HAL open science

Global sensitivity analysis and Wasserstein spaces

Jean-Claude Fort, Thierry Klein, Agnès Lagnoux

To cite this version:

Jean-Claude Fort, Thierry Klein, Agnès Lagnoux. Global sensitivity analysis and Wasserstein spaces. 2020. hal-02905753v1

HAL Id: hal-02905753

https://hal.science/hal-02905753v1

Preprint submitted on 23 Jul 2020 (v1), last revised 15 Jan 2021 (v2)

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.

Global sensitivity analysis and Wasserstein spaces

Jean-Claude Fort[1], Thierry Klein[2], and Agnès Lagnoux[3]

[1]MAP5 Université Paris Descartes, SPC, 45 rue des Saints Pères, 75006 Paris, France.
[2]Institut de Mathématiques de Toulouse; UMR5219. Université de Toulouse; ENAC - Ecole Nationale de l'Aviation Civile , Université de Toulouse, France
[3]Institut de Mathématiques de Toulouse; UMR5219. Université de Toulouse; CNRS. UT2J, F-31058 Toulouse, France.

July 23, 2020

Abstract

Sensitivity indices are commonly used to quantity the relative influence of any specific group of input variables on the output of a computer code. In this paper, we focus both on computer codes the output of which is a cumulative distribution function and on stochastic computer codes. We propose a way to perform a global sensitivity analysis for these kinds of computer codes. In the first setting, we define two indices: the first one is based on Wasserstein Fréchet means while the second one is based on the Hoeffding decomposition of the indicators of Wasserstein balls. Further, when dealing with the stochastic computer codes, we define an "ideal version" of the stochastic computer code thats fits into the frame of the first setting. Finally, we deduce a procedure to realize a second level global sensitivity analysis, namely when one is interested in the sensitivity related to the input distributions rather than in the sensitivity related to the inputs themselves. Several numerical studies are proposed as illustrations in the different settings.

Keywords: Global sensitivity indices, functional computer codes, stochastic computer codes, second level uncertainty, Fréchet means, Wasserstein spaces.

AMS subject classification 62G05, 62G20, 62G30, 65C60, 62E17.

1 Introduction

The use of complex computer models for the analysis of applications from sciences, engineering and other fields is by now routine. For instance, in the area of marine submersion, complex computer codes have been developed to simulate submersion events (see, e.g., [3, 34] for more details). In the context of aircraft design, sensitivity analysis and metamodelling are intensively used to optimize the design of an airplane (see, e.g., [51]). Several other concrete examples of stochastic computer codes can be found in [42].

Often, the models are expensive to run in terms of computational time. Thus it is crucial to understand the global influence of one or several inputs on the output of the system under study with a moderate number of runs afforded [53]. When these inputs are regarded as random elements, this problem is generally called (global) sensitivity analysis. We refer to [16, 52, 56] for an overview of the practical aspects of global sensitivity analysis.

A classical tool to perform global sensitivity analysis consists in computing the Sobol indices. These indices were first introduced in [50] and then considered by [55]. They are well tailored when the output space is $\mathbb{R}$. The Sobol indices compare, using the Hoeffding decomposition [33], the conditional variance of the output knowing some of the input variables to the total variance of the output. Many different estimation procedures of the Sobol indices have been proposed and studied in the literature. Some are based on Monte-Carlo or quasi Monte-Carlo design of experiments (see [38, 47] and references therein for more details). More recently a method based on nested Monte-Carlo [28] has been developed. In particular, an efficient estimation of the Sobol indices can be performed through the so-called Pick-Freeze method. For the description of this method and its theoretical study (consistency, central limit theorem, concentration inequalities and Berry- Esseen bounds), we refer to [36, 25] and references therein. Some

other estimation procedures are based on different designs of experiments using for example polynomial chaos expansions (see [57] and the reference therein for more details).

Since Sobol indices are variance based, they only quantify the influence of the inputs on the mean behaviour of the code. Many authors proposed other criteria to compare the conditional distribution of the output knowing some of the inputs to the distribution of the output. In [47, 49, 48], the authors use higher moments to define new indices while, in [6, 7, 15], the use of divergences or distances between measures allows to define new indices. In [20], the authors use contrast functions to build indices that are goal oriented. Although these works define nice theoretical indices, the existence of a relevant statistical estimation procedure is still in most cases an open question. The case of vectorial-valued computer codes is considered in [26] where a sensitivity index based on the whole distribution utilizing the Cramér-von-Mises distance is defined. Within this framework, the authors show that the Pick-Freeze estimation procedure provides an asymptotically Gaussian estimator of the index. The definition of the Cramér-von-Mises indices has been extended to computer codes valued in general metric spaces in [21, 27].

Nowadays, the computer code output is often no longer a real-valued multidimensional variable but rather a function computed at various locations. In that sense, it can be considered as a functional output. Some other times, the computer code is stochastic in the sense that the same inputs can lead to different outputs. When the output of the computer code is a function (for instance, a cumulative distribution function) or when the computer code is stochastic, Sobol indices are no longer well tailored. It is then crucial to define indices adapted to the functional or random aspect of the output. When the output is vectorial or valued in an Hilbert space some generalizations of Sobol indices are available [39, 24]. Nevertheless, these indices are still based on the Hoeffding decomposition of the output; so that they only quantify the relative influence of an input through the variance. More recently, indices based on the whole distribution have been developed [15, 8, 6]. In particular, the method relying on Cramér-von-Mises distance [26] compares the conditionnal cumulative distribution function with the unconditional one by considering the Hoeffding decomposition of half-space indicators (rather than the Hoeffding decomposition of the output itself) and by integrating them. This method was then extend to codes taking values in a Riemannian manifold [21] and then in general metric spaces [27].

In this work, we focus on two kinds of computer codes: 1) computer codes the output of which is the cumulative distribution function of a real random variable and 2) real-valued stochastic computer codes. A first step will consist in performing global sensitivity analysis for these kinds of computer codes. Further, we will deduce how to perform second level sensitivity analysis using the tools developed in the first step. A code with cumulative distribution function as output can be seen as a code taking values in the space of all probability measures on $\mathbb{R}$. This space can be endowed with a metric (for example, the Wasserstein metric [59]). This point of view allows to define at least two different indices for this kind of codes, generalizing the framework of [27]. The first one is based on Wasserstein Fréchet means while the second one is based on the Hoeffding decomposition of the indicators of Wasserstein balls. Further, stochastic codes (see Section 5 for a bibliographical study) can be seen as a "discrete approximation" of codes having cumulative distribution functions as values. Then it is possible to define "natural" indices for such stochastic codes. Finally, second level sensitivity analysis aims at considering uncertainties on the type of the input distributions and/or on the parameters of the input distributions (see Section 6 for a bibliographical study). Actually, this kind of problem can be embedded in the framework of stochastic codes.

The article is organized as follows. In Section 2, we introduce and precisely define a general class of global sensitivity indices. We also present statistical methods to estimate these indices. In Section 3, we recall some basic facts on Wasserstein distances, Wasserstein costs and Fréchet means. In Section 4, we define and study the statistical properties of two new global sensitivity indices for computer codes valued in general Wasserstein spaces. Further, in Section 5, we study the case of stochastic computer codes. Finally, Section 6 is dedicated to the sensitivity analysis with respect to the distributions of the input variables.

2 Sensitivity indices for codes valued in general metric spaces

We consider a black-box code f defined on a product of measurable spaces $E = E_1 \times E_2 \times \ldots \times E_p$ ($p \in \mathbb{N}^*$) taking its values in a metric space $\mathcal{X}$. The output denoted by Z is then given by

$$Z = f(X_1, \ldots, X_p). \tag{1}$$

We denote by $\mathbb{P}$ the distribution of the output code Z.

The aim of this work is to give some partial answers to the following questions.

Question 1 How can we perform Global Sensitivity Analysis (GSA) when the output space is the space of probability distribution functions (p.d.f.) on $\mathbb{R}$ or the space of cumulative distribution functions (c.d.f.)?

Question 2 How can we perform GSA for stochastic computer codes?

Question 3 How can we perform GSA with respect to the choice of the distributions of the input variables?

2.1 The general metric spaces sensitivity index

In [27], the authors performed GSA for codes f taking values in general metric spaces. To do so, they consider a family of test functions parameterized by $m \in \mathbb{N}^*$ elements of $\mathcal{X}$ and defined by

$$\begin{array}{ccc} \mathcal{X}^m \times \mathcal{X} & \to & \mathbb{R} \\ (a, x) & \mapsto & T_a(x). \end{array}$$

Let $\mathbf{u} \subset \{1, \ldots, p\}$ and $X_{\mathbf{u}} = (X_i, i \in \mathbf{u})$. Assuming that the test functions T_a are L^2-functions with respect to the product measure $\mathbb{P}^{\otimes m} \otimes \mathbb{P}$ (where $\mathbb{P}^{\otimes m}$ is the product m-times of the distribution of the output code Z) on $\mathcal{X}^m \times \mathcal{X}$, they allow to defined the general metric space sensitivity index with respect to $X_{\mathbf{u}}$ by

$$S^{\mathbf{u}}_{2,\mathrm{GMS}} = \frac{\int_{\mathcal{X}^m} \mathbb{E}\left[\left(\mathbb{E}[T_a(Z)] - \mathbb{E}[T_a(Z)|X_{\mathbf{u}}]\right)^2\right] d\mathbb{P}^{\otimes m}(a)}{\int_{\mathcal{X}^m} \mathrm{Var}(T_a(Z)) d\mathbb{P}^{\otimes m}(a)} = \frac{\int_{\mathcal{X}^m} \mathrm{Var}\left(\mathbb{E}[T_a(Z)|X_{\mathbf{u}}]\right) d\mathbb{P}^{\otimes m}(a)}{\int_{\mathcal{X}^m} \mathrm{Var}(T_a(Z)) d\mathbb{P}^{\otimes m}(a)}. \quad (2)$$

Roughly speaking, there are two parts in the previous indices. First, for any value of a, we consider the numerator $\mathbb{E}\left[\left(\mathbb{E}[T_a(Z)] - \mathbb{E}[T_a(Z)|X_{\mathbf{u}}]\right)^2\right]$ and the denominator $\mathrm{Var}(T_a(Z))$ of the classical Sobol index of $T_a(Z)$. We call this part the Sobol' part. Second, we integrate each part with respect to the measure $\mathbb{P}^{\otimes m}$; this will be called the integration part.

As explained in [27], by construction, the indices $S^{\mathbf{u}}_{2,\mathrm{GMS}}$ lie in $[0, 1]$ and share the same properties as their Sobol counterparts:

1. the different contributions sum to 1;
2. they are invariant by translation, by any isometric and by any non-degenerated scaling of Z.

Estimation Three different estimation procedures are available in this context. The two first methods are based on the so-called Pick-Freeze scheme. More precisely, the Pick-Freeze scheme, considered in [36], is a well tailored design of experiment. Namely, let $X^{\mathbf{u}}$ be the random vector such that $X^{\mathbf{u}}_i = X_i$ if $i \in \mathbf{u}$ and $X^{\mathbf{u}}_i = X'_i$ if $i \notin \mathbf{u}$ where X'_i is an independent copy of X_i. We then set

$$Z^{\mathbf{u}} := f(X^{\mathbf{u}}). \quad (3)$$

Further, the procedure consists in rewriting the variances of the conditional expectation in terms of covariances as follows:

$$\mathrm{Var}(\mathbb{E}[Z|X^{\mathbf{u}}]) = \mathrm{Cov}(Z, Z^{\mathbf{u}}). \quad (4)$$

Alternatively, the third estimation procedure that can be seen as an ingenious and effective approximation of the Pick-Freeze scheme is based on rank statistics. Until now, it is unfortunately only available to estimate first order indices in the case of real-valued inputs.

- <u>First method - Pick-Freeze</u>. Introduced in [26], this procedure is based on a double Monte-Carlo scheme to estimate the Cramér-von-Mises indices $S^{\mathbf{u}}_{2,\mathrm{CVM}}$. More precisely, to estimate $S^{\mathbf{u}}_{2,\mathrm{GMS}}$ in our context, one considers the following design of experiment consisting in:
  1. a classical Pick-Freeze N-sample, that is two N-samples of Z: $(Z_j, Z^{\mathbf{u}}_j)$, $1 \leqslant j \leqslant N$;

2. m another N-samples of Z independent of $(Z_j, Z_j^{\mathbf{u}})_{1\leqslant j\leqslant N}$: $W_{l,k}$, $1\leqslant l\leqslant m$, $1\leqslant k\leqslant N$.

The empirical estimator of the numerator of $S^{\mathbf{u}}_{2,\mathrm{GMS}}$ is then given by

$$\begin{aligned}\widehat{N}^{\mathbf{u}}_{2,\mathrm{GMS,PF}} =& \frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left[\frac{1}{N}\sum_{j=1}^{N} T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j)T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j^{\mathbf{u}})\right]\\ &-\frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left[\frac{1}{2N}\sum_{j=1}^{N}\left(T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j)+T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j^{\mathbf{u}})\right)\right]^2\end{aligned}$$

while the one of the denominator is

$$\begin{aligned}\widehat{D}^{\mathbf{u}}_{2,\mathrm{GMS,PF}} =& \frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left[\frac{1}{2N}\sum_{j=1}^{N}\left(T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j)^2+T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j^{\mathbf{u}})^2\right)\right]\\ &-\frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left[\frac{1}{2N}\sum_{j=1}^{N}\left(T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j)+T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j^{\mathbf{u}})\right)\right]^2.\end{aligned}$$

For $\mathcal{X}=\mathbb{R}^k$, $m=1$, and T_a given by $T_a(x)=\mathbb{1}_{\{x\leqslant a\}}$, the index $S^{\mathbf{u}}_{2,\mathrm{GMS,PF}}$ is nothing more than the index $S^{\mathbf{u}}_{2,\mathrm{CVM}}$ defined in [26] based on the Cramér-von-Mises distance and the whole distribution of the output. Its estimator $\widehat{S}^{\mathbf{u}}_{2,\mathrm{CVM}}$ defined as the ratio of $\widehat{N}^{\mathbf{u}}_{2,\mathrm{GMS,PF}}$ and $\widehat{D}^{\mathbf{u}}_{2,\mathrm{GMS,PF}}$ with $T_a(x)=\mathbb{1}_{\{x\leqslant a\}}$ has been proved to be asymptotically Gaussian [26, Theorem 3.8]. The proof relies on Donsker's theorem and the functional delta method [58, Theorem 20.8]. Hence, in the general case of $S^{\mathbf{u}}_{2,\mathrm{GMS}}$, the central limit theorem will be still valid as soon as the collection $(T_a)_{a\in\mathcal{X}^m}$ forms a Donsker's class of functions.

- Second method - U-statistics. As done in [27], this method allows the practitioner to get rid of the additional random variables $(W_{l,k})$ for $l\in\{1,\dots,m\}$ and $k\in\{1,\dots,N\}$. The estimator is now based on U-statistics and deals simultaneously with the Sobol part and the integration part with respect to $d\mathbb{P}^{\otimes m}(a)$. It suffices to rewrite $S^{\mathbf{u}}_{2,\mathrm{GMS}}$ as

$$S^{\mathbf{u}}_{2,\mathrm{GMS}} = \frac{I(\Phi_1)-I(\Phi_2)}{I(\Phi_3)-I(\Phi_4)}, \tag{5}$$

where,

$$\begin{aligned}\Phi_1(\mathbf{z}_1,\dots,\mathbf{z}_{m+1}) &= T_{z_1,\dots,z_m}(z_{m+1})T_{z_1,\dots,z_m}(z^{\mathbf{u}}_{m+1}),\\ \Phi_2(\mathbf{z}_1,\dots,\mathbf{z}_{m+2}) &= T_{z_1,\dots,z_m}(z_{m+1})T_{z_1,\dots,z_m}(z^{\mathbf{u}}_{m+2}),\\ \Phi_3(\mathbf{z}_1,\dots,\mathbf{z}_{m+1}) &= T_{z_1,\dots,z_m}(z_{m+1})^2,\\ \Phi_4(\mathbf{z}_1,\dots,\mathbf{z}_{m+2}) &= T_{z_1,\dots,z_m}(z_{m+1})T_{z_1,\dots,z_m}(z_{m+2}),\end{aligned} \tag{6}$$

denoting by $\mathbf{z}_i$ the pair $(z_i, z_i^{\mathbf{u}})$ and, for $l=1,\dots,4$,

$$I(\Phi_l) = \int_{\mathcal{X}^{m(l)}} \Phi_l(\mathbf{z}_1,\dots,\mathbf{z}_{m(l)})d\mathbb{P}_2^{u,\otimes m(l)}(\mathbf{z}_1\dots,\mathbf{z}_{m(l)}), \tag{7}$$

with $m(1)=m(3)=m+1$ and $m(2)=m(4)=m+2$. Finally, one considers the empirical version of (5) as estimator of $S^{\mathbf{u}}_{2,\mathrm{GMS}}$:

$$\widehat{S}^{\mathbf{u}}_{2,\mathrm{GMS,Ustat}} = \frac{U_{1,N}-U_{2,N}}{U_{3,N}-U_{4,N}}, \tag{8}$$

where, for $l=1,\dots,4$,

$$U_{l,N} = \binom{N}{m(l)}^{-1}\sum_{1\leqslant i_1<\cdots<i_{m(l)}\leqslant N}\Phi_l^s\left(\mathbf{Z}_{i_1},\dots,\mathbf{Z}_{i_{m(l)}}\right) \tag{9}$$

and the function:

$$\Phi_l^s(\mathbf{z}_1, \ldots, \mathbf{z}_{m(l)}) = \frac{1}{(m(l))!} \sum_{\tau \in \mathcal{S}_{m(l)}} \Phi_l(\mathbf{z}_{\tau(1)}, \ldots, \mathbf{z}_{\tau(m(l))})$$

is the symmetrized version of Φ_l. In [27, Theorem 2.3], the estimator $\widehat{S}^{\mathbf{u}}_{2,\mathrm{GMS}}$ has been proved to be consistent and asymptotically Gaussian.

