HAL
open science

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

Kevin Elie-Dit-Cosaque, Véronique Maume-Deschamps

To cite this version:

Kevin Elie-Dit-Cosaque, Véronique Maume-Deschamps. Goal-Oriented Shapley Effects with Special Attention to the Quantile-Oriented Case. SIAM/ASA Journal on Uncertainty Quantification, 2022, 10 (3), pp.1037-1069. 10.1137/21M1395247 . hal-03137402v2

HAL Id: hal-03137402

https://hal.science/hal-03137402v2

Submitted on 4 Feb 2023

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.

Goal-oriented Shapley effects with a special attention to the quantile-oriented case

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

[1]Institut Camille Jordan - ICJ, Univ Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5208, F-69622 Villeurbanne, France
[2]Model Analysis, SCOR, Paris, France

February 25, 2022

Abstract

We propose to study quantile-oriented sensitivity indices (QOSA indices) and show some of their theoretical properties. These have a number of shortcomings, when dealing both with independent and dependent inputs, which lead us to define new generic indices based on the Shapley values named goal-oriented Shapley effects (GOSE). In particular, we focus on quantile-oriented Shapley effects (QOSE) and subsequently perform several calculations of QOSA indices and QOSE in order to better understand the behaviour and the respective interest of each.

Keywords: Quantile-oriented sensitivity analysis; Goal-oriented 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, 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 first focus on indices defined in Fort et al. (2016) and generalized later in Borgonovo et al. (2021) who have established that they can be interpreted as information value in a reporting problem under a proper scoring rule.

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

For the sake of completeness, Section 2 shortly recalls variance-based methods dealing both with independent and dependent inputs.
Section 3 introduces QOSA indices which allow to quantify the sensitivity over a quantile. It should be noted that QOSA indices are closely related to the information value when the decision problem is to report the α-quantile of the output (see Proposition 12 of Borgonovo et al. (2021)), which gives them a natural interpretation. Other quantile-oriented indices were proposed in Kucherenko et al. (2019) based on a quadratic scoring rule. Nevertheless, it is well-known that such a rule is relevant to deal with the problem of choosing a point estimate of the output. Accordingly, there is a mismatch with the goal to accurately assess the impact of the inputs on a α-quantile. Shortcomings of such indices regarding the interpretation of the influence of the inputs on the α-quantile of the output will be highlighted on an example 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 in order to understand the impact of the statistical dependence between the inputs over these indices.
Facing some interpretation issues, new indices based on Shapley values (Shapley, 1953) named GOSE (goal-oriented Shapley effects) are proposed in Section 5 to quantify the importance of an input on a specific quantity of interest of the output. In particular, QOSE (quantile-oriented Shapley effects) are introduced to assess the influence of inputs on the α-quantile of the output and seem to be a promising alternative.
A natural extension of Cramér-von Mises indices, defined in Gamboa et al. (2018), is also proposed in this paper. Cramér-von Mises indices rest on the Hoeffding decomposition which requires the independence assumption between the inputs which could lead to some intepretation problems when dealing with dependent inputs. To overcome this limitation, we take advantage of the allocation rule of Shapley values and introduce, in Section 6, new indices named Cramér-von Mises distance-based Shapley effects, valid for both independent and dependent inputs. These new indices complement GOSE as they allow to the practitioner to measure the importance of inputs on the whole distribution of the output and not to a specific feature of interest.
Finally, a first estimator for QOSE is introduced and studied in Section 7 while Section 8 presents some further perspectives of research.

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

2 Variance-based sensitivity indices

We shall briefly recall the framework of Sobol' indices and variance-based 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). Let us mention that Pearson (1905) discussed the correlation ratio and Kolmogoroff (1933) recognized this as an estimator of the variance of a conditional expectation.
The theoretical basis for the definition of Sobol' indices is a functional ANOVA expansion of the

model output η:

$$
\begin{aligned}
Y = \eta(\mathbf{X}) &= \eta_0 + \sum_{i=1}^{d} \eta_i(X_i) + \sum_{1 \leqslant i < j \leqslant d} \eta_{ij}(X_i, X_j) + \cdots + \eta_{1,\ldots,d}(\mathbf{X}) \\
&= \sum_{u \subseteq \mathcal{D}} \eta_u\left(\mathbf{x}_u\right) \ .
\end{aligned}
\tag{2.1}
$$

This decomposition is unique under orthogonality constraints (see Sobol (1993)), that are verified when dealing with independent inputs, and therefore each term writes as follows,

$$
\eta_u\left(\mathbf{x}_u\right) = \sum_{v \subseteq u} (-1)^{|u|-|v|}\, \mathbb{E}\left[Y \,|\, \mathbf{X}_u = \mathbf{x}_u\right], \qquad u \subseteq \mathcal{D} \ .
$$

The orthogonality of the terms η_u leads us to the following decomposition of the global variance

$$
\begin{aligned}
\mathrm{Var}(Y) &= \sum_{i=1}^{d} V_i + \sum_{1 \leqslant i < j \leqslant d} V_{ij} + \cdots + V_{1,\ldots,d} \\
&= \sum_{u \subseteq \mathcal{D}} \mathrm{Var}\left(\eta_u\left(\mathbf{X}_u\right)\right) \ .
\end{aligned}
\tag{2.2}
$$

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

$$
S_u = \frac{\mathrm{Var}\left(\eta_u\left(\mathbf{X}_u\right)\right)}{\mathrm{Var}(Y)} = \frac{\sum_{v \subseteq u} (-1)^{|u|-|v|}\, \mathrm{Var}\left(\mathbb{E}\left[Y \,|\, \mathbf{X}_u = \mathbf{x}_u\right]\right)}{\mathrm{Var}(Y)}, \qquad u \subseteq \mathcal{D} \ . \tag{2.3}
$$

The sum of Sobol' indices over all subsets of inputs being equal to one, they can be directly interpreted as the contribution of each subset of inputs to the output variance. In particular, 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 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, the computation of all Sobol indices requires the evaluation of $2^d - 1$ indices, which can be too computationally demanding. 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

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

where the notation $\mathbf{X}_{-i}$ indicates the set of all variables except X_i. The following property can easily be deduced (the independence assumption is crucial here):

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

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 that inputs have a statistical dependence structure: $\mathbf{X}$ may have support in a non-rectangular region or physical constraints may impose a probabilistic dependence structure. If the independence assumption is not satisfied, 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); Owen and Prieur (2017); Benoumechiara and Elie-Dit-Cosaque (2019); 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).

In cooperative game theory, 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. Then, thanks to an attribution method, denoted v, the amount $c(\mathcal{D})$ earned by the entire team is fairly shared between all the players. The latter should satisfy some intuitive properties and the following ones are classical.

- **Efficiency:** $\sum_{i=1}^{d} v_i = c(\mathcal{D})$. 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 $\mathcal{J} \subseteq \mathcal{D} \setminus \{i\}$. It has to be noted that there are alternative representations also satisfying the properties.

In the framework of GSA, 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) first 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 later in their Theorem 1 that 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. 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. 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, 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 inputs involved, making them condensed and easy-to-interpret indices.

3 Quantile-oriented sensitivity indices

In the previous section, variance-based measures have been reviewed. 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, in order to compare conditional expectations with the unconditional expectation of the output distribution. 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); Luyi et al. (2012); Gamboa et al. (2018) 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:

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

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

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

For $\alpha \in]0,1[$ and if Y admits a positive density on its support, then the denominator in (3.1) does not vanish. 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 $\theta = q^\alpha(Y)$. Hence 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

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

Firstly, it should be noted that the properties of the term $\mathbb{E}\left[\left(\mathbb{E}\left[Y\right] - Y\right)\mathbb{1}_{\{X_i\leqslant q^\alpha(X_i)\}}\right]$ instead of $\mathbb{E}\left[\left(\mathbb{E}\left[Y\right] - Y\right)\mathbb{1}_{\{Y\leqslant q^\alpha(Y|X_i)\}}\right]$, used in the equation above, have been studied as cusunoro curves in Plischke (2012). Secondly, first-order QOSA indices are closely related to the recent quantile sensitivity measures defined in Proposition 12 of Borgonovo et al. (2021), i.e., the information value when the decision problem is to report the α-quantile of the output. It supports the use of QOSA indices as sensitivity measures when the goal of the practitioner is to assess the impact of the inputs on the α-quantile of the output.

