leqslant q^\alpha(Y|X_i)\}}\right] \; ,$$

For any $i = 1, \ldots, d$, let $Xs_{(-i)} = \sum\limits_{\substack{1\leqslant j\leqslant d \\ j\neq i}} m_j(X_j)$, thanks to the independence, we have

$$q^\alpha\left(Y|X_i\right) = m_0 + m_i\left(X_i\right) + q^\alpha\left(Xs_{(-i)}\right), \text{ and } \{Y\leqslant q^\alpha\left(Y|X_i\right)\} = \{Xs_{(-i)}\leqslant q^\alpha\left(Xs_{(-i)}\right)\} \; ,$$

$$q^\alpha\left(Y|\mathbf{X}_{-i}\right) = m_0 + Xs_{(-i)} + q^\alpha\left(m_i\left(X_i\right)\right), \text{ and } \{Y\leqslant q^\alpha\left(Y|\mathbf{X}_{-i}\right)\} = \{m_i\left(X_i\right)\leqslant q^\alpha\left(m_i\left(X_i\right)\right)\} \; .$$

Then,

$$\begin{aligned}
g(\alpha) &= \alpha\mathbb{E}\left[Y\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y\leqslant q^\alpha(Y|\mathbf{X}_{-i})\}}\right] + \mathbb{E}\left[Y\mathbb{1}_{\{Y\leqslant q^\alpha(Y)\}}\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y\leqslant q^\alpha(Y|X_i)\}}\right] \\
&= \alpha\mathbb{E}\left[Y\right] - \alpha m_0 - \mathbb{E}\left[m_i\left(X_i\right)\mathbb{1}_{\{m_i(X_i)\leqslant q^\alpha(m_i(X_i))\}}\right] - \mathbb{E}\left[Xs_{(-i)}\mathbb{1}_{\{m_i(X_i)\leqslant q^\alpha(m_i(X_i))\}}\right] \\
&\quad + \mathbb{E}\left[Y\mathbb{1}_{\{Y\leqslant q^\alpha(Y)\}}\right] - \alpha m_0 - \mathbb{E}\left[Xs_{(-i)}\mathbb{1}_{\{Xs_{(-i)}\leqslant q^\alpha(Xs_{(-i)})\}}\right] \\
&\quad - \mathbb{E}\left[m_i\left(X_i\right)\mathbb{1}_{\{Xs_{(-i)}\leqslant q^\alpha(Xs_{(-i)})\}}\right] \; .
\end{aligned}$$

Now, the independence of the X_i's implies that $\mathbb{E}\left[Xs_{(-i)}\mathbb{1}_{\{m_i(X_i)\leqslant q^\alpha(m_i(X_i))\}}\right] = \alpha\mathbb{E}\left[Xs_{(-i)}\right]$ and $\mathbb{E}\left[m_i(X_i)\mathbb{1}_{\{Xs_{(-i)}\leqslant q^\alpha(Xs_{(-i)})\}}\right] = \alpha\mathbb{E}\left[m_i(X_i)\right]$. Thus,

$$\begin{aligned}
g(\alpha) &= \mathbb{E}\left[\left(\sum_{j=1}^{d} m_j\left(X_j\right)\right)\mathbb{1}_{\{Y\leqslant q^\alpha(Y)\}}\right] - \mathbb{E}\left[m_i\left(X_i\right)\mathbb{1}_{\{m_i(X_i)\leqslant q^\alpha(m_i(X_i))\}}\right] \\
&\quad - \mathbb{E}\left[Xs_{(-i)}\mathbb{1}_{\{Xs_{(-i)}\leqslant q^\alpha(Xs_{(-i)})\}}\right] \\
&= \left(\mathbb{E}\left[m_i\left(X_i\right)\mathbb{1}_{\{Y\leqslant q^\alpha(Y)\}}\right] - \mathbb{E}\left[m_i\left(X_i\right)\mathbb{1}_{\{m_i(X_i)\leqslant q^\alpha(m_i(X_i))\}}\right]\right) \\
&\quad + \left(\mathbb{E}\left[Xs_{(-i)}\mathbb{1}_{\{Y\leqslant q^\alpha(Y)\}}\right] - \mathbb{E}\left[Xs_{(-i)}\mathbb{1}_{\{Xs_{(-i)}\leqslant q^\alpha(Xs_{(-i)})\}}\right]\right) \; .
\end{aligned}$$

Now, the two last terms are positive according to Lemma 3.1 which concludes the proof. ∎

Several calculations below will help to better understand the behaviour, the usefulness and the limitations of QOSA indices.

4   QOSA calculations on special cases

In this section, we compute for some distributions the first-order and total QOSA indices for which we obtain the expressions in a closed or nearly closed-form. In particular, the examples with Gaussian inputs allow to investigate the behavior of the indices when there is some statistical dependence between the inputs.

4.1 Product of two exponential input variables

Let $Y = X_1 \cdot X_2$, with $X_1 \sim \mathcal{E}(\lambda)$, $X_2 \sim \mathcal{E}(\delta)$, these two variables being independent. After simple calculations, we get the first-order QOSA indices

$$S_1^\alpha = S_2^\alpha = 1 - \frac{(\alpha - 1) \log (1 - \alpha)}{\alpha - \lambda \delta \cdot \mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right]} \ , \tag{4.1}$$

and the total QOSA indices

$$ST_1^\alpha = ST_2^\alpha = \frac{(\alpha - 1) \log (1 - \alpha)}{\alpha - \lambda \delta \cdot \mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right]} \ . \tag{4.2}$$

The term $\mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right]$ can be approximated by using a Monte-Carlo estimation or a numerical integration.

The equality of the first-order and total QOSA indices for both inputs is a particular case due to the exponential distribution. Indeed, let $X_2 \overset{\mathcal{L}}{=} \frac{\lambda}{\delta} X_1'$ with X_1' an independent copy of X_1. Then, the model writes

$$Y \overset{\mathcal{L}}{=} \frac{\lambda}{\delta} X_1 \cdot X_1'$$
$$\overset{\mathcal{L}}{=} kZ$$

with $Z = X_1 \cdot X_1'$ and $k = \frac{\lambda}{\delta} > 0$. As the inputs X_1 and X_1' have the same distribution, their impact over the α-quantile of the model output Z is identical. Therefore, by using item 2. of Propositions 3.2 and 3.4 (because Y is just a homothety with strictly positive ratio of Z), that explains why both first-order and total QOSA indices of the inputs are equal.

Figure 1 below presents the behavior of the indices as a function of the level α for the model computed with $\lambda = 1/10$ and $\delta = 1$. The truncated expectation is estimated with a Monte-Carlo algorithm and a sample of size $n = 10^9$. We observe that the first-order and total QOSA indices vary in opposite directions. The first-order QOSA indices go to 1 when α tends to 1 while the total ones go to 0 when α tends to 1. It is interesting to notice that from $\alpha \approx 0.96$ the total QOSA indices are lower than the first-order ones and the sum of the first-order ones is greater than 1. That corroborates that Propositions 3.2 and 3.4 are not verified outside the additive model context.

4.2 Linear model with Gaussian input variables

We study in this subsection a linear model with Gaussian inputs which implies that the resulting output is also Gaussian. This framework facilitates calculations to obtain the analytical values given below.