Even if the Pick-Freeze procedure is quite general, it presents some drawbacks. First of all, the Pick-Freeze design of experiment is peculiar and may not be available in real applications. Moreover, it can be unfortunately very time consuming in practice. For instance, estimating all the first order Sobol indices requires $(p+1)N$ calls to the computer code.

- Third method - Rank-based. In [14], Chatterjee proposes an efficient way based on ranks to estimate a new coefficient of correlation. This estimation procedure can be seen as an approximation of the Pick-Freeze scheme and then has been exploited in [23] to perform a more efficient estimation of $S^{\mathbf{u}}_{2,\mathrm{GMS}}$. Anyway, this method is only well tailored for estimating first order indices i.e. in the case of $\mathbf{u} = \{i\}$ for some $i \in \{1, \ldots, p\}$ and when the input $X_i \in \mathbb{R}$.

  More precisely, an i.i.d. sample of pairs of real-valued random variables $(X_{i,j}, Y_j)_{1 \leqslant j \leqslant N}$ ($i \in \{1, \cdots, p\}$) is considered, assuming for simplicity that the laws of X_i and Y are both diffuse (ties are excluded). The pairs $(X_{i,(1)}, Y_{(1)}), \ldots, (X_{i,(N)}, Y_{(N)})$ are rearranged in such a way that

$$X_{i,(1)} < \ldots < X_{i,(N)}$$

  and, for any $j = 1, \ldots, N$, $Y_{(j)}$ is the output computed from $X_{i,(j)}$. Let r_j be the rank of $Y_{(j)}$, that is,

$$r_j = \#\{j' \in \{1, \ldots, N\},\ Y_{(j')} \leqslant Y_{(j)}\}.$$

  The new correlation coefficient is then given by

$$\xi_N(X_i, Y) = 1 - \frac{3 \sum_{j=1}^{N-1} |r_{j+1} - r_j|}{N^2 - 1}. \tag{10}$$

  In [14], it is proved that $\xi_N(X, Y)$ converges almost surely to a deterministic limit $\xi(X, Y)$ which is actually equal to $S^i_{2,\mathrm{CVM}}$ when $Y = Z = f(X_1, \cdots, X_p)$. Further, the author also proves a central limit theorem when X_i and Y are independent, which is clearly not relevant in the context of sensitivity analysis (where X_i and Y are assumed to be dependent through the computer code).

  In our context, recall that $\mathbf{u} = \{i\}$ and let $Y = Z$. Let also $\pi_i(j)$ be the rank of $X_{i,j}$ in the sample $(X_{i,1}, \ldots, X_{i,N})$ of X_i and define

$$N_i(j) = \begin{cases} \pi_i^{-1}(\pi_i(j) + 1) & \text{if } \pi_i(j) + 1 \leqslant N, \\ \pi_i^{-1}(1) & \text{if } \pi_i(j) = N. \end{cases} \tag{11}$$

  Then the empirical estimator $\widehat{S}^i_{2,\mathrm{GMS,Rank}}$ of $S^i_{2,\mathrm{GMS}}$ only requires a N-sample $(Z_j)_{1 \leqslant j \leqslant N}$ of Z and is given by the ratio between

$$\begin{aligned} \widehat{N}^i_{2,\mathrm{GMS,Rank}} = & \frac{1}{N^m} \sum_{1 \leqslant i_1, \ldots, i_m \leqslant N} \left[\frac{1}{N} \sum_{j=1}^N T_{Z_{i_1}, \cdots, Z_{i_m}}(Z_j) T_{Z_{i_1}, \cdots, Z_{i_m}}(Z_{N_i(j)}) \right] \\ & - \frac{1}{N^m} \sum_{1 \leqslant i_1, \ldots, i_m \leqslant N} \left[\frac{1}{N} \sum_{j=1}^N T_{Z_{i_1}, \cdots, Z_{i_m}}(Z_j) \right]^2 \end{aligned}$$

  and $\widehat{D}_{2,\mathrm{GMS,Rank}}$

$$\frac{1}{N^m} \sum_{1 \leqslant i_1, \ldots, i_m \leqslant N} \left[\frac{1}{N} \sum_{j=1}^N T_{Z_{i_1}, \cdots, Z_{i_m}}(Z_j)^2 \right] - \frac{1}{N^m} \sum_{1 \leqslant i_1, \ldots, i_m \leqslant N} \left[\frac{1}{N} \sum_{j=1}^N T_{Z_{i_1}, \cdots, Z_{i_m}}(Z_j) \right]^2.$$

It is worth mentioning that $Z_{N_i(j)}$ plays the same role as Z_j^i (the Pick-Freeze version of Z_j) in the Pick-Freeze estimation procedure. Anyway, the strength of the rank-based estimation procedure lies in the fact that only one N-sample of Z is required while $(m+2)$ samples of size N are necessary in the Pick-Freeze estimation of a single index (worse, $(m+1+p)$ samples of size N are required when one wants to estimates p indices).

Comparison of the estimation procedures

First, the Pick-Freeze estimation procedure allows the estimation of several sensitivity indices: the classical Sobol indices for real-valued outputs, as well as their generalization for vectorial-valued codes, but also the indices based on higher moments [49] and the Cramér-von-Mises indices which take into account on the whole distribution [26, 21]. Moreover, the Pick-Freeze estimators have desirable statistical properties. More precisely, this estimation scheme has been proved to be consistent and asymptotically normal (i.e. the rate of convergence is $\sqrt{N}$) in [36, 25, 27]. The limiting variances can be computed explicitly, allowing the practitioner to build confidence intervals. In addition, for a given sample size N, exponential inequalities have been established. Last but not least, the sequence of estimators is asymptotically efficient from such a design of experiment (see, [58] for the definition of the asymptotic efficiency and [25] for more details on the result).

However, the Pick-Freeze estimators have two major drawbacks. First, they rely on a particular experimental design that may be unavailable in practice. Second, the number of model calls to estimate all first order Sobol indices grows linearly with the number of input parameters. For example, if we consider $p = 99$ input parameters and only $n = 1000$ calls are allowed, then only a sample of size $n/(p+1) = 10$ is available to estimate each single first order Sobol index.

Secondly, the estimation procedure based on U-statistics has the same kind of asymptotic guarantees as consistency and asymptotic normality. Furthermore, the estimation scheme is reduced to $2N$ evaluations of the code. Last but not least, using the results of Hoeffding [33] on U-statistics, the asymptotic normality is proved straightforwardly.

Finally, embedding Chatterjee's method in the GSA framework (called rank-based method in this framework) thereby eliminates the two drawbacks of the classical Pick-Freeze estimation. In addition, the rank-based method allows for the estimation of a large class of GSA indices which include the Sobol indices and the higher order moment indices proposed by Owen [47, 49, 48]. Using a single sample of size N, it is now possible to estimate at the same time all the first order Sobol indices , first order Cramér-von-Mises indices, and other useful first order sensitivity indices as soon as all inputs are real valued.

2.2 The universal sensitivity index

Formula (2) can be generalized in the following ways.

1. The point a in the definition of the test functions is allowed to belong to another measurable space than $\mathcal{X}^m$.

2. The probability measure $\mathbb{P}^{\otimes m}$ in (2) can be replaced by any "admissible" probability measure.

Such generalizations lead to the definition of a universal sensitivity index and its procedures of estimation.

Definition 2.1. Let a belongs to some measurable space Ω endowed with some probability measure $\mathbb{Q}$. For any $\mathbf{u} \subset \{1, \cdots, p\}$, we define the *universal sensitivity index* with respect to $X_{\mathbf{u}}$ by

$$S^{\mathbf{u}}_{2,\mathrm{Univ}}(T_a, \mathbb{Q}) = \frac{\int_\Omega \mathbb{E}\left[\left(\mathbb{E}[T_a(Z)] - \mathbb{E}[T_a(Z)|X_{\mathbf{u}}]\right)^2\right] d\mathbb{Q}(a)}{\int_\Omega \mathrm{Var}(T_a(Z)) d\mathbb{Q}(a)} = \frac{\int_\Omega \mathrm{Var}\left(\mathbb{E}[T_a(Z)|X_{\mathbf{u}}]\right) d\mathbb{Q}(a)}{\int_\Omega \mathrm{Var}(T_a(Z)) d\mathbb{Q}(a)}. \tag{12}$$

Notice that the index $S^{\mathbf{u}}_{2,\mathrm{Univ}}(T_a, \mathbb{Q})$ is obtained by the integration over a with respect to $\mathbb{Q}$ of the Hoeffding decomposition of $T_a(Z)$. Hence, by construction, this index lies in $[0, 1]$ and shares the same properties as its Sobol counterparts:

1. the different contributions sum to 1;

2. it is invariant by translation, by any isometric and by any non-degenerated scaling of Z.

The universality is twofold. First, it allows to consider more general relevant indices. Secondly, this definition encompasses, as particular cases, the classical sensitivity indices. Indeed,

- the so-called Sobol index $S^{\mathbf{u}}$ with respect to $X_{\mathbf{u}}$ is $S^{\mathbf{u}}_{2,\mathrm{Univ}}(\mathrm{Id},\mathbb{P})$;
- the Cramér-von-Mises index $S^{\mathbf{u}}_{2,\mathrm{CVM}}$ with respect to $X_{\mathbf{u}}$ is $S^{\mathbf{u}}_{2,\mathrm{Univ}}(\mathbb{1}_{\cdot\leqslant a},\mathbb{P}^{\otimes d})$ where $\mathcal{X}=\mathbb{R}^d$ and $\Omega=\mathcal{X}$;
- the general metric space sensitivity index $S^{\mathbf{u}}_{2,\mathrm{GMS}}$ with respect to $X_{\mathbf{u}}$ is $S^{\mathbf{u}}_{2,\mathrm{Univ}}(T_a,\mathbb{P}^{\otimes m})$ where $\Omega=\mathcal{X}^m$.

An example where $\mathbb{Q}$ is different from $\mathbb{P}$ will be considered in Section 4.

Estimation Here, we assume that $\mathbb{Q}$ is different from $\mathbb{P}^{\otimes m}$ and we follow the same tracks as for the estimation of $S^{\mathbf{u}}_{2,\mathrm{GMS}}$ in Section 2.1.

- <u>First method - Pick-Freeze</u>. We use the same design of experiment as in the First method of Section 2.1 but instead of considering that the m additional N-samples $(W_{l,k})$ for $l\in\{1,\dots,m\}$ and $k\in\{1,\dots,N\}$ are drawn with respect to the distribution $\mathbb{P}$ of the output, they are now drawn with respect to the law $\mathbb{Q}$. More precisely, one considers the following design of experiment consisting in:

  1. a classical Pick-Freeze sample, that is two N-samples of Z: $(Z_j,Z_j^{\mathbf{u}})$, $1\leqslant j\leqslant N$;
  2. m $\mathbb{Q}$-distributed N-samples $W_{l,k}$, $l\in\{1,\dots,m\}$ and $k\in\{1,\dots,N\}$ that are independent of $(Z_j,Z_j^{\mathbf{u}})$ for $1\leqslant j\leqslant N$.

  The empirical estimator of the numerator of $S^{\mathbf{u}}_{2,\mathrm{Univ}}$ is then given by

$$\begin{aligned}\widehat{N}^{\mathbf{u}}_{2,\mathrm{Univ,PF}}=&\frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left[\frac{1}{N}\sum_{j=1}^N T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j)T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j^{\mathbf{u}})\right]\\&-\frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left[\frac{1}{2N}\sum_{j=1}^N\left(T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j)+T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j^{\mathbf{u}})\right)\right]^2\end{aligned}$$

  while the one of the denominator is

$$\begin{aligned}\widehat{D}^{\mathbf{u}}_{2,\mathrm{Univ,PF}}=&\frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left[\frac{1}{2N}\sum_{j=1}^N\left(T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j)^2+T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j^{\mathbf{u}})^2\right)\right]\\&-\frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left[\frac{1}{2N}\sum_{j=1}^N\left(T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j)+T_{W_{1,i_1},\cdots,W_{m,i_m}}(Z_j^{\mathbf{u}})\right)\right]^2.\end{aligned}$$

  As previously, it is straightforward (as soon as the collection $(T_a)_{a\in\mathcal{X}^m}$ forms a Donsker's class of functions) to adapt the proof of Theorem [26, Theorem 3.8] to prove the asymptotic normality of the estimator.

- <u>Second method - U-statistics</u>. This method is not relevant in this case since $\mathbb{Q}\neq\mathbb{P}^{\otimes d}$.

- <u>Third method - Rank-based</u>. Here, the design of experiment reduces to:

  1. a N-sample of Z: Z_j, $1\leqslant j\leqslant N$;
  2. a N-sample of W that is $\mathbb{Q}$-distributed: W_k, $1\leqslant k\leqslant N$, independent of Z_j, $1\leqslant j\leqslant N$.

  The empirical estimator $\widehat{S}^{\mathbf{u}}_{2,\mathrm{Univ,Rank}}$ of $S^{\mathbf{u}}_{2,\mathrm{Univ}}$ is then given by the ratio between

$$\begin{aligned}\widehat{N}^{\mathbf{u}}_{2,\mathrm{Univ,Rank}}=&\frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left[\frac{1}{N}\sum_{j=1}^N T_{W_{i_1},\cdots,W_{i_m}}(Z_j)T_{W_{i_1},\cdots,W_{i_m}}(Z_{N(j)})\right]\\&-\frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left[\frac{1}{N}\sum_{j=1}^N T_{W_{i_1},\cdots,W_{i_m}}(Z_j)\right]^2\end{aligned}$$

and $\widehat{D}_{2,\mathrm{Univ,Rank}}$

$$\frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left[\frac{1}{N}\sum_{j=1}^{N}T_{W_{i_1},\cdots,W_{i_m}}(Z_j)^2\right]-\frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left[\frac{1}{N}\sum_{j=1}^{N}T_{W_{i_1},\cdots,W_{i_m}}(Z_j)\right]^2.$$

We recall that this last method only applies for first order sensitivity indices and real-valued input variables.

2.3 A sketch of answer to Questions 1 to 3

In the sequel, we discuss how pertinent choices of the metric, of the class of functions T_a and of the probability measure $\mathbb{Q}$ can provide some partial answers to Questions 1 to 3 raised at the beginning of Section 2. For instance, in order to answer to Question 1, we can consider that $\mathcal{X}=\mathcal{M}_q(\mathbb{R})$ is the space of probability measures μ on $\mathbb{R}$ that are L^q-functions and we endow this space with the Wasserstein metric W_q (see Section 3.1 for some recalls on Wasserstein metrics). We will propose two possible approaches to define interesting sensitivity indices in this framework.

- In Section 4.1, we use (2) with $m=2$, $a=(\mu_1,\mu_2)$ and $T_a(Z)=\mathbb{1}_{\{Z\in B(\mu_1,\mu_2)\}}$ where $B(\mu_1,\mu_2)$ is the open ball: $\{\mu\in\mathcal{M}_q(\mathbb{R}), W_q(\mu,\mu_1)<W_q(\mu_1,\mu_2)\}$.
- In Section 4.2, we use the notion of Fréchet means on Wasserstein spaces (see Section 3.2) and the index defined in (12) with appropriate choices of a, T_a, and $\mathbb{Q}$.

The case of stochastic computer computer codes raised in Question 2 will be addressed as follows. A computer code (to be defined) valued in $\mathcal{M}_q(\mathbb{R})$ will be seen as an ideal case of stochastic computer codes. Finally, it will be possible to treat Question 3 using the framework of Question 2.

3 Wasserstein spaces and random distributions

3.1 Definition

For any $q\geqslant 1$, we define the q-Wasserstein distance between two probability distributions that are L^q-integrable and characterized by their c.d.f.'s F and G on $\mathbb{R}^d$ by:

$$W_q(F,G)=\min_{X\sim F,Y\sim G}\left(\mathbb{E}[\|X-Y\|^q]^{1/q}\right),$$

where $X\sim F$ and $Y\sim G$ mean that X and Y are random variables with respective c.d.f.'s F and G. We define the Wasserstein space $\mathcal{W}_q(\mathbb{R}^d)$ as the space of all L^q-integrable measures defined on $\mathbb{R}^d$ endowed with the q-Wasserstein distance W_q. In the sequel, any measure is identified to its c.d.f. or in some cases to its p.d.f. In the unidimensional case ($d=1$), it is a well known fact that $W_q(F,G)$ has an explicitly expression given by

$$W_q(F,G)=\left(\int_0^1|F^-(v)-G^-(v)|^q dv\right)^{1/q}=\mathbb{E}[|F^-(U)-G^-(U)|^q]^{1/q}. \tag{13}$$

Here F^- and G^- are the generalized inverses of the increasing functions F and G and U is a random variable uniformly distributed on $[0,1]$. Of course, $F^-(U)$ and $G^-(U)$ have c.d.f.'s F and G. The representation (13) of the q-Wasserstein distance when $d=1$ can be generalized to a wider class of "contrast functions". For more details on Wasserstein spaces, one can refer to [59] and [5] and the references therein.

Definition 3.1. We call *contrast function* any application c from $\mathbb{R}^2$ to $\mathbb{R}$ satisfying the "measure property" $\mathcal{P}$ defined by

$$\mathcal{P}:\forall x\leqslant x' \text{ and } \forall y\leqslant y', c(x',y')-c(x',y)-c(x,y')+c(x,y)\leqslant 0,$$

meaning that c defines a negative measure on $\mathbb{R}^2$.