Another quantile-oriented index has also been proposed in Kucherenko et al. (2019) 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

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

They also provide the normalized versions as follows

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

These measures thereby quantify the mean distance between quantiles $q^{\alpha}(Y)$ and $q^{\alpha}(Y|X_i)$ 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 questionable, 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. Hence 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, the behaviour of QOSA, $\bar{q}_{i,1}^{\alpha}$ and $\bar{q}_{i,2}^{\alpha}$ indices is shown in Figure 1. 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.

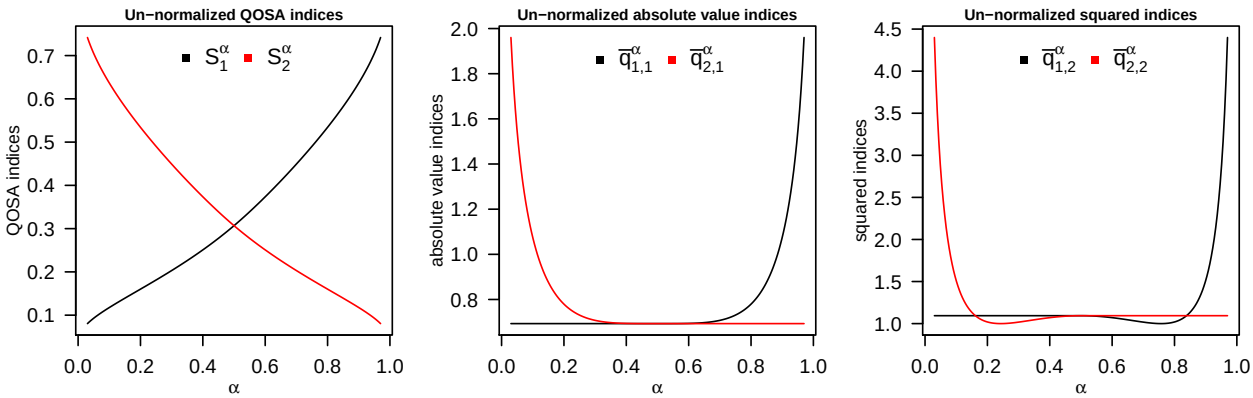

Figure 1: Un-normalized first-order QOSA, absolute value and squared indices as a function of the α level for the model $Y = X_1 - X_2$

The normalized indices, presented in Figure 2 below, show that the interpretation for $Q_{i,j}^{\alpha}$ remains questionable while the normalized QOSA indices keep the interpretation of the un-normalized ones.

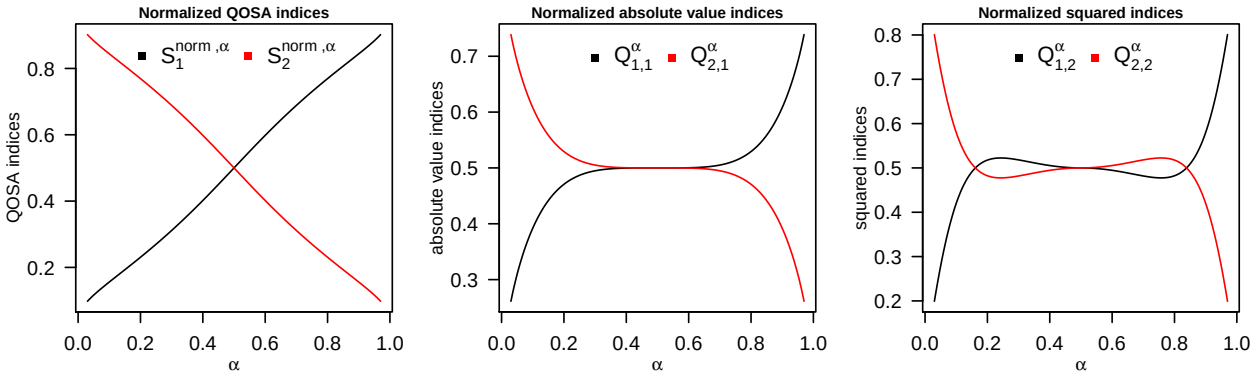

Figure 2: Normalized first-order QOSA, absolute value and squared indices as a function of the α level for the model $Y = X_1 - X_2$

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[1] in risk theory (see e.g. Marceau (2013)).

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:

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

■

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

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

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

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

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]Recall that the definition of TVaR of level $\alpha \in]0,1[$ for a continuous random variable X is given by: $\mathrm{TVaR}^\alpha(X) = \mathbb{E}(X \mid X > q^\alpha(X))$.

1. *S_i^α is invariant with respect to translations of the output Y.*
2. *S_i^α is invariant by scale change with positive ratio of the output Y.*
3. *A scale change with 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 $S_i'^\alpha = S_i^{1-\alpha}$.

■

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 shown 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 $Xs_{(-i)} = \sum_{\substack{1 \leqslant j \leqslant d \\ j \neq i}} m_j(X_j)$. Due to the independence between 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

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

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

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

Now, we use Lemma 3.1 which gives:

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

As a consequence,

$$
\begin{aligned}
g(\alpha) &\geqslant \sum_{i=1}^{d-1}\mathbb{E}\left[m_i(X_i)\mathbb{1}_{\{Y\leqslant q^\alpha(Y)\}}\right] - \mathbb{E}\left[Xs_{(-d)}\mathbb{1}_{\left\{Xs_{(-d)}\leqslant q^\alpha\left(Xs_{(-d)}\right)\right\}}\right] \\
&= \mathbb{E}\left[Xs_{(-d)}\mathbb{1}_{\{Y\leqslant q^\alpha(Y)\}}\right] - \mathbb{E}\left[Xs_{(-d)}\mathbb{1}_{\left\{Xs_{(-d)}\leqslant q^\alpha\left(Xs_{(-d)}\right)\right\}}\right] \\
&\geqslant 0 \text{ using again Lemma 3.1.}
\end{aligned}
$$

■

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:

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

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

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

Although this decomposition appears similar to that of Sobol' indices, it is quite different since it does not stem from an unique 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 scale change with positive ratio of the output Y.*
3. *A scale change with 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 independent inputs. 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 Sobol' indices $ST_i = S_i$, for all $i \in \mathcal{D}$, which is not the case for the QOSA indices. It therefore appears that the total QOSA index captures some interactions between the inputs when using an additive model. The origin of this phenomenon is not yet understood at this stage but could be due to the presence of the indicator random variable $\mathbb{1}_{\{Y\leqslant q^\alpha(Y|\mathbf{X}_{-i})\}}$ in the index. Indeed, when studying, in reliability oriented sensitivity analysis (ROSA), the importance of inputs on the occurence of critical failure event given by $\mathbb{P}\left(Y > t\right)$, where $t \in \mathbb{R}$ represents a threshold characterizing the system, it is common to notice interactions between the variables even if the underlying model is linear. This happens, for instance, in Example 1 of Wei et al. (2012) or in the example considered in the Section 5.1 of Il Idrissi et al. (2021) when focusing on the indicator variable of interest $\mathbb{1}_{\{Y>t\}}$ instead of the model output Y.
Besides, 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* $Y = m_0 + \sum_{i=1}^{d} m_i(X_i)$ *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,*

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

Proof of Proposition 3.5.

We have:

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

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

Let

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

and for any $i = 1, \ldots, d$, let $Xs_{(-i)} = \sum_{\substack{1\leqslant j\leqslant d\\ j\neq i}} m_j(X_j)$. Due to the independence between the X_i's, we have

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

Then,

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

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 $\mathbb{E}\left[m_i(X_i)\mathbb{1}_{\left\{Xs_{(-i)}\leqslant q^\alpha\left(Xs_{(-i)}\right)\right\}}\right] = \alpha\mathbb{E}[m_i(X_i)]$. Thus,

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

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 behaviour 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

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

and the total QOSA indices

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

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

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

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 3 below presents the behaviour 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.

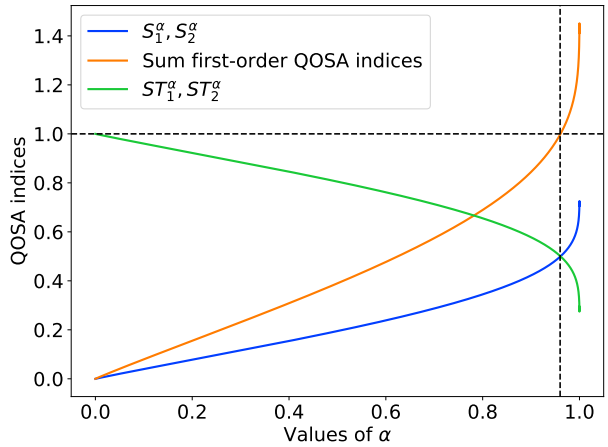

Figure 3: Evolution of the first-order and total QOSA indices at different levels α for the product of two exponentials with $\lambda = 1/10$ for the first input and $\delta = 1$ for the second one.