Proposition 4.1.
Let $Y = \eta(\mathbf{X}) = \beta_0 + \boldsymbol{\beta}^\mathsf{T} \mathbf{X}$ with $\beta_0 \in \mathbb{R}$, $\boldsymbol{\beta} \in \mathbb{R}^d$ and $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$ is a positive-definite matrix, then the first-order and total QOSA indices for the variable i at the α-level are

$$S_i^\alpha = 1 - \frac{\sqrt{\boldsymbol{\beta}_{-i}^\mathsf{T} \left(\boldsymbol{\Sigma}_{-i,-i} - \boldsymbol{\Sigma}_{-i,i} \boldsymbol{\Sigma}_{i,i}^{-1} \boldsymbol{\Sigma}_{i,-i} \right) \boldsymbol{\beta}_{-i}}}{\sigma_Y} \ , \tag{4.3}$$

$$ST_i^\alpha = \frac{|\boldsymbol{\beta}_i| \sqrt{\boldsymbol{\Sigma}_{i,i} - \boldsymbol{\Sigma}_{i,-i} \boldsymbol{\Sigma}_{-i,-i}^{-1} \boldsymbol{\Sigma}_{-i,i}}}{\sigma_Y} \ , \tag{4.4}$$
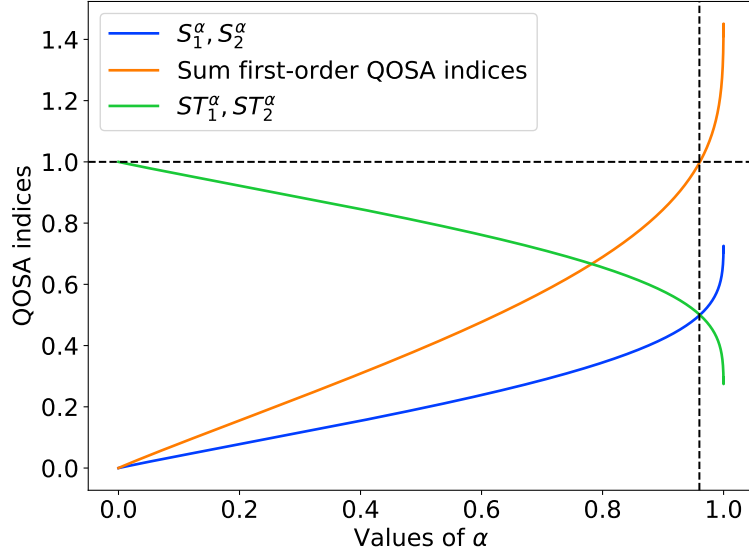
Figure 1: Evolution of the first-order and total QOSA indices at different levels α for the product of two exponentials with $\lambda = 1/10$ for the fisrt input and $\delta = 1$ for the second one.

with $\sigma_Y^2 = \text{Var}(Y) = \boldsymbol{\beta}^{\mathsf{T}} \boldsymbol{\Sigma} \boldsymbol{\beta}$.

We observe that as β_0 and $\boldsymbol{\mu}$ are translation parameters, they do not have any influence. Nevertheless, no general conclusion can be drawn from Equations (4.3) and (4.4) except that the values of the first-order and total QOSA indices are the same for all levels α. This phenomenon is specific to the Gaussian linear model and will be detailed hereafter in dimension 2. Indeed, a look at the case $d = 2$ may help to understand why QOSA indices does not depend on α-level in the Gaussian framework. This feature will be analyzed by studying only the impact of the variable X_1.

We have that

$$Y|X_1 \sim \mathcal{N}\left(\beta_1 X_1 + \beta_2 \mathbb{E}\left[X_2| X_1\right],\ \beta_2^2 \text{Var}\left(X_2| X_1\right)\right) \ .$$

As we work with Gaussian distributions, the conditional variance $\text{Var}\left(X_2| X_1\right)$ does not depend on the specific value of X_1 and is $\text{Var}\left(X_2| X_1\right) = \sigma_2^2 \left(1 - \rho^2\right)$. The conditional quantile of Y given X_1 has the following expression

$$q^\alpha\left(Y| X_1\right) = \beta_1 X_1 + \beta_2 \mathbb{E}\left[X_2| X_1\right] + |\beta_2| \sqrt{\text{Var}\left(X_2| X_1\right)} \Phi^{-1}\left(\alpha\right)$$

with Φ the standard normal distribution function. One way to assess the impact of the variable X_1 on the quantile $q^\alpha\left(Y\right)$ would be to calculate the ratio

$$\frac{\mathbb{E}\left[q^\alpha\left(Y| X_1\right)\right] - \mathbb{E}\left[Y\right]}{q^\alpha\left(Y\right) - \mathbb{E}\left[Y\right]} \ .$$

Thus, by using that $q^\alpha\left(Y\right) = \mathbb{E}\left[Y\right] + \sigma_Y \Phi^{-1}\left(\alpha\right)$, the previous ratio equals

$$\frac{\mathbb{E}\left[q^\alpha\left(Y| X_1\right)\right] - \mathbb{E}\left[Y\right]}{q^\alpha\left(Y\right) - \mathbb{E}\left[Y\right]} = \frac{|\beta_2| \sigma_2 \sqrt{(1 - \rho^2)}}{\sigma_Y} \ .$$

A simple calculation shows that $S_1^\alpha = 1 - \dfrac{|\beta_2| \sigma_2 \sqrt{1 - \rho^2}}{\sigma_Y}$ and gives the relation with QOSA indices. In a more general way, the following equality holds for all variables $i = 1, \ldots, d$ when

13

using a **linear Gaussian model** and explains why the first-order and total QOSA indices do not depend on the α-level:

$$\frac{\alpha \mathbb{E}\left[Y\right] - \mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y|X_i)\}}\right]}{\alpha \mathbb{E}\left[Y\right] - \mathbb{E}\left[Y \mathbb{1}_{\{Y \leqslant q^\alpha(Y)\}}\right]} = \frac{\mathbb{E}\left[q^\alpha\left(Y \mid X_i\right)\right] - \mathbb{E}\left[Y\right]}{q^\alpha\left(Y\right) - \mathbb{E}\left[Y\right]} \ .$$

We now study the particular case $\mu_1 = \mu_2 = 0, \beta_1 = \beta_2 = 1, \sigma_1 = 1$ and $\sigma_2 = 2$. The analytical values of the indices are depicted in Figure 2 on the left-hand graph for independent inputs and on the right-hand plot as a function of the correlation coefficient between the two inputs in order to investigate the influence of the dependence.
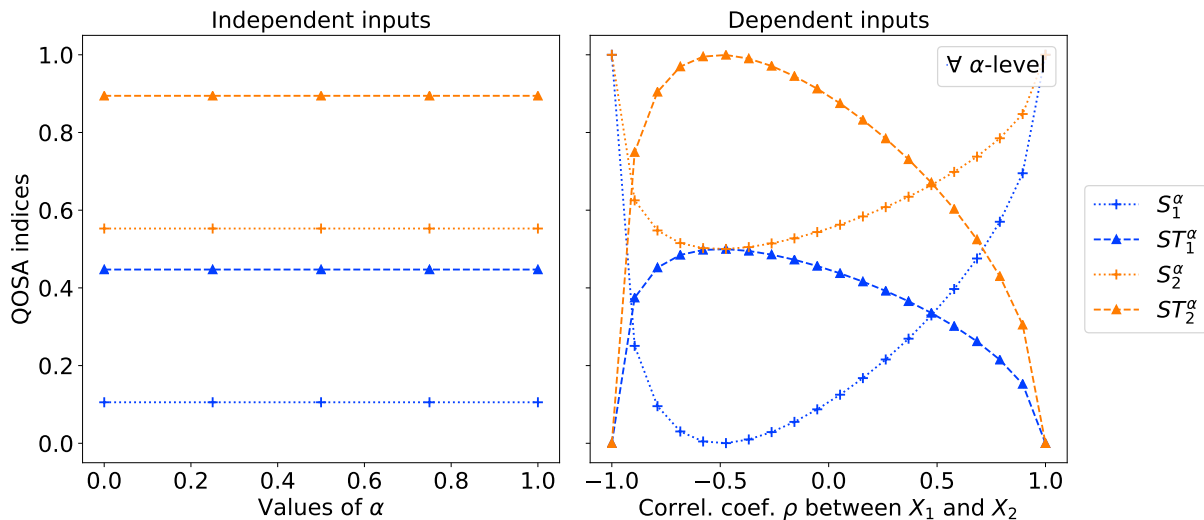

Figure 2: First-order and total QOSA indices with independent (resp. dependent) inputs on the left (resp. right) graph.