For instance, $c(x,y) = -xy$ satisfies $\mathcal{P}$. If c satisfies $\mathcal{P}$ then any function of the form $a(x)+b(y)+c(x,y)$ also satisfies $\mathcal{P}$. If C is a convex real function, $c(x,y) = C(x-y)$ satisfies $\mathcal{P}$. In particular, $c(x,y) = (x-y)^2 = x^2+y^2-2xy$ satisfy $\mathcal{P}$ and actually so does $c(x,y) = |x-y|^p$ as soon as $p \geqslant 1$.

Definition 3.2. We define the *Skorohod space* $\mathcal{D} := \mathcal{D}([0,1])$ of all distribution functions as the space of all non-decreasing functions from $\mathbb{R}$ to $[0,1]$ that are càd-làg with limit 0 (resp. 1) in $-\infty$ (resp. $+\infty$) equipped with the supremum norm.

Definition 3.3. For any $F \in \mathcal{D}$, any $G \in \mathcal{D}$, and any positive contrast function c, we define the *c-Wasserstein cost* by

$$W_c(F,G) = \min_{X\sim F, Y\sim G} \mathbb{E}\left[c(X,Y)\right] < +\infty.$$

Obviously, $W_q^q = W_c$ with $c(x,y) = |x-y|^q$. The following theorem can be found in ([11]).

Theorem 3.4 (Cambanis, Simon, Stout [11])**.** *Let c be a contrast function. Then*

$$W_c(F,G) = \int_0^1 c(F^-(v), G^-(v))dv = \mathbb{E}[c(F^-(U), G^-(U))],$$

where U is a random variable uniformly distributed on $[0,1]$.

3.2 Extension of the Fréchet mean to contrast functions

Definition 3.5. We call a *loss function* any positive and measurable function l. Then, we define a *Fréchet feature* $\mathcal{E}_l[X]$ of a random variable X taking values in a measurable space $\mathcal{M}$ as (whenever it exists):

$$\mathcal{E}_l[X] \in \operatorname*{Argmin}_{\theta\in\mathcal{M}} \mathbb{E}[l(X,\theta)]. \tag{14}$$

When $\mathcal{M}$ is a metric space endowed with a distance $d = l$, the Fréchet feature corresponds to the classical Fréchet mean (see [22]). In particular, $\mathcal{E}_d[X]$ minimizes $\mathbb{E}[d^2(X,\theta)]$ which is an extension of the definition of the classical mean in $\mathbb{R}^d$ that minimizes $\mathbb{E}[\|X-\theta\|^2]$.

Now we consider $\mathcal{M} = \mathcal{D}$ and $l = W_c$. Further, Equation (14) becomes

$$\mathcal{E}_{W_c}[\mathbb{F}] \in \operatorname*{Argmin}_{G\in\mathcal{D}} \mathbb{E}\left[W_c(\mathbb{F},G)\right].$$

where $\mathbb{F}$ is a measurable function from a measurable space Ω to $\mathcal{D}$.

Theorem 3.6. *Let c be a positive contrast function. Assume that the application defined by $(\omega,v) \in \Omega\times(0,1) \mapsto \mathbb{F}^-(\omega,v) \in \mathbb{R}$ is measurable. In addition, assume that $\mathcal{E}_c[\mathbb{F}]$ exists and is unique. Then there exists a unique Fréchet mean of $\mathbb{E}[c(\mathbb{F}^-(v),s)]$ denoted by $\mathcal{E}_c[\mathbb{F}^-](v)$ and we have*

$$(\mathcal{E}_c[\mathbb{F}])^-(v) = \mathcal{E}_c[\mathbb{F}^-](v) = \operatorname*{Argmin}_{s\in\mathbb{R}} \mathbb{E}[c(\mathbb{F}^-(v),s)].$$

Proof of Theorem 3.6. Since c satisfies $\mathcal{P}$, we have

$$\mathbb{E}[W_c(\mathbb{F},G)] = \mathbb{E}\left[\int_0^1 c(\mathbb{F}^-(v), G^-(v))dv\right] = \int_0^1 \mathbb{E}[c(\mathbb{F}^-(v), G^-(v))]dv,$$

by Fubini's theorem. Now, for all $v \in (0,1)$, the quantity $\mathbb{E}[c(\mathbb{F}^-(v), G^-(v))]$ is minimum for $G^-(v) = \mathcal{E}_c[\mathbb{F}^{-1}](v)$.

$$\int_0^1 \mathbb{E}[c(\mathbb{F}^-(v), \mathcal{E}_c[\mathbb{F}^{-1}](v))]dv \leqslant \int_0^1 \mathbb{E}[c(\mathbb{F}^-(v), G^-(v))]dv$$

and, in particular, for $G^- = \mathcal{E}_c[\mathbb{F}]^{-1}$, one gets

$$\int_0^1 \mathbb{E}[c(\mathbb{F}^-(v), \mathcal{E}_c[\mathbb{F}^{-1}](v))]dv \leqslant \int_0^1 \mathbb{E}[c(\mathbb{F}^-(v), \mathcal{E}_c[\mathbb{F}]^{-1}(v))]dv.$$

Conversely, by the definition of $\mathcal{E}_c[\mathbb{F}]^{-1}$, we have for all G,

$$\int_0^1 \mathbb{E}[c(\mathbb{F}^-(v), \mathcal{E}_c[\mathbb{F}]^{-1}(v))]dv \leqslant \int_0^1 \mathbb{E}[c(\mathbb{F}^-(v), G^-(v))]dv$$

and, in particular, for $G^- = \mathcal{E}_c[\mathbb{F}^{-1}]$, one gets

$$\int_0^1 \mathbb{E}[c(\mathbb{F}^-(v), \mathcal{E}_c[\mathbb{F}]^{-1}(v))]dv \leqslant \int_0^1 \mathbb{E}[c(\mathbb{F}^-(v), \mathcal{E}_c[\mathbb{F}^{-1}](v))]dv.$$

The theorem then follows by the uniqueness of the minimizer. □

In the previous theorem, we propose a very genereal non parametric framework for the random element $\mathbb{F}$ together with some assumptions on existence and uniqueness of the Fréchet feature and measurability of the map $(\omega, v) \mapsto \mathbb{F}^-(\omega, v)$. Nevertheless, it is possible to construct explicit parametric models for $\mathbb{F}$ for whom theses assumptions are satisfied. For instance, the authors of [4] ensures measurability for some parametric models on $\mathbb{F}$ using results of [18]. Notice that in [19] a new sensitivity indice is defined for each feature associated to a contrast function. In section 4.2 we will restrict our analysis to Fréchet means, hence to Sobol indices.

3.3 Examples

The Fréchet mean in the $\mathcal{W}_2(\mathbb{R})$ space is the inverse of the function $v \mapsto \mathbb{E}\ [\mathbb{F}^-(v)]$. Another example is the Fréchet median. Since the median in $\mathbb{R}$ is related to the L^1 cost, the Fréchet $\mathcal{W}_1(\mathbb{R})$ median of a random c.d.f. is

$$(\mathrm{Med}(\mathbb{F})^-(v) \in \mathrm{Med}(\mathbb{F}^-(v)).$$

More generally, we recall that, for $\alpha \in (0,1)$, the α-quantile in $\mathbb{R}$ is the Fréchet mean associated to the contrast function $c_\alpha(x,y) = (1-\alpha)(y-x)\mathbb{1}_{x-y<0} + \alpha(x-y)\mathbb{1}_{x-y\geqslant 0}$, also called the *pinball function*. Then we can define an α-quantile $q_\alpha(\mathbb{F})$ of a random c.d.f. as:

$$(q_\alpha(\mathbb{F}))^-(v) \in q_\alpha(\mathbb{F}^-(v)),$$

where $q_\alpha(X)$ is the set of the α-quantiles of a random variable X taking values in $\mathbb{R}$. Naturally, taking $\alpha = 1/2$ leads to the median.

Let us illustrate the previous definitions on an example. Let X be a random variable with c.d.f. F_0 which is assumed to be increasing and continuous. Let also m and σ two real random variables such that $\sigma > 0$. Then we consider the random c.d.f. $\mathbb{F}$ of $\sigma X + m$:

$$\mathbb{F}(x) = F_0\left(\frac{x-m}{\sigma}\right) \quad \text{and} \quad \mathbb{F}^{-1}(v) = \sigma F_0^{-1}(v) + m.$$

Naturally, the Fréchet mean of $\mathbb{F}$ is $\mathcal{E}[\mathbb{F}](x) = F_0\left(x - \mathbb{E}[m]/\mathbb{E}[\sigma]\right)$ and its α-quantile is given by

$$(q_\alpha(\mathbb{F}))^{-1}(v) = q_\alpha(\sigma F_0^{-1}(v) + m).$$

4 Sensitivity analysis in general Wasserstein spaces

In this section, we consider that our computer code is $\mathcal{W}_q(\mathbb{R})$-valued; namely, the output of an experiment is the c.d.f. or the p.d.f. of a measure $\mu \in \mathcal{W}_q(\mathbb{R})$. For instance, in [9], [40] and [46], the authors deal with p.d.f.-valued computer codes (and stochastic computer codes). In other words, they define the following application:

$$\begin{aligned} f: \quad E &\to \mathcal{F} \\ x &\mapsto f(x) \end{aligned} \tag{15}$$

where $\mathcal{F}$ is the set of p.d.f.'s:

$$\mathcal{F} = \left\{ g \in L^1(\mathbb{R});\ g \geqslant 0,\ \int_{\mathbb{R}} g(x)dx = 1 \right\}.$$

Here, we choose to identify any element of $\mathcal{W}_q(\mathbb{R})$ with its c.d.f. In this framework, the output of the cmoputer code is then a c.d.f. denoted by

$$\mathbb{F} = f(X_1, \dots, X_p). \tag{16}$$

Here, $\mathbb{P}$ denotes the law of the c.d.f. $\mathbb{F}$. In addition, we set $q = 2$. The case of a general q can be handled analogously.

4.1 Sensitivity anlaysis using Equation (2) and Wasserstein balls

Consider F, F_1, and F_2 three elements of $\mathcal{W}_2(\mathbb{R})$ and, for $a = (F_1, F_2)$, the family of test functions

$$T_a(F) = T_{(F_1,F_2)}(F) = \mathbb{1}_{W_2(F_1,F)\leqslant W_2(F_1,F_2)}. \tag{17}$$

Then, for all $\mathbf{u} \subset \{1, \cdots, p\}$, (2) becomes

$$\begin{aligned} S^{\mathbf{u}}_{2,W_2} &= S^{\mathbf{u}}_{2,\mathrm{Univ}}((F_1, F_2, F) \mapsto T_{F_1,F_2}(F), \mathbb{P}^{\otimes 2}) \\ &= \frac{\int_{\mathcal{W}_2(\mathbb{R})\times\mathcal{W}_2(\mathbb{R})} \mathbb{E}\left[\left(\mathbb{E}[\mathbb{1}_{W_2(F_1,\mathbb{F})\leqslant W_2(F_1,F_2)}] - \mathbb{E}[\mathbb{1}_{W_2(F_1,\mathbb{F})\leqslant W_2(F_1,F_2)}|X^{\mathbf{u}}]\right)^2\right] d\mathbb{P}^{\otimes 2}(F_1, F_2)}{\int_{\mathcal{W}_2(\mathbb{R})\times\mathcal{W}_2(\mathbb{R})} \mathrm{Var}(\mathbb{1}_{W_2(F_1,\mathbb{F})\leqslant W_2(F_1,F_2)}) d\mathbb{P}^{\otimes 2}(F_1, F_2)} \\ &= \frac{\int_{\mathcal{W}_2(\mathbb{R})\times\mathcal{W}_2(\mathbb{R})} \mathrm{Var}\left(\mathbb{E}[\mathbb{1}_{W_2(F_1,\mathbb{F})\leqslant W_2(F_1,F_2)}|X^{\mathbf{u}}]\right) d\mathbb{P}^{\otimes 2}(F_1, F_2)}{\int_{\mathcal{W}_2(\mathbb{R})\times\mathcal{W}_2(\mathbb{R})} \mathrm{Var}(\mathbb{1}_{W_2(F_1,\mathbb{F})\leqslant W_2(F_1,F_2)}) d\mathbb{P}^{\otimes 2}(F_1, F_2)}. \end{aligned} \tag{18}$$

As explained in Section 2.1, $S^{\mathbf{u}}_{2,W_2}$ is obtained by integration over a of the Hoeffding decomposition of $T_a(\mathbb{F})$. Hence, by construction, this index lies in $[0, 1]$ and shares the same properties as its Sobol counterparts:

1. the different contributions sum to 1;
2. it is invariant by translation, by any isometric and by any non-degenerated scaling of $\mathbb{F}$.

4.2 Sensitivity analysis using Equation (12) and Fréchet means

In the classical framework where the output Z is real, we recall that the Sobol index with respect to $X_{\mathbf{u}}$ is defined by

$$S^{\mathbf{u}} = \frac{\mathrm{Var}(\mathbb{E}[Z|X_{\mathbf{u}}])}{\mathrm{Var}(Z)} = \frac{\mathrm{Var}(Z) - \mathbb{E}[\mathrm{Var}(Z|X_{\mathbf{u}})]}{\mathrm{Var}(Z)}, \tag{19}$$

by the property of the conditional expectation. In one hand, one may extend this formula to the framework of this section where the output of interest is the c.d.f. $\mathbb{F}$:

$$S^{\mathbf{u}}(\mathbb{F}) = \frac{\mathrm{Var}(\mathbb{F}) - \mathbb{E}[\mathrm{Var}(\mathbb{F}|X_{\mathbf{u}}))]}{\mathrm{Var}(\mathbb{F})},$$

where $\mathrm{Var}(\mathbb{F}) = \mathbb{E}[W_2^2(\mathbb{F}, \mathcal{E}_{W_2}(\mathbb{F}))]$ with $\mathcal{E}_{W_2}(\mathbb{F})$ the Fréchet mean of $\mathbb{F}$. From Theorem 3.6, we get

$$\mathrm{Var}(\mathbb{F}) = \mathbb{E}\left[\int_0^1 |\mathbb{F}^-(v) - \mathcal{E}(\mathbb{F})^-(v)|^2 dv\right] = \mathbb{E}\left[\int_0^1 |\mathbb{F}^-(v) - \mathbb{E}[\mathbb{F}^-(v)]|^2 dv\right] = \int_0^1 \mathrm{Var}(\mathbb{F}^-(v)) dv$$

leading to

$$S^{\mathbf{u}}(\mathbb{F}) = \frac{\int_0^1 \mathrm{Var}(\mathbb{F}^-(v)) dv - \int_0^1 \mathbb{E}[\mathrm{Var}(\mathbb{F}^-(v)|X_{\mathbf{u}})] dv}{\int_0^1 \mathrm{Var}(\mathbb{F}^-(v)) dv} = \frac{\int_0^1 \mathrm{Var}(\mathbb{E}[\mathbb{F}^-(v)|X_{\mathbf{u}}]) dv}{\int_0^1 \mathrm{Var}(\mathbb{F}^-(v)) dv}. \tag{20}$$

In the other hand, one can consider (12), with $m = 1$,

$$T_v(\mathbb{F}) = \mathbb{F}^-(v) \tag{21}$$

and $\mathbb{Q}$ the uniform probability measure on $[0, 1]$. In that case,

$$\mathrm{Var}(\mathbb{F}) = \mathbb{E}\left[\int_0^1 |\mathbb{F}^-(v) - \mathcal{E}_{W_2}(\mathbb{F})^-(v)|^2 dv\right] = \int_0^1 \mathrm{Var}(\mathbb{F}^-(v)) dv = \mathbb{E}[W_2^2(\mathbb{F}, \mathcal{E}_{W_2}(\mathbb{F}))].$$

Then

$$S^{\mathbf{u}}_{2,\mathrm{Univ}}(T_v,\mathcal{U}([0,1])) = \frac{\int_0^1 \mathbb{E}\left[\left(\mathcal{E}_{W_2}(\mathbb{F})^-(v) - \mathcal{E}_{W_2}(\mathbb{F}|X_{\mathbf{u}})^-(v)\right)^2\right] dv}{\int_0^1 \mathrm{Var}(\mathbb{F}^-(v))dv} = \frac{\mathbb{E}\left[W_2^2(\mathcal{E}_{W_2}(\mathbb{F}|X_{\mathbf{u}}),\mathcal{E}_{W_2}(\mathbb{F}))\right]}{\mathbb{E}\left[W_2^2(\mathbb{F},\mathcal{E}_{W_2}(\mathbb{F}))\right]}.$$

is exactly the same as $S^{\mathbf{u}}(\mathbb{F})$ in (20). Thus, as explained in Section 2.2, $S^{\mathbf{u}}(\mathbb{F})$ lies in $[0,1]$ and:

1. the different contributions sum to 1;
2. it is invariant by translation, by any isometric and by any non-degenerated scaling of $\mathbb{F}$.

4.3 Estimation procedure

As noticed in the previous section, both

$$S^{\mathbf{u}}_{2,W_2} = S^{\mathbf{u}}_{2,Univ}(T_a,\mathbb{P}^{\otimes 2})$$

with T_a defined in (17) and

$$S^{\mathbf{u}}(\mathbb{F}) = S^{\mathbf{u}}_{2,\mathrm{Univ}}(T_v,\mathcal{U}([0,1]))$$

with T_v defined in (21), are particular cases of indices of the form (12).