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

$$S_i^\alpha = 1 - \frac{\sqrt{\boldsymbol{\beta}_{-i}^\mathsf{T}\left(\boldsymbol{\Sigma}_{-i,-i} - \boldsymbol{\Sigma}_{-i,i}\boldsymbol{\Sigma}_{i,i}^{-1}\boldsymbol{\Sigma}_{i,-i}\right)\boldsymbol{\beta}_{-i}}}{\sigma_Y}\ , \tag{4.1}$$

$$ST_i^\alpha = \frac{|\boldsymbol{\beta}_i|\sqrt{\boldsymbol{\Sigma}_{i,i} - \boldsymbol{\Sigma}_{i,-i}\boldsymbol{\Sigma}_{-i,-i}^{-1}\boldsymbol{\Sigma}_{-i,i}}}{\sigma_Y}\ , \tag{4.2}$$

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

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 (4.1) and (4.2) 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

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

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

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

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

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

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

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

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

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

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 4 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.

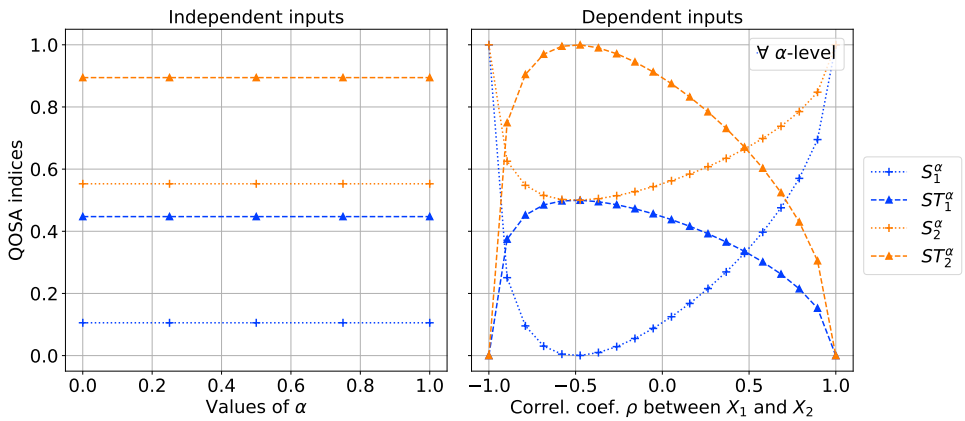

Figure 4: First-order and total QOSA indices with independent (resp. dependent) inputs on the left (resp. right) graph.

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

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

$$ST_i^\alpha = \frac{\alpha - \Phi\left(\Phi^{-1}(\alpha) - |\boldsymbol{\beta}_i|\sqrt{\boldsymbol{\Sigma}_{i,i} - \boldsymbol{\Sigma}_{i,-i}\boldsymbol{\Sigma}_{-i,-i}^{-1}\boldsymbol{\Sigma}_{-i,i}}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)}, \tag{4.4}$$

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 (4.3) and (4.4)

$$\begin{aligned} S_1^\alpha &= 1 - \frac{\alpha - \Phi\left(\Phi^{-1}(\alpha) - |\beta_2|\,\sigma_2\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)},\\ S_2^\alpha &= 1 - \frac{\alpha - \Phi\left(\Phi^{-1}(\alpha) - |\beta_1|\,\sigma_1\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)}, \end{aligned} \tag{4.5}$$

and

$$\begin{aligned} ST_1^\alpha &= \frac{\alpha - \Phi\left(\Phi^{-1}(\alpha) - |\beta_1|\,\sigma_1\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)},\\ ST_2^\alpha &= \frac{\alpha - \Phi\left(\Phi^{-1}(\alpha) - |\beta_2|\,\sigma_2\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)}. \end{aligned} \tag{4.6}$$

In all further tests, we take $\mu_1 = \mu_2 = 0, \beta_1 = \beta_2 = 1, \sigma_1 = 1$ and $\sigma_2 = 2$.
Figure 5 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, $ST_i^\alpha \leqslant S_i^\alpha,\ i = 1, 2$. This supports that Proposition 3.5 is not true outside the additive framework with independent inputs.
The behaviour 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 influential variable. Indeed, the index S_1^α increases faster than in independent case. On the contrary, the

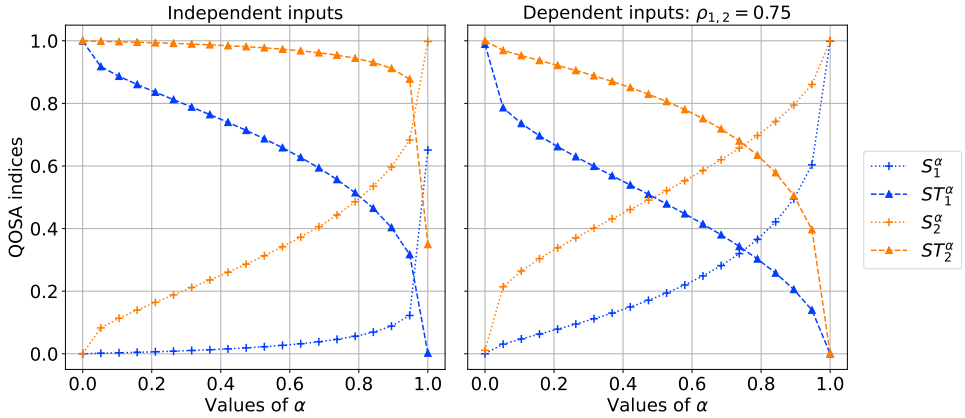

Figure 5: First-order and total QOSA indices with independent (resp. dependent) inputs on the left (resp. right) graph.

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 6, 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.

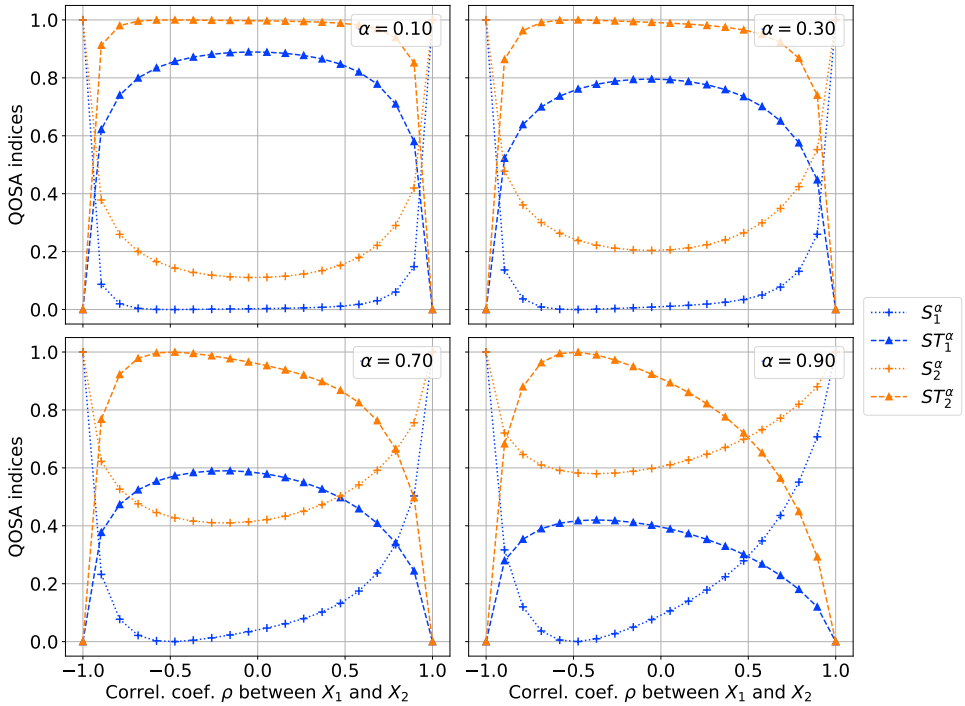

Figure 6: Evolution of the first-order and total QOSA indices at different values of ρ for several levels α.

