



# Generalized Hoeffding-Sobol Decomposition for Dependent Variables -Application to Sensitivity Analysis

Gaëlle Chastaing, Fabrice Gamboa, Clémentine Prieur

## ► To cite this version:

Gaëlle Chastaing, Fabrice Gamboa, Clémentine Prieur. Generalized Hoeffding-Sobol Decomposition for Dependent Variables -Application to Sensitivity Analysis. 2011. hal-00649404v2

**HAL Id: hal-00649404**

**<https://hal.science/hal-00649404v2>**

Preprint submitted on 15 Dec 2011 (v2), last revised 9 Mar 2012 (v3)

**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.

# Generalized Hoeffding-Sobol Decomposition for Dependent Variables -Application to Sensitivity Analysis

Gaëlle Chastaing<sup>†</sup>

Fabrice Gamboa<sup>‡</sup>

Clémentine Prieur<sup>†</sup>

## Abstract

In this paper, we consider a regression model built on dependent variables. This regression modelizes an input output relationship. Under boundedness assumptions on the joint distribution function of the input variables, we show that a generalized Hoeffding-Sobol decomposition is available. This leads to new indices measuring the sensitivity of the output with respect to the input variables. We also study and discuss the estimation of these new indices.

## List of abbreviations

- SA = Sensitivity Analysis
- HOFD = Hierarchically Orthogonal Functional Decomposition
- IPDV = Independent Pairs of Dependent Variables
- DVP = Da Veiga Procedure

## 1 Introduction

Sensitivity analysis (SA) aims to identify the variables that most contribute to the variability into a non linear regression model. Global SA is a stochastic approach whose objective is to determine a global criterion based on density of the joint probability distribution function of the output and the inputs of a deterministic model. The most usual quantification is the variance-based method, widely studied in SA literature. Hoeffding decomposition [1] (see also Owen [2]) states that the variance of the output can be uniquely decomposed into summands of increasing dimensions under orthogonality constraints. Following this approach, Sobol [3] introduces variability measures, the so called *Sobol sensitivity indices*, to estimate the contribution of each input on a system.

Different methods have been exploited to estimate Sobol indices. The Monte Carlo algorithm was proposed by Sobol [4], and has been later improved by the Quasi Monte Carlo technique, performed by Owen [5]. FAST methods are also widely used to estimate Sobol indices. Introduced earlier by Cukier & al. [6] [7], they are well known to reduce the computational cost of multidimensional integrals thanks to Fourier transformations. Later, Tarantola & al. [8] adapted the Random Balance Designs (RBD) to FAST method for SA (see also recent advances on the subject by Tissot & al. [9]).

---

<sup>†</sup>Université Joseph Fourier, LJK/MOISE BP 53, 38041 Grenoble Cedex, France, [gaelle.chastaing@imag.fr](mailto:gaelle.chastaing@imag.fr), [clementine.prieur@imag.fr](mailto:clementine.prieur@imag.fr)

<sup>‡</sup>Université Paul Sabatier, IMT-EPS, 118, Route de Narbonne, 31062 Toulouse Cedex 9, France, [fabrice.gamboa@univ-tlse.fr](mailto:fabrice.gamboa@univ-tlse.fr)

However, these indices are constructed on the hypothesis that input variables are independent, which seems unrealistic for many real life phenomena. If correlations are not zero, the analytical Hoeffding decomposition is not true anymore. So that, the Sobol indices are irrelevant. If we use them, conclusions on variables sensitivity can be completely wrong even on very simple models (see, for example, Da Veiga's illustration [10]). In terms of estimation, useful methods also rely on the inputs independence. Hence, when they are used on dependent variables, they lead to an important source of errors. It is then necessary to take dependency into account in analytical indices and in estimation methods.

Recent researches have been led to deal with this issue : Jacques & al [11] proposed a model with groups of correlated incomes independent of one another, and expanded classical Sobol indices on these groups. Mara [12] studied the particular case of additive dependence. In order to use Polynomial Chaos as meta models, he orthogonalized inputs by the Gram-Schmidt process to come back to independent variables. Approximating the output by a linear model, Xu & Gertner [13] have proposed to decompose the partial variance due to an input into a correlated and uncorrelated parts. The estimation of these new quantities was made by linear regression. More generally, Li & Rabitz [14] exploited this approach with Generalized Additive Models (GAM). GAM components were approximated by suitable basis functions. Thus, global variance were decomposed and estimated accordingly. An extension of the FAST method has been proposed by Xu & Gertner [15]: The extension is based on the reordering of the independent sample in the traditional FAST. Sensitivity indices are then derived by FAST. The construction of a new generalized moment free sensitivity index was initiated by Borgonovo & al. [16] [17]. Based on geometrical consideration, these indices measure the shift area between the outcome density and this same density conditionally to a parameter.

Thus, none of these works gave an exact and unambiguous definition of the functional ANOVA for correlated inputs as the one provided by Hoeffding-Sobol decomposition when inputs are independent. Consequently, sensitivity measures based on functional ANOVA could neither be deduced in the dependent frame.

In a pionnering work, Hooker [18], inspired by Stone [19], shed new lights on hierarchically orthogonal function decomposition. Nevertheless, the right assumption to get this kind of decomposition is not completely clear in [18], as there is a gap to show the uniqueness of the decomposition. In the present paper, under suitable conditions on the joint distribution function of the input variables, we give a hierarchically orthogonal functional decomposition (HOFD) of the model. The summands of this decomposition are functions depending only on a subset of input variables and are hierarchically decorrelated. This means that two of these components are orthogonal whenever all the variables involved in one of the summands also appear in the other. This decomposition leads to the construction of generalized sensitivity indices well tailored to perform global SA when the input variables are dependent. In the case of independent inputs, this decomposition is nothing more than the Hoeffding one. Furthermore, our generalized sensitivity indices are in this case the classical Sobol ones. In the general case, the computation of the summands of the HOFD involves a minimization problem under constraints (see Proposition 1). This optimization procedure will be studied for statistical purposes in the future. Here, we will focus on the special case where the inputs are independent pairs of dependent variables (IPDV). As a matter of fact, in the simplest case of a single pair of dependent variables, the HOFD may be performed by solving a functional linear system of equations involving suitable projection operators (see Procedure 1). In the more general IPDV case, the HOFD is then obtained in two steps (see Procedure 2). The first step is a classical Hoeffding-Sobol decomposition of the output on the input pairs, as developped in Jacques [11]. The second step is the HOFDs of all the pairs. In practical situations, the non parametric regression function of the model is

generally not exactly known. In this case, one can only have at hand some realizations of the model and have to estimate, with this information, the HOFD. Here, we study this statistical problem in the IPDV case. We build estimators of the generalized sensitivity indices and study numerically their properties. We also compare these generalized indices to the classical Sobol ones (independent case). One of the main conclusion is that the generalized indices have a total normalized sum. This is not true for classical Sobol indices in the frame of dependent variables.

The paper is organized as follows.

In Section 2, we give and discuss general results on the HOFD. The main result is Theorem 1. We show here that a HOFD is available under a boundedness assumption on density of the joint distribution function of the inputs. Further, we introduce the generalized indices.

In Section 3, we give examples of distribution functions that satisfy the boundedness assumption. A sufficient condition is proposed for any model, and more specific conditions are given in the IPDV case. In Section 4, the HOFD is performed by using adapted projection operators in the case of IPDV models. In the case of a single pair of dependent variables, the components of the HOFD are found by solving a functional linear system. Otherwise, a preliminary step of classical decomposition on independent pairs is necessary to come back to the resolution of linear systems. Section 5 is devoted to numerical applications. Through three toy functions, we estimate generalized indices and compare their performances with the analytical values. In Section 6, we give conclusions and discuss upcoming work. Technical proofs and further details are postponed to Section 7.

## 2 Generalized Hoeffding decomposition-Application to SA

To begin with, let introduce some notation. We briefly recall the usual functional ANOVA decomposition, and Sobol indices. We then state a generalization of this decomposition, allowing to deal with correlated inputs.

### 2.1 Notation and first assumptions

We denote by  $\subset$  the strict inclusion, that is  $A \subset B \Rightarrow A \cap B \neq B$ , whereas we use  $\subseteq$  when equality is possible.

Let  $(\Omega, \mathcal{A}, P)$  be a probability space and let  $Y$  be the output of a deterministic model  $\eta$ . Suppose that  $\eta$  is a function of a random vector  $\mathbf{X} = (X_1, \dots, X_p) \in \mathbb{R}^p$ ,  $p \geq 1$  and that  $P_X$  is the pushforward measure of  $P$  by  $\mathbf{X}$ ,

$$Y : \begin{array}{ccccc} (\Omega, \mathcal{A}, P) & \rightarrow & (\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p), P_X) & \rightarrow & (\mathbb{R}, \mathcal{B}(\mathbb{R})) \\ \omega & \mapsto & \mathbf{X}(\omega) & \mapsto & \eta(\mathbf{X}(\omega)) \end{array}$$

Let  $\nu$  be a measure on  $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$ . Assume that  $P_X \ll \nu$  and let  $p_X$  be the density of  $P_X$  with respect to  $\nu$ , that is  $p_X = \frac{dP_X}{d\nu}$ .