When a belongs to the same space as the output and when $\mathbb{Q}$ is equal to $\mathbb{P}^{\otimes m}$, one may first use the Pick-Freeze estimations of the indices given in (18) and (20). To do so, it is convenient once again to use (4) leading to

$$S^{\mathbf{u}}_{2,W_2} = \frac{\int_{\mathcal{W}_2(\mathbb{R})\times\mathcal{W}_2(\mathbb{R})} \mathrm{Cov}\left(\mathbb{1}_{W_2(F_1,\mathbb{F})\leqslant W_2(F_1,F_2)}, \mathbb{1}_{W_2(F_1,\mathbb{F}^{\mathbf{u}})\leqslant W_2(F_1,F_2)}\right) d\mathbb{P}^{\otimes 2}(F_1,F_2)}{\int_{\mathcal{W}_2(\mathbb{R})\times\mathcal{W}_2(\mathbb{R})} \mathrm{Var}(\mathbb{1}_{W_2(F_1,\mathbb{F})\leqslant W_2(F_1,F_2)})d\mathbb{P}^{\otimes 2}(F_1,F_2)} \tag{22}$$

and

$$S^{\mathbf{u}}(\mathbb{F}) = \frac{\int_0^1 \mathrm{Cov}\left(\mathbb{F}^-(v),\mathbb{F}^{-,\mathbf{u}}(v)\right) dv}{\int_0^1 \mathrm{Var}(\mathbb{F}^-(v))dv} \tag{23}$$

where $\mathbb{F}^{\mathbf{u}}$ and $\mathbb{F}^{-,\mathbf{u}}$ are respectively the Pick-Freeze versions of $\mathbb{F}$ and $\mathbb{F}^-$. Secondly, one may resort to the estimations based on U-statistics together on the Pick-Freeze design of experiment. Thirdly, it is also possible and easy to obtain rank-based estimations in the vein of (10).

4.4 Numerical comparison of both indices

Example 4.1 (Toy model)**.** Let X_1, X_2, X_3 be 3 independent and positive random variables. We consider the c.d.f.-valued code f, the output of which is given by

$$\mathbb{F}(t) = \frac{t}{1+X_1+X_2+X_1X_3}\mathbb{1}_{0\leqslant t\leqslant 1+X_1+X_2+X_1X_3} + \mathbb{1}_{1+X_1+X_2+X_1X_3<t}, \tag{24}$$

so that

$$\mathbb{F}^{-1}(v) = v\Big(1+X_1+X_2+X_1X_3\Big). \tag{25}$$

In addition, one gets

$$\begin{aligned}\mathrm{Var}\left(\mathbb{F}^{-1}(v)\right) &= v^2\left(\mathrm{Var}(X_1(1+X_3)) + \mathrm{Var}(X_2)\right)\\ &= v^2\left(\mathrm{Var}(X_1)\mathrm{Var}(X_3) + \mathrm{Var}(X_1)(1+\mathbb{E}[X_3])^2 + \mathrm{Var}(X_3)\mathbb{E}[X_1]^2 + \mathrm{Var}(X_2)\right)\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\left[\mathbb{F}^{-1}(v)|X_1\right] &= v\Big(1+X_1(1+\mathbb{E}[X_3]) + \mathbb{E}[X_2]\Big),\\ \mathbb{E}\left[\mathbb{F}^{-1}(v)|X_2\right] &= v\Big(1+\mathbb{E}[X_1](1+\mathbb{E}[X_3]) + X_2\Big),\\ \mathbb{E}\left[\mathbb{F}^{-1}(v)|X_3\right] &= v\Big(1+\mathbb{E}[X_1](1+X_3) + \mathbb{E}[X_2]\Big),\\ \mathbb{E}\left[\mathbb{F}^{-1}(v)|X_1X_3\right] &= v\Big(1+X_1(1+X_3) + \mathbb{E}[X_2]\Big),\end{aligned}$$

and finally

$$\begin{aligned}
&\mathrm{Var}\left(\mathbb{E}\left[\mathbb{F}^{-1}(v)|X_1\right]\right) = v^2(1+\mathbb{E}[X_3])^2\mathrm{Var}(X_1),\\
&\mathrm{Var}\left(\mathbb{E}\left[\mathbb{F}^{-1}(v)|X_2\right]\right) = v^2\mathrm{Var}(X_2),\\
&\mathrm{Var}\left(\mathbb{E}\left[\mathbb{F}^{-1}(v)|X_3\right]\right) = v^2\mathbb{E}[X_1]^2\mathrm{Var}(X_3),\\
&\mathrm{Var}\left(\mathbb{E}\left[\mathbb{F}^{-1}(v)|X_1,X_3\right]\right) = v^2\left(\mathrm{Var}(X_1)\mathrm{Var}(X_3)+\mathrm{Var}(X_1)(1+\mathbb{E}[X_3])^2+\mathrm{Var}(X_3)\mathbb{E}[X_1]^2\right).
\end{aligned}$$

For $\mathbf{u} = i \in 1,2,3$ or $\mathbf{u} = \{1,3\}$, it remains to plug the previous formulas in (20) to get the explicit expressions of the indices $S^{\mathbf{u}}(\mathbb{F})$.

Now, in order to get a closed formula for the indices defined in (18), we assume X_i is Bernoulli distributed with parameter $0 < p_i < 1$. In (18), the distributions F_1 and F_2 can be either $\mathcal{U}([0,1])$, $\mathcal{U}([0,2])$, $\mathcal{U}([0,3])$, or $\mathcal{U}([0,4])$ with respective probabilities $q_1 = (1-p_1)(1-p_2)$, $q_2 = (1-p_1)p_2 + p_1(1-p_2)(1-p_3)$, $q_3 = p_1((1-p_2)p_3 + p_2(1-p_3))$, and $q_4 = p_1p_2p_3$. In the sequel, we give, for all sixteen possibilities for the distribution of (F_1, F_2), the corresponding contributions for the numerator and for the denominator of (18).

With probability $p_{1,1} = (1-p_1)^2(1-p_2)^2$, F_1 and $F_2 \sim \mathcal{U}([0,1])$. Then $W_2^2(F_1,F_2) = 0$, $W_2^2(F_1,\mathbb{F}) = \frac{1}{3}(X_1+X_2+X_1X_3)^2$, and $W_2^2(F_1,\mathbb{F}) \leqslant W_2^2(F_1,F_2)$ if and only if $X_1+X_2+X_1X_3 = 0$. Since $\mathbb{P}(X_1+X_2+X_1X_3 = 0) = (1-p_1)(1-p_2)$, the contribution $d_{1,1}$ of this case to the denominator is thus

$$d_{1,1} = q_{1,1}(1-q_{1,1}) \quad \text{with } q_{1,1} = (1-p_1)(1-p_2).$$

Moreover,

$$\mathbb{E}[\mathbb{1}_{X_1+X_2+X_1X_3=0}|X_1] = \mathbb{P}\Big(X_1+X_2+X_1X_3 = 0|X_1\Big) = \mathbb{1}_{X_1=0}\mathbb{P}(X_2=0) = (1-p_2)\mathbb{1}_{X_1=0}.$$

so that, the contribution to the numerator is here given by

$$n_{1,1}^1 = \mathrm{Var}(\mathbb{E}[\mathbb{1}_{X_1+X_2+X_1X_3=0}|X_1]) = p_1(1-p_1)(1-p_2)^2.$$

Similarly, one gets

$$n_{1,1}^2 = \mathrm{Var}(\mathbb{E}[\mathbb{1}_{X_1+X_2+X_1X_3=0}|X_2]) = p_2(1-p_2)(1-p_1)^2 \quad \text{and} \quad n_{1,1}^3 = 0.$$

Moreover, regarding the indices with respect to X_1 and X_3,

$$\mathbb{E}[\mathbb{1}_{X_1+X_2+X_1X_3=0}|X_1,X_3] = \mathbb{P}\Big(X_1+X_2+X_1X_3 = 0|X_1,X_3\Big) = \mathbb{1}_{X_1=0}\mathbb{P}(X_2=0) = (1-p_2)\mathbb{1}_{X_1=0}$$

and the contribution to the numerator is given by

$$n_{1,1}^{1,3} = \mathrm{Var}(\mathbb{E}[\mathbb{1}_{X_1+X_2+X_1X_3=0}|X_1,X_3]) = p_1(1-p_1)(1-p_2)^2.$$

The remaining fifteen cases can be treated similarly and are gathered (with the first case developed above) in the following table. Finally, for $k = 1,\dots,3$, one may compute the explicit expression of S_{2,W_2}^k:

$$S_{2,W_2}^k = \frac{\int_{\mathcal{W}_2(\mathbb{R})\times\mathcal{W}_2(\mathbb{R})} \mathrm{Cov}\left(\mathbb{1}_{W_2(F_1,\mathbb{F})\leqslant W_2(F_1,F_2)}, \mathbb{1}_{W_2(F_1,\mathbb{F}^{\mathbf{u}})\leqslant W_2(F_1,F_2)}\right) d\mathbb{P}^{\otimes 2}(F_1,F_2)}{\int_{\mathcal{W}_2(\mathbb{R})\times\mathcal{W}_2(\mathbb{R})} \mathrm{Var}(\mathbb{1}_{W_2(F_1,\mathbb{F})\leqslant W_2(F_1,F_2)}) d\mathbb{P}^{\otimes 2}(F_1,F_2)} = \frac{\sum_{i,j} p_{i,j} n_{i,j}^k}{\sum_{i,j} p_{i,j} d_{i,j}}.$$

Some numerical values have not been explicited in the table but given below:

Case 2 $\quad \mathrm{Var}(\mathbb{1}_{X_1=1}(1-(1-p_2)\mathbb{1}_{X_3=0})) = p_1(1-p_1)(1-(1-p_2)(1-p_3))^2 + p_1(1-p_2)^2p_3(1-p_3),$

Case 6 $\quad \mathrm{Var}(\mathbb{1}_{X_1=1}(p_2-(1-p_2)\mathbb{1}_{X_3=0})) = p_1(1-p_1)(p_2-(1-p_2)(1-p_3))^2 + p_1(1-p_2)^2p_3(1-p_3),$

Case 11 $\quad \mathrm{Var}(\mathbb{1}_{X_1=1}(p_2+(1-2p_2)\mathbb{1}_{X_3=1})) = p_1(1-p_1)(p_2+(1-2p_2)p_3)^2 + p_1(1-2p_2)^2p_3(1-p_3),$

Case 15 $\quad \mathrm{Var}(\mathbb{1}_{X_1=1}(p_2+(1-p_2)\mathbb{1}_{X_3=1})) = p_1(1-p_1)(p_2+(1-p_2)p_3)^2 + p_1(1-p_2)^2p_3(1-p_3).$

Case 1	$F_1 \sim \mathcal{U}([0,1])$, $F_2 \sim \mathcal{U}([0,1])$
Prob.	q_1^2
Num. 1	$p_1(1-p_1)(1-p_2)^2$
Num. 2	$(1-p_1)^2 p_2(1-p_2)$
Num. 3	0
Num. 1,3	$p_1(1-p_1)(1-p_2)^2$
q Den.	$(1-p_1)(1-p_2)$

Case 2	$F_1 \sim \mathcal{U}([0,1])$, $F_2 \sim \mathcal{U}([0,2])$
Prob.	$q_1 q_2$
Num. 1	$p_1(1-p_1)(p_2+p_3-p_2p_3)^2$
Num. 2	$p_1^2 p_2(1-p_2)(1-p_3)^2$
Num. 3	$p_1^2(1-p_2)^2 p_3(1-p_3)$
Num. 1,3	$\mathrm{Var}(\mathbb{1}_{X_1=1}(1-(1-p_2)\mathbb{1}_{X_3=0})$
q Den.	$(1-p_1)+p_1(1-p_2)(1-p_3)$

Case 3	$F_1 \sim \mathcal{U}([0,1])$, $F_2 \sim \mathcal{U}([0,3])$
Prob.	$q_1 q_3$
Num. 1	$p_1(1-p_1)p_2^2 p_3^2$
Num. 2	$p_1^2 p_2(1-p_2)p_3^2$
Num. 3	$p_1^2 p_2^2 p_3(1-p_3)$
Num. 1,3	$p_1 p_2^2 p_3(1-p_1p_3)$
q Den.	$1-p_1p_2p_3$

Case 4	$F_1 \sim \mathcal{U}([0,1])$, $F_2 \sim \mathcal{U}([0,4])$
Prob.	$q_1 q_4$
Num. 1	0
Num. 2	0
Num. 3	0
Num. 1,3	0
q Den.	0

Case 5	$F_1 \sim \mathcal{U}([0,2])$, $F_2 \sim \mathcal{U}([0,1])$
Prob.	$q_1 q_2$
Num. 1	$p_1(1-p_1)p_2^2 p_3^2$
Num. 2	$p_1^2 p_2(1-p_2)p_3^2$
Num. 3	$p_1^2 p_2^2 p_3(1-p_3)$
Num. 1,3	$p_1 p_2^2 p_3(1-p_1p_3)$
q Den.	$1-p_1p_2p_3$

Case 6	$F_1 \sim \mathcal{U}([0,2])$, $F_2 \sim \mathcal{U}([0,2])$
Prob.	q_2^2
Num. 1	$p_1(1-p_1)(p_2-(1-p_2)(1-p_3))^2$
Num. 2	$p_2(1-p_2)(p_1(1-p_3)-(1-p_1))^2$
Num. 3	$p_1^2(1-p_2)^2 p_3(1-p_3)$
Num. 1,3	$\mathrm{Var}(\mathbb{1}_{X_1=1}(p_2-(1-p_2)\mathbb{1}_{X_3=0}))$
q Den.	$(1-p_1)p_2+p_1(1-p_2)(1-p_3)$

Case 7	$F_1 \sim \mathcal{U}([0,2])$, $F_2 \sim \mathcal{U}([0,3])$
Prob.	$q_2 q_3$
Num. 1	$p_1(1-p_1)p_2^2 p_3^2$
Num. 2	$p_1^2 p_2(1-p_2)p_3^2$
Num. 3	$p_1^2 p_2^2 p_3(1-p_3)$
Num. 1,3	$p_1 p_2^2 p_3(1-p_1p_3)$
q Den.	$1-p_1p_2p_3$

Case 8	$F_1 \sim \mathcal{U}([0,2])$, $F_2 \sim \mathcal{U}([0,4])$
Prob.	$q_2 q_4$
Num. 1	0
Num. 2	0
Num. 3	0
Num. 1,3	0
q Den.	0

Case 9	$F_1 \sim \mathcal{U}([0,3])$, $F_2 \sim \mathcal{U}([0,1])$
Prob.	$q_1 q_3$
Num. 1	0
Num. 2	0
Num. 3	0
Num. 1,3	0
q Den.	0

Case 10	$F_1 \sim \mathcal{U}([0,3])$, $F_2 \sim \mathcal{U}([0,2])$
Prob.	$q_2 q_3$
Num. 1	$p_1(1-p_1)(1-p_2)^2$
Num. 2	$(1-p_1)^2 p_2(1-p_2)$
Num. 3	0
Num. 1,3	$p_1(1-p_1)(1-p_2)^2$
q Den.	$(1-p_1)p_2+p_1$

Case 11	$F_1 \sim \mathcal{U}([0,3])$, $F_2 \sim \mathcal{U}([0,3])$
Prob.	q_3^2
Num. 1	$p_1(1-p_1)(p_2(1-p_3)+(1-p_2)p_3)^2$
Num. 2	$p_1^2 p_2(1-p_2)(2p_3-1)^2$
Num. 3	$p_1^2(2p_2-1)^2 p_3(1-p_3)$
Num. 1,3	$\mathrm{Var}(\mathbb{1}_{X_1=1}(p_2+(1-2p_2)\mathbb{1}_{X_3=1})$
q Den.	$p_1(p_2(1-p_3)+(1-p_2)p_3)$

Case 12	$F_1 \sim \mathcal{U}([0,3])$, $F_2 \sim \mathcal{U}([0,4])$
Prob.	$q_3 q_4$
Num. 1	$p_1(1-p_1)(1-p_2)^2$
Num. 2	$(1-p_1)^2 p_2(1-p_2)$
Num. 3	0
Num. 1,3	$p_1(1-p_1)(1-p_2)^2$
q Den.	$(1-p_1)p_2+p_1$

Case 13	$F_1 \sim \mathcal{U}([0,4])$, $F_2 \sim \mathcal{U}([0,1])$
Prob.	$q_1 q_4$
Num. 1	0
Num. 2	0
Num. 3	0
Num. 1,3	0
q Den.	0

Case 14	$F_1 \sim \mathcal{U}([0,4])$, $F_2 \sim \mathcal{U}([0,2])$
Prob.	$q_2 q_4$
Num. 1	$p_1(1-p_1)(1-p_2)^2$
Num. 2	$(1-p_1)^2 p_2(1-p_2)$
Num. 3	0
Num. 1,3	$p_1(1-p_1)(1-p_2)^2$
q Den.	$(1-p_1)p_2+p_1$

Case 15	$F_1 \sim \mathcal{U}([0,4])$, $F_2 \sim \mathcal{U}([0,3])$
Prob.	$q_3 q_4$
Num. 1	$p_1(1-p_1)(p_2+(1-p_2)p_3)^2$
Num. 2	$p_1^2 p_2(1-p_2)(1-p_3)^2$
Num. 3	$p_1^2(1-p_2)^2 p_3(1-p_3)$
Num. 1,3	$\mathrm{Var}(\mathbb{1}_{X_1=1}(p_2+(1-p_2)\mathbb{1}_{X_3=1})$
q Den.	$p_1(p_2+(1-p_2)p_3)$

Case 16	$F_1 \sim \mathcal{U}([0,4])$, $F_2 \sim \mathcal{U}([0,4])$
Prob.	q_4^2
Num. 1	$p_1(1-p_1)p_2^2 p_3^2$
Num. 2	$p_1^2 p_2(1-p_2)p_3^2$
Num. 3	$p_1^2 p_2^2 p_3(1-p_3)$
Num. 1,3	$p_1 p_2^2 p_3(1-p_1p_3)$
q Den.	$p_1p_2p_3$

In Figure 1, we have represented the indices $S^1(\mathbb{F})$, $S^2(\mathbb{F})$, $S^3(\mathbb{F})$, and $S^{13}(\mathbb{F})$ given by (20) with respect to the values of p_1 and p_2 varying from 0 to 1 for a fixed value of p_3. We have considered three different values of p_3: $p_3 = 0.01$ (first row), 0.5, (second row) and 0.99 (third row). Analogously, the same kind of illustration for the indices S^1_{2,W_2}, S^2_{2,W_2}, S^3_{2,W_2}, and S^{13}_{2,W_2} given by (18) is provided in Figure 2 . In addition, the regions of predominance of each index $S^{\mathbf{u}}(\mathbb{F})$ are plotted in Figure 3. The values of p_1 and p_2 still vary from 0 to 1 and the fixed values of p_3 considered are: $p_3 = 0.01$ (first row), 0.5, (second row) and 0.99 (third row). Finally, the same kind of illustration for the indices $S^{\mathbf{u}}_{2,W_2}$ is given in Figure 4.