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. These observations led us to turn to Shapley values which present good properties for both independent and dependent inputs. Indeed, they allocate fairly to each input its contribution due to dependence and interaction.

5 Goal- and 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

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

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 $\mathrm{Var}\left(Y\right)$ 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 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 $\mathcal{J}$:

$$\tilde{c}(\mathcal{J}) = \mathrm{Var}\left(\mathbb{E}\left[Y|\mathbf{X}_{\mathcal{J}}\right]\right) \text{ and } c(\mathcal{J}) = \mathbb{E}\left[\mathrm{Var}\left(Y|\mathbf{X}_{-\mathcal{J}}\right)\right] \ . \tag{5.1}$$

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 (5.1), the feature of interest of the output considered is the expectation denoted by $\theta^*\left(Y\right) = \mathbb{E}\left[Y\right]$. 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 $\theta^*\left(Y|\, X_{\mathcal{J}}\right) = \mathbb{E}\left[Y|\, \mathbf{X}_{\mathcal{J}}\right]$ for the first cost function and $\theta^*\left(Y|\, X_{-\mathcal{J}}\right) = \mathbb{E}\left[Y|\, \mathbf{X}_{-\mathcal{J}}\right]$ for the second one.

Let us now use $\theta^*\left(Y\right),\ \theta^*\left(Y|\, X_{\mathcal{J}}\right)$ and $\theta^*\left(Y|\, X_{-\mathcal{J}}\right)$ 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 $\theta^*\left(Y\right)$ is therefore assessed by measuring their contribution to the average contrast function $\mathbb{E}\left[\psi\left(Y,\ \theta^*\left(Y\right)\right)\right]$. 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 $\mathcal{J} \subseteq \mathcal{D}$

$$\tilde{c}(\mathcal{J}) = \mathbb{E}\left[\psi\left(\theta^*\left(Y|\, X_{\mathcal{J}}\right),\ \theta^*\left(Y\right)\right)\right] \text{ and } c(\mathcal{J}) = \mathbb{E}\left[\psi\left(Y,\ \theta^*\left(Y|\, X_{-\mathcal{J}}\right)\right)\right] \ . \tag{5.2}$$

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)*. The relevance of these new indices is, for instance, highlighted in Il Idrissi et al. (2021) in order to quantify the importance of

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.

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

independent or dependent inputs on a failure probability in a reliability assessment. Indeed, the (ℓ^2)-target Shapley effect $T - Sh_i^{\ell^2}, i = 1, \ldots, d$, proposed by the authors is a particular case of GOSE when considering the random variable $Z = \mathbb{1}_{\{Y>t\}}$ as model output instead of Y directly where $t \in \mathbb{R}$ represents a threshold characterizing the failure of the system. In this scheme, the feature of interest is the expectation of Z, i.e., $\theta^*(Z) = \mathbb{E}[Z] = \mathbb{P}(Y > t)$ and the use of the associated contrast functions (i.e., $\tilde{c}(\mathcal{J}) = \mathrm{Var}(\mathbb{E}[Z\vert\,\mathbf{X}_{\mathcal{J}}])$ or $c(\mathcal{J}) = \mathbb{E}[\mathrm{Var}(Z\vert\,\mathbf{X}_{-\mathcal{J}})]$) therefore allows to measure the contribution of each input to the failure probability of the system.

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 $c(\mathcal{J} \cup \{i\}) - c(\mathcal{J})$ is non-negative. This is a desirable property for cost functions which implies that the index Sh_i^α is non-negative and can be interpreted as a percentage of the global cost $\mathbb{E}[\psi(Y,\ \theta^*(Y))]$ allocated to the input $X_i, i = 1, \ldots, d$. Indeed, for $\mathcal{J} \subseteq \mathcal{D}\backslash\{i\}$, we have

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

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 $\Upsilon(Y)$ 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

$$Sh_i^\alpha = \frac{1}{d\cdot\sigma_Y}\sum_{\mathcal{J}\subseteq\mathcal{D}\setminus\{i\}}\binom{d-1}{|\mathcal{J}|}^{-1}\left[\sqrt{\boldsymbol{\beta}_{\mathcal{J}+i}^\mathsf{T}\left(\boldsymbol{\Sigma}_{\mathcal{J}+i,\mathcal{J}+i} - \boldsymbol{\Sigma}_{\mathcal{J}+i,-\mathcal{J}-i}\boldsymbol{\Sigma}_{-\mathcal{J}-i,-\mathcal{J}-i}^{-1}\boldsymbol{\Sigma}_{-\mathcal{J}-i,\mathcal{J}+i}\right)\boldsymbol{\beta}_{\mathcal{J}+i}}\right.$$
$$\left.-\sqrt{\boldsymbol{\beta}_{\mathcal{J}}^\mathsf{T}\left(\boldsymbol{\Sigma}_{\mathcal{J},\mathcal{J}} - \boldsymbol{\Sigma}_{\mathcal{J},-\mathcal{J}}\boldsymbol{\Sigma}_{-\mathcal{J},-\mathcal{J}}^{-1}\boldsymbol{\Sigma}_{-\mathcal{J},\mathcal{J}}\right)\boldsymbol{\beta}_{\mathcal{J}}}\right] \tag{5.3}$$

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

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

We have $\sigma_Y^2 = Var(Y) = \beta_1^2\sigma_1^2 + 2\rho\beta_1\beta_2\sigma_1\sigma_2 + \beta_2^2\sigma_2^2$ and obtain from(5.3)

$$\begin{aligned} Sh_1^\alpha &= \frac{1}{2} - \frac{|\beta_2|\,\sigma_2\sqrt{1-\rho^2}}{2\cdot\sigma_Y} + \frac{|\beta_1|\,\sigma_1\sqrt{1-\rho^2}}{2\cdot\sigma_Y}\ ,\\ Sh_2^\alpha &= \frac{1}{2} - \frac{|\beta_1|\,\sigma_1\sqrt{1-\rho^2}}{2\cdot\sigma_Y} + \frac{|\beta_2|\,\sigma_2\sqrt{1-\rho^2}}{2\cdot\sigma_Y}\ . \end{aligned} \tag{5.4}$$

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 QOSE are equal when the correlation is maximum (i.e. $|\rho| = 1$).

Figure 7 presents the first-order and total QOSA indices as well as 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 QOSE are constant and they are brackected 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, QOSE lie between the first-order

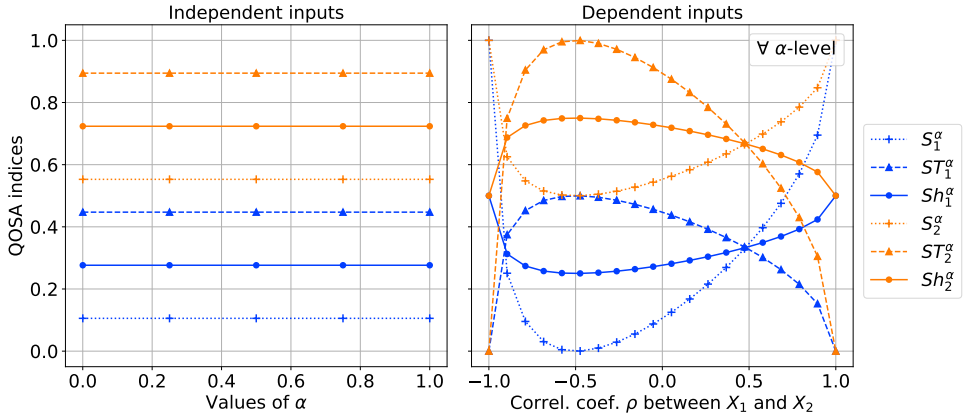

Figure 7: First-order and total QOSA indices as well as the QOSE with independent (resp. dependent) inputs on the left (resp. right) graph.

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.