Also, assume that  $\eta \in L^2_{\mathbb{R}}(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p), P_X)$ . The associated inner product of this Hilbert space is:

$$\langle h_1, h_2 \rangle = \int h_1(\mathbf{x}) h_2(\mathbf{x}) p_X d\nu(\mathbf{x}) = \mathbb{E}(h_1(\mathbf{X}) h_2(\mathbf{X}))$$

Here  $\mathbb{E}(\cdot)$  denotes the expectation. The corresponding norm will be classically denoted by  $\|\cdot\|$ . Further,  $V(\cdot) = \mathbb{E}[(\cdot - \mathbb{E}(\cdot))^2]$  denotes the variance, and  $\text{Cov}(\cdot, *) = \mathbb{E}[(\cdot - \mathbb{E}(\cdot))(\cdot - \mathbb{E}(\cdot))]$  the covariance.

Let  $\mathcal{P}_p := \{1, \dots, p\}$  and  $S$  be the collection of all subsets of  $\mathcal{P}_p$ . Define  $S^- := S \setminus \mathcal{P}_p$  as the collection of all subsets of  $\mathcal{P}_p$  except  $\mathcal{P}_p$  itself.

Further, let  $X_u := (X_l)_{l \in u}$ ,  $u \in S \setminus \{\emptyset\}$ . We introduce the subspaces of  $L^2_{\mathbb{R}}(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p), P_X)$   $(H_u)_{u \in S}$ ,  $(H_u^0)_{u \in S}$  and  $H^0$ .  $H_u$  is the set of all measurable and square integrable functions depending only on  $X_u$ .  $H_{\emptyset}$  is the set of constants and is identical to  $(H_{\emptyset}^0)_{u \in S}$ .  $H_u^0$ ,  $u \in S \setminus \emptyset$ , and  $H^0$  are defined as follows:

$$H_u^0 = \{h_u(X_u) \in H_u, \langle h_u, h_v \rangle = 0, \forall v \subset u, \forall h_v \in H_v^0\}$$

$$H^0 = \left\{ h(X) = \sum_{u \in S} h_u(X_u), h_u \in H_u^0 \right\}$$

At this stage, we do not make assumptions on the support of  $\mathbf{X}$ . For  $u \in S \setminus \emptyset$ , the support of  $X_u$  is denoted by  $\mathcal{X}_u$ .

## 2.2 Sobol sensitivity indices

In this section, we recall the classical Hoeffding-Sobol decomposition, and the Sobol sensitivity indices if the inputs are independent, that is when  $P_X = P_{X_1} \otimes \dots \otimes P_{X_p}$ .

The usual presentation is done when  $\mathbf{X} \sim \mathcal{U}([0, 1]^p)$  [3], but the Hoeffding decomposition remains true in general case [20].

Let  $\mathbf{x} = (x_1, \dots, x_p) \in \mathbb{R}^p$  and assume that  $\eta \in \mathbb{L}^2(\mathbb{R}^p, P_X)$ . The decomposition consists in writing  $\eta(\mathbf{x}) = \eta(x_1, \dots, x_p)$  as the sum of increasing dimension functions:

$$\begin{aligned} \eta(\mathbf{x}) &= \eta_0 + \sum_{i=1}^p \eta_i(x_i) + \sum_{1 \leq i < j \leq p} \eta_{i,j}(x_i, x_j) + \dots + \eta_{1, \dots, p}(\mathbf{x}) \\ &= \sum_{u \subseteq \{1 \dots p\}} \eta_u(x_u) \end{aligned} \tag{1}$$

The expansion (1) exists and is unique under one of the hypothesis:

$$\left\{ \begin{array}{ll} \text{i)} & \int \eta_u(x_u) dP_{X_i} = 0 \quad \forall i \in u, \forall u \subseteq \{1 \dots p\} \\ \text{or} & \\ \text{ii)} & \int \eta_u(x_u) \eta_v(x_v) dP_X = 0 \quad \forall u, v \subseteq \{1 \dots p\}, \quad u \neq v \end{array} \right.$$

Equation (1) tells us that the model function  $Y = \eta(\mathbf{X})$  can be expanded in a functional ANOVA. The independence of the inputs and the orthogonality properties ensure the global variance decomposition of the output as  $V(Y) = \sum_{u \in S} V(\eta_u(X_u))$ .

Moreover, by integration, each term  $\eta_u$  has an explicit expression, given by:

$$\eta_0 = \mathbb{E}(X), \quad \eta_i = \mathbb{E}(Y/X_i) - \mathbb{E}(Y), \quad i = 1, \dots, p, \quad \eta_u = \mathbb{E}(Y/X_u) - \sum_{v \subset u} \eta_v, \quad |u| \geq 2 \tag{2}$$

Hence, the contribution of a group of variables  $X_u$  in the model can be quantified in the fluctuations of  $Y$ . The Sobol indices expressions are defined by:

$$S_u = \frac{V(\eta_u)}{V(Y)} = \frac{V[\mathbb{E}(Y/X_u)] - \sum_{v \subset u} V[\mathbb{E}(Y/X_v)]}{V(Y)}, \quad u \subseteq \mathcal{P}_p \tag{3}$$

Furthermore,

$$\sum_{u \in S} S_u = 1$$

However, the main assumption is that the input parameters are independent. This is unrealistic in many cases. The use of expressions previously set up is not excluded in case of inputs' dependence, but they could lead to an unobvious and sometimes a wrong interpretation. Also, technics exploited to estimate them may mislead final results because most of them are built on the hypothesis of independence.

For these reasons, the objective of the upcoming work is to show that the construction of sensitivity indices under dependence condition can be done into a mathematical frame.

In the next section, we propose a generalization of the Hoeffding decomposition under suitable conditions on the joint distribution function of the inputs. This decomposition consists of summands of increasing dimension, like in Hoeffding one. But this time, the components are hierarchically orthogonal instead of being mutually orthogonal. The hierarchical orthogonality will be mathematically defined further. Thus, the global variance of the output could be decomposed as a sum of covariance terms depending on the summands of the HOFD. It leads to the construction of generalized sensitivity indices summed to 1 to perform well tailored SA in case of dependence.

### 2.3 Generalized decomposition for dependent inputs

We no more assume that  $P_X$  is a product measure. Nevertheless, we assume:

$$\boxed{\begin{array}{c} \text{where} \\ P_X \ll \nu \\ \nu(dx) = \nu_1(dx_1) \otimes \cdots \otimes \nu_p(dx_p) \end{array}} \quad (\text{C.1})$$

Our main assumption is :

$$\boxed{\exists 0 < M \leq 1, \forall u \subseteq \mathcal{P}_p, \quad p_X \geq M \cdot p_{X_u} p_{X_{u^c}} \quad \nu\text{-a.e.}} \quad (\text{C.2})$$

where  $u^c$  denotes the complement set of  $u$  in  $\mathcal{P}_p$ .  $p_{X_u}$  and  $p_{X_{u^c}}$  are respectively the marginal densities of  $X_u$  and  $X_{u^c}$ .

The section is organized as follows: a preliminary lemma gives the main result to show that  $H^0$  is a complete space. Then, this ensures the existence and the uniqueness of the projection of  $\eta$  onto  $H^0$ . The generalized decomposition of  $\eta$  is finally obtained by adding a residual term orthogonal to every summand, as suggested in [18]. The first part of the reasoning is mostly inspired by Stone's work [19], except that our assumptions are more general. Indeed, we have a relaxed condition on the inputs distribution function. Moreover, the support  $\mathcal{X}$  of  $\mathbf{X}$  is general.

To begin with, let us state some definitions. In the usual ANOVA context, a model is said to be hierarchical if for every term involving some inputs, all lower-order terms involving a subset of these inputs also appear in the model. Correspondingly, a hierarchical collection  $T$  of subsets of  $\mathcal{P}_p$  is defined as follows:

**Definition 1.** A collection  $T \subset S$  is hierarchical if for  $u \in T$  and  $v$  a subset of  $u$ , one has  $v \in T$ .

The next Lemma is a generalization of the Lemma 3.1 of [19]. As already mentioned, it will be the key to show the hierarchical decomposition.

**Lemma 1.** Let  $T \subset S$  be hierarchical. Suppose that (C.1) and (C.2) hold. Set  $\delta = 1 - \sqrt{1 - M} \in ]0, 1]$ . Then, for any  $h_u \in H_u^0$ ,  $u \in T$ , we have:

$$\mathbb{E}[(\sum_{u \in T} h_u(\mathbf{X}))^2] \geq \delta^{\#(T)-1} \sum_{u \in T} \mathbb{E}[h_u^2(\mathbf{X})] \quad (4)$$

The proof of Lemma 1 is postponed to the Appendix. Our main theorem follows:

**Theorem 1.** Let  $\eta$  be any function in  $L_{\mathbb{R}}^2(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p), P_X)$ . Then, under (C.1) and (C.2), there exist functions  $\eta_0, \eta_1, \dots, \eta_{\mathcal{P}_p} \in H_{\emptyset} \times H_1^0 \times \dots \times H_{\mathcal{P}_p}^0$  such that the following equality holds :

$$\begin{aligned} \eta(X_1, \dots, X_p) &= \sum_i \eta_i(X_i) + \sum_{i,j} \eta_{ij}(X_i, X_j) + \dots + \eta_{\mathcal{P}_p}(X_1, \dots, X_p) \\ &= \sum_{u \in S} \eta_u(X_u) \end{aligned} \quad (5)$$

Moreover, this decomposition is unique.