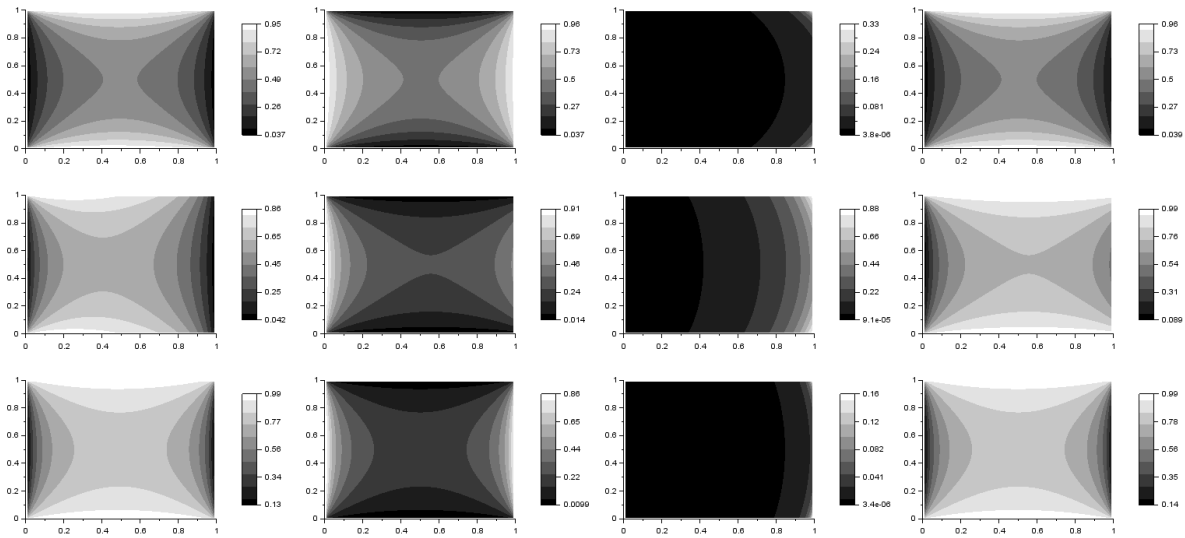

Figure 1: Model (24). Values of the indices $S^1(\mathbb{F})$, $S^2(\mathbb{F})$, $S^3(\mathbb{F})$, and $S^{13}(\mathbb{F})$ given by (20) (from left to right) with respect to the values of p_1 and p_2 (varying from 0 to 1). In the first row (resp. second and third), p_3 is fixed to $p_3 = 0.01$ (resp. 0.5 and 0.99).

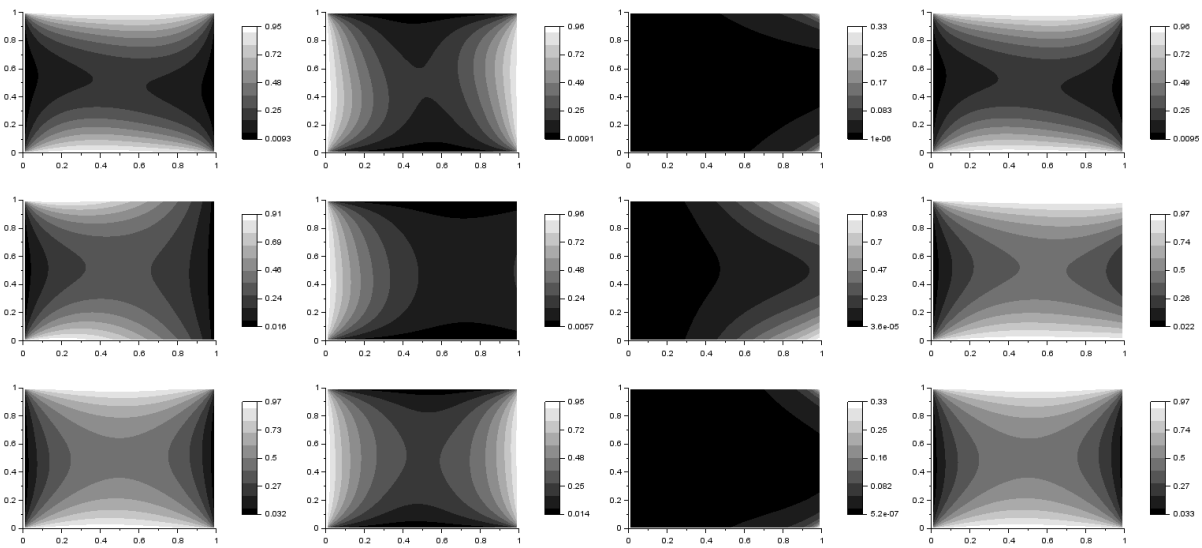

Figure 2: Model (24). Values of the indices S^1_{2,W_2}, S^2_{2,W_2}, S^3_{2,W_2}, and S^{13}_{2,W_2} given by (18) (from left to right) with respect to the values of p_1 and p_2 (varying from 0 to 1). In the first row (resp. second and third), p_3 is fixed to $p_3 = 0.01$ (resp. 0.5 and 0.99).

In order to compare the estimation accuracy of the Pick-Freeze method and the rank-based method at a fixed size, we assume that only $N = 450$ calls of the computer code are allowed to estimate the indices

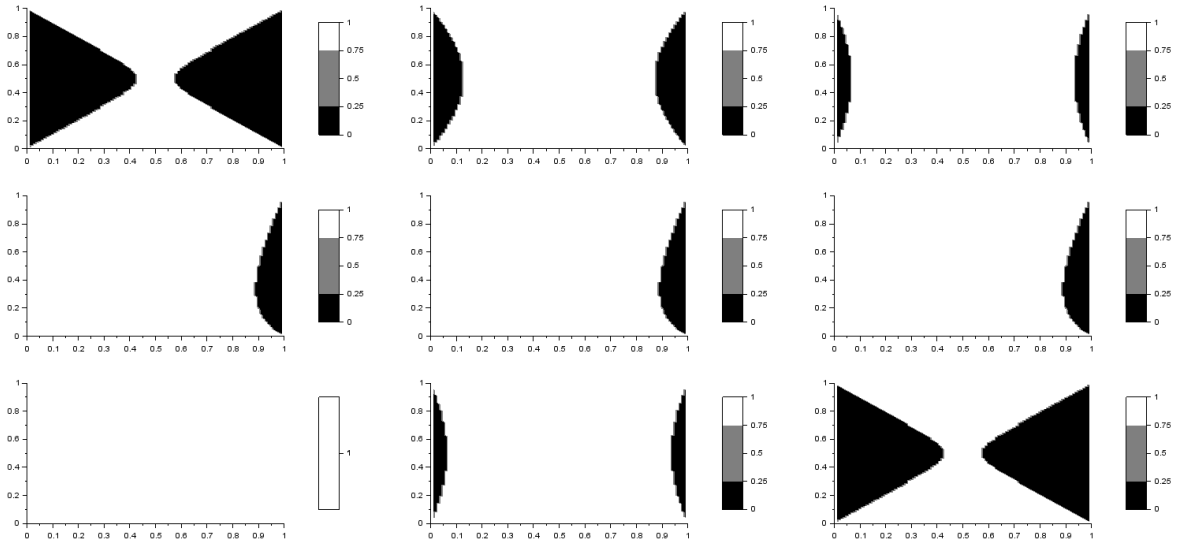

Figure 3: Model (24). In the first row of the figure, regions where $S^1(\mathbb{F}) \geqslant S^2(\mathbb{F})$ (black), $S^1(\mathbb{F}) \leqslant S^2(\mathbb{F})$ (white), and $S^1(\mathbb{F}) = S^2(\mathbb{F})$ (gray) with respect to p_1 and p_2 varying from 0 to 1 and, from left to right, $p_3 = 0.01$, 0.5, and 0.99. Analogously, the second (resp. last) row considers the regions with $S^1(\mathbb{F})$ and $S^3(\mathbb{F})$ (resp. $S^2(\mathbb{F})$ and $S^3(\mathbb{F})$) with respect to p_1 and p_3 (resp. p_2 and p_3) varying from 0 to 1 and, from left to right, $p_2 = 0.01$, 0.5, and 0.99 (resp. $p_1 = 0.01$, 0.5, and 0.99).

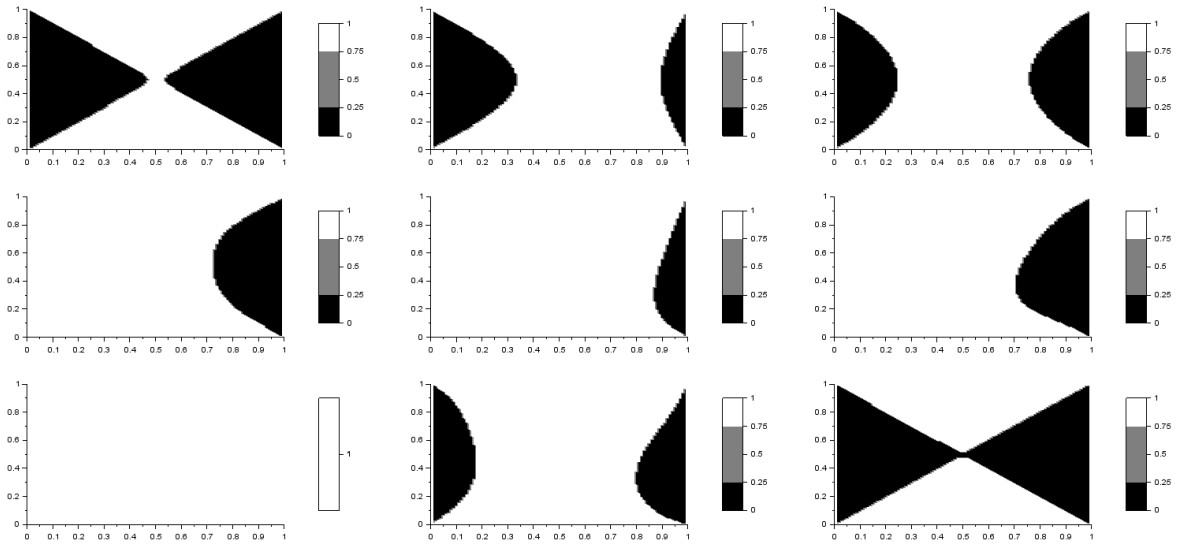

Figure 4: Model (24). In the first row of the figure, regions where $S^1_{2,W_2} \geqslant S^2_{2,W_2}$ (black), $S^1_{2,W_2} \leqslant S^2_{2,W_2}$ (white), and $S^1_{2,W_2} = S^2_{2,W_2}$ (gray) with respect to p_1 and p_2 varying from 0 to 1 and, from left to right, $p_3 = 0.01$, 0.5, and 0.99. Analogously, the second (resp. last) row considers the regions with S^1_{2,W_2} and S^3_{2,W_2} (resp. S^2_{2,W_2} and S^3_{2,W_2}) with respect to p_1 and p_3 (resp. p_2 and p_3) varying from 0 to 1 and, from left to right, $p_2 = 0.01$, 0.5, and 0.99 (resp. $p_1 = 0.01$, 0.5, and 0.99).

$S^{\mathbf{u}}(\mathbb{F})$ and $S^{\mathbf{u}}_{2,W_2}$ for $\mathbf{u} = \{1\}$, $\{2\}$, and $\{3\}$. We only focus on the first order indices since, as explained previously, the rank-based procedure has not been developed yet for higher order indices. We repeat the estimation procedure 500 times. The boxplots of the mean square errors for the estimation of the Fréchet indices $S^{\mathbf{u}}(\mathbb{F})$ and the Wasserstein indices $S^{\mathbf{u}}_{2,W_2}$ have been plotted in Figure 5. We observe that, for a fixed sample size $N = 450$ (corresponding to a Pick-Freeze sample size $N = 64$), the rank-based estimation procedure performs much better than the Pick-Freeze method with significantly lower mean errors.

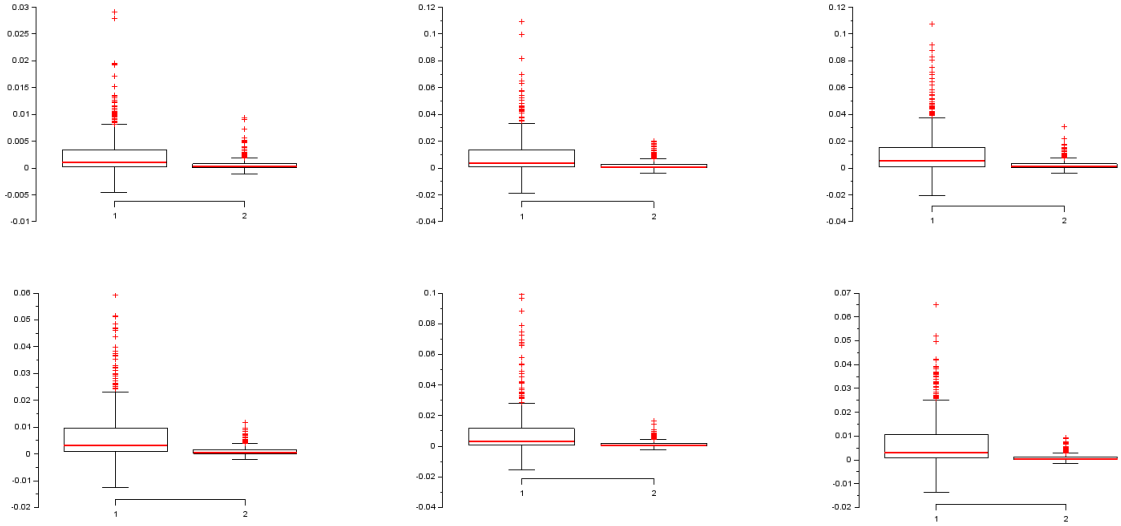

Figure 5: Model (24) with $p_1 = 1/3$, $p_2 = 2/3$, and $p_3 = 3/4$. Boxplots of the mean square errors of the estimation of the Fréchet indices $S^{\mathbf{u}}(\mathbb{F})$ (top row) and the Wasserstein indices $S^{\mathbf{u}}_{2,W_2}$ (bottom row) with a fixed sample size and 500 replications. The indices with respect to $\mathbf{u} = \{1\}$, $\{2\}$, and $\{3\}$ are displayed from left to right. The results of the Pick-Freeze estimation procedure with $N = 64$ are provided in the left side of each graphic. The results of the rank-based methodology with $N = 450$ are provided in the right side of each graphic.

5 Sensitivity analysis for stochastic computer codes

This section deals with stochastic computer codes in the sense that two evaluations of the code for the same input lead to different outputs.

5.1 State of the art

A first natural way to handle stochastic computer codes is definitely to consider the expectation of the output code. Indeed, as mentioned in [9], previous works dealing with stochastic simulators together with robust design or optimization and sensitivity analysis consist mainly in approximating the mean and variance of the stochastic output [17, 10, 37, 2] and then performing a global sensitivity analysis on the expectation of the output code [42].

As pointed out by [35], another approach is to consider that the stochastic code is of the form $f(X, D)$ where the random element X contains the classical input variables and the variable D is an extra unobserved random input.

Such an idea was exploited in [36] to compare the estimation of the Sobol indices in an "exact" model to the estimation of the Sobol indices in an associated metamodel.

Analogously, the author of [43] assumes the existence of an extra random variable D which is not chosen by the practitioner but rather generated at each computation of the output Y independently of X. In this framework, he builds two different indices. The first index is obtained by substituting $f(X, D)$ for $f(X)$ in the classical definition of the first order Sobol index $S^i = \mathrm{Var}(\mathbb{E}[f(X)|X_i])/\mathrm{Var}(f(X))$. In this case, D is considered as another input, even though it is not observable. The second index is obtained by substituting $\mathbb{E}[f(X, D)|X]$ for $f(X)$ in the Sobol index. The noise is then smoothed out.