For the independent case, it appears that the variable X_2 has the higher impact over the α-quantile, which is consistent with the setting established. Besides, we have $S_i^\alpha \leqslant ST_i^\alpha$, $i = 1, 2$ as proved in Proposition 3.5.
Regarding the dependent case, we observe that the total QOSA indices tend to zero as $|\rho| \to 1$. It is also interesting to notice that $ST_i^\alpha \leqslant S_i^\alpha$ for some correlation coefficients. The behaviour of these indices is similar to that of the Sobol indices in the context of dependent inputs as studied in Kucherenko et al. (2012); Iooss and Prieur (2019). Indeed, by making an analogy with the method proposed by Mara et al. (2015) based on four Sobol indices, we could say that in the case of dependent inputs:

- the first-order QOSA index describes the influence of a variable including its dependence with other variables,

- the total QOSA index describes the influence of a variable without its dependence with other variables.

4.3 Gaussian input variables, **Lognormal output**

We analyze in this subsection a model with Gaussian inputs whose output is a Log-normal distribution so that we no longer have identical indices for any α-level. Using Gaussian inputs makes calculations possible and we obtain the following analytical values.

Proposition 4.2.

Let $Y = \eta(\mathbf{X}) = \exp\left(\beta_0 + \boldsymbol{\beta}^\mathsf{T}\mathbf{X}\right)$ *with* $\beta_0 \in \mathbb{R}$, $\boldsymbol{\beta} \in \mathbb{R}^d$ *and* $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *where* $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$ *is a positive-definite matrix, then the first-order and total QOSA indices for the variable i at the α-level are*

$$S_i^\alpha = 1 - \frac{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sqrt{\boldsymbol{\beta}_{-i}^\mathsf{T}\left(\boldsymbol{\Sigma}_{-i,-i} - \boldsymbol{\Sigma}_{-i,i}\boldsymbol{\Sigma}_{i,i}^{-1}\boldsymbol{\Sigma}_{i,-i}\right)\boldsymbol{\beta}_{-i}}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)} \, , \qquad (4.5)$$

$$ST_i^\alpha = \frac{\alpha - \Phi\left(\Phi^{-1}(\alpha) - |\boldsymbol{\beta}_i|\sqrt{\boldsymbol{\Sigma}_{i,i} - \boldsymbol{\Sigma}_{i,-i}\boldsymbol{\Sigma}_{-i,-i}^{-1}\boldsymbol{\Sigma}_{-i,i}}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)} \, , \qquad (4.6)$$

with $\sigma^2 = \boldsymbol{\beta}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{\beta}$ *and* Φ *the standard normal distribution function.*

We observe that β_0 and $\boldsymbol{\mu}$ do not play any role as these are scale parameters in this example. Let us consider the particular case $d = 2$ with

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \ \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \text{ and } \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \ -1 \le \rho \le 1, \ \sigma_1 > 0, \ \sigma_2 > 0.$$

We have $\sigma^2 = \beta_1^2\sigma_1^2 + 2\rho\beta_1\beta_2\sigma_1\sigma_2 + \beta_2^2\sigma_2^2$ and obtain from Equations (4.5) and (4.6)

$$\begin{aligned} S_1^\alpha &= 1 - \frac{\alpha - \Phi\left(\Phi^{-1}(\alpha) - |\beta_2|\sigma_2\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)} \, , \\ S_2^\alpha &= 1 - \frac{\alpha - \Phi\left(\Phi^{-1}(\alpha) - |\beta_1|\sigma_1\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)} \, , \end{aligned} \qquad (4.7)$$

and

$$\begin{aligned} ST_1^\alpha &= \frac{\alpha - \Phi\left(\Phi^{-1}(\alpha) - |\beta_1|\sigma_1\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)} \, , \\ ST_2^\alpha &= \frac{\alpha - \Phi\left(\Phi^{-1}(\alpha) - |\beta_2|\sigma_2\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)} \, . \end{aligned} \qquad (4.8)$$

In all further tests, we take $\mu_1 = \mu_2 = 0, \beta_1 = \beta_2 = 1, \sigma_1 = 1$ and $\sigma_2 = 2$.
Figure 3 presents the analytical values of the first-order and total QOSA indices for both independent inputs and correlated inputs with $\rho_{1,2} = 0.75$. In the independent setting, the influence of the variable X_1 is close to 0 except for large values of α. We also note that the first-order and total QOSA indices vary in reverse direction and from some α-level, $ST_i^\alpha \le S_i^\alpha$, $i = 1, 2$. This supports that Proposition 3.5 is not true outside the additive framework with independent inputs.
The behavior of the indices is similar in the dependent case. However, the influence of the input X_1 is reinforced in this scheme due to its large correlation with X_2 that is an influent variable. Indeed, the index S_1^α increases faster than in independent case. On the contrary, the index ST_2^α decreases to 0 quicker than in the independent case because of its high dependence with X_1.

To get another perspective on the impact of the dependence over the indices, we plot in Figure 4, for several levels α, the evolution of the latter as a function of the correlation coefficient. As for the linear Gaussian model, we observe that the total QOSA indices tend to zero as $|\rho| \to 1$ and they are lower than the first-order ones for some correlation coefficients.
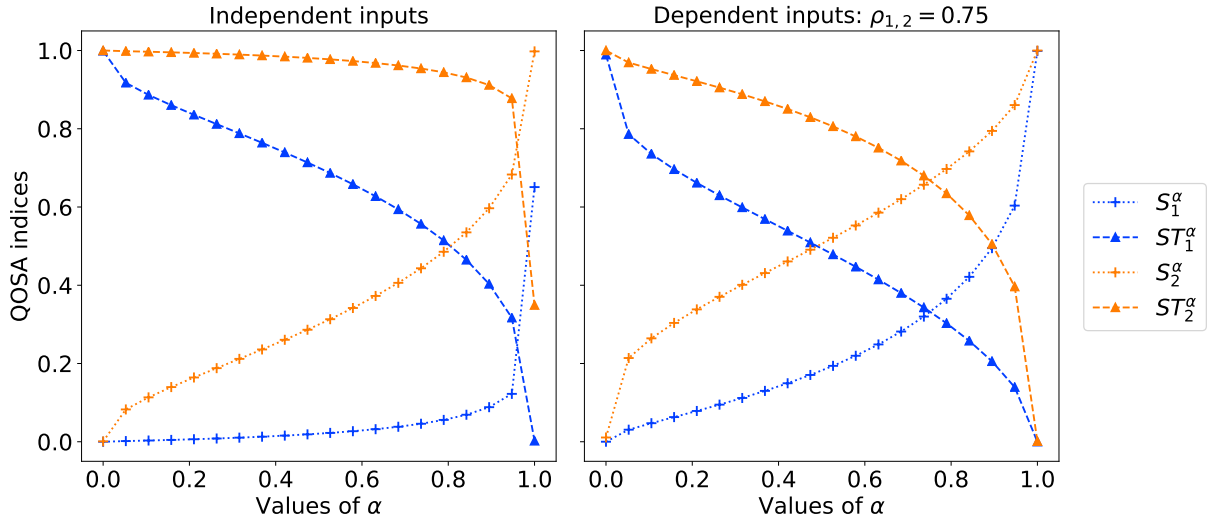
Figure 3: First-order and total QOSA indices with independent (resp. dependent) inputs on the left (resp. right) graph.

Hence, we have $S_i \leqslant ST_i^{\alpha}$, $i = 1, \ldots, d$ for additive models with independent inputs. But this context is far from reality for many concrete examples and this inequality is no longer valid outside this framework as outlined by examples presented in Subsections 4.1 and 4.3. This therefore makes the interpretation of the indices complicated.