The proof is given in the Appendix.

Notice that, in case where the input variables  $X_1, \dots, X_p$  are independent,  $\delta = 1$  and Inequality (4) of Lemma 1 is an equality. Indeed, in this case, this equality is directly obtained by orthogonality of the summands.

The variational counterpart of Theorem 1 is a minimization problem under conditional constraints.

**Proposition 1.** Suppose that (C.1) and (C.2) hold. Let  $(\mathcal{P})$  be the minimization problem under constraints:

$$(\mathcal{P}) \begin{cases} \min_{(\tilde{\eta}_u)_{u \in S}} \mathbb{E}[(Y - \sum_{u \in S} \tilde{\eta}_u(X_u))^2] \\ \mathbb{E}(\tilde{\eta}_u(X_u)/X_{u \setminus i}) = 0, \forall i \in u, \forall u \in S \setminus \emptyset \end{cases}$$

Then  $(\mathcal{P})$  admits a unique solution  $\eta^* = (\eta_u)_{u \in S}$ .

Proof of Proposition 1 is postponed to the Appendix. Notice that a similar result for the Lebesgue measure is given in [18]. Its purpose was to provide diagnostics for high-dimensional functions. Here, we will no more exploit this idea. This will be done in a forthcoming work. Instead, we are going to construct stochastic sensitivity indices based on the new decomposition (5) and focus on a specific estimation method for IPDV models.

## 2.4 Generalized sensitivity indices

As stated in Theorem 1, under (C.1) and (C.2), the output  $Y$  of the model can be uniquely decomposed as a sum of hierarchically orthogonal terms. Thus, the global variance has a simplified decomposition into a sum of covariance terms. So, we can define generalized sensitivity indices.

**Definition 2.** The sensitivity index  $S_u$  of order  $|u|$  measuring the contribution of  $X_u$  into the model is given by :

$$S_u = \frac{V(\eta_u(X_u)) + \sum_{\substack{v \neq \emptyset \\ u \cap v \neq u, v}} \text{Cov}(\eta_u(X_u), \eta_v(X_v))}{V(Y)} \quad (6)$$

More specifically, the first order sensitivity index  $S_i$  is given by :

$$S_i = \frac{V(\eta_i(X_i)) + \sum_{\substack{v \neq \emptyset \\ i \notin v}} \text{Cov}(\eta_i(X_i), \eta_v(X_v))}{V(Y)} \quad (7)$$

An immediate consequence is given in Proposition 2 (see proof in the Appendix) :

**Proposition 2.** Under (C.1) and (C.2), the sensitivity indices  $S_u$  previously defined are summed to 1, i.e.

$$\sum_{u \in S \setminus \{\emptyset\}} S_u = 1 \quad (8)$$

Thus, sensitivity indices are summed to 1. Furthermore, the covariance terms included in these new indices allow to take into account the inputs dependence. Thus, we are now able to measure the influence of a variable on the model, especially when a part of its variability is embedded into the one of other dependent terms. We can distinguish the full contribution of a variable and its contribution into another correlated income.

Note that for independent inputs, the summands  $\eta_u$  are mutually orthogonal, so  $\text{Cov}(\eta_u, \eta_v) = 0$ ,  $u \neq v$ , and we recover the well known Sobol indices. Hence, these new sensitivity indices can be seen as a generalization of Sobol indices.

However, the HOFD and subsequent indices are only obtained under constraints (C.1) and (C.2). In the following, we give illustrations of distribution functions satisfying these main assumptions.

### 3 Examples of distribution function

This section is devoted to examples of distribution function satisfying (C.1) and (C.2). The first hypothesis only implies that the reference measure is a product of measures, whereas the second is trickier to obtain.

In the first part, we give a sufficient condition to get (C.2) for any number  $p$  of input variables. The second part deals with the case  $p = 2$ , for which we give equivalences of (C.2) in terms of copulas.

#### 3.1 Boundedness of the inputs density function

The difficulty of Condition (C.2) is that the inequality has to be true for any splitting of the set  $(X_1, \dots, X_p)$  into two disjoint blocks. We give a sufficient condition for (C.2) to hold in Proposition 3 (the proof is postponed to the Appendix):

**Proposition 3.** Assume that there exist  $M_1, M_2 > 0$  with



$$\boxed{M_1 \leq p_X \leq M_2} \quad (C.3)$$

Then, Condition (C.2) holds.

Let give now an example where (C.3) is satisfied.

**Example 1:** Let  $\nu$  be the multidimensional gaussian distribution  $N_p(m, \Sigma)$  with

$$m = \begin{pmatrix} m_1 \\ \vdots \\ m_p \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & \sigma_p^2 \end{pmatrix}$$

Assume that  $P_X$  is a Gaussian mixture  $\alpha \cdot N_p(m, \Sigma) + (1 - \alpha) \cdot N_p(\mu, \Omega)$ ,  $\alpha \in ]0, 1[$  with

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}, \quad \Omega = \begin{pmatrix} \varphi_1^2 & \rho_{12} & \cdots & \rho_{1p} \\ & \ddots & & \\ \rho_{1p} & \cdots & \cdots & \varphi_p^2 \end{pmatrix}$$

Then, (C.3) holds iff the matrix  $(\Omega^{-1} - \Sigma^{-1})$  is positive definite.

In the next section, we will see that (C.2) has a copula version when  $p = 2$ . We will give some examples of distribution satisfying one of these conditions.

### 3.2 Examples of distribution of two inputs

Here, we consider the simpler case of inputs  $\mathbf{X} = (X_1, X_2)$ . Also, until Section 4, we will assume that  $\nu$  is absolutely continuous with respect to Lebesgue measure. The structure of dependence of  $X_1$  and  $X_2$  can be modeled by copulas. Copulas [21] give a relationship between a joint distribution and its marginals. Sklar's theorem [22] ensures that for any distribution function  $F(x_1, x_2)$  with marginal distributions  $F_1(x_1)$  and  $F_2(x_2)$ ,  $F$  has the copula representation,

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$$

where the measurable function  $C$  is unique whenever  $F_1$  and  $F_2$  are absolutely continuous.

The next corollary gives in the absolutely continuous case the relationship between a joint density and its marginal:

**Corollary 1.** *In terms of copulas, the joint density of  $\mathbf{X}$  is given by:*

$$p_X(x_1, x_2) = c(F_1(x_1), F_2(x_2))p_{X_1}(x_1)p_{X_2}(x_2) \quad (9)$$

Furthermore,

$$c(u, v) = \frac{\partial^2 C}{\partial u \partial v}(u, v), \quad (u, v) \in [0, 1]^2 \quad (10)$$

Now, Condition (C.2) may be rephrased in terms of copulas:

**Proposition 4.** *For a two-dimensional model, the three following conditions are equivalent:*

$$1. \quad \boxed{p_X \geq M \cdot p_{X_1} p_{X_2} \quad \nu\text{-a.e.} \quad \text{for some } 0 < M < 1} \quad (C.4)$$

$$2. \quad \boxed{c(u, v) \geq M, \quad \forall (u, v) \in [0, 1]^2} \quad (C.5)$$

$$3. \quad \boxed{C(u, v) = Muv + (1 - M)\tilde{C}(u, v), \forall (u, v) \in [0, 1]^2 \text{ for some 2-D copula } \tilde{C}} \quad (C.6)$$

The proof of Proposition 4 is postponed to the Appendix.

Hence, the generalized Hoeffding decomposition holds for a wide class of examples. The first example is the Morgenstern copulas [23]:

**Example 2:** The expression of the Morgenstern copulas is given by:

$$C_\theta(u, v) = uv[1 + \theta(1 - u)(1 - v)], \quad \theta \in [-1, 1]$$

For  $\theta \in ]-1, 1[$ , (C.6) holds, and

$$C_\theta(u, v) = (1 - |\theta|)uv + |\theta|uv[1 + \frac{\theta}{|\theta|}(1 - u)(1 - v)]$$

Let now consider the class of Archimedian copulas,

$$C(u, v) = \varphi^{-1}[\varphi(u) + \varphi(v)], \quad u, v \in [0, 1] \quad (11)$$

where the generator  $\varphi$  is a non negative two times differentiable function defined on  $[0, 1]$  with  $\varphi(1) = 0$ ,  $\varphi'(u) < 0$  and  $\varphi''(u) > 0$ ,  $\forall u \in [0, 1]$ .

A sufficient condition for (C.5) is given in Proposition 5:

**Proposition 5.** *If there exist  $M_1, M_2 > 0$  such that:*

$$-\varphi'(u) \geq M_1 \quad \forall u \in [0, 1] \quad (12)$$

$$\frac{d}{du} \left( \frac{1}{2} \frac{1}{\varphi'(u)^2} \right) \geq M_2, \quad \forall u \in [0, 1] \quad (13)$$

Then, Condition (C.5) holds.