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

$$Sh_i^\alpha = \frac{1}{d \cdot A} \sum_{\mathcal{J} \subseteq \mathcal{D}\setminus\{i\}} \binom{d-1}{|\mathcal{J}|}^{-1} \left[\Phi\left(\Phi^{-1}(\alpha) - B(\mathcal{J})\right) - \Phi\left(\Phi^{-1}(\alpha) - C(\mathcal{J}, i)\right)\right] \tag{5.5}$$

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 (5.5). Accordingly, we consider the particular case $d = 2$ with

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

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 (5.5)

$$\begin{aligned} Sh_1^\alpha &= \frac{1}{2} + \frac{1}{2} \cdot \frac{\Phi\left(\Phi^{-1}(\alpha) - |\beta_2|\,\sigma_2\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)} - \frac{1}{2} \cdot \frac{\Phi\left(\Phi^{-1}(\alpha) - |\beta_1|\,\sigma_1\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)}, \\ Sh_2^\alpha &= \frac{1}{2} + \frac{1}{2} \cdot \frac{\Phi\left(\Phi^{-1}(\alpha) - |\beta_1|\,\sigma_1\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)} - \frac{1}{2} \cdot \frac{\Phi\left(\Phi^{-1}(\alpha) - |\beta_2|\,\sigma_2\sqrt{1-\rho^2}\right)}{\alpha - \Phi\left(\Phi^{-1}(\alpha) - \sigma\right)}. \end{aligned} \tag{5.6}$$

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 QOSE are illustrated in Figure 8 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, 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 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.

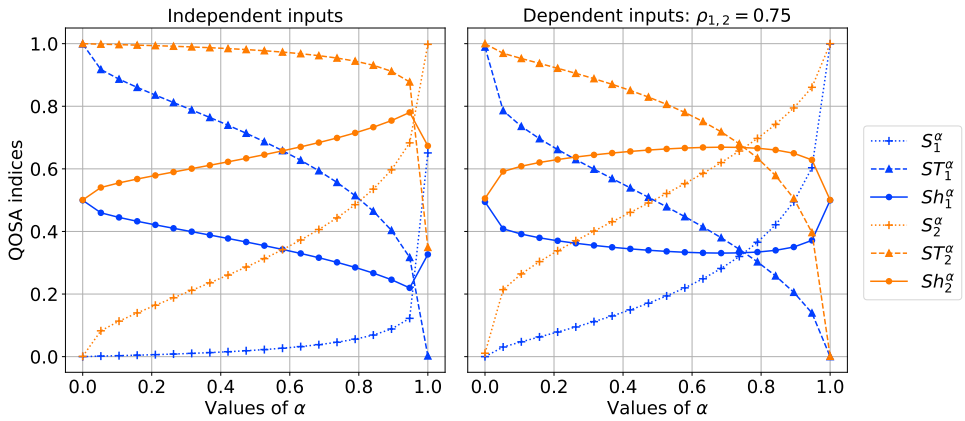

Figure 8: First-order and total QOSA indices as well as QOSE with independent (resp. dependent) inputs on the left (resp. right) graph.

Again, to get another insight on the impact of the dependence over the indices, we plot in Figure 9, for several levels α, the evolution of the latter as a function of the correlation coefficient. As explained before, 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.

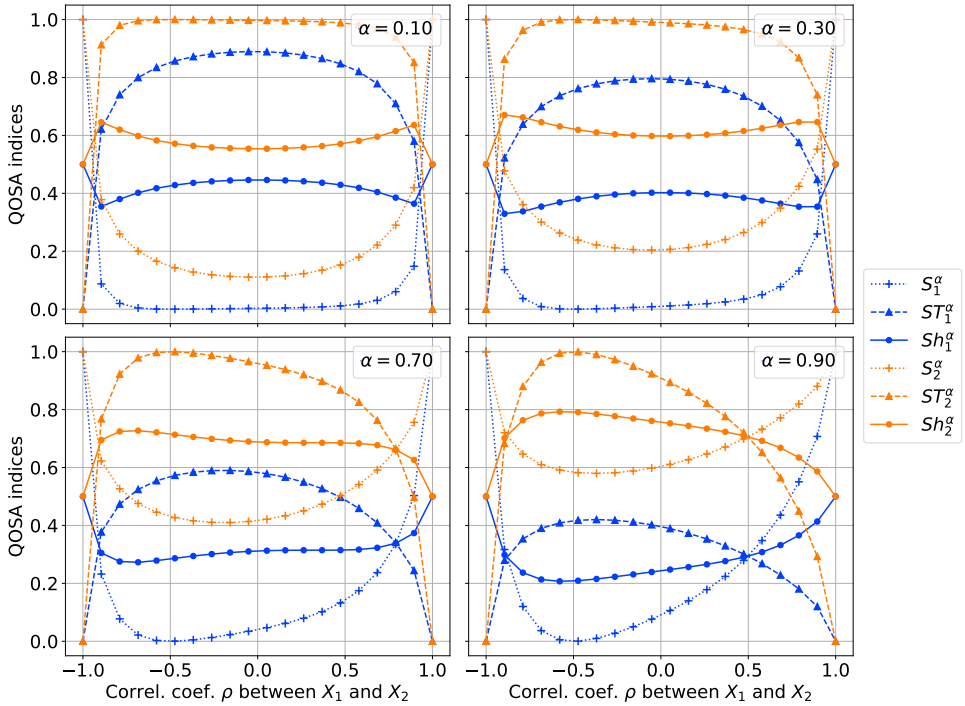

Figure 9: Evolution of the first-order and total QOSA indices as well as the quantile-oriented Shapley effects at different values of ρ for several levels α.

6 Cramér-von Mises distance-based Shapley effects

Various indices presented in the previous sections only allow to quantify the influence of the inputs on a feature of interest of the output Y such as the mean, a quantile, the median, etc. In order to measure the importance of the inputs on the whole distribution of the output, various moment independent importance measures based on either the density, the cumulative distribution function or the Csizár f-divergence, just to name a few, were developed in Borgonovo (2007); Borgonovo et al. (2011); Luyi et al. (2012); Da Veiga (2015). One can also mention the first-order indices recently introduced in Gamboa et al. (2018) based on the Cramér-von Mises distance to compare the distribution of the output Y to its conditional one which are defined by

$$S_{2,CVM}^{i} = \frac{\int_{\mathbb{R}} \mathbb{E}\left[\left(F^{i}(t) - F(t)\right)^{2}\right] dF(t)}{\int_{\mathbb{R}} F(t)\left(1 - F(t)\right) dF(t)}, \qquad i = 1, \ldots, d,$$

where F is the cumulative distribution function of Y

$$F(t) = \mathbb{P}\left(Y \leqslant t\right) = \mathbb{E}\left[\mathbb{1}_{\{Y \leqslant t\}}\right], \qquad \forall t \in \mathbb{R},$$

and F^{i} is the conditonal distribution function of Y given X_i

$$F^{i}(t) = \mathbb{P}\left(Y \leqslant t | X_i\right) = \mathbb{E}\left[\mathbb{1}_{\{Y \leqslant t\}} \middle| X_i\right] \qquad \forall t \in \mathbb{R}.$$

These indices arise from the application of Sobol' indices to the set of random variables $\left(Z(t) = \mathbb{1}_{\{Y \leqslant t\}}\right)_{t \in \mathbb{R}}$. However, as detailed in subsection 2.1, Sobol' indices rest on the Hoeffding decomposition which requires the independence assumption between the inputs. Therefore, Cramér-von Mises indices are well-designed and easy to interpret when the inputs are independent. When dealing with

dependent inputs, the latter can still be computed but their interpretation becomes difficult as noted in Remark 3.3 of Gamboa et al. (2018). Thus, it is proposed to overcome the limitation of Cramér-von Mises indices by defining a new index based on Shapley values which will allow to assess the contribution of independent or dependent inputs to the whole distribution of Y.