Furthermore, in the case of dependent inputs, the behaviour of the QOSA indices should be compared to that of Sobol indices. Indeed, whatever the model (additive or not), it may happen in this scheme that the first-order QOSA indices are higher than the total ones depending on the correlation level. We have also observed that total indices tend to zero as the absolute value of the correlation goes to 1.

We could establish a strategy similar to Mara et al. (2015) in order to better understand the impact of inputs in case of statistical dependence over the α-quantile, i.e., if their contribution derives from their marginal importance or their dependence with another variable. But, we prefer to turn to Shapley values which present good properties for both independent and dependent inputs. Indeed, they allocate fairly to each input the interaction and/or dependency effect in which it is involved.

5 Quantile oriented Shapley effects

In this section, we propose to use Shapley values defined in Equation (2.5), and recalled below, in order to quantify the impact of each input over the α-quantile of the output distribution

$$v^i = \sum_{\mathcal{J} \subseteq \mathcal{D} \backslash \{i\}} \frac{(d - |\mathcal{J}| - 1)!|\mathcal{J}|!}{d!} \left(c\left(\mathcal{J} \cup \{i\}\right) - c\left(\mathcal{J}\right) \right) \ , \tag{5.1}$$

with $c(\cdot)$ a generic cost function which maps the exploratory power generated by each subset $\mathcal{J} \subseteq \mathcal{D}$.

Shapley value was first adapted within the framework of variance-based sensitivity measures to measure how much of Var (Y) can be attributed to each X_i. Indeed, Owen (2014) and Song et al. (2016) proposed to use the two following unnormalized cost functions to measure the
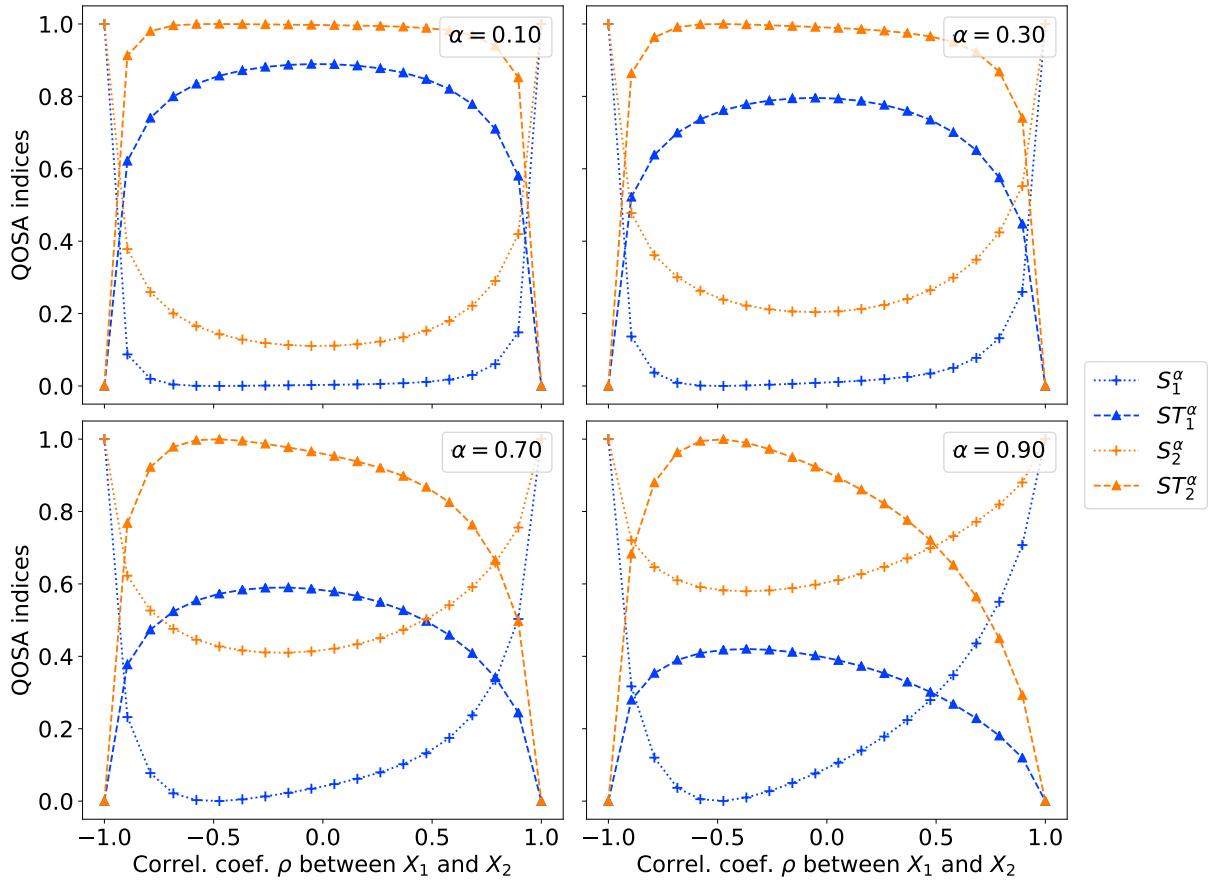
Figure 4: Evolution of the first-order and total QOSA indices at different values of ρ for several levels α.

variance of Y caused by the uncertainty of the inputs in the subset $\mathcal{J} \subseteq \mathcal{D}$ also named as being the explanatory power created by $\mathcal{J}$:

$$\tilde{c}(\mathcal{J}) = \text{Var}\left(\mathbb{E}\left[Y|\mathbf{X}_{\mathcal{J}}\right]\right) \text{ and } c(\mathcal{J}) = \mathbb{E}\left[\text{Var}\left(Y|\mathbf{X}_{-\mathcal{J}}\right)\right] . \tag{5.2}$$

Measuring the variance of Y caused by the uncertainty of the inputs in $\mathcal{J}$ is equivalent to assess the impact of the inputs over the expected output. Thus, when using the cost functions given in (5.2), the feature of interest of the output considered is the expectation denoted by $\theta^*(Y) = \mathbb{E}[Y]$. We show in the left-hand column in Table 1 that both cost functions may be rewritten according to the contrast function related to the expectation as well as the conditional feature $\theta^*(Y | X_{\mathcal{J}}) = \mathbb{E}[Y | \mathbf{X}_{\mathcal{J}}]$ for the first cost function and $\theta^*(Y | X_{-\mathcal{J}}) = \mathbb{E}[Y | \mathbf{X}_{-\mathcal{J}}]$ for the second one.

Thus, in order to define indices for another feature of interest, the idea is to substitute the contrast function of the expectation by that associated with the feature of interest required.