The proof is straightforward. Now, we will see three illustrative Archimedian copulas satisfying (C.5).

**Example 3:** The Frank copula is characterized by the generator:

$$\varphi_1(x) = \log \left( \frac{e^{-\theta x} - 1}{e^{-\theta} - 1} \right), \quad \theta \in \mathbb{R} \setminus \{0\}$$

Condition (C.5) holds, and  $c(u, v) \geq -\theta(e^{-\theta} - 1)e^{-2\theta}$  if  $\theta > 0$ ,  $c(u, v) \geq -\theta(e^{-\theta} - 1)$  elsewhere.

The next two examples also satisfy (C.5) by the intermediate Proposition 5.

**Example 4:** Let  $\alpha < 0$ ,  $\theta > 0$  and  $\beta$  with  $\beta < -\alpha e^{-\theta}$ . Set

$$\varphi_2(x) = -\frac{\alpha}{\theta} e^{-\theta x} + \beta x + \left( \frac{\alpha}{\theta} e^{-\theta} - \beta \right), \quad x \in [0, 1] \quad (14)$$

**Example 5:** Let  $C < 0$  and set

$$\varphi_3(x) = x \ln x + (C - 1)x + (1 - C), \quad x \in [0, 1] \quad (15)$$

Leaving the class of copulas, we now directly work with the joint density function. Proposition 6 gives a general form of distribution for our framework:

**Proposition 6.** *If  $p_X$  has the form*

$$p_X(x_1, x_2) = \alpha \cdot f_{X_1}(x_1)f_{X_2}(x_2) + (1 - \alpha) \cdot g_X(x_1, x_2), \quad \alpha \in ]0, 1[ \quad (16)$$

where  $f_{X_1}$ ,  $f_{X_2}$  are univariate density functions, and  $g_X$  is any density function (with respect to  $\nu$ ) with marginals  $f_{X_1}$  and  $f_{X_2}$ , then  $p_X$  satisfies (C.5).

The proof is straightforward.

**Example 6:** As an illustration of Proposition 6, take  $\nu = \nu_L$ ,  $f_X = f_{X_1}f_{X_2}$  a normal density with a diagonal covariance matrix  $\Sigma$ , and  $g_X$  a normal density of covariance matrix  $\Omega$ , with  $\Omega_{ii} = \Sigma_{ii}$ ,  $i = 1, 2$ . Notice that because a copula of Gaussian mixture distribution is a mixture of Gaussian copulas (see [24]), this example can be directly recovered by the copula approach.

**Example 7:** Let generalize Example 6. If  $P_X$  is a Gaussian mixture

$$\alpha \cdot N_2(m, \Sigma) + (1 - \alpha) \cdot N_2(\mu, \Omega), \quad \alpha \in ]0, 1[$$

with

$$m = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}, \quad \Omega = \begin{pmatrix} \omega_1^2 & \rho\omega_1\omega_2 \\ \rho\omega_1\omega_2 & \omega_2^2 \end{pmatrix}, \quad \rho \in ]-1, 1[$$

then (C.4) holds iff  $\omega_1^2 \leq \sigma_1^2$  and  $\omega_2^2 \leq \sigma_2^2$ .

Thus, for many distributions, the generalized decomposition holds, and generalized sensitivity indices may thus be defined.

For the remaining part of the paper, we will assume that the set of inputs is an IPDV. If  $p$  is odd, we will assume that an input variable is independent to all the others.

The next section is devoted to the estimation of HOFD components. The simplest case of a single pair of dependent variables is first discussed. Then, the more general IPDV case is studied. In this last part, first and second order indices are defined to measure the contribution of each pair of dependent variables and each of its components in the model. Indices of order greater than one involving variables from different pairs will not be studied here.

## 4 Estimation

Using the property of hierarchical orthogonality ( $H_u^0 \perp H_v^0, \forall v \subset u$ ), we will see that the summands of the decomposition are solution of a functional linear system. For  $u \in S$ , the projection operator onto  $H_u^0$  is denoted by  $P_{H_u^0}$ .

In this section, we present the HOFD terms computation, based on the resolution of a functional linear system. The result relies on projection operators previously set up. Further, we expose the linear system estimation for practice.

#### 4.1 Models of $p = 2$ input variables

This part is devoted to the simple case of bidimensional models. Let

$$Y = \eta(X_1, X_2)$$

Assuming that Conditions (C.1) and (C.2) both hold, we proceed as follows:

##### **Procedure 1**

1. *HOFD of the output:*

$$Y = \eta_0 + \eta_1(X_1) + \eta_2(X_2) + \eta_{12}(X_1, X_2) \quad (17)$$

2. *Projection of  $Y = \eta(\mathbf{X})$  on  $H_u^0$ ,  $\forall u \subseteq \{1, 2\}$ . As  $H_u^0 \perp H_v^0$ ,  $\forall v \subset u$ , we obtain:*

$$\begin{pmatrix} Id & 0 & 0 & 0 \\ 0 & Id & P_{H_1^0} & 0 \\ 0 & P_{H_2^0} & Id & 0 \\ 0 & 0 & 0 & Id \end{pmatrix} \begin{pmatrix} \eta_0 \\ \eta_1 \\ \eta_2 \\ \eta_{12} \end{pmatrix} = \begin{pmatrix} P_{H_\emptyset}(\eta) \\ P_{H_1^0}(\eta) \\ P_{H_2^0}(\eta) \\ P_{H_{12}^0}(\eta) \end{pmatrix} \quad (18)$$

3. *Computation of the right-hand side vector of (18):*

$$\begin{pmatrix} P_{H_\emptyset}(\eta) \\ P_{H_1^0}(\eta) \\ P_{H_2^0}(\eta) \\ P_{H_{12}^0}(\eta) \end{pmatrix} = \begin{pmatrix} \mathbb{E}(\eta) \\ \mathbb{E}(\eta/X_1) - \mathbb{E}(\eta) \\ \mathbb{E}(\eta/X_2) - \mathbb{E}(\eta) \\ \eta - \mathbb{E}(\eta/X_1) - \mathbb{E}(\eta/X_2) + \mathbb{E}(\eta) \end{pmatrix} \quad (19)$$

In this frame, we have:

**Proposition 7.** *Let  $\eta$  be any function of  $L_{\mathbb{R}}^2(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p), P_X)$ . Then, under (C.1) and (C.2), the linear system*

$$(\mathcal{S}) \begin{pmatrix} Id & 0 & 0 & 0 \\ 0 & Id & P_{H_1^0} & 0 \\ 0 & P_{H_2^0} & Id & 0 \\ 0 & 0 & 0 & Id \end{pmatrix} \begin{pmatrix} h_0 \\ h_1 \\ h_2 \\ h_{12} \end{pmatrix} = \begin{pmatrix} P_{H_\emptyset}(\eta) \\ P_{H_1^0}(\eta) \\ P_{H_2^0}(\eta) \\ P_{H_{12}^0}(\eta) \end{pmatrix} \quad (20)$$

admits in  $h = (h_0, \dots, h_{12}) \in H_\emptyset \times \dots \times H_{12}^0$  the unique solution  $h^* = (\eta_0, \eta_1, \eta_2, \eta_{12})$ .

4. *Reduction of the system (18). As the constant term corresponds to the expected value of  $\eta$ , and the residual one can be deduced from the others, the dimension of the system (20) can even be reduced to:*

$$\underbrace{\begin{pmatrix} Id & P_{H_1^0} \\ P_{H_2^0} & Id \end{pmatrix}}_{A_2} \underbrace{\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}}_{\Delta} = \underbrace{\begin{pmatrix} \mathbb{E}(\eta/X_1) - \mathbb{E}(\eta) \\ \mathbb{E}(\eta/X_2) - \mathbb{E}(\eta) \end{pmatrix}}_B \quad (21)$$

5. *Practical resolution:* The numerical resolution of (21) is achieved by an iterative Gauss Seidel algorithm [25] which consists first in decomposing  $A_2$  as a sum of lower triangular ( $L_2$ ) and strictly upper triangular ( $U_2$ ) matrices.

Further, the technique uses an iterative scheme to compute  $\Delta$ . At step  $k + 1$ , we have:

$$\Delta^{(k+1)} := \begin{pmatrix} \Delta_1^{(k+1)} \\ \Delta_2^{(k+1)} \end{pmatrix} = L_2^{-1}(B - U_2 \cdot \Delta^{(k)}) \quad (22)$$

Using expression of  $A_2$ , we get:

$$\Delta^{(k+1)} = \begin{pmatrix} \mathbb{E}(Y - \Delta_2^{(k)}/X_1) - \mathbb{E}(Y - \Delta_2^{(k)}) \\ \mathbb{E}(Y - \Delta_1^{(k+1)}/X_2) - \mathbb{E}(Y - \Delta_1^{(k+1)}) \end{pmatrix} \quad (23)$$

6. *Convergence of the algorithm:* now, we hope that the Gauss Seidel algorithm converges to the true solution. Looking back at (18), we see that we only have to consider  $P_{H_1^0}$  (respectively  $P_{H_2^0}$ ) restricted to  $H_2^0$  (respectively to  $H_1^0$ ).