Similarly to the approach followed in Gamboa et al. (2018), the contribution of each input X_i will be first assessed for any point $t \in \mathbb{R}$ of the output distribution, i.e. on the following feature of interest:

$$\theta^* \left(Z\left(t\right)\right) = \mathbb{E}\left[Z\left(t\right)\right] = \mathbb{E}\left[\mathbb{1}_{\{Y \leqslant t\}}\right] = F(t), \text{ for all } t \in \mathbb{R} \ ,$$

by applying the framework presented in Section 5. In more detail, the influence of each input on the expectation of the random variable $Z\left(t\right) = \mathbb{1}_{\{Y \leqslant t\}}$ will be measured by decomposing $\mathrm{Var}\left(Z\left(t\right)\right) = F\left(t\right)\left(1 - F\left(t\right)\right)$ by means of the following cost functions associated to this feature of interest :

$$\tilde{c}\left(\mathcal{J}\right) = \mathrm{Var}\left(\mathbb{E}\left[Z\left(t\right)|\, \mathbf{X}_{\mathcal{J}}\right]\right) \qquad \text{and} \qquad c(\mathcal{J}) = \mathbb{E}\left[\mathrm{Var}\left(Z\left(t\right)|\, \mathbf{X}_{-\mathcal{J}}\right)\right] \ .$$

Hence, by plugging, for instance, the first cost function $\tilde{c}(\mathcal{J})$ within the Shapley values defined in (2.5), the following decomposition of $\mathrm{Var}\left(Z\left(t\right)\right)$ over the inputs is obtained:

$$\begin{aligned}
\mathrm{Var}\left(Z\left(t\right)\right) &= F\left(t\right)\left(1 - F\left(t\right)\right) \\
&= \sum_{i=1}^{d} \left[\frac{1}{d} \sum_{\mathcal{J} \subseteq \mathcal{D} \setminus \{i\}} \binom{d-1}{|\mathcal{J}|}^{-1} \left(\mathrm{Var}\left(\mathbb{E}\left[Z\left(t\right)|\, \mathbf{X}_{\mathcal{J} \cup \{i\}}\right]\right) - \mathrm{Var}\left(\mathbb{E}\left[Z\left(t\right)|\, \mathbf{X}_{\mathcal{J}}\right]\right)\right)\right] \\
&= \sum_{i=1}^{d} \left[\frac{1}{d} \sum_{\mathcal{J} \subseteq \mathcal{D} \setminus \{i\}} \binom{d-1}{|\mathcal{J}|}^{-1} \left(\mathbb{E}\left[\left(F^{\mathcal{J} \cup \{i\}}\left(t\right) - F\left(t\right)\right)^2\right] - \mathbb{E}\left[\left(F^{\mathcal{J}}\left(t\right) - F\left(t\right)\right)^2\right]\right)\right] \ .
\end{aligned} \tag{6.1}$$

Then, in order to quantify the impact of each input on the whole distribution of Y, the terms in (6.1) are integrated in $t \in \mathbb{R}$ with respect to the distribution of Y in the same way as in Gamboa et al. (2018):

$$\begin{aligned}
&\int_{\mathbb{R}} F\left(t\right)\left(1 - F\left(t\right)\right) dF\left(t\right) \\
&= \sum_{i=1}^{d} \left[\frac{1}{d} \sum_{\mathcal{J} \subseteq \mathcal{D} \setminus \{i\}} \binom{d-1}{|\mathcal{J}|}^{-1} \left(\int_{\mathbb{R}} \mathbb{E}\left[\left(F^{\mathcal{J} \cup \{i\}}\left(t\right) - F\left(t\right)\right)^2\right] dF\left(t\right) - \int_{\mathbb{R}} \mathbb{E}\left[\left(F^{\mathcal{J}}\left(t\right) - F\left(t\right)\right)^2\right] dF\left(t\right)\right)\right]
\end{aligned} \tag{6.2}$$

It is then recognized, in the right-hand side of (6.2), the Cramér-von Mises type distance of order 2 between the distribution of Y and the distribution of Y given $\mathbf{X}_{\mathcal{J} \cup \{i\}}$ and $\mathbf{X}_{\mathcal{J}}$ on which the indices introduced in Gamboa et al. (2018) are based. To ease the interpretation of the new indices introduced thereafter, the previous equation is normalized by $\int_{\mathbb{R}} F\left(t\right)\left(1 - F\left(t\right)\right) dF\left(t\right)$. That leads us to the following definition of the Cramér-von Mises distance-based Shapley effect.

Definition 6.1.
The Cramér-von Mises distance-based Shapley effect related to the input X_i is given by

$$Sh_i^{CVM} = \frac{1}{d} \sum_{\mathcal{J} \subseteq \mathcal{D} \setminus \{i\}} \binom{d-1}{|\mathcal{J}|}^{-1} \left(\tilde{c}\left(\mathcal{J} \cup \{i\}\right) - \tilde{c}\left(\mathcal{J}\right)\right)$$

with respect to the cost function

$$\tilde{c}\left(\mathcal{J}\right)=\frac{\int_{\mathbb{R}}\mathrm{Var}\left(\mathbb{E}\left[Z\left(t\right)|\,\mathbf{X}_{\mathcal{J}}\right]\right)dF\left(t\right)}{\int_{\mathbb{R}}\mathrm{Var}\left(Z\left(t\right)\right)dF\left(t\right)}=\frac{\int_{\mathbb{R}}\mathbb{E}\left[\left(F^{\mathcal{J}}\left(t\right)-F\left(t\right)\right)^{2}\right]dF\left(t\right)}{\int_{\mathbb{R}}F\left(t\right)\left(1-F\left(t\right)\right)dF\left(t\right)}\,.$$

These indices therefore complete GOSE previously introduced as they allow to quantify the importance of inputs on the whole distribution of the output and not to a specific feature of interest. It should be noted that these indices share several properties of the variance-based Shapley effects. Firstly, they are non-negative and sum up to one so that they can be interpreted as the percentage of the expectation on t of the variance of the indicator random variable $Z\left(t\right)=\mathbb{1}_{\{Y\leqslant t\}}$, i.e. $\mathbb{E}_{Y'}\left[\mathrm{Var}_{Y}\left(\mathbb{1}_{\{Y\leqslant Y'\}}\right)\right]$ where Y' is an independent copy of Y, allocated to the input X_i. It can also be shown that the Cramér-von Mises distance-based Shapley effects presented in Definition 6.1 above are equivalent to those defined by using the following cost function

$$c\left(\mathcal{J}\right)=\frac{\int_{\mathbb{R}}\mathbb{E}\left[\mathrm{Var}\left(Z\left(t\right)|\,\mathbf{X}_{-\mathcal{J}}\right)\right]dF\left(t\right)}{\int_{\mathbb{R}}\mathrm{Var}\left(Z\left(t\right)\right)dF\left(t\right)}=\frac{\int_{\mathbb{R}}\mathbb{E}\left[\left(\mathbb{1}_{\{Y\leqslant t\}}-F^{-\mathcal{J}}\left(t\right)\right)^{2}\right]dF\left(t\right)}{\int_{\mathbb{R}}F\left(t\right)\left(1-F\left(t\right)\right)dF\left(t\right)}\,.$$

by simply adapting the proof of Theorem 1 of Song et al. (2016). In addition, based on the results of Owen (2014), one can also show that these indices are bracketed by the first-order and total Cramér-von Mises indices introduced in Gamboa et al. (2018) when dealing with independent inputs. At last, it is worth noting that the previous definition of the Cramér-von Mises distance-based Shapley effects naturally extends to the case of multivariate outputs.

7 Estimation methods of QOSE and numerical experiments

An estimator of QOSE, Sh_i^{α}, defined in Section 5 is first proposed in this section followed by a numerical study of its performance on toy examples.

7.1 Estimation methods of QOSE

The estimators developed below are based on those defined in Song et al. (2016) for the variance-based Shapley effects tailored to our quantile-oriented context. Therefore, the former are briefly recalled below and their transposition to our framework will be naturally established.

First of all, a generic Shapley value with respect to a cost function c can also be expressed in terms of all possible permutations of the inputs. Let us denote by $\Pi\left(\mathcal{D}\right)$ the set of all possible permutations with the inputs set $\mathcal{D}$. Given a permutation $\pi\in\Pi\left(\mathcal{D}\right)$, define the set $P_i\left(\pi\right)$ as the inputs that precede the input i in π. Thus, a generic Shapley value can also be rewritten in the following way

$$v_i=\frac{1}{d!}\sum_{\pi\in\Pi(\mathcal{D})}\left(c\left(P_i\left(\pi\right)\cup\{i\}\right)-c\left(P_i\left(\pi\right)\right)\right)\,.$$

From this formula, Castro et al. (2009) proposed to estimate v_i with $\widehat{v}_i$ by drawing randomly m permutations in $\Pi\left(\mathcal{D}\right)$ and thus we have

$$\widehat{v}_i=\frac{1}{m}\sum_{l=1}^{m}\left(c\left(P_i\left(\pi\right)\cup\{i\}\right)-c\left(P_i\left(\pi\right)\right)\right)\,.$$

Section 4 of Song et al. (2016) proposed some improvements on the Castro's algorithm by including the Monte-Carlo estimation $\widehat{c}$ of the cost function $c(\mathcal{J}) = \mathbb{E}\left[\mathrm{Var}\left[Y|\,\mathbf{X}_{-\mathcal{J}}\right]\right]/\mathrm{Var}\left(Y\right)$ in order to estimate the normalized variance-based Shapley effects. The estimator writes

$$\widehat{Sh}_i = \frac{1}{m}\sum_{l=1}^{m}\left(\widehat{c}\left(P_i\left(\pi_l\right)\cup\{i\}\right) - \widehat{c}\left(P_i\left(\pi_l\right)\right)\right),$$

where m refers to the number of permutations. Song et al. (2016) proposed the following two algorithms, the main features of which are spelled out below:

- The *exact permutation method* if d is small, one does all possible permutations between the inputs (i.e. $m = d!$);
- The *random permutation method* which consists in randomly sampling m permutations of the inputs in $\Pi(\mathcal{D})$.