Let us now use $\theta^*(Y)$, $\theta^*(Y | X_{\mathcal{J}})$ and $\theta^*(Y | X_{-\mathcal{J}})$ for $\mathcal{J} \subseteq \mathcal{D}$ as generic expressions to designate a feature of interest and the conditional ones related to a contrast function ψ. The impact of the inputs over $\theta^*(Y)$ is therefore assessed by measuring their contribution to the averaged contrast function $\mathbb{E}\left[\psi\left(Y, \theta^*(Y)\right)\right]$. This one can be seen as a relevant distance allowing to quantify the variability around the feature of interest. The contributions of the inputs are then calculated with the following cost functions measuring the explanatory power of the subset

17

Feature of interest	
$\theta^*\left(Y\right) = \mathbb{E}\left[Y\right]$	$\theta^*\left(Y\right) = q^\alpha\left(Y\right)$
Contrast function	
$\psi\left(y,\theta\right) = \left(y-\theta\right)^2$	$\psi\left(y,\theta\right) = \left(y-\theta\right)\left(\alpha - \mathbb{1}_{\{y\leqslant\theta\}}\right)$
Average contrast function	
$\mathrm{Var}\left(Y\right) = \mathbb{E}\left[\psi\left(Y,\ \mathbb{E}\left[Y\right]\right)\right]$	$\Upsilon\left(Y\right) = \mathbb{E}\left[\psi\left(Y,\ q^\alpha\left(Y\right)\right)\right]$
$= \mathbb{E}\left[\psi\left(Y,\ \theta^*\left(Y\right)\right)\right]$	$= \mathbb{E}\left[\psi\left(Y,\ \theta^*\left(Y\right)\right)\right]$
First cost function	
$\tilde{c}(\mathcal{J}) = \mathrm{Var}\left(\mathbb{E}\left[Y\vert\mathbf{X}_\mathcal{J}\right]\right)$	$\tilde{c}(\mathcal{J}) = \mathbb{E}\left[\psi\left(\theta^*\left(Y\vert X_\mathcal{J}\right),\ \theta^*\left(Y\right)\right)\right]$
$= \mathbb{E}\left[\psi\left(\mathbb{E}\left[Y\vert\mathbf{X}_\mathcal{J}\right],\ \mathbb{E}\left[Y\right]\right)\right]$	$= \mathbb{E}\left[\psi\left(q^\alpha\left(Y\vert\mathbf{X}_\mathcal{J}\right),\ q^\alpha\left(Y\right)\right)\right]$
$= \mathbb{E}\left[\psi\left(\theta^*\left(Y\vert X_\mathcal{J}\right),\ \theta^*\left(Y\right)\right)\right]$	
$\tilde{c}(\emptyset) = 0 \qquad \tilde{c}(\mathcal{D}) = \mathrm{Var}\left(Y\right)$	$\tilde{c}(\emptyset) = 0 \qquad \tilde{c}(\mathcal{D}) = \Upsilon\left(Y\right)$
Second cost function	
$c(\mathcal{J}) = \mathbb{E}\left[\mathrm{Var}\left(Y\vert\mathbf{X}_{-\mathcal{J}}\right)\right]$	$c(\mathcal{J}) = \mathbb{E}\left[\psi\left(Y,\ \theta^*\left(Y\vert X_{-\mathcal{J}}\right)\right)\right]$
$= \mathbb{E}\left[\psi\left(Y,\ \mathbb{E}\left[Y\vert\mathbf{X}_{-\mathcal{J}}\right]\right)\right]$	$= \mathbb{E}\left[\psi\left(Y,\ q^\alpha\left(Y\vert\mathbf{X}_{-\mathcal{J}}\right)\right)\right]$
$= \mathbb{E}\left[\psi\left(Y,\ \theta^*\left(Y\vert X_{-\mathcal{J}}\right)\right)\right]$	
$c(\emptyset) = 0 \qquad c(\mathcal{D}) = \mathrm{Var}\left(Y\right)$	$c(\emptyset) = 0 \qquad c(\mathcal{D}) = \Upsilon\left(Y\right)$

Table 1: Analogy of the cost functions used for quantifying the impact of the inputs over the expectation for the case where the quantile is the feature of interest.

$\mathcal{J} \subseteq \mathcal{D}$

$$\tilde{c}(\mathcal{J}) = \mathbb{E}\left[\psi\left(\theta^*\left(Y\vert X_\mathcal{J}\right),\ \theta^*\left(Y\right)\right)\right] \text{ and } c(\mathcal{J}) = \mathbb{E}\left[\psi\left(Y,\ \theta^*\left(Y\vert X_{-\mathcal{J}}\right)\right)\right] . \tag{5.3}$$

These cost functions are valid choices if they satisfy that the empty set creates no value, and that all inputs generate $\mathbb{E}\left[\psi\left(Y,\ \theta^*\left(Y\right)\right)\right]$. This is, for example, verified for all contrast functions listed in Fort et al. (2016) which allow to propose new indices named **_Goal oriented Shapley effects (GOSE)_**.

In our context, this property is verified in particular for cost functions related to quantiles presented in Table 1. Hence we may propose **_Shapley effects subordinated to quantiles_**. However, we define the **_Quantile oriented Shapley effects (QOSE)_** denoted by Sh_i^α with the second cost function because it verifies that the incremental cost $c\left(\mathcal{J}\cup\{i\}\right) - c\left(\mathcal{J}\right)$ is non negative, which is also a desirable property for cost functions. Indeed, for $\mathcal{J} \subseteq \mathcal{D}\backslash\{i\}$, we have

$$c\left(\mathcal{J}\cup i\right) - c\left(\mathcal{J}\right) = \left(\alpha\mathbb{E}\left[Y\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y\leqslant q^\alpha(Y\vert\mathbf{X}_{-\mathcal{J}\cup i})\}}\right]\right) - \left(\alpha\mathbb{E}\left[Y\right] - \mathbb{E}\left[Y\mathbb{1}_{\{Y\leqslant q^\alpha(Y\vert\mathbf{X}_{-\mathcal{J}})\}}\right]\right)$$
$$= \mathbb{E}\left[\left(Y - q^\alpha\left(Y\vert\mathbf{X}_{-\mathcal{J}\cup i}\right)\right)\left(\mathbb{1}_{\{Y\leqslant q^\alpha(Y\vert X_{-\mathcal{J}})\}} - \mathbb{1}_{\{Y\leqslant q^\alpha(Y\vert\mathbf{X}_{-\mathcal{J}\cup i})\}}\right)\right] \geqslant 0 .$$

At this stage, this property has not yet been demonstrated for the first cost function $\tilde{c}$.

We study in the next subsections examples whose analytical values of the index Sh_i^α are computed by using the cost function normalized by the quantity $\Upsilon\left(Y\right)$ introduced in Table 1, so that $\sum_{i=1}^{d} Sh_i^\alpha = 1$. Our aim is to show that these new indices give sensible answers compared to the classical QOSA indices defined in Section 3.

5.1 Linear model with Gaussian input variables

We obtain the following analytical values for the linear model with Gaussian inputs.

Proposition 5.1.
If $Y = \eta(\mathbf{X}) = \beta_0 + \boldsymbol{\beta}^{\mathsf{T}}\mathbf{X}$ with $\beta_0 \in \mathbb{R}$, $\boldsymbol{\beta} \in \mathbb{R}^d$ and $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$ is a positive-definite matrix, then the Quantile oriented Shapley effect for the variable i at the α-level is

$$Sh_i^\alpha = \frac{1}{d \cdot \sigma_Y} \sum_{\mathcal{J} \subseteq \mathcal{D} \setminus \{i\}} \binom{d-1}{|\mathcal{J}|}^{-1} \left[\sqrt{\boldsymbol{\beta}_{\mathcal{J}+i}^{\mathsf{T}} \left(\boldsymbol{\Sigma}_{\mathcal{J}+i,\mathcal{J}+i} - \boldsymbol{\Sigma}_{\mathcal{J}+i,-\mathcal{J}-i}\boldsymbol{\Sigma}_{-\mathcal{J}-i,-\mathcal{J}-i}^{-1}\boldsymbol{\Sigma}_{-\mathcal{J}-i,\mathcal{J}+i} \right) \boldsymbol{\beta}_{\mathcal{J}+i}} \right.$$

$$\left. - \sqrt{\boldsymbol{\beta}_{\mathcal{J}}^{\mathsf{T}} \left(\boldsymbol{\Sigma}_{\mathcal{J},\mathcal{J}} - \boldsymbol{\Sigma}_{\mathcal{J},-\mathcal{J}}\boldsymbol{\Sigma}_{-\mathcal{J},-\mathcal{J}}^{-1}\boldsymbol{\Sigma}_{-\mathcal{J},\mathcal{J}} \right) \boldsymbol{\beta}_{\mathcal{J}}} \right]$$

(5.4)

with $\sigma_Y^2 = \mathrm{Var}(Y) = \boldsymbol{\beta}^{\mathsf{T}}\boldsymbol{\Sigma}\boldsymbol{\beta}$, and $\mathcal{J}+i$ (resp. $-\mathcal{J}-i$), a notational compression for $\mathcal{J} \cup \{i\}$ (resp. $-\mathcal{J} \cup \{i\}$).