Under this restriction, let us define the associated norm operator as :

$$\|P_{H_i^0}\|^2 := \sup_{\substack{\mathbb{E}(U^2)=1 \\ U \in H_j^0}} \mathbb{E}[P_{H_i^0}(U)^2], \quad i, j = 1, 2, j \neq i$$

As explained in [26], Gauss Seidel algorithm converges to the true solution  $\Delta$  if  $A_2$  is strictly diagonally dominant, which is implied by :

$$\|P_{H_i^0}\| < \|Id\|, \quad i = 1, 2 \quad (24)$$

i.e.

$$\sup_{\substack{\mathbb{E}(U^2)=1 \\ U \in H_j^0}} \mathbb{E}[P_{H_i^0}(U)^2] < 1, \quad i = 1, 2, j \neq i \quad (25)$$

Equality (19) reveals that  $P_{H_i^0}(U) = \mathbb{E}(U/X_i) - \mathbb{E}(U)$ . Hence, by the Jensen inequality [27] for conditional expectations,  $\|P_{H_i^0}\|, i = 1, 2$  admits an upper bound:

Take  $U \in H_1^0$  :

$$\begin{aligned} \|P_{H_2^0}\| &= \sup_{\substack{\mathbb{E}(U^2)=1 \\ U \in H_1^0}} \mathbb{E}[(\mathbb{E}(U/X_2) - \mathbb{E}(U))^2] = \sup_{\substack{\mathbb{E}(U^2)=1 \\ U \in H_1^0}} \mathbb{E}[\mathbb{E}(U/X_2)^2] \quad \text{as } U \in H_1^0 \\ &\leq \sup_{\substack{\mathbb{E}(U^2)=1 \\ U \in H_1^0}} \mathbb{E}[\mathbb{E}(U^2/X_2)] = 1 \quad \text{by Jensen} \end{aligned}$$

The same holds for  $U \in H_2^0$ , and we also have  $\|P_{H_2^0}\| \leq 1$ .

Moreover, the Jensen's inequality is strict if  $U$  is not  $X_i$ -measurable. As  $U$  is a function of  $X_j$  (that is  $j = 2$  if  $i = 1$  and conversely), the condition of convergence holds if  $X_1$  is not a measurable function of  $X_2$ . Hence, Gauss Seidel algorithm converges if  $X_1$  is not a function of  $X_2$ .

7. *Estimation procedure:* Suppose that we get a sample of  $n$  observations  $(Y_k, \mathbf{X}_k)_{k=1, \dots, n}$ .

- *Estimation of the components of the HOFD:* the iterative scheme (23) requires to estimate conditional expectations. As extended in Da Veiga & al [28], we propose to estimate them by local polynomial regression at each point of observation. Then, we use the leave-one-out technique to set the learning sample and the test sample. Moreover, as the local polynomial method can be summed up to a generalized least squares (see Fan & Gijbels [29]), the Sherman-Morrison formula [30] is applied to reduce the computational time.

A more detailed procedure is given in the Appendix. The iterative algorithm is easy to implement. We stop when  $\|\Delta^{(k+1)} - \Delta^{(k)}\| \leq \varepsilon$ , for a small positive  $\varepsilon$ .

Once  $(\eta_1, \eta_2)$  have been estimated, we estimate  $\eta_0$  by the empirical mean of the output. Then, an estimation of  $\eta_{12}$  is obtained by subtraction.

- We use empirical variance and covariance estimation to estimate sensitivity indices  $S_1, S_2$  and  $S_{12}$ .

## 4.2 Generalized IPDV models

Assume that the number of inputs is even, so  $p = 2k$ ,  $k \geq 2$ . We note each group of dependent variables as  $X^{(i)} := (X_1^{(i)}, X_2^{(i)})$ ,  $i = 1, \dots, k$ . By rearrangement, we may assume that:

$$\mathbf{X} = (\underbrace{X_1, X_2, \dots, X_{2k-1}, X_{2k}}_{X^{(1)}}, \dots, \underbrace{X_{2k-1}, X_{2k}}_{X^{(k)}})$$

SA for IPDV models has already been treated in [11]. Indeed, they proposed to estimate usual sensitivity indices on groups of variables via Monte Carlo estimation. Thus, they have interpreted the influence of every group of variables on the global variance. Here, we will go further by trying to measure the influence of each variable on the output, but also the effects of the independent pairs.

To begin with, as a slight generalization of [3] and used in [11], let apply the Sobol decomposition on groups of dependent variables,

$$\eta(\mathbf{X}) = \eta_0 + \eta_1(X^{(1)}) + \dots + \eta_k(X^{(k)}) + \sum_{|u|=2}^k \eta_u(X^{(u)})$$

where for  $u = \{u_1, \dots, u_t\}$  and  $t = |u|$ , we set  $X^{(u)} = (X^{(u_1)}, \dots, X^{(u_t)})$ . Furthermore,  $\langle \eta_u, \eta_v \rangle = 0$ ,  $\forall u \neq v$ .

Thus, we obtain independent components of IPDV. Under the assumptions discussed in the previous section, we can apply the HOFD of each of these components, that is,

$$\eta_i(X^{(i)}) = \eta_i(X_1^{(i)}, X_2^{(i)}) = \varphi_{i0} + \varphi_{i,1}(X_1^{(i)}) + \varphi_{i,2}(X_2^{(i)}) + \varphi_{i,12}(X^{(i)})$$

with  $\langle \varphi_{i,u}, \varphi_{i,v} \rangle = 0$ ,  $\forall v \subset u \subseteq \{1, 2\}$ . In this way, let define some new generalized indices for IPDV models:

**Definition 3.** For  $i = 1, \dots, k$ , the first order sensitivity index measuring the contribution of  $X_1^{(i)}$  (respectively  $X_2^{(i)}$ ) on the output of the model is:

$$S_{i,1} = \frac{V(\varphi_{i,1}) + \text{Cov}(\varphi_{i,1}, \varphi_{i,2})}{V(Y)}, \quad \left( \text{resp. } S_{i,2} = \frac{V(\varphi_{i,2}) + \text{Cov}(\varphi_{i,1}, \varphi_{i,2})}{V(Y)} \right) \quad (26)$$

The second order sensitivity index for the pair  $X^{(i)}$ ,  $i = 1, \dots, k$ , is defined as:

$$S_{i,12} = \frac{V(\varphi_{i,12})}{V(Y)} \quad (27)$$

The estimation procedure for these new indices is quite similar to Procedure 1:

**Procedure 2**

1. Estimation of  $(\eta_i)_{i=1,\dots,k}$ : as reminded in Part 2.2 with Equations (2),  $\eta_i = \mathbb{E}(Y/X^{(i)}) - \mathbb{E}(Y)$ . Step 7 of Procedure 1 gives method to estimate the conditional expectations. So that, we will have estimations of  $\eta_i$ ,  $i = 1, \dots, k$ .
2. For  $i = 1, \dots, k$ , we apply step 2 to step 7 of Procedure 1, considering  $\eta_i$  as the output.

If  $p$  is odd, the procedure is the same except that the influence of the independent variable is measured by a Sobol index, as it is independent from all the others.

The next part is devoted to numerical examples.

## 5 Numerical examples

In this section, we study three examples with dependent input variables. We consider IPDV models and a Gaussian mixture distribution on the input variables. We choose covariance matrices of the mixture satisfying conditions of Example 1.

We give estimations of our new indices, and compare them to the analytical ones, computed from expressions (6). We also compute dispersions of the estimated new indices.

In [28], Da Veiga & al. proposed to estimate the classical Sobol indices  $S_u = (V[\mathbb{E}(Y/X_u)] - \sum_{v \subset u} V[\mathbb{E}(Y/X_v)]) / V(Y)$ ,  $u \subseteq \mathcal{P}_p$ , by nonparametric tools. Indeed, the local polynomial regression were used to estimate conditional moments  $\mathbb{E}(Y/X_u)$ ,  $u \subseteq \mathcal{P}_p$ . This method, used further, will be called Da Veiga procedure (DVP). Results given by DVP are compared with the ones given by our method. The goal is to show that the usual sensitivity indices are not appropriate in the dependence frame, even if a relevant estimation method is used.

### 5.1 Two-dimensional IPDV model

Let consider the model

$$Y = X_1 + X_2 + X_1 X_2$$

Here,  $\nu$  and  $P_X$  are of the form given by Example 1, with  $m = \mu = 0$ .

Thus, the analytical decomposition of  $Y$  is

$$\eta_0 = \mathbb{E}(X_1 X_2), \quad \eta_1 = X_1, \quad \eta_2 = X_2 \quad \eta_{12} = X_1 X_2 - \mathbb{E}(X_1 X_2)$$

For the application, we implement Procedure 1 in Matlab software. We proceed to  $L = 50$  simulations and  $n = 1000$  observations. Parameters were fixed at  $\sigma_1 = \sigma_2 = 1$ ,  $\varphi_1^2 = \varphi_2^2 = 0.5$ ,  $\rho_{12} = 0.4$  and  $\alpha = 0.2$ .