Similarly, the authors of [31] traduces the randomness of the computer code using such an extra random variable D. In practice, their algorithm returns m realizations of the first order Sobol indices S^i for $i = 1, \ldots, p$, denoted by $\hat{S}^i_j(d_j)$ for $j = 1, \ldots, m$ and $i = 1, \ldots, p$. Then, for any $i = 1, \ldots, p$, they approximate the statistical properties of S^i by considering the sample r-th moments given by

$$\hat{\mu}^i_r = \frac{1}{m} \sum_{j=1}^{m} (\hat{S}^i_j(d_j))^r$$

noticing that

$$\mathbb{E}_D[\hat{\mu}_r^i] = \mathbb{E}[(\hat{S}^i)^r] \quad \text{and} \quad \mathrm{Var}_D(\hat{\mu}_r^i) = \frac{1}{M}\mathrm{Var}_D((\hat{S}^i)^r).$$

5.2 The space $\mathcal{W}_q$ as an ideal version of stochastic computer codes

When dealing with stochastic computer codes, the practitioner is generally interested in the distribution μ_x of the output for a given x. As previously seen, one can translate this type of codes in terms of a deterministic code by considering an extra input which is not chosen by the practitioner itself but which is a latent variable generated randomly by the computer code and independently of the classical input. As usual in the framework of sensitivity analysis, one considers the input as a random variable. All the random variables (the one chosen by the practitioner and the one generated by the computer code) are built on the same probability space, leading to the function f_s:

$$\begin{aligned} f_s : \quad & E \times \mathcal{D} \quad \to \mathbb{R} \\ & (x, D) \quad \mapsto f_s(x, D), \end{aligned} \tag{26}$$

where D is the extra random variable lying in $\mathcal{D}$. We naturally denote the output random variable $f_s(x, \cdot)$ by $f_s(x)$.

Hence, one may define another (deterministic) computer code associated with f_s whose output is the probability measure:

$$\begin{aligned} f : \quad & E \quad \to \mathcal{W}_q(E) \\ & x \quad \mapsto \mu_x. \end{aligned} \tag{27}$$

The framework of (27) is exactly the one of Section 4.1 and has already been handled. Obviously, in practice, one does not assess the output of the code f but one can only obtain an empirical approximation of the measure μ_x given by n evaluations of f_s at x, namely,

$$\mu_{x,n} = \frac{1}{n}\sum_{k=1}^{n} \delta_{f_s(x,D_k)}.$$

Further, (27) can be seen as an ideal version of (26). Concretely, for a single random input $X \in E = E_1 \times \cdots \times E_p$, we will evaluate n times the code f_s defined by (26) (so that the code will generate independently n hidden variables $D_1, \ldots, D_n$) and one may observe

$$f_s(X, D_1), \ldots, f_s(X, D_n)$$

leading to the measure $\mu_{X,n} = \sum_{k=1}^{n} \delta_{f_s(X,D_k)}/n$ approximating the distribution of $f_s(X)$. We emphasize on the fact that the random variables $D_1, \ldots, D_n$ are not observed.

5.3 Sensitivity analysis

Let us now present the methodology we adopt. In order to study the sensitivity of the distribution μ_x, one can use the framework introduced in Section 4.1 and the index $S^{\mathbf{u}}_{2,W_q}$ given by (18).

In an ideal scenario which corresponds to the framework of (27), one may asses to the probability measure μ_x for any x. Then following the estimation procedure of Section 4.3, one gets an estimation of the sensitivity index $S^{\mathbf{u}}_{2,W_q}$ with good asymptotic properties [27, Theorem 2.3].

In the more realistic framework presented above in (26), we only have access to the approximation $\mu_{x,n}$ of μ_x rendering more complex the estimation procedure and the study of the asymptotic properties. In this case, the general design of experiments is the following:

$$\begin{array}{ccc} (X_1, D_{1,1}, \ldots, D_{1,n}) & \to & f_s(X_1, D_{1,1}), \ldots, f_s(X_1, D_{1,n}), \\ (X_1^{\mathbf{u}}, D'_{1,1}, \ldots, D'_{1,n}) & \to & f_s(X_1^{\mathbf{u}}, D'_{1,1}), \ldots, f_s(X_1^{\mathbf{u}}, D'_{1,n}), \\ & \vdots & \\ (X_N, D_{N,1}, \ldots, D_{N,n}) & \to & f_s(X_N, D_{N,1}), \ldots, f_s(X_N, D_{N,n}), \\ (X_N^{\mathbf{u}}, D'_{N,1}, \ldots, D'_{N,n}) & \to & f_s(X_N^{\mathbf{u}}, D'_{N,1}), \ldots, f_s(X_N^{\mathbf{u}}, D'_{N,n}), \end{array}$$

where $2 \times N \times n$ is the total number of evaluations of the stochastic computer code (26). Then we construct the approximations of μ_j (standing for μ_{X_j}) for any $j = 1, \ldots, N$ given by

$$\mu_{j,n} = \frac{1}{n} \sum_{k=1}^{n} \delta_{f_s(X_j, D_{j,k})}. \tag{28}$$

From there, one may use one of the three estimation procedures presented in Section 2.1.

- First method - Pick-Freeze. It suffices to plug the empirical version μ_n of each measure μ under concern in (22).
- Second method - U-statistics. For $l = 1, \ldots, 4$, let

$$U_{l,N,n} = \binom{N}{m(l)}^{-1} \sum_{1 \leqslant i_1 < \cdots < i_{m(l)} \leqslant N} \Phi_l^s \left(\boldsymbol{\mu}_{i_1,n}, \ldots, \boldsymbol{\mu}_{i_{m(l)},n} \right) \tag{29}$$

  where as previously seen $\Phi_\cdot^s$ is the symmetrized version of $\Phi_\cdot$ defined in (6) and $\boldsymbol{\mu} = (\mu, \mu^{\mathbf{u}})$. Then we estimate $S_{2,W_q}^{\mathbf{u}}$ by

$$\widehat{S}_{2,W_q,\text{Ustat},n}^{\mathbf{u}} = \frac{U_{1,N,n} - U_{2,N,n}}{U_{3,N,n} - U_{4,N,n}}. \tag{30}$$

- Third method - Rank-based. The rank-based estimation procedure may also easily extend to this context by using the empirical version μ_n of each measure μ under concern instead of the true one μ, as explained into more details in the numerical study developed in Section 5.1.

Actually, these estimators are easy to compute since for two discrete measures supported on a same number of points and given by

$$\nu_1 = \frac{1}{n} \sum_{k=1}^{n} \delta_{x_k}, \ \nu_2 = \frac{1}{n} \sum_{k=1}^{n} \delta_{y_k},$$

the Wasserstein distance between ν_1 and ν_2 simply writes

$$W_q^q(\nu_1, \nu_2) = \frac{1}{n} \sum_{k=1}^{n} (x_{(k)} - y_{(k)})^q, \tag{31}$$

where $x_{(k)}$ is the k-th order statistics of x.

5.4 Central limit theorem for the estimator based on U-statistics

Proposition 5.1. *Consider three i.i.d. copies X_1, X_2 and X_3 of a random variable X. Let $\delta(N)$ be a sequence tending to 0 as N goes to infinity and such that*

$$\mathbb{P}\left(|W_q(\mu_{X_1}, \mu_{X_3}) - W_q(\mu_{X_1}, \mu_{X_2})| \leqslant \delta(N)\right) = o\left(\frac{1}{\sqrt{N}}\right).$$

Let n such that $\mathbb{E}[W_q(\mu_X, \mu_{X,n})] = o(\delta(N)/\sqrt{N})$. Under the assumptions of [27, Theorem 2.3], we get, for any $\boldsymbol{u} \subset \{1, \cdots, p\}$,

$$\sqrt{N}\left(\widehat{S}_{2,W_q,Ustat,n}^{\boldsymbol{u}} - S_{2,W_q}^{\boldsymbol{u}}\right) \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2) \tag{32}$$

where the asymptotic variance σ^2 is given by Equation (13) in the proof of Theorem 2.3 in [27].

In some particular frameworks, one may derive easily a suitable value of $\delta(N)$. Two examples are given in the following.

Example 5.2. If the inverse of the random variable $W = |W_q(\mu_{X_1}, \mu_{X_3}) - W_q(\mu_{X_1}, \mu_{X_2})|$ has a finite expectation, then, by Markov inequality,

$$\mathbb{P}\left(W \leqslant \delta(N)\right) = \mathbb{P}\left(W^{-1} \geqslant \delta(N)^{-1}\right) \leqslant \frac{1}{\delta(N)} \mathbb{E}\left[\frac{1}{W}\right]$$

and it suffices to choose $\delta(N)$ so that $\delta(N)^{-1} = o\left(N^{-1/2}\right)$ as N goes to infinity.

Example 5.3 (Uniform example)**.** Assume that X is uniformly distributed on $[0,1]$ and that μ_X is a Gaussian distribution centered at X with unit variance. Then the Wasserstein distance $W_2(\mu_{X_1}, \mu_{X_2})$ rewrites as $(X_1 - X_2)^2$ so that the random variable $W = |W_2(\mu_{X_1}, \mu_{X_3}) - W_2(\mu_{X_1}, \mu_{X_2})|$ is given by

$$\left|(X_1 - X_3)^2 - (X_1 - X_2)^2\right| = |(X_3 - X_2)(X_2 + X_3 - 2X_1)|\,.$$

Consequently,

$$\mathbb{P}(W \leqslant \delta(N)) \leqslant \mathbb{P}(|X_3 - X_2| \leqslant \sqrt{\delta(N)}) + \mathbb{P}(|X_2 + X_3 - 2X_1| \leqslant \sqrt{\delta(N)}).$$

Notice that $|X_3 - X_2|$ is triangular distributed with parameter $a = 0$, $b = 1$, and $c = 0$ leading to

$$\mathbb{P}(|X_3 - X_2| \leqslant \alpha) = \alpha(2 - \alpha), \quad \text{for all } \alpha \in [0,1].$$

In addition,

$$\begin{aligned}\mathbb{P}(|X_2 + X_3 - 2X_1| \leqslant \sqrt{\delta(N)}) &\leqslant \mathbb{P}(||X_2 - X_1| - |X_3 - X_1|| \leqslant \sqrt{\delta(N)}) \\ &= \int_0^1 \mathbb{P}(||X_2 - u| - |X_3 - u|| \leqslant \sqrt{\delta(N)})du.\end{aligned}$$

Now, $X_2 - u$ and $X_3 - u$ are two independent random variables uniformly distributed on $[-u, -u]$. Then (see Figure 6), one has

$$\mathbb{P}(||X_2 - u| - |X_3 - u|| \leqslant \alpha) \leqslant 4\alpha,$$

whence

$$\mathbb{P}(|X_2 + X_3 - 2X_1| \leqslant \sqrt{\delta(N)}) \leqslant 4\sqrt{\delta(N)}.$$

Thus it turns out that $\mathbb{P}(W \leqslant \delta(N)) = O\Big(\sqrt{\delta(N)}\Big)$. Consequently, a suitable choice for $\delta(N)$ is $\delta(N) = o(1/N)$.

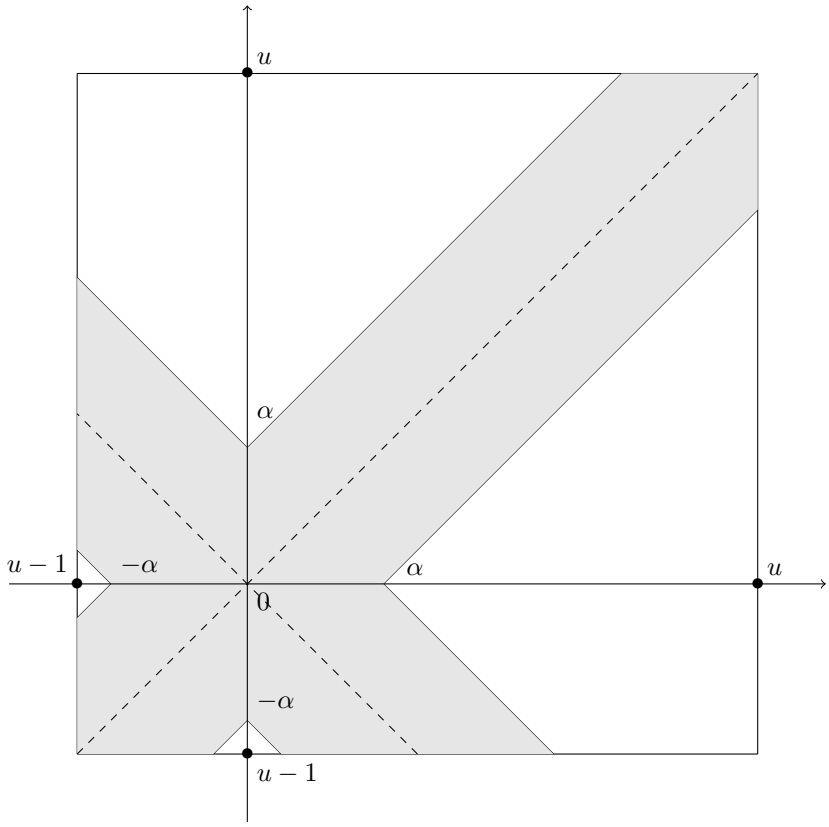

Figure 6: Domain $\Gamma_{u,\alpha} = \{(x_1, x_2) \in [0,1];\ ||x_1 - u| - |x_2 - u|| \leqslant \alpha\}$ (in grey).

Analogously, one may derive suitable choices for n in some particular cases. For instance, we refer the reader to [5] to get upper bounds on $\mathbb{E}[W_q(\mu_X, \mu_{X,n})]$ for several values of $q \geqslant 1$ and several assumptions on the distribution on μ_X: general, uniform, Gaussian, beta, log concave... Here are some results.

- In the general framework, the upper bound for $q \geqslant 1$ relies on the functional

$$J_q(\mu_X) = \int_{\mathbb{R}} \frac{(F_{\mu_X}(x)(1 - F_{\mu_X}(x)))^{q/2}}{f_{\mu_X}(x)^{q-1)}} dx$$

  where F_{μ_X} is the c.d.f. associated to μ_X and f_{μ_X} its p.d.f. See Cf. [5, Theorems 3.2, 5.1 and 5.3].

- Assume that μ_X is uniformly distributed on $[0, 1]$. Then by [5, Theorems 4.7, 4.8 and 4.9], for any $n \geqslant 1$,

$$\mathbb{E}[W_2(\mu_X, \mu_{X,n})^2] \leqslant \frac{1}{6n},$$

  for any $q \geqslant 1$ and for any $n \geqslant 1$,

$$\mathbb{E}[W_q(\mu_X, \mu_{X,n})^q]^{1/q} \leqslant (Const)\sqrt{\frac{q}{n}}.$$

  and for any $n \geqslant 1$,

$$\mathbb{E}[W_\infty(\mu_X, \mu_{X,n})] \leqslant \frac{(Const)}{n}.$$

  E.g. $(Const) = \sqrt{\pi/2}$.

- Assume that μ_X is a log-concave distribution with standard deviation σ. Then by [5, Corollaries 6.10 and 6.12], for any $1 \leqslant q < 2$ and for any $n \geqslant 1$,

$$\mathbb{E}[W_q(\mu_X, \mu_{X,n})^q] \leqslant \frac{(Const)}{2-q}\left(\frac{\sigma}{\sqrt{n}}\right)^q,$$

  for any $n \geqslant 1$,

$$\mathbb{E}[W_2(\mu_X, \mu_{X,n})^2] \leqslant \frac{(Const)\sigma^2 \log n}{n},$$

  and for any $q > 2$ and for any $n \geqslant 1$,

$$\mathbb{E}[W_q(\mu_X, \mu_{X,n})^q] \leqslant \frac{C_q \sigma^q}{n},$$

  where C_q depends on q, only. Furthermore, if μ_X supported on $[a, b]$, then for any $n \geqslant 1$,

$$\mathbb{E}[W_2(\mu_X, \mu_{X,n})^2] \leqslant \frac{(Const)(b-a)^2}{n+1}.$$

  E.g. $(Const) = 4/\ln 2$. Cf. [5, Corollary 6.11].

Example 5.3 - continued. We consider that X is uniformly distributed on $[0, 1]$ and μ_X is a Gaussian distribution centered at X with unit variance. Then, by [5, Corollary 6.14], we have for any $n \geqslant 3$,

$$\mathbb{E}[W_2(\mu_X, \mu_{X,n})^2] \leqslant \frac{(Const)\log\log n}{n},$$

and for any $q > 2$ and for any $n \geqslant 3$,

$$\mathbb{E}[W_q(\mu_X, \mu_{X,n})^q] \leqslant \frac{C_q}{n(\log n)^{q/2}},$$

where C_q depends only on q. Since we have already chosen $\delta(N) = o(N^{-1})$, it remains to take n so that $\log\log n/n = o(N^{-2})$ to fulfill the condition $\mathbb{E}[W_2(\mu_X, \mu_{X,n})] = o(\delta(N)/\sqrt{N})$.

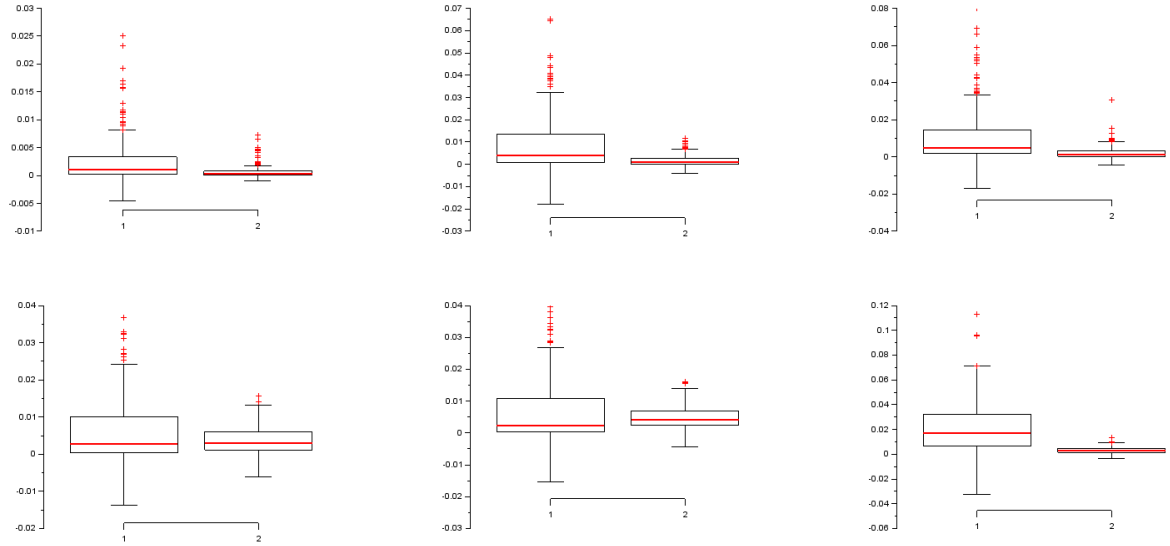

Figure 7: Model (33) with $p_1 = 1/3$, $p_2 = 2/3$, and $p_3 = 3/4$. Boxplot of the mean square errors of the estimation of the Fréchet indices $S^{\mathbf{u}}(\mathbb{F})$ (top row) and the Wasserstein indices $S^{\mathbf{u}}_{2,W_2}$ (bottom row) with a fixed sample size and 200 replications. The indices with respect to $\mathbf{u} = \{1\}$, $\{2\}$, and $\{3\}$ are displayed from left to right. The results of the Pick-Freeze estimation procedure with $N = 64$ are provided in the left side of each graphic. The results of the rank-based methodology with $N = 450$ are provided in the right side of each graphic. The approximation size n is fixed at 500.