For each iteration of this loop on the inputs' permutations, a conditional variance expectation must be computed. The cost C of these algorithms is the following $C = N_v + m(d-1)N_oN_i$ with N_v the sample size for the variance computation of Y, N_o the outer loop size for the expectation, N_i the inner loop size for the conditional variance of Y and m the number of permutations according to the selected method.

In our quantile-oriented context, the cost C of these algorithms remains unchanged, i.e., $C = N_v + m(d-1)N_oN_i$ with m the number of permutations according to the selected method. But, now as illustrated in Table 1 from Section 5,

- the sample of size N_v is used to estimate the α-quantile of Y and compute the average contrast function $\Upsilon\left(Y\right)$,
- N_o the outer loop size for the expectation of the second cost function,
- and N_i the inner loop size to estimate the conditional quantile of Y given $\mathbf{X}_{-\mathcal{J}}$ as well as the conditional average contrast function $\mathbb{E}\left[\psi\left(Y,\ q^{\alpha}\left(Y|\,\mathbf{X}_{-\mathcal{J}}\right)\right)|\,\mathbf{X}_{-\mathcal{J}}\right]$.

The estimators of QOSE obtained from these tailored algorithms will be denoted $\widehat{Sh}_i^{\alpha}$ in the sequel. Likewise, the strategies based on a bootstrap sampling implemented in Benoumechiara and Elie-Dit-Cosaque (2019) to estimate confidence intervals for the variance-based Shapley effects remain valid for QOSE.

Optimal values, based on theoretical results, for the parameters of the exact and random permutation methods were given in Song et al. (2016) when estimating variance-based Shapley effects. Unfortunately, we do not have such theoretical results at this stage and it seems reasonable to think that these are obsolete in our framework as the complexity is definitely not the same to estimate an (resp. conditional) expectation and a (resp. conditional) quantile. Accordingly, by using a toy example, we empirically study how the algorithm settings can influence the estimation of QOSE in the next subsection.

7.2 Numerical experiments

As defined previously, three parameters of the algorithm govern the estimation accuracy of QOSE: N_v, N_o and N_i. The average contrast functions are estimated through Monte Carlo procedures. $\Upsilon\left(Y\right)$ is computed from a sample $\{Y^j = \eta\left(\mathbf{X}^j\right)\}_{j=1,\ldots,N_v}$. Because N_v is a small proportion of the overall cost $C = N_v + m(d-1)N_oN_i$, especially when d is large, N_v can be selected as large as possible in order to reach the smallest possible estimation error of $\Upsilon\left(Y\right)$. However, it is more difficult to choose N_o and N_i to estimate the expectation of the conditional

average contrast functions. These choices also depend on the used algorithm: exact or random permutations.

Therefore, we empirically study the influence of N_o and N_i on the estimation error and the coverage probability for several levels α. The Probability Of Coverage (POC) is defined as the probability to have the true value of the index inside the estimated confidence intervals. We consider the bidimensional model with Gaussian input variables and a lognormal output of Section 5.2 as a toy example with independent inputs, $\beta_1 = \beta_2 = 1$, $\sigma_1 = 1$ and $\sigma_2 = 2$. The POC is estimated with 100 independent algorithm runs and for a 90 % confidence interval. The confidence intervals are estimated with 500 bootstrap sampling. We also set a large value of $N_v = 10^5$ for all the experiments.

First experiments aim to assess the influence of N_o on the estimation accuracy and the POC for the exact permutation algorithm. The Figure 10 shows the variation of the POC (solid lines) and the absolute error (dashed lines) for three α levels, averaged over the two indices, as a function of the product mN_oN_i, where only N_o is varying and for three values of N_i at 200, 1500 and 3000. Because the errors are computed for 100 independent runs, we show in color areas the 95% quantiles.

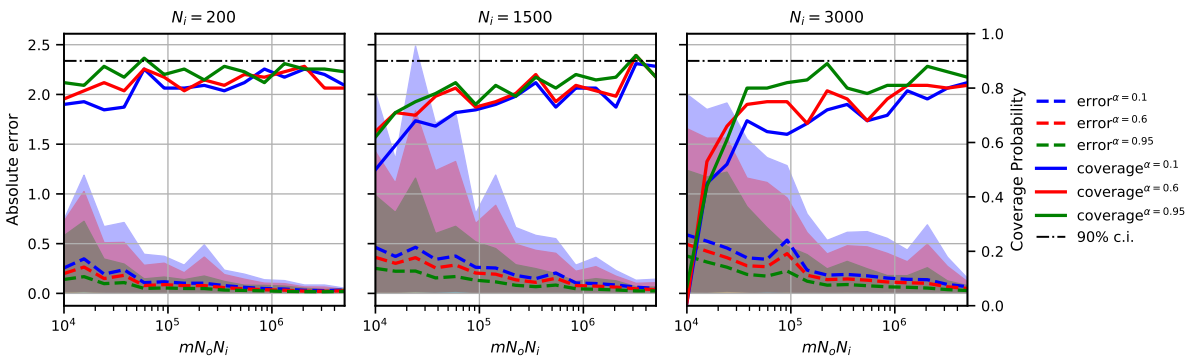

Figure 10: Variation of the absolute error and the POC with N_o for three values of $N_i = 200, 1500, 3000$ for the exact permutation algorithm ($m = d! = 2$).

We observe that the estimation error is the smallest for the three α levels with $N_i = 200$. One can also notice that the POC tends, at different rates, to the true probability: 90 %. Nonetheless, when looking at thoroughly the graph with $N_i = 200$, it appears that the POC is a little underestimated compared to the plot with $N_i = 1500$ (with the largest calculation budget). This could be due to the too small value of N_i leading to somewhat biased indices. In our scheme, for a same computational cost mN_oN_i, the smaller the value of N_i and the larger the value of N_o. Thus, these results show that in order to have a correct confidence interval, it is important to have both a sufficiently high value of N_i but also a large value of N_o compared to N_i. Indeed, exploring multiple conditional average contrast functions with a reasonable precision (large N_o and low N_i) is more important than having less conditional average contrast functions with a large precision (low N_o and large N_i).

The Figure 11 is similar to Figure 10 but for the random permutation algorithm and by fixing $N_o = 10$ and by varying the number of permutations.

As for the exact permutation algorithm, we can see that the estimation errors are the smallest with $N_i = 200$ and the difference is shown for the POC. We observe that the lower N_i and the faster the POC converges to the true probability. Indeed, for a same computational cost, the

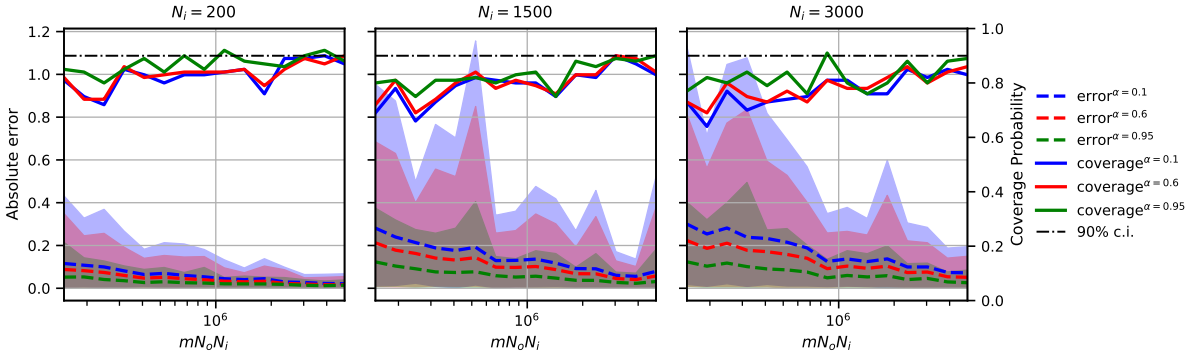

Figure 11: Variation of the absolute error and the POC with m for three values of $N_i = 200, 1500, 3000$ and $N_o = 10$ for the random permutation algorithm.

lower N_i and the larger the number of permutations m can be.