In Table 1, we give the estimation of our indices and their standard deviation (indicated by  $\pm \cdot$ ) on  $L$  simulations. In comparison, we give the analytical value of each index.

The analytical classical Sobol indices are difficult to obtain, but we give estimators of the classical Sobol indices with DVP.

		$S_1$	$S_2$	$S_{12}$	$\sum_u S_u$
New indices	Estimation	$0.42 \pm 0.041$	$0.41 \pm 0.043$	$0.17 \pm 0.026$	$1 \pm 9.10^{-16}$
	Analytical	0.39	0.39	0.22	1
DVP indices	Estimation	$0.64 \pm 0.045$	$0.65 \pm 0.044$	$0.41 \pm 0.038$	$1.7 \pm 0.09$

Table 1: Estimation of the new and DVP indices with  $\rho_{12} = 0.4$

We notice that estimations with our method give quite good results in comparison with their analytical values. The estimation error of the interaction term is due to the fact that the component  $\hat{\eta}_{12}$  is obtained by difference between the output and the other estimated components. The DVP indices are difficult to interpret as the sum is higher than 1.

In our method, it would be relevant to separate the variance part to the covariance one in the first order indices. Indeed, in this way, we would be able to get the part of variability explained by  $X_i$  alone in  $S_i$ , and its contribution hidden in the dependence with  $X_j$ . We note  $S_i^v$  the variance contribution alone, and  $S_i^c$  the covariance contribution, that is

$$S_i = \underbrace{\frac{V(X_i)}{V(Y)}}_{S_i^v} + \underbrace{\frac{\text{Cov}(X_i, X_j)}{V(Y)}}_{S_i^c}, \quad i = 1, 2, j \neq i$$

The new indices estimations given in Table 1 are decomposed in Table 2. As previously, the number at the right of  $\pm$  indicates the standard deviation on  $L$  simulations.

	$S_i^v$	$S_i^c$	$S_i$
$X_1$	$0.28 \pm 0.04$	$0.14 \pm 0.01$	$0.42 \pm 0.041$
$X_2$	$0.27 \pm 0.043$	$0.14 \pm 0.01$	$0.41 \pm 0.043$
Analytical	0.25	0.14	0.39

Table 2: Estimation of  $S_i^v$  and  $S_i^c$  with  $\rho_{12} = 0.4$

For each index, the covariate itself explains 28% (in estimation, 25% in reality) of the part of the total variability. However, the contribution embedded in the correlation is not negligible as it represents 14% of the total variance. Considering the shape of the model, and coefficients of parameters distribution, it is quite natural to get the same contribution of  $X_1$  and  $X_2$  into the global variance. Also, as their dependence is quite important with a covariance term equals to 0.4, we are not surprised by the relatively high value of  $S_1^c$  (resp.  $S_2^c$ ).

From now, we take  $\rho_{12} = 0$ , i.e. we assume that the inputs are independent. Let compare our new estimated indices with their analytical values in Table 3. We again decompose new indices into a variance ( $S_i^v$ ) and a covariance ( $S_i^c$ ) contribution.



	New indices			Theoretical
	$S_i^v$	$S_i^c$	$S_i$	$S_i$
$X_1$	$0.39 \pm 0.039$	$-0.01 \pm 0.01$	$0.38 \pm 0.036$	0.375
$X_2$	$0.38 \pm 0.045$	$-0.01 \pm 0.024$	$0.37 \pm 0.037$	0.375
$X_1X_2$	$0.26 \pm 0.038$	$-0.01 \pm 0.01$	$0.25 \pm 0.027$	0.25

Table 3: Comparison between analytical and estimated indices with  $\rho_{12} = 0$

Thus, the new indices are well tailored if we have a small idea on inputs dependence in a system. Indeed, Table 3 shows that our new indices take dependence into account if it exists, and the covariance contribution is estimated by 0 if not. New indices recover the classical Sobol indices in case of independence.

## 5.2 Linear Four-dimensional model

The test model is

$$Y = 5X_1 + 4X_2 + 3X_3 + 2X_4$$

Actually, Condition (C.2) only needs to be satisfied on groups of correlated variables. Let consider the two blocks  $X^{(1)} = (X_1, X_3)$  and  $X^{(2)} = (X_2, X_4)$  of correlated variables. The previous form of density can be taken for  $X^{(1)}$  and  $X^{(2)}$ .  $P_{X^{(i)}}$  is then the Gaussian mixture  $\alpha_i \cdot N_2(0, I_2) + (1 - \alpha_i) \cdot N_2(0, \Omega_i)$ ,  $i = 1, 2$ . The analytical sensitivity indices are given by (26) & (27).

For  $L = 50$  simulations and  $n = 1000$  observations, we took  $\varphi_1^{2(1)} = \varphi_2^{2(1)} = 0.5$ ,  $\varphi_1^{2(2)} = 0.7$ ,  $\varphi_2^{2(2)} = 0.3$ ,  $\rho_{12}^{(1)} = 0.4$ ,  $\rho_{12}^{(2)} = 0.37$  and  $\alpha_1 = \alpha_2 = 0.2$ .

Figure 1 displays the dispersion of indices of first order for all variables and second order for grouped variables. We compare them to their analytical values. In the same figure, we also represented the estimators of classical Sobol indices with DVP.

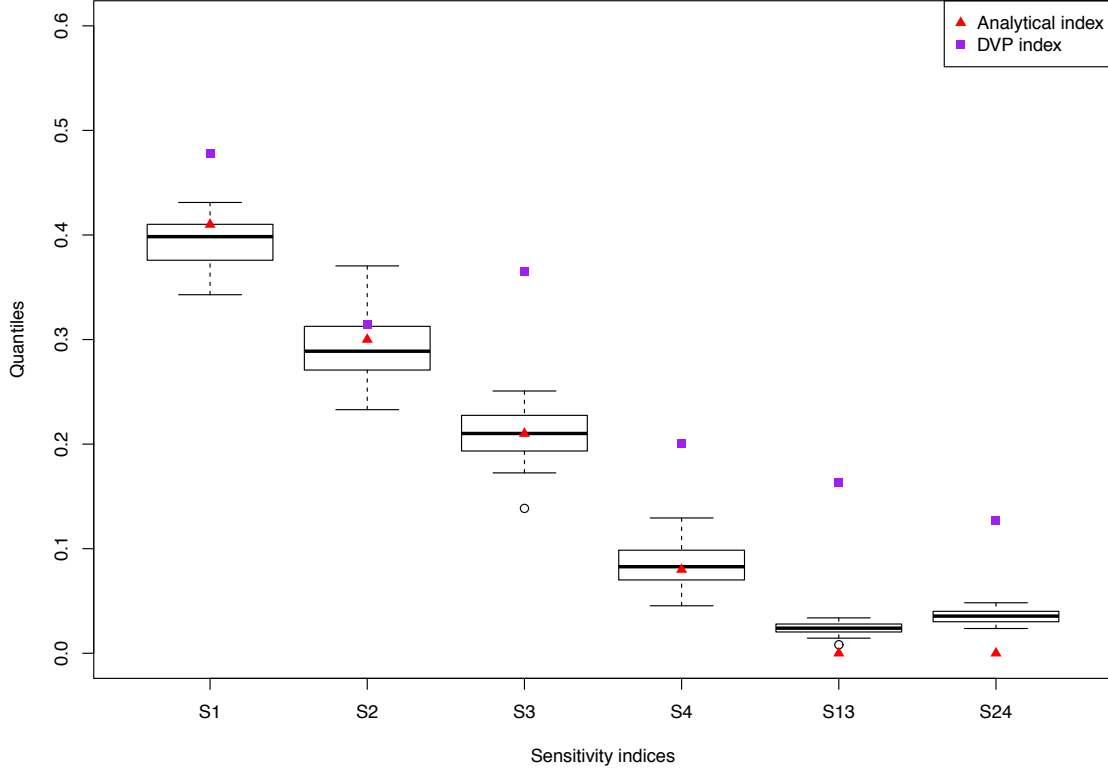


Figure 1: Boxplots representation of new indices-Comparison with analytical and DVP indices

We see that  $X_1$  has the biggest contribution, whereas the influence of  $X_4$  is very low. It reflects well the model if we look at the coefficients of  $X_i$ ,  $i = 1, \dots, 4$ . Also, interaction terms are well estimated, as they are closed to 0. For each case, the dispersion on 50 simulations is very low. As for the DVP estimation, it is once again very high compared with the true indices values.

### 5.3 The Ishigami function

This function is well known in SA ([31]). It is defined by:

$$Y = \sin(X_1) + a \sin^2(X_2) + bX_3^3 \sin(X_1)$$

We assume that  $X_3$  is the independent variable, and that  $X_1$  and  $X_2$  are correlated.  $P_X$  is again the Gaussian mixture  $\alpha \cdot N_3(0, I_3) + (1 - \alpha) \cdot N_3(0, \Omega)$ .