5.5 Numerical study

Example 4.1 - continued. Here, we consider again the code given by (24). Having in mind the notation of Section 5.2, we consider the ideal version of the code:

$$\begin{array}{rccl} f: & E & \to & \mathcal{W}_q(E) \\ & (X_1, X_2, X_3) & \mapsto & \mu_{(X_1,X_2,X_3)} \end{array}$$

where $\mu_{(X_1,X_2,X_3)}$ is the uniform distribution on $[0, 1 + X_1 + X_2 + X_1X_3]$, the c.d.f. of which is $\mathbb{F}$ given by (24) and its stochastic counterpart:

$$\begin{array}{rccl} f_s: & E \times D & \to & \mathbb{R} \\ & (X_1, X_2, X_3, D) & \mapsto & f_s(X_1, X_2, X_3, D) \end{array} \tag{33}$$

where $f_s(X_1, X_2, X_3, D)$ is a realization of $\mu_{(X_1,X_2,X_3)}$.

Hence, we no longer assume that one may observe N realizations of $\mathbb{F}$ associated to the N initial realizations of (X_1, X_2, X_3). Instead, for any of the N initial realizations of (X_1, X_2, X_3), we assess n realizations of a uniform random variable on $[0, 1 + X_1 + X_2 + X_1X_3]$.

In order to compare the estimation accuracy of the Pick-Freeze method and the rank-based method at a fixed size, we assume that only $N = 450$ calls of the computer code f are allowed to estimate the indices $S^{\mathbf{u}}(\mathbb{F})$ and $S^{\mathbf{u}}_{2,W_2}$ for $\mathbf{u} = \{1\}$, $\{2\}$, and $\{3\}$. We only focus on the first order indices since, as explained previously, the rank-based procedure has not been developed yet for higher order indices. The empirical c.d.f. based on the empirical measures $\mu_{i,n}$ for $i = 1, \ldots, n$ in (28) are constructed with $n = 100$ evaluations. We repeat the estimation procedure 500 times. The boxplots of the mean square errors for the estimation of the Fréchet indices $S^{\mathbf{u}}(\mathbb{F})$ and the Wasserstein indices $S^{\mathbf{u}}_{2,W_2}$ have been plotted in Figure 7. We observe that, for a fixed sample size $N = 450$ (corresponding to a Pick-Freeze sample size $N = 64$), the rank-based estimation procedure performs much better than the Pick-Freeze method with significantly lower mean errors.

Another numerical study in the particular setting of stochastic computer codes and inspired by [32] will be considered in Section 6.3.

6 Sensitivity analysis with respect to the law of the inputs

6.1 State of the art

The paper [44] is devoted to second level uncertainty which corresponds to the uncertainty on the type of the input distributions and/or on the parameters of the input distributions. As mentioned by the authors, such uncertainties can be handled in two different manners: (1) aggregating them with no distinction [13, 12] or (2) separating them [44]. In [13], e.g., the uncertainty concerns the parameters of the input distributions. The authors study the expectation with respect to the distribution of the parameters of the conditional output. In [12], the second level uncertainties are transformed into first level uncertainties considering the aggregated vector containing the input random variables vector together with the vector of uncertain parameters. Alternatively, in [44], the uncertainty brought by the lack of knowledge of the input distributions and the uncertainty of the random inputs are treated separately. A double Monte-Carlo algorithm is first considered. In the outer loop, a Monte-Carlo sample of input distribution is generated, while the inner loop proceeds to a global sensitivity analysis associated to each distribution. A more efficient algorithm is also proposed with a unique Monte-Carlo loop. The sensitivity analysis is then performed using the so-called Hilbert-Schmidt dependence measures (HSIC indices) on the input distributions rather than the input random variables themselves. See, e.g., [29] for the definition of the HSIC indices and more details on the algorithms.

In [45], a different approach is adopted. A failure probability is studied while the uncertainty concerns the parameters of the input distributions. An algorithm with low computational cost is proposed to handle such uncertainty together with the rare event setting. A single initial sample allows to compute the failure probabilities associated to different parameters of the input distributions. A similar idea is exploited in [41] in which the authors consider input perturbations. and Perturbed-Law based Indices that are used to quantify the impact of a perturbation of an input p.d.f. on a failure probability. Analogously, the authors of [30, 32] are interested in (marginal) p.d.f. perturbations and the aim is to study the "robustness of the Sobol indices to distributional uncertainty and to marginal distribution uncertainty" which correspond to second level uncertainty. For instance, the basic idea of the approach proposed in [30] is to view the total Sobol index as an operator which inputs the p.d.f. and returns the Sobol index. Then the analysis of robustness is done computing and studying the Fréchet derivative of this operator. The same principle is used in [32] to treat the robustness with respect to the marginal distribution uncertainty.

Last but not least, it is worth mentioning the classical approach of epistemic global sensitivity analysis of Dempster-Shafer theory (see, e.g., [54, 1]). This theory describes the random variables together with an epistemic uncertainty traduced in terms of an associated epistemic variable Z on a set A, a mass function representing a probability measure on the set $\mathbb{P}(A)$ of all subsets A. This lack of knowledge leads to an upper and lower bound of the c.d.f. and can be viewed as a second level uncertainty.

6.2 Link with stochastic computer codes

We propose a new procedure that stems from the the methodology in the context of stochastic computer codes described in Section 5. We still denote by μ_i $(i = 1, \ldots, p)$ the distribution of the input X_i $(i = 1, \ldots, p)$ in the model given by (1). There are several ways to model the uncertainty with respect to the choice of each μ_i. Here we adopt the following framework. We assume that each μ_i belongs to some family $\mathcal{P}_i$ of probability measures endowed with the probability measure $\mathbb{P}_{\mu_i}$. In general, there might be measurability issues and the question of how to define a $\sigma-$field on some general spaces $\mathcal{P}_i$ can be tricky. We will restrict our study to the simple case where the existence of the probability measure $\mathbb{P}_{\mu_i}$ on $\mathcal{P}_i$ is given by the construction of the set $\mathcal{P}_i$. More precisely, we proceed as follows.

- First, for $1 \leqslant i \leqslant p$, let d_i be an integer and let $\Theta_i \subset \mathbb{R}^{d_i}$. Then consider the probability space $(\Theta_i, \mathcal{B}(\Theta_i), \nu_{\Theta_i})$ where $\mathcal{B}(\Theta_i)$ is the Borel $\sigma-$field and ν_{Θ_i} is a probability measure on $(\Theta_i, \mathcal{B}(\Theta_i))$.
- Second, for $1 \leqslant i \leqslant p$, we consider an identifiable parametric set of probability measure $\mathcal{P}_i$ on E_i: $\mathcal{P}_i := \{\mu_\theta, \theta \in \Theta_i\}$. Let us denote by π_i the one-to-one mapping from Θ_i to $\mathcal{P}_i$ defined by $\pi_i(\theta) := \mu_\theta \in \mathcal{P}_i$ and define the $\sigma-$field $\mathcal{F}_i$ on $\mathcal{P}_i$ by

$$A \in \mathcal{F}_i \iff \exists B \in \mathcal{B}(\Theta_i),\ A = \pi_i(B).$$

  Then we endow this measurable space with the probability Π_i defined, for any $A \in \mathcal{F}_i$, by

$$\Pi_i(A) = \nu_{\Theta_i}\left(\pi_i^{-1}(A)\right).$$

- Third, in order to perform a second level sensitivity analysis on (1), we introduce the stochastic mapping f_s from $\mathcal{P}_1 \times \ldots \times \mathcal{P}_p$ to $\mathcal{X}$ defined by

$$f_s(\mu_1, \ldots, \mu_p) = f(X_1, \ldots, X_p) \tag{34}$$

  where $X_1, \ldots, X_p$ are independently drawn according to the distribution $\mu_1 \times \ldots \times \mu_p$. Hence f_s is a stochastic computer code from $\mathcal{P}_1 \times \ldots \times \mathcal{P}_p$ to $\mathcal{X}$ and once the probability measures $\mathbb{P}_{\mu_i}$ on each $\mathcal{P}_i$ are defined, we can perform sensitivity analysis using the framework of Section 5.

6.3 Numerical study

As in [32], let us consider the synthetic example defined on $[0,1]^3$ by

$$f(X_1, X_2, X_3) = 2X_2 e^{-2X_1} + X_3^2. \tag{35}$$

We are interested in the uncertainty in the support of the random variables X_1, X_2 and X_3. To do so, we follow the notation and framework of [32]. For $i = 1$, 2, and 3, we assume that X_i is uniformly distributed on the interval $[A_i, B_i]$, where A_i and B_i are themselves uniformly distributed on $[0, 0.1]$ and $[0.9, 1]$ respectively. As remarked in [32], it seems natural that f will vary more in the X_2-direction when X_1 is close to 0 and less when X_1 is close to 1.

As mentioned in Section 6.1, the authors of [32] view the total Sobol index as an operator which inputs the p.d.f. and returns the total Sobol index. Then they study the Fréchet derivative of this operator and determine the most influential p.d.f., which depends on a parameter denoted by δ. Finally, they make vary this parameter δ.

Here, we adopt the methodology explained in the previous section (Section 6.2). Namely, we consider the stochastic computer code given by:

$$f_s(\mu_1, \mu_2, \mu_3) = 2X_2 e^{-2X_1} + X_3^2, \tag{36}$$

where the X_i's are independently drawn according to the uniform measure μ_i on $[A_i, B_i]$ with A_i and B_i themselves uniformly distributed on $[0, 0.1]$ and $[0.9, 1]$ respectively. Then to estimate the indices S^i_{2,W_2}, for $i = 1$, 2, and 3, we proceed as follows.

1. For $i = 1$, 2, and 3, we produce a N-sample $([A_{i,j}, B_{i,j}])_{j=1,\ldots,N}$ of intervals $[A_i, B_i]$.

2. For $i = 1$, 2, and 3, and, for $j = 1, \ldots, N$, we generate a n-sample $(X_{i,j,k})_{k=1,\ldots,n}$ of X_i, where $X_{i,j,k}$ is uniformly distributed on $[A_{i,j}, B_{i,j}]$.

3. For $j = 1, \ldots, N$, we compute the n-sample $(Y_{j,k})_{k=1,\ldots,n}$ of the output using

$$Y = f(X_1, X_2, X_3) = 2X_2 e^{-2X_1} + X_3^2.$$

   Thus we get a N-sample of the empirical measures of the distribution of the output Y given by:

$$\mu_{j,n} = \frac{1}{n} \sum_{k=1}^{n} \delta_{Y_{j,k}}, \quad \text{for } j = 1, \ldots, N.$$

4. For $i = 1$, 2, and 3, we order the intervals $([A_{i,j}, B_{i,j}])_{j=1,\ldots,N}$ and get the Pick-Freeze versions of Y to treat the sensitivity analysis regarding the i-th input.

5. Finally, it remains to compute the indicators of the empirical version of (22) using (31) and their means to get the Pick-Freeze estimators of $S^{\mathbf{u}}_{2,W_2}$, for $\mathbf{u} = \{1\}$, $\{2\}$, $\{3\}$, $\{1,2\}$, $\{1,3\}$, and $\{2,3\}$.

Notice that we only consider the estimators based on the Pick-Freeze method since we allow for both bounds of the interval to vary and, as explained previously, the rank-based procedure has not been developed yet neither for higher order indices nor in higher dimensions.

First, we compute the estimators of $S^{\mathbf{u}}_{2,W_2}$ following the previous procedure with a sample size $N = 500$ and an approximation size $n = 500$. We also perform another batch of simulations allowing for higher variability on the bounds: A_i is now uniformly distributed on $[0, 0.45]$ while B_i is now uniformly distributed on $[0.55, 1]$. The results are displayed in Table 1. Second, we run another simulations allowing

	$\mathbf{u}$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$
$A_i \in [0, 0.1]$ $B_i \in [0.9, 1]$	$\hat{S}^{\mathbf{u}}_{2,W_2}$	0.07022	0.08791	0.09236	0.14467	0.21839	0.19066
$A_i \in [0, 0.45]$ $B_i \in [0.55, 1]$	$\hat{S}^{\mathbf{u}}_{2,W_2}$	0.11587	0.06542	0.169529	0.22647	0.40848	0.34913

Table 1: Model (35). GSA on the parameters of the input distributions. Estimations of $S^{\mathbf{u}}_{2,W_2}$ with a sample size $N = 500$ and an approximation size $n = 500$. In the first row, A_i is uniformly distributed on $[0, 0.1]$ while B_i is uniformly distributed on $[0.55, 1]$. In the second row, we allow for more variability: A_i is uniformly distributed on $[0, 0.45]$ while B_i is uniformly distributed on $[0.55, 1]$.

$\mathbf{u}$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$
$\hat{S}^{\mathbf{u}}_{2,W_2}$	0.01196	0.06069	0.56176	-0.01723	0.63830	0.59434

Table 2: Model (35). GSA on the parameters of the input distributions. Estimations of $S^{\mathbf{u}}_{2,W_2}$ with a sample size $N = 500$ and an approximation size $n = 500$ and more variability on B_3, now uniformly distributed on $[0.5, 1]$.

for more variability on the upper bound related to the third input X_3 only: B_3 is uniformly distributed on $[0.5, 1]$ (instead of $[0.9, 1]$). The results are displayed in Table 2. We still use a sample size $N = 500$ and an approximation size $n = 500$. Third, we perform a classical GSA on the inputs rather than on the parameters of their distributions. Namely, we estimate the index $S^{\mathbf{u}}_{2,CVM}$ with a sample size $N = 10^4$. The reader is referred to [26, Section 3] for the definition of the index $S^{\mathbf{u}}_{2,CVM}$ and its Pick-Freeze estimator together with their properties. The results are displayed in Table 3.

$\mathbf{u}$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$
$\hat{S}^{\mathbf{u}}_{2,W_2}$	0.13717	0.15317	0.33889	0.33405	0.468163	0.53536

Table 3: Model (35). Direct GSA on the inputs. Estimations of $S^{\mathbf{u}}_{2,CVM}$ with a sample size $N = 10^4$. The reader is referred to [26, Section 3] for the definition of the index $S^{\mathbf{u}}_{2,CVM}$ and its Pick-Freeze estimator together with their properties.

In Table 3, we see that the Cramér-von-Mises index related to X_3 is more than twice as important as X_1 and X_2 (when considering only first order effects). Nevertheless, when one is interested in the choice of the input distributions of X_1, X_2, and X_3, the first row in Table 1 shows that each choice is equally important. Now, if one give more freedom to the space where the distribution lives, the relative importance may change as one can see in Table 2 (second row) and in Table 3. More precisely, in Table 2, the variability of the third input distribution (namely, the variability of its upper bound) is five times bigger than the other variabilities. Not surprisingly, it results that the importance of the choice of the third input distribution is then much more important than the choices of the distributions of the two first inputs.

7 Conclusion

In this article, we present a very general way to perform sensitivity analysis when the output Z of a computer code lives in a metric space. The main idea is to consider real-valued squared integrable test functions $(T_a(Z))_{a\in\Omega}$ parameterized by a finite number of elements of a probability space. Then Hoeffding decomposition of the test functions $T_a(Z)$ is computed and integrated with respect to the parameter a. This very general and flexible definition allows, in one hand, to recover a lot of classical indices (namely, the Sobol indices and the Cramér-von-Mises indices) and, in the other hand, to perform a well tailored and interpretable sensitivity analysis. Furthermore, a sensitivity analysis is also made possible for computer

codes the output of which is a c.d.f., for stochastic computer codes (that are seen as an approximation of c.d.f.-valued computer codes). Last but not least, it enables also to perform second level sensitivity analysis by embedding second level sensitivity analysis as a particular case of stochastic computer codes.

Acknowledgment

A Proofs

A.1 Notation

It is convenient to have short expressions for terms that converge in probability to zero. We follow [58]. The notation $o_{\mathbb{P}}(1)$ (respectively $O_{\mathbb{P}}(1)$) stands for a sequence of random variables that converges to zero in probability (resp. is bounded in probability) as $n \to \infty$. More generally, for a sequence of random variables R_n,

$$\begin{aligned} X_n = o_{\mathbb{P}}(R_n) \quad &\text{means} \quad X_n = Y_n R_n \quad \text{with} \quad Y_n \xrightarrow{\mathbb{P}} 0 \\ X_n = O_{\mathbb{P}}(R_n) \quad &\text{means} \quad X_n = Y_n R_n \quad \text{with} \quad Y_n = O_{\mathbb{P}}(1). \end{aligned}$$

For deterministic sequences X_n and R_n, the stochastic notation reduce to the usual o and O. Finally, c stands for a generic constant that may differ from one line to another.