As for the QOSA index, we may notice that β_0 and $\boldsymbol{\mu}$ do not play any role as translation parameters and that the index does not depend on the α-level which is a specificity of the linear Gaussian model as explained previously. Let us consider the case $d = 2$ with

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \text{ and } \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad -1 \le \rho \le 1, \ \sigma_1 > 0, \ \sigma_2 > 0.$$

We have $\sigma_Y^2 = Var(Y) = \beta_1^2\sigma_1^2 + 2\rho\beta_1\beta_2\sigma_1\sigma_2 + \beta_2^2\sigma_2^2$ and obtain from (5.4)

$$\begin{aligned} Sh_1^\alpha &= \frac{1}{2} - \frac{|\beta_2|\,\sigma_2\sqrt{1-\rho^2}}{2 \cdot \sigma_Y} + \frac{|\beta_1|\,\sigma_1\sqrt{1-\rho^2}}{2 \cdot \sigma_Y}, \\ Sh_2^\alpha &= \frac{1}{2} - \frac{|\beta_1|\,\sigma_1\sqrt{1-\rho^2}}{2 \cdot \sigma_Y} + \frac{|\beta_2|\,\sigma_2\sqrt{1-\rho^2}}{2 \cdot \sigma_Y}. \end{aligned}$$

(5.5)

We observe that the correlation effects on the first-order QOSA indices (e.g. $\sigma_Y - |\beta_2|\,\sigma_2\sqrt{1-\rho^2}$ for X_1) and on the total QOSA indices (e.g. $|\beta_1|\,\sigma_1\sqrt{1-\rho^2}$ for X_1) are allocated half to the Quantile oriented Shapley effects - QOSE. We also see that the Shapley effects are equal when the correlation is maximum (i.e. $|\rho| = 1$).

Figure 5 presents the first-order and total QOSA indices as well as the QOSE for the particular case $\mu_1 = \mu_2 = 0$, $\beta_1 = \beta_2 = 1$, $\sigma_1 = 1$ and $\sigma_2 = 2$.
On the left-hand graph of the figure, we see that the Shapley effects are also constant and they are bracketed by the first-order and total QOSA indices : $S_i^\alpha \leqslant Sh_i^\alpha \leqslant ST_i^\alpha$, $i = 1, 2$.

We illustrate on the right-hand graph the evolution of the indices as a function of the correlation between the two inputs. As X_2 is the more uncertain variable, its sensitivity indices are larger than those of X_1. Then, although the values are not identical, we can note that the shape of the curves is exactly the same as that observed for the variance-based Shapley effects calculated for the two-dimensional Gaussian linear model (with the same setting) in Iooss and Prieur (2019). Indeed, we observe that in the presence of correlation, the QOSE lie between the first-order QOSA indices and the total ones with either $S_i^\alpha \leqslant Sh_i^\alpha \leqslant ST_i^\alpha$ or $ST_i^\alpha \leqslant Sh_i^\alpha \leqslant S_i^\alpha$, $i = 1, 2$. This phenomenon is called the "sandwich effect" within the variance framework in Iooss and Prieur (2019). Finally, as for the variance-based Shapley effects, it also seems that the dependence between the two inputs lead to a rebalancing of their respective QOSE.
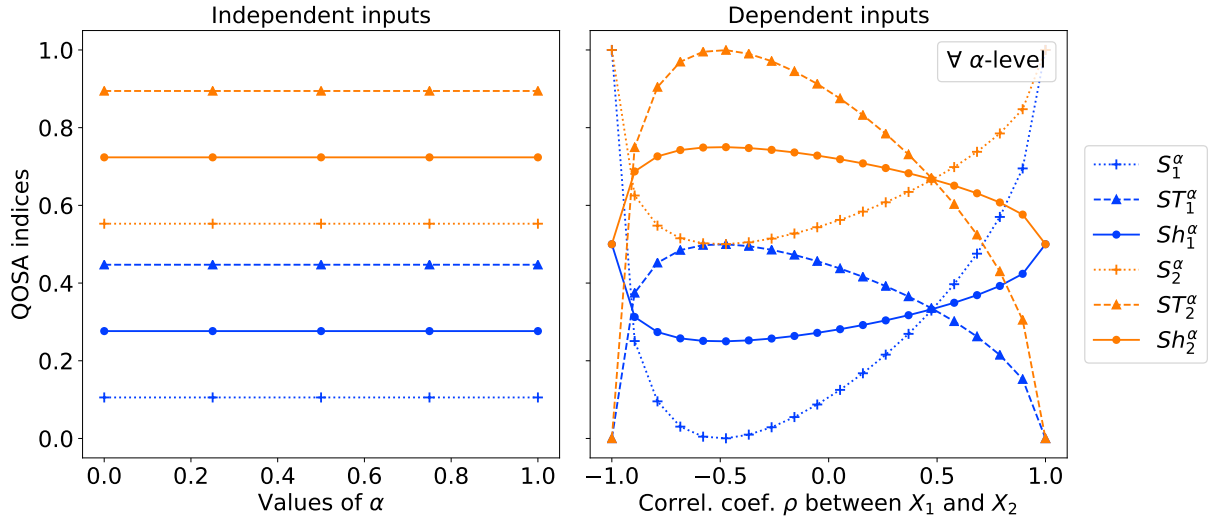
Figure 5: First-order and total QOSA indices as well as the QOSE with independent (resp. dependent) inputs on the left (resp. right) graph.

5.2 Gaussian input variables, Lognormal output

We analyze in this subsection the analytical values below for the model with Gaussian inputs and the resulting output Log-normal distributed.

Proposition 5.2.
If $Y = \eta(\mathbf{X}) = \exp\left(\beta_0 + \boldsymbol{\beta}^{\mathsf{T}}\mathbf{X}\right)$ with $\beta_0 \in \mathbb{R}$, $\boldsymbol{\beta} \in \mathbb{R}^d$ and $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$ is a positive-definite matrix, then the Quantile oriented Shapley effect for the variable i at the α-level is

$$Sh_i^\alpha = \frac{1}{d \cdot A} \sum_{\mathcal{J} \subseteq \mathcal{D} \setminus \{i\}} \binom{d-1}{|\mathcal{J}|}^{-1} \left[\Phi\left(\Phi^{-1}(\alpha) - B(\mathcal{J})\right) - \Phi\left(\Phi^{-1}(\alpha) - C(\mathcal{J}, i)\right) \right] \qquad (5.6)$$

with

$$A = \alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right) \text{ and } \sigma^2 = \boldsymbol{\beta}^{\mathsf{T}} \boldsymbol{\Sigma} \boldsymbol{\beta} ,$$

$$B(\mathcal{J}) = \sqrt{\boldsymbol{\beta}_{\mathcal{J}}^{\mathsf{T}} \left(\boldsymbol{\Sigma}_{\mathcal{J},\mathcal{J}} - \boldsymbol{\Sigma}_{\mathcal{J},-\mathcal{J}} \boldsymbol{\Sigma}_{-\mathcal{J},-\mathcal{J}}^{-1} \boldsymbol{\Sigma}_{-\mathcal{J},\mathcal{J}} \right) \boldsymbol{\beta}_{\mathcal{J}}} ,$$

$$C(\mathcal{J}, i) = \sqrt{\boldsymbol{\beta}_{\mathcal{J}+i}^{\mathsf{T}} \left(\boldsymbol{\Sigma}_{\mathcal{J}+i,\mathcal{J}+i} - \boldsymbol{\Sigma}_{\mathcal{J}+i,-\mathcal{J}-i} \boldsymbol{\Sigma}_{-\mathcal{J}-i,-\mathcal{J}-i}^{-1} \boldsymbol{\Sigma}_{-\mathcal{J}-i,\mathcal{J}+i} \right) \boldsymbol{\beta}_{\mathcal{J}+i}} ,$$

where $\mathcal{J} + i$ (resp. $-\mathcal{J} - i$) is a notational compression for $\mathcal{J} \cup \{i\}$ (resp. $-\mathcal{J} \cup \{i\}$).