With  $L = 50$  simulations of  $n = 1000$  observations, we fixed parameters of distribution at  $\varphi_1^2 = 0.15$ ,  $\varphi_2^2 = 0.85$ ,  $\varphi_3^2 = 0.75$ ,  $\rho_{12} = 0.3$  and  $\alpha = 0.2$ .

In Figure 2, the dispersion of the new measures is represented for fixed  $b = 0.1$  and different values of  $a$ . In addition, the analytical new indices and estimated classical Sobol indices with DVP are displayed.

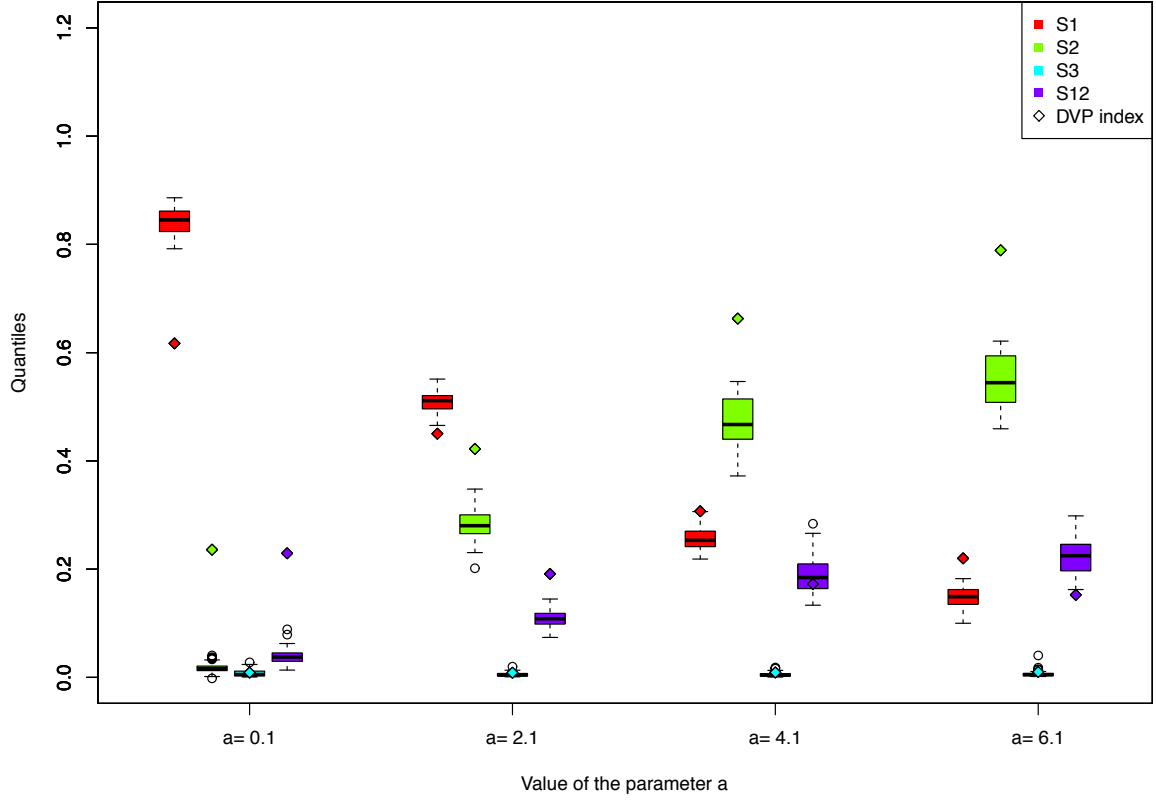


Figure 2: Boxplots of new indices for different values of  $a$  for Ishigami function

Every boxplot shows that there is a small dispersion. For all estimations, DVP indices are larger than the new ones. The figure clearly shows that, for all values of  $a$ , the sum of these four indices is greater than 1. It again shows their non adaptation to a situation of dependence.

If we have a look on the values taken by our new sensitivity indices, we see that, for small values of  $a$ , the variable  $X_1$  contributes the most to the model's variability. This role decreases as  $a$  increases, and  $X_2$  then gets the biggest contribution. For any value of  $a$ , the input  $X_3$  plays a very negligible role, which seems realistic as  $b$  is a small fixed value. As for the interaction index  $S_{12}$ , it is getting bigger with the increasing importance of  $a$ , but the contribution remains low.

## 6 Conclusions and Perspectives

The Hoeffding decomposition and associated Sobol indices have been widely studied in SA over past years. Recently, a literature appears to treat the case of dependent input variables, in which authors propose different ways to deal with dependence. The goal of this paper is to conciliate the problem of inputs dependence with the Hoeffding decomposition. Indeed, we study a functional ANOVA decomposition in a generalized inputs distribution frame. Thus, we show that a model can be uniquely decomposed as a sum of hierarchically orthogonal functions of increasing dimension. Also, this approach generalizes the Hoeffding's one, as we recover it in case of independence.

Similarly to the classical Sobol decomposition, this leads to the construction of new sensitivity indices. They consist of a variance and a covariance contribution able to take into account the possible correlation among variables. In case of independence, these indices are the classical Sobol indices. However, the indices construction is only possible under specific assumptions on the joint distribution function of the inputs. We expose few cases that satisfy these assumptions for any  $p$ -dimensional models. More specifically, for two-dimensional models, the required assumption is equivalent to assumptions on copulas. In this context, we give examples satisfying one of these assumptions .

Focused on the IPDV models, summands of the decomposition are estimated thanks to projection operations. This leads to the numerical resolution of functional linear systems. The strength of this method is that it does not require to make assumptions on the form of the model or on the structure of dependence. We neither use meta-modelling and avoid in this way many sources of errors.

Through three applications on test models, we observe the importance of considering the inputs correlation, and show how our method could catch it. The comparison with estimators of classical indices with DVP shows that the Sobol indices are not appropriate in case of correlations, even when using nonparametric method. Also, when inputs independence holds, the new indices remain well suited to measure sensitivity into a model.

Nevertheless, only considering IPDV models for estimation is restrictive. The perspective is to explore other estimation methods suitable for more general models. Also, we intend to lead a systematic study on copulas satisfying or not our assumptions.

## 7 Appendix

### A Generalized Hoeffding decomposition-Application to SA

#### A.1 Generalized decomposition for dependent inputs

The upcoming proof follows the guideline of the proof of Lemma 3.1 in Stone [19].

*Proof of Lemma 1*

By induction on the cardinal of  $T$ , let show that

$$\mathcal{H}(n) : \forall T/\#(T) = n, \quad \mathbb{E}[(\sum_{u \in T} h_u(\mathbf{X}))^2] \geq \delta^{\#(T)-1} \sum_{u \in T} \mathbb{E}[h_u^2(\mathbf{X})]$$

- $\mathcal{H}(1)$  is obviously true, as  $T$  is reduced to a singleton
- Let  $n \in \mathbb{N}^*$ . Suppose that  $\mathcal{H}(n')$  is true for all  $1 \leq n' \leq n$ . Let  $T$  such that  $\#(T) = n + 1$ . We want to prove  $\mathcal{H}(n + 1)$ .

Choose a maximal set  $r$  of  $T$ , i.e.  $r$  is not a proper subset of any set  $u$  in  $T$ . We show first that

$$\mathbb{E}[(\sum_{u \in T} h_u(\mathbf{X}))^2] \geq M \cdot \mathbb{E}(h_r^2(\mathbf{X})) \quad (28)$$

- If  $\#(r) = p$ , by definition of  $H_r^0$ , we get  $\mathbb{E}[(\sum_{u \in T} h_u(\mathbf{X}))^2] \geq \mathbb{E}(h_r^2(\mathbf{X})) \geq M\mathbb{E}(h_r^2(\mathbf{X}))$  as  $M \leq 1$ .
- If  $1 \leq \#(r) \leq p - 1$ , set  $\mathbf{X} = (X_1, X_2)$ , where  $X_1 = (X_l)_{l \notin r}$  and  $X_2 = (X_l)_{l \in r}$ . By Condition (C.2), it follows that

$$p_X \geq M \cdot p_{X_1} p_{X_2}$$

As a consequence,

$$\begin{aligned} \mathbb{E}[(\sum_{u \in T} h_u(\mathbf{X}))^2] &= \int_{\mathcal{X}_1} \int_{\mathcal{X}_2} [h_r(x_2) + \sum_{u \neq r} h_u(x_1, x_2)]^2 p_X \nu(dx_1, dx_2) \\ &\geq M \int_{\mathcal{X}_1} \int_{\mathcal{X}_2} [h_r(x_2) + \sum_{u \neq r} h_u(x_1, x_2)]^2 p_{X_1} p_{X_2} \nu_1(dx_1) \nu_2(dx_2) \\ &\geq M \int_{\mathcal{X}_1} \mathbb{E}[(h_r(X_2) + \sum_{u \neq r} h_u(x_1, X_2))^2] p_{X_1} \nu_1(dx_1) \end{aligned}$$

By maximality of  $r$  and by definition of  $H_r^0$ ,

- \* If  $u \subset r$ ,  $h_u$  only depends on  $X_2$  and by orthogonality,

$$\mathbb{E}(h_u(X_2)h_r(X_2)) = 0$$

- \* If  $u \not\subset r$ ,  $h_u$  depends on  $X_1$  fixed at  $x_1$ , and  $X_2^u = (X_l)_{l \in r \cap u}$ , so  $h_u \in H_{r \cap u}^0$ , with  $r \cap u \subset r$ , it comes then

$$\mathbb{E}(h_u(x_1, X_2)h_r(X_2)) = 0$$

Thus,

$$\begin{aligned} \mathbb{E}[(\sum_{u \in T} h_u(\mathbf{X}))^2] &\geq M \int_{\mathcal{X}_1} \mathbb{E}(h_r^2(X_2)) p_{X_1} \nu_1(dx_1) \\ &= M \cdot \mathbb{E}(h_r^2(\mathbf{X})) \end{aligned}$$

So (28) holds for any size of any maximal sets of  $T$ .