A.2 Proof of Proposition 5.1

One has

$$\sqrt{N}\left(\widehat{S}^{\mathbf{u}}_{2,W_q,\mathrm{Ustat},n} - S^{\mathbf{u}}_{2,GMS}\right) = \sqrt{N}\left(\widehat{S}^{\mathbf{u}}_{2,W_q,\mathrm{Ustat},n} - \widehat{S}^{\mathbf{u}}_{2,GMS,\mathrm{Ustat}}\right) + \sqrt{N}\left(\widehat{S}^{\mathbf{u}}_{2,GMS,\mathrm{Ustat}} - S^{\mathbf{u}}_{2,GMS}\right).$$

By [27, Theorem 2.3], the second term in the right hand side of the previous equation is asymptotically Gaussian. If we prove that the first term in the right hand side is $o_{\mathbb{P}}(1)$, then by Slutsky's Lemma [58, Lemma 2.8], $\sqrt{N}\left(\widehat{S}^{\mathbf{u}}_{2,GMS,\mathrm{Ustat},n} - S^{\mathbf{u}}_{2,GMS}\right)$ is asymptotically Gaussian.

Now we prove that $\sqrt{N}\left(\widehat{S}^{\mathbf{u}}_{2,GMS,\mathrm{Ustat},n} - \widehat{S}^{\mathbf{u}}_{2,GMS,\mathrm{Ustat}}\right) = o_{\mathbb{P}}(1)$. We write

$$\begin{aligned} &\widehat{S}^{\mathbf{u}}_{2,W_q,\mathrm{Ustat},n} - \widehat{S}^{\mathbf{u}}_{2,GMS,\mathrm{Ustat}} = \Psi(U_{1,N,n}, U_{2,N,n}, U_{3,N,n}, U_{4,N,n}) - \Psi(U_{1,N}, U_{2,N}, U_{3,N}, U_{4,N}) \\ &= \frac{\left[(U_{1,N,n} - U_{1,N}) - (U_{2,N,n} - U_{2,N})\right](U_{3,N} - U_{4,N})}{\left[(U_{3,N,n} - U_{3,N}) - (U_{4,N,n} - U_{4,N}) + (U_{3,N} - U_{4,N})\right](U_{3,N} - U_{4,N})} \\ &\quad - \frac{\left[(U_{3,N,n} - U_{3,N}) - (U_{4,N,n} - U_{4,N})\right](U_{1,N} - U_{2,N})}{\left[(U_{3,N,n} - U_{3,N}) - (U_{4,N,n} - U_{4,N}) + (U_{3,N} - U_{4,N})\right](U_{3,N} - U_{4,N})}. \end{aligned}$$

Since $(U_{l,N,n} - U_{l,N,n})$, for $l = 3$ and 4 and $(U_{3,N} - U_{4,N})$ converges almost surely respectively to 0 and $I(\Phi_3) - I(\Phi_4)$, the denominator converges almost surely. Thus it suffices to prove that the numerator is $o_{\mathbb{P}}(1/\sqrt{N})$ which reduces to prove that $\sqrt{N}\left(U_{l,N,n} - U_{l,N}\right) = o_{\mathbb{P}}(1)$ for $l = 1, \ldots, 4$, where $U_{l,N,n}$ (respectively $U_{l,N}$) has been defined in (29) (resp. (9)). Let $l = 1$ for example. The other terms can be treated analogously. Here, $m(1) = 3$. We write

$$\begin{aligned} &\mathbb{E}\left[|U_{1,N,n} - U_{1,N}|\right] \\ &\leqslant \binom{N}{3}^{-1} (3!)^{-1} \sum_{\substack{1 \leqslant i_1 < i_2 < i_3 \leqslant N \\ \tau \in \mathcal{S}_3}} \mathbb{E}\left[\left|\Phi_1\left(\boldsymbol{\mu}_{\tau(i_1),n}, \boldsymbol{\mu}_{\tau(i_2),n}, \boldsymbol{\mu}_{\tau(i_3),n}\right) - \Phi_1\left(\boldsymbol{\mu}_{\tau(i_1)}, \boldsymbol{\mu}_{\tau(i_2)}, \boldsymbol{\mu}_{\tau(i_3)}\right)\right|\right] \\ &= \mathbb{E}\left[\left|\Phi_1\left(\boldsymbol{\mu}_{1,n}, \ldots \boldsymbol{\mu}_{2,n}, \boldsymbol{\mu}_{3,n}\right) - \Phi_1\left(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3\right)\right|\right] \\ &\leqslant 2\mathbb{E}\left[\left|\mathbb{1}_{W_q(\mu_1,\mu_3) \leqslant W_q(\mu_1,\mu_2)} - \mathbb{1}_{W_q(\mu_{1,n},\mu_{3,n}) \leqslant W_q(\mu_{1,n},\mu_{2,n})}\right|\right] \\ &=: 2\mathbb{E}\left[B_n\right] \end{aligned}$$

where the random variable B_n in the expectation in the right hand side of the previous inequality is a Bernoulli random variable whose distribution does not depend on $(1, 2, 3)$. Let $\Delta(N)$ be the following event

$$\Delta(N) = \left\{\left|W_q(\mu_{\tau(1)}, \mu_{\tau(3)}) - W_q(\mu_{\tau(1)}, \mu_{\tau(2)})\right| \geqslant \delta(N)\right\}.$$

Obviously, we get $\mathbb{E}\left[B_n \mathbb{1}_{\Delta(N)^c}\right] \leqslant \mathbb{P}(\Delta(N)^c)$, where A^c stands for the complementary of A in Ω. Furthermore,

$$\begin{aligned}\mathbb{E}\left[B_n \mathbb{1}_{\Delta(N)}\right] &\leqslant \mathbb{E}\left[B_n|\Delta(N)\right] = \mathbb{P}\left(B_n = 1|\Delta(N)\right)\\ &\leqslant \sum_{r=1}^{3} \mathbb{P}\left(W_q(\mu_r, \mu_{r,n}) \geqslant \frac{\delta(N)}{4}\right)\\ &\leqslant \frac{12}{\delta(N)}\mathbb{E}[W_q(\mu_1, \mu_{1,n})].\end{aligned}$$

Finally, we introduce $\varepsilon > 0$ and we study:

$$\begin{aligned}\mathbb{P}\left(\sqrt{N}\,|U_{1,N,n} - U_{1,N}| \geqslant \varepsilon\right) &\leqslant \frac{\sqrt{N}}{\varepsilon}\mathbb{E}\left[|U_{1,N,n} - U_{1,N}|\right]\\ &\leqslant 2\frac{\sqrt{N}}{\varepsilon}\mathbb{E}\left[B_n\right]\\ &\leqslant \frac{\sqrt{N}}{\varepsilon}\frac{24}{\delta(N)}\mathbb{E}[W_q(\mu_1, \mu_{1,n})] + 2\frac{\sqrt{N}}{\varepsilon}\mathbb{P}(\Delta(N)^c).\end{aligned}$$

It remains to choose first, $\delta(N)$ so that $\mathbb{P}(\Delta(N)^c) = o\left(1/\sqrt{N}\right)$ and second, n such that $\mathbb{E}[W_q(\mu_1, \mu_{1,n})] = o(\delta(N)/\sqrt{N})$. Consequently, $\sqrt{N}(U_{1,N,n} - U_{1,N}) = o_{\mathbb{P}}(1)$. Analogously, one gets $\sqrt{N}(U_{l,N,n} - U_{l,N}) = o_{\mathbb{P}}(1)$ for l=2, 3 and 4.

References

[1] D. A. Alvarez. Reduction of uncertainty using sensitivity analysis methods for infinite random sets of indexable type. *International journal of approximate reasoning*, 50(5):750–762, 2009.

[2] B. Ankenman, B. L. Nelson, and J. Staum. Stochastic kriging for simulation metamodeling. In *2008 Winter Simulation Conference*, pages 362–370. IEEE, 2008.

[3] J. D. Betancourt, F. Bachoc, T. Klein, D. Idier, R. Pedreros, and J. Rohmer. Gaussian process metamodeling of functional-input code for coastal flood hazard assessment. *Reliability Engineering and System Safety*, 198, June 2020.

[4] J. Bigot and T. Klein. Consistent estimation of a population barycenter in the Wasserstein space. *ArXiv e-prints*, Dec. 2012.

[5] S. Bobkov and M. Ledoux. One-dimensional empirical measures, order statistics, and kantorovich transport distances. *Memoirs of the American Mathematical Society*, To appear, 2019.

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

[7] E. Borgonovo, W. Castaings, and S. Tarantola. Moment independent importance measures: New results and analytical test cases. *Risk Analysis*, 31(3):404–428, 2011.

[8] E. Borgonovo and B. Iooss. *Moment Independent Importance Measures and a Common Rationale.* Manuscript.

[9] T. Browne, B. Iooss, L. Le Gratiet, J. Lonchampt, and E. Remy. Stochastic simulators based optimization by gaussian process metamodels - application to maintenance investments planning issues. *Quality and Reliability Engineering International*, 32:2067–2080, 2016.

[10] D. Bursztyn and D. M. Steinberg. Screening experiments for dispersion effects. In *Screening*, pages 21–47. Springer, 2006.

[11] S. Cambanis, G. Simons, and W. Stout. Inequalities for $Ek(X, Y)$ when the marginals are fixed. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 36(4):285–294, 1976.

[12] V. Chabridon. *Reliability-oriented sensitivity analysis under probabilistic model uncertainty.* PhD thesis, Université Clermont-Auvergne, 2018.

[13] V. Chabridon, M. Balesdent, J.-M. Bourinet, J. Morio, and N. Gayton. Evaluation of failure probability under parameter epistemic uncertainty: application to aerospace system reliability assessment. *Aerospace Science and Technology*, 69:526–537, 2017.

[14] S. Chatterjee. A new coefficient of correlation. *arXiv e-prints*, page arXiv:1909.10140, Sep 2019.

[15] S. Da Veiga. Global sensitivity analysis with dependence measures. *J. Stat. Comput. Simul.*, 85(7):1283–1305, 2015.

[16] E. De Rocquigny, N. Devictor, and S. Tarantola. *Uncertainty in industrial practice.* Wiley Chisterter England, 2008.

[17] G. Dellino and C. Meloni. *Uncertainty management in simulation-optimization of complex systems.* Springer, 2015.

[18] J. Fontbona, H. Guérin, and S. Méléard. Measurability of optimal transportation and strong coupling of martingale measures. *Electron. Commun. Probab.*, 15:124–133, 2010.

[19] J.-C. Fort, T. Klein, and N. Rachdi. New sensitivity analysis subordinated to a contrast. May 2013.

[20] J.-C. Fort, T. Klein, and N. Rachdi. New sensitivity analysis subordinated to a contrast. *Comm. Statist. Theory Methods*, 45(15):4349–4364, 2016.

[21] R. Fraiman, F. Gamboa, and L. Moreno. Sensitivity indices for output on a Riemannian manifold. *arXiv e-prints*, page arXiv:1810.11591, Oct 2018.

[22] M. Fréchet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. H.Poincaré, Sect. B, Prob. et Stat.*, 10:235–310, 1948.

[23] F. Gamboa, P. Gremaud, T. Klein, and A. Lagnoux. Global Sensitivity Analysis: a new generation of mighty estimators based on rank statistics. Working paper or preprint, Feb. 2020.

[24] F. Gamboa, A. Janon, T. Klein, and A. Lagnoux. Sensitivity analysis for multidimensional and functional outputs. *Electronic Journal of Statistics*, 8:575–603, 2014.

[25] F. Gamboa, A. Janon, T. Klein, A. Lagnoux, and C. Prieur. Statistical inference for Sobol pick-freeze Monte Carlo method. *Statistics*, 50(4):881–902, 2016.

[26] F. Gamboa, T. Klein, and A. Lagnoux. Sensitivity analysis based on Cramér–von Mises distance. *SIAM/ASA J. Uncertain. Quantif.*, 6(2):522–548, 2018.

[27] F. Gamboa, T. Klein, A. Lagnoux, and L. Moreno. Sensitivity analysis in general metric spaces. Working paper or preprint, Feb. 2020.

[28] T. Goda. Computing the variance of a conditional expectation via non-nested monte carlo. *Operations Research Letters*, 45(1):63 – 67, 2017.

[29] A. Gretton, O. Bousquet, A. Smola, and B. Schölkopf. Measuring statistical dependence with Hilbert-Schmidt norms. In *International conference on algorithmic learning theory*, pages 63–77. Springer, 2005.

[30] J. Hart and P. A. Gremaud. Robustness of the Sobol'indices to distributional uncertainty. *International Journal for Uncertainty Quantification*, 9(5), 2019.

[31] J. L. Hart, A. Alexanderian, and P. A. Gremaud. Efficient computation of Sobol'indices for stochastic models. *SIAM Journal on Scientific Computing*, 39(4):A1514–A1530, 2017.

[32] J. L. Hart and P. A. Gremaud. Robustness of the Sobol'indices to marginal distribution uncertainty. *SIAM/ASA Journal on Uncertainty Quantification*, 7(4):1224–1244, 2019.

[33] W. Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics*, 19:293–325, 1948.

[34] D. Idier, A. Aurouet, F. Bachoc, A. Baills, J. D. Betancourt, J. Durand, R. Mouche, J. Rohmer, F. Gamboa, T. Klein, J. Lambert, G. Le Cozannet, S. Leroy, J. Louisor, R. Pedreros, and A.-L. Véron. Toward a User-Based, Robust and Fast Running Method for Coastal Flooding Forecast, Early Warning, and Risk Prevention. *Journal of Coastal Research, Special Issue*, 95:11–15, 2020.

[35] B. Ioss, T. Klein, and A. Lagnoux. Sobol' sensitivity analysis for stochastic numerical codes. In *Proceedings of the SAMO 2016 Conference, Reunion Island, France*, pages 48–49, 2016.

[36] A. Janon, T. Klein, A. Lagnoux, M. Nodet, and C. Prieur. Asymptotic normality and efficiency of two Sobol index estimators. *ESAIM: Probability and Statistics*, 18:342–364, 1 2014.

[37] J. P. Kleijnen. Design and analysis of simulation experiments. In *International Workshop on Simulation*, pages 3–22. Springer, 2015.

[38] S. Kucherenko and S. Song. Different numerical estimators for main effect global sensitivity indices. *Reliability Engineering & System Safety*, 165:222–238, 2017.

[39] M. Lamboni, H. Monod, and D. Makowski. Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models. *Reliability Engineering & System Safety*, 96(4):450–459, 2011.

[40] L. Le Gratiet. Asymptotic normality of a Sobol index estimator in gaussian process regression framework. *Preprint*, 2013.

[41] P. Lemaître, E. Sergienko, A. Arnaud, N. Bousquet, F. Gamboa, and B. Iooss. Density modification-based reliability sensitivity analysis. *Journal of Statistical Computation and Simulation*, 85(6):1200–1223, 2015.

[42] A. Marrel, B. Iooss, S. Da Veiga, and M. Ribatet. Global sensitivity analysis of stochastic computer models with joint metamodels. *Statistics and Computing*, 22(3):833–847, 2012.

[43] G. Mazo. An optimal tradeoff between explorations and repetitions in global sensitivity analysis for stochastic computer models. 2019.

[44] A. Meynaoui, A. Marrel, and B. Laurent. New statistical methodology for second level global sensitivity analysis. working paper or preprint, Feb. 2019.

[45] J. Morio. Influence of input pdf parameters of a model on a failure probability estimation. *Simulation Modelling Practice and Theory*, 19(10):2244–2255, 2011.

[46] V. Moutoussamy, S. Nanty, and B. Pauwels. Emulators for stochastic simulation codes. In *CEMRACS 2013—modelling and simulation of complex systems: stochastic and deterministic approaches*, volume 48 of *ESAIM Proc. Surveys*, pages 116–155. EDP Sci., Les Ulis, 2015.

[47] A. Owen. Better estimation of small Sobol'sensitivity indices. *arXiv preprint arXiv:1204.4763*, 2012.

[48] A. Owen. Variance components and generalized Sobol' indices. *SIAM/ASA Journal on Uncertainty Quantification*, 1(1):19–41, 2013.

[49] A. Owen, J. Dick, and S. Chen. Higher order Sobol' indices. *Information and Inference*, 3(1):59–81, 2014.

[50] K. Pearson. On the partial correlation ratio. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 91(632):492–498, 1915.

[51] N. Peteilh, T. Klein, T. Druot, N. Bartoli, and R. P. Liem. Challenging Top Level Aircraft Requirements based on operations analysis and data-driven models, application to take-off performance design requirements. In *AIAA AVIATION 2020 FORUM*, AIAA AVIATION 2020 FORUM, Reno, NV, United States, June 2020. American Institute of Aeronautics and Astronautics, American Institute of Aeronautics and Astronautics.

[52] A. Saltelli, K. Chan, and E. Scott. *Sensitivity analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 2000.

[53] T. J. Santner, B. Williams, and W. Notz. *The Design and Analysis of Computer Experiments*. Springer-Verlag, 2003.

[54] P. Smets et al. What is Dempster-Shafer's model. *Advances in the Dempster-Shafer theory of evidence*, pages 5–34, 1994.

[55] I. Sobol. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55(1-3):271–280, 2001.

[56] I. M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Math. Modeling Comput. Experiment*, 1(4):407–414, 1993.

[57] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, 93(7):964–979, 2008.

[58] A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.

[59] C. Villani. *Topics in Optimal Transportation*. American Mathematical Society, 2003.