As for the QOSA indices, we observe that β_0 and $\boldsymbol{\mu}$ do not play any role and that the indices depend on α compared to the linear Gaussian model. However, it is difficult to reach a conclusion from Equation (5.6). Accordingly, we consider the particular case $d = 2$ with

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \ \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \text{ and } \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \ -1 \leq \rho \leq 1, \ \sigma_1 > 0, \ \sigma_2 > 0.$$

20

We have $\sigma^2 = \beta_1^2\sigma_1^2 + 2\rho\beta_1\beta_2\sigma_1\sigma_2 + \beta_2^2\sigma_2^2$ and obtain from (5.6)

$$
\begin{aligned}
Sh_1^\alpha &= \frac{1}{2} + \frac{1}{2}\cdot\frac{\Phi\left(\Phi^{-1}\left(\alpha\right) - |\beta_2|\,\sigma_2\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}\left(\alpha\right) - \sigma\right)} - \frac{1}{2}\cdot\frac{\Phi\left(\Phi^{-1}\left(\alpha\right) - |\beta_1|\,\sigma_1\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}\left(\alpha\right) - \sigma\right)} \ , \\
Sh_2^\alpha &= \frac{1}{2} + \frac{1}{2}\cdot\frac{\Phi\left(\Phi^{-1}\left(\alpha\right) - |\beta_1|\,\sigma_1\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}\left(\alpha\right) - \sigma\right)} - \frac{1}{2}\cdot\frac{\Phi\left(\Phi^{-1}\left(\alpha\right) - |\beta_2|\,\sigma_2\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}\left(\alpha\right) - \sigma\right)} \ .
\end{aligned}
\tag{5.7}
$$

We adopt the next settings in all further tests: $\mu_1 = \mu_2 = 0, \beta_1 = \beta_2 = 1, \sigma_1 = 1$ and $\sigma_2 = 2$.

The analytical values of the first-order, total QOSA indices and the QOSE are illustrated in Figure 6 for both independent inputs and correlated inputs with $\rho_{1,2} = 0.75$. The "sandwich effect" which was noticed in the linear Gaussian model in the presence of correlation is also observed here. Indeed, both in the dependent and independent cases and for all the levels α, the QOSE lie between the first-order and total QOSA indices.

Besides, with the three indices, we obtain the same ranking of the inputs for all α-levels but the QOSE is easier to interpret because it properly condenses all the information (dependence and interaction effects). For instance, let us focus over the input X_1 on the right-hand graph at the level $\alpha = 0.2$. If we use the first-order QOSA index S_1^α, we conclude that the impact of the input X_1 is low, but not so small because, conversely its total QOSA index is high enough. But, ultimately, it is difficult to quantify precisely on the basis of these two indices the contribution of the input X_1 at level $\alpha = 0.2$. The Shapley index, in contrast, contains the marginal contribution of the variable but also those due to dependence and interaction effects that are correctly allocated to it. It therefore makes easier to express an opinion on the impact of the variable by taking into account all possible contributions. This observation is valid for all the levels α.
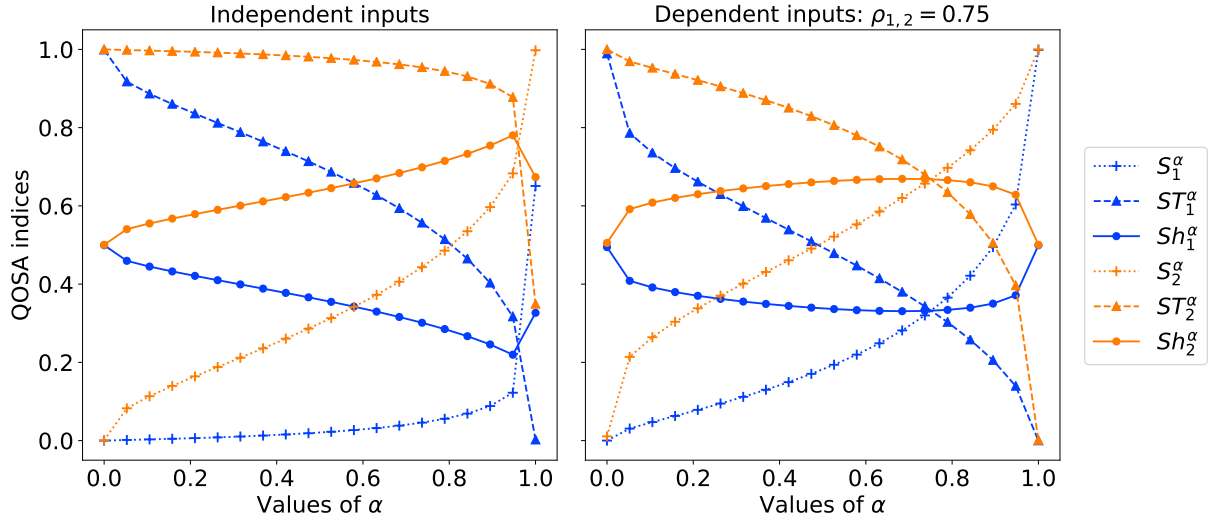

Figure 6: First-order and total QOSA indices as well as the Quantile oriented Shapley effects with independent (resp. dependent) inputs on the left (resp. right) graph.

Again, to get another insight on the impact of the dependence over the indices, we plot in Figure 7, for several levels α, the evolution of the latter as a function of the correlation coefficient. As explained before, the QOSE give a condensed information of all contributions. That explains why we observe that the Shapley effects of both variables are almost equal for small values of α.