To show the influence of N_o with the random permutation algorithm, the Figure 12 is the same as Figure 11 but with $N_o = 50$. We observe that the convergence rates of the POC are slower than the ones for $N_o = 10$. Thus, it shows that having a lower value of N_o and a large value of m is more important to have consistent confidence intervals.

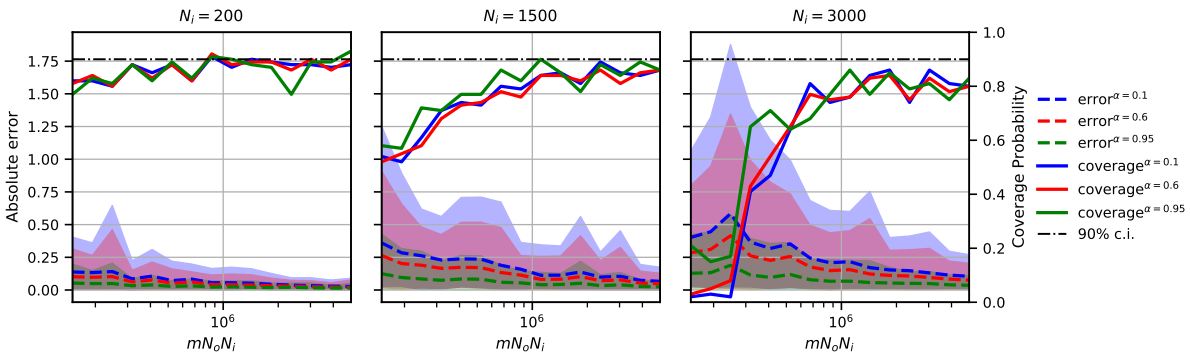

Figure 12: Variation of the absolute error and the POC with m for three values of $N_i = 200, 1500, 3000$ and $N_o = 50$ for the random permutation algorithm.

From these experiments, we can conclude that the parametrization does not significantly influence the estimation error but has a strong influence on the POC. Therefore, in order to have consistent confidence intervals, we can suggest:

- for the exact permutation algorithm to consider $N_i = 400$ and to take N_o as large as possible,
- for the random permutation algorithm to consider $N_i = 300$, $N_o = 10$ and take m as large as possible.

8 Conclusion

This paper presents a set of new indices named GOSE allowing to assess the contribution of the dependent or independent inputs on various features of the output such as quantile, threshold exceedance probabilities, etc. These are based on two main concepts. On the one hand, the

Shapley values developed in cooperative game theory which enables to fairly allocate the interaction and dependence effects between the inputs. On the other hand, the use of a cost function, within the Shapley values, based on the contrast function related to the feature of interest in order to properly measure the influence of the inputs. When using the contrast function attached to the expectation of the output, GOSE equal the variance-based Shapley effects developed in GSA which quantify the impact of the inputs on the expectation. Accordingly, GOSE can be seen as a generalization of the latter which enable to quantify the contribution of the inputs on any quantity of interest.

When considering the quantile of the output as feature of interest, 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. 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.
A first estimator has been proposed to compute these new indices. Although it performs well with the right settings, it is quite cumbersome to handle in practice. Indeed, it requires the ability to sample from the conditional distributions of all possible subsets of inputs but also a large calculation budget to properly estimate the conditional quantiles as well as the conditional average contrast functions. These shortcomings prevent the use of such a method in practice because most of the numerical models used in industrial studies are costly to evaluate and the inputs distribution is not known. To overcome these limitations, a first step could be to extend the estimator of the variance-based Shapley effects proposed in Broto et al. (2020) which uses a single input-output sample without requiring the knowledge of the inputs distribution for the conditional terms. Then, trying to tailor to our quantile-oriented context the efficient algorithm developed in Plischke et al. (2021) to reduce the run time of QOSE.

In addition of GOSE and QOSE, Cramér-von Mises distance-based Shapley effects were also introduced to measure the contribution of the dependent or independent inputs to the whole distribution of the output and not anymore on a specific quantity of interest. These are an extension of the first-order indices defined in Gamboa et al. (2018) which can be difficult to interpret in the case of dependent inputs. Although these indices seem promising, it would be interesting to perform studies similar to those carried out for the variance-based Shapley effects (Song et al., 2016; Owen and Prieur, 2017; Iooss and Prieur, 2019; Benoumechiara and Elie-Dit-Cosaque, 2019) in order to understand exactly how interaction and dependency effects are taken into account and therefore, interpret them correctly. Regarding their estimation, leveraging the efficient estimator, based on the Chatterjee's notion of correlation, of the Cramér-von Mises indices presented in Gamboa et al. (2020) would be an attractive starting point.

At last, Rabitti and Borgonovo (2019) have recently introduced Shapley-Owen interaction effects which are a generalization of the variance-based Shapley effects to quantify the interaction effects between the inputs. A future work could be to adapt these results to the goal-oriented framework as well as to the Cramér-von Mises distance-based Shapley effects.

Acknowledgments.

We thank the two anonymous reviewers as well as the associate editor for their fruitful and detailed suggestions and comments which led to an improvement of our paper. We are also grateful to Andrés Cuberos, Ecaterina Nisipasu and Mathieu Poulin from SCOR for their valuable comments and support.

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. (2017). Sensitivity analysis. *An Introduction for the Management Scientist. International Series in Operations Research and Management Science. Cham, Switzerland: Springer*.

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., Hazen, G. B., Jose, V. R. R., and Plischke, E. (2021). Probabilistic sensitivity measures as information value. *European Journal of Operational Research*, 289(2):595–610.

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

Broto, B., Bachoc, F., and Depecker, M. (2020). Variance reduction for estimation of Shapley effects and adaptation to unknown input distribution. *SIAM/ASA Journal on Uncertainty Quantification*, 8(2):693–716.

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.

Castro, J., Gómez, D., and Tejada, J. (2009). Polynomial calculation of the Shapley value based on sampling. *Computers & Operations Research*, 36(5):1726–1730.

Chastaing, G., Gamboa, F., and Prieur, C. (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. A. and Mackenzie, W. A. (1923). Studies in crop variation. II. The manurial response of different potato varieties. *The Journal of Agricultural Science*, 13(3):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.

Gamboa, F., Gremaud, P., Klein, T., and Lagnoux, A. (2020). Global Sensitivity Analysis: a new generation of mighty estimators based on rank statistics. *arXiv preprint arXiv:2003.01772*.

Gamboa, F., Klein, T., and Lagnoux, A. (2018). Sensitivity Analysis Based on Cramér–von Mises Distance. *SIAM/ASA Journal on Uncertainty Quantification*, 6(2):522–548.

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 non-linear models. *Reliability Engineering & System Safety*, 52(1):1–17.

Il Idrissi, M., Chabridon, V., and Iooss, B. (2021). Developments and applications of Shapley effects to reliability-oriented sensitivity analysis with correlated inputs. *Environmental Modelling & Software*, 143:105115.

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.

Kolmogoroff, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer Verlag, Berlin.

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.

Luyi, L., Zhenzhou, L., Jun, F., and Bintuan, W. (2012). Moment-independent importance measure of basic variable and its state dependent parameter solution. *Structural Safety*, 38:40–47.

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 évaluation 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.

Pearson, K. (1905). *On the General Theory of Skew Correlation and Non-linear Regression*, volume XIV of *Mathematical Contributions to the Theory of Evolution, Drapers' Company Research Memoirs*. Dulau & Co., London.

Plischke, E. (2012). An adaptive correlation ratio method using the cumulative sum of the reordered output. *Reliability Engineering & System Safety*, 107:149–156.

Plischke, E., Rabitti, G., and Borgonovo, E. (2021). Computing Shapley Effects for Sensitivity Analysis. *SIAM/ASA Journal on Uncertainty Quantification*, 9(4):1411–1437.

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*, 22(1):118–171.

Wei, P., Lu, Z., Hao, W., Feng, J., and Wang, B. (2012). Efficient sampling methods for global reliability sensitivity analysis. *Computer Physics Communications*, 183(8):1728–1743.