Conversely, for large values, the variable X_2 is the most influential overall except when $|\rho| \to 1$ where both inputs have the same contribution.
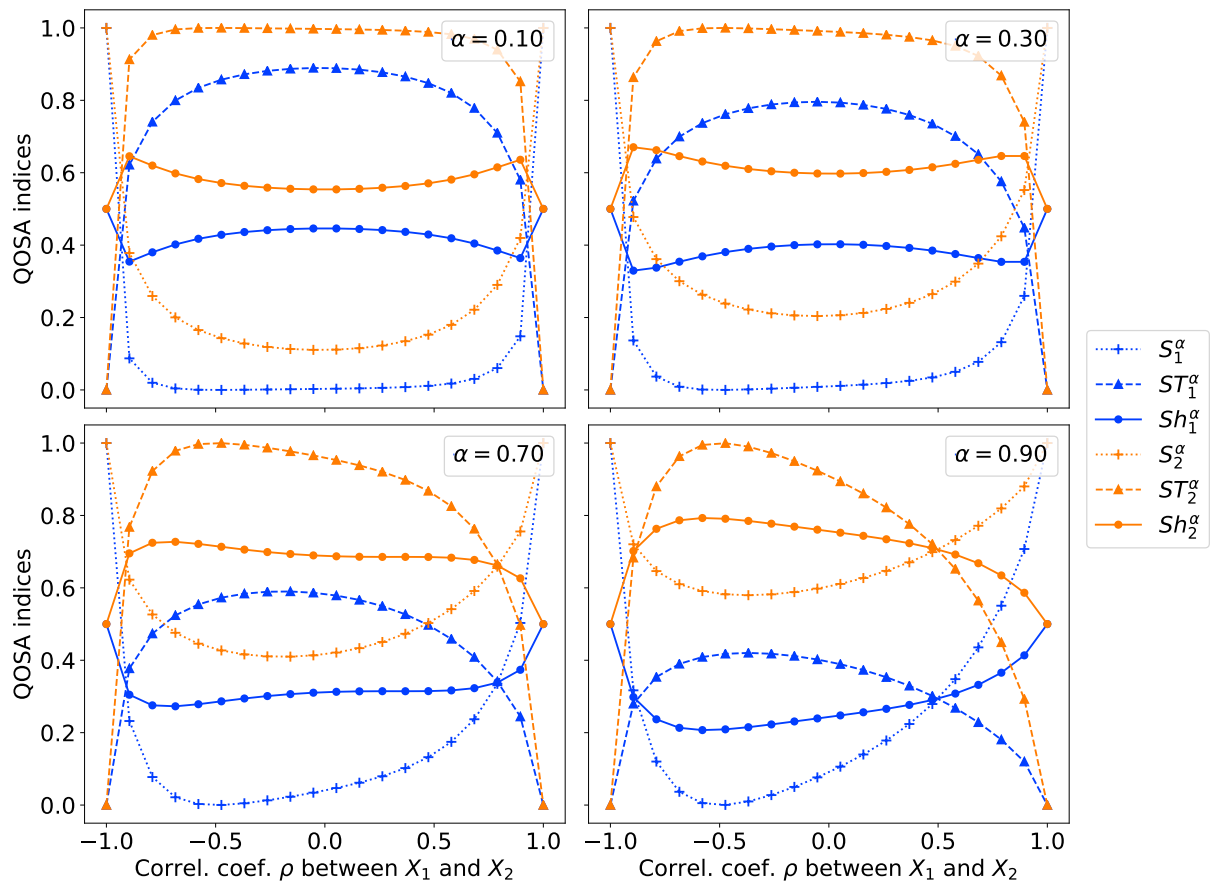

Figure 7: Evolution of the first-order and total QOSA indices as well as the Quantile oriented Shapley effects at different values of ρ for several levels α.

6 Perspective

As a conclusion of this preliminary work, QOSE appear to be a good alternative to the classical QOSA indices. Indeed, they make possible to overcome the various problems encountered with the QOSA indices such as $ST_i^\alpha \leqslant S_i^\alpha$ outside the additive framework or when using dependent inputs. Besides, the Sobol indices stem, for example, from the functional ANOVA decomposition but there is no such a decomposition for the quantiles. Hence, as an allocation method, these indices therefore allow to quantify precisely the contribution of each input at the α-quantile while taking into account interaction and dependency effects.

An additional work would be to study the QOSE for Gaussian examples in higher dimension. Last but not least, the development of an estimation algorithm to compute these new indices would be a significant contribution.

References

Benoumechiara, N. and Elie-Dit-Cosaque, K. (2019). Shapley effects for sensitivity analysis with dependent inputs: bootstrap and kriging-based algorithms. *ESAIM: Proceedings and Surveys*, 65:266–293.

Borgonovo, E. (2006). Measuring uncertainty importance: investigation and comparison of alternative approaches. *Risk analysis*, 26(5):1349–1361.

Borgonovo, E. (2007). A new uncertainty importance measure. *Reliability Engineering & System Safety*, 92(6):771–784.

Borgonovo, E., Castaings, W., and Tarantola, S. (2011). Moment independent importance measures: new results and analytical test cases. *Risk Analysis: An International Journal*, 31(3):404–428.

Borgonovo, E. et al. (2017). Sensitivity analysis. *Number*, 251:93–100.

Borgonovo, E. and Plischke, E. (2016). Sensitivity analysis: a review of recent advances. *European Journal of Operational Research*, 248(3):869–887.

Browne, T., Fort, J.-C., Iooss, B., and Le Gratiet, L. (2017). Estimate of quantile-oriented sensitivity indices. Technical Report, hal-01450891.

Caniou, Y. (2012). *Global sensitivity analysis for nested and multiscale modelling.* PhD thesis, Université Blaise Pascal - Clermont-Ferrand II.

Chastaing, G., Gamboa, F., Prieur, C., et al. (2012). Generalized hoeffding-sobol decomposition for dependent variables-application to sensitivity analysis. *Electronic Journal of Statistics*, 6:2420–2448.

Da Veiga, S. (2015). Global sensitivity analysis with dependence measures. *Journal of Statistical Computation and Simulation*, 85(7):1283–1305.

Efron, B. and Stein, C. (1981). The jackknife estimate of variance. *The Annals of Statistics*, pages 586–596.

Elie-Dit-Cosaque, K. and Maume-Deschamps, V. (2021). Random forest-based qosa index estimation.

Faivre, R., Iooss, B., Mahévas, S., Makowski, D., and Monod, H. (2016). *Analyse de sensibilité et exploration de modèles: application aux sciences de la nature et de l'environnement.* Editions Quae.

Fisher, R. and Mackenzie, W. (1923). Studies in crop variation: The manurial response of different potato varieties. *Journal of Agricultural Sciences*, 13:311–320.

Fort, J.-C., Klein, T., and Rachdi, N. (2016). New sensitivity analysis subordinated to a contrast. *Communications in Statistics-Theory and Methods*, 45(15):4349–4364.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics*, 19(3):293–325.

Homma, T. and Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1):1–17.

Iooss, B. and Lemaître, P. (2015). A review on global sensitivity analysis methods. In *Uncertainty Management in Simulation-Optimization of Complex Systems*, pages 101–122. Springer.

Iooss, B. and Prieur, C. (2019). Shapley effects for sensitivity analysis with correlated inputs: comparisons with sobol'indices, numerical estimation and applications. *International Journal for Uncertainty Quantification*, 9(5).

Kala, Z. (2019). Quantile-oriented global sensitivity analysis of design resistance. *Journal of Civil Engineering and Management*, 25(4):297–305.

Kucherenko, S., Song, S., and Wang, L. (2019). Quantile based global sensitivity measures. *Reliability Engineering & System Safety*, 185:35–48.

Kucherenko, S., Tarantola, S., and Annoni, P. (2012). Estimation of global sensitivity indices for models with dependent variables. *Computer Physics Communications*, 183(4):937–946.

Mara, T. A. and Tarantola, S. (2012). Variance-based sensitivity indices for models with dependent inputs. *Reliability Engineering & System Safety*, 107:115–121.

Mara, T. A., Tarantola, S., and Annoni, P. (2015). Non-parametric methods for global sensitivity analysis of model output with dependent inputs. *Environmental Modelling & Software*, 72:173–183.

Marceau, E. (2013). *Modélisation et évaltuation quantitative des risques en actuariat*. Springer Berlin.

Maume-Deschamps, V. and Niang, I. (2018). Estimation of quantile oriented sensitivity indices. *Statistics & Probability Letters*, 134:122–127.

Owen, A. B. (2014). Sobol'indices and shapley value. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):245–251.

Owen, A. B. and Prieur, C. (2017). On shapley value for measuring importance of dependent inputs. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):986–1002.

Rabitti, G. and Borgonovo, E. (2019). A shapley–owen index for interaction quantification. *SIAM/ASA Journal on Uncertainty Quantification*, 7(3):1060–1075.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008). *Global sensitivity analysis: the primer*. John Wiley & Sons.

Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M. (2004). *Sensitivity analysis in practice: a guide to assessing scientific models*. John Wiley & Sons.

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games*, 2(28):307–317.

Sobol, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*, 1(4):407–414.

Song, E., Nelson, B. L., and Staum, J. (2016). Shapley effects for global sensitivity analysis: Theory and computation. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):1060–1083.

Stone, C. J. (1994). The use of polynomial splines and their tensor products in multivariate function estimation. *The Annals of Statistics*, pages 118–171.

Winter, E. et al. (2002). The shapley value. *Handbook of game theory with economic applications*, 3(2):2025–2054.