By using (28) with  $\tilde{h}_r = h_r$  and  $\tilde{h}_u = -\beta h_u$ ,  $\forall u \neq r$ , we get

$$\mathbb{E}[(h_r(\mathbf{X}) - \beta \sum_{u \neq r} h_u(\mathbf{X}))^2] \geq M\mathbb{E}(h_r^2(\mathbf{X})) \quad (29)$$

Taking  $\beta = \frac{\mathbb{E}[h_r(\mathbf{X}) \sum_{u \neq r} h_u(\mathbf{X})]}{\mathbb{E}[(\sum_{u \neq r} h_u(\mathbf{X}))^2]}$ , it follows that:

$$\begin{aligned} \mathbb{E}[(h_r(\mathbf{X}) - \beta \sum_{u \neq r} h_u(\mathbf{X}))^2] &\geq M \mathbb{E}(h_r^2(\mathbf{X})) \\ \text{that is } \mathbb{E}[h_r^2(\mathbf{X})] - \frac{[\mathbb{E}(h_r(\mathbf{X}) \sum_{u \neq r} h_u(\mathbf{X}))]^2}{\mathbb{E}[(\sum_{u \neq r} h_u(\mathbf{X}))^2]} &\geq M \mathbb{E}(h_r^2(\mathbf{X})) \end{aligned}$$

Hence,

$$[E(h_r(\mathbf{X}) \sum_{u \neq r} h_u(\mathbf{X}))]^2 \leq (1 - M) \cdot \mathbb{E}(h_r^2(\mathbf{X})) \cdot \mathbb{E}[(\sum_{u \neq r} h_u(\mathbf{X}))^2] \quad (30)$$

This implies

$$\mathbb{E}[(\sum_u h_u(\mathbf{X}))^2] \geq (1 - \sqrt{1 - M}) \left[ \mathbb{E}(h_r^2(\mathbf{X})) + \mathbb{E}[(\sum_{u \neq r} h_u(\mathbf{X}))^2] \right] \quad (31)$$

Set  $x = h_r(\mathbf{X})$  and  $y = \sum_{u \neq r} h_u(\mathbf{X})$ . (31) is rephrased as

$$\|x + y\|^2 \geq (1 - \sqrt{1 - M})\{\|x\|^2 + \|y\|^2\} \quad (32)$$

Further, (30) is  $\langle x, y \rangle \geq -\sqrt{1 - M}\|x\| \cdot \|y\|$ . Thus,

$$\begin{aligned} \|x\|^2 + \|y\|^2 &\geq 2\langle x, y \rangle \\ &\geq -\frac{2}{\sqrt{1 - M}}\langle x, y \rangle \quad \text{by (32)} \end{aligned}$$

So  $\|x + y\|^2 \geq (1 - \sqrt{1 - M})\{\|x\|^2 + \|y\|^2\}$ .

As  $\mathcal{H}(n)$  is supposed to be true and (31) holds, it follows that:

$$\begin{aligned} \mathbb{E}[(\sum_u h_u(\mathbf{X}))^2] &\geq \delta \left[ \mathbb{E}(h_r^2(\mathbf{X})) + \delta^{n-1} \sum_{u \neq r} \mathbb{E}(h_u^2(\mathbf{X})) \right] \\ &\geq \delta^n \sum_u \mathbb{E}(h_u^2(\mathbf{X})) \quad \text{as } \delta \in ]0, 1] \\ &= \delta^{\#(T)-1} \sum_u \mathbb{E}(h_u^2(\mathbf{X})) \end{aligned}$$

Hence,  $\mathcal{H}(n+1)$  holds.

We can deduce that  $\mathcal{H}(n)$  is true for any collection  $T$  of  $\mathcal{P}_p$ . ■

### Proof of Theorem 1

Let define the vector space  $K^0 = \{\sum_{u \in S^-} h_u(X_u), h_u \in H_u^0, \forall u \in S^-\}$ .

In the first step, we will prove that  $K^0$  is a complete space to prove the existence and uniqueness of the projection of  $\eta$  in  $K^0$ , thanks to the projection theorem [32].

Secondly, we will show that  $\eta$  is exactly equal to the decomposition into  $H^0$ , and finally, we will see that each term of the summand is unique.

- We show that  $H_u^0$  is closed into  $H_u$  (as  $H_u$  is a Hilbert space).

Let  $(h_{n,u})_n$  be a convergent sequence of  $H_u^0$  with  $h_{n,u} \rightarrow h_u$ . As  $(h_{n,u})_n \in H_u^0 \subset H_u$  complete,  $h_u \in H_u$ . Let  $v \subset u$ , and  $h_v \in H_v^0$  :

$$\begin{array}{ccc}
\langle h_u - h_{n,u}, h_v \rangle & = & \langle h_u, h_v \rangle - \langle h_{n,u}, h_v \rangle \\
\downarrow & & \parallel \\
0 & & 0 \quad \text{as } H_u^0 \perp H_v^0
\end{array}$$

Thus,  $\langle h_u, h_v \rangle = 0$ , so that  $h_u \in H_u^0$ .  $H_u^0$  is then a complete space.

Let  $(h_n)_n$  be a Cauchy sequence in  $K^0$  and we show that each component is of Cauchy and that  $h_n \rightarrow h \in K^0$ .

As  $h_n \in K^0$ ,  $h_n = \sum_{u \in S^-} h_{n,u}$ ,  $h_{n,u} \in H_u^0$ . It follows that :

$$\begin{aligned}
\|h_n - h_m\|^2 &= \|\sum_u (h_{n,u} - h_{m,u})\|^2 \\
&\geq \delta^{\#(S^-)-1} \sum_{u \in S^-} \|h_{n,u} - h_{m,u}\|^2 \quad \text{by Inequality (4)}
\end{aligned}$$

As  $(h_n)_n$  is a Cauchy sequence, by the above inequality,  $(h_{n,u})_n$  is also Cauchy. As  $h_{n,u} \rightarrow h_u \in H_u^0$ , we deduce that  $h_n \xrightarrow{n \rightarrow \infty} \sum_{u \in S^-} h_u = h \in K^0$ .

Thus,  $K^0$  is complete. By the projection theorem, we can deduce there exists a unique element into  $K^0$  such that :

$$\|\eta - \sum_{u \in S^-} \eta_u\|^2 \leq \|\eta - h\|^2 \quad \forall h \in K^0$$

- *Decomposition of  $\eta$ : following Hooker [18], we introduce the residual term as*

$$\eta_{\mathcal{P}_p}(X_1, \dots, X_p) = \eta(X_1, \dots, X_p) - \sum_{u \in S^-} \eta_u(X_u)$$

By projection,  $\langle \eta - \sum_{v \in S^-} \eta_v, h_u \rangle = 0 \quad \forall u \in S^-, \forall h_u \in H_u^0$ . Hence,  $\eta(\mathbf{X}) = \sum_{u \in S} \eta_u(X_u)$ ,  $\eta_u \in H_u^0$ ,  $\forall u \in S$ , and this decomposition is well defined.

- *Terms of the summand are unique: assume that  $\eta = \sum_{u \in S} \eta_u = \sum_{u \in S} \tilde{\eta}_u$ ,  $\tilde{\eta}_u \in H_u^0$ . By Lemma 1, it follows that*

$$\left. \begin{aligned} \sum_{u \in S} (\eta_u - \tilde{\eta}_u) &= 0 \\ \|\sum_{u \in S} (\eta_u - \tilde{\eta}_u)\|^2 &\geq \delta^{\#(S)-1} \sum_{u \in S} \|\eta_u - \tilde{\eta}_u\|^2 \end{aligned} \right\} \Rightarrow \|\eta_u - \tilde{\eta}_u\|^2 = 0 \quad \forall u \in S$$

■

### Proof of Proposition 1

Let first prove the following equivalence :

$$\begin{aligned}
\int \eta_u(x_u) \eta_v(x_v) p_X(x) d\nu(x) &= 0 \quad \forall v \subset u, \forall \eta_v \\
&\Updownarrow \\
\int \eta_u(x_u) p_X(x) d\nu_i(x_i) d\nu_{u^c}(x_{u^c}) &= 0 \quad \forall u \in S, \forall i \in u
\end{aligned}$$

Let  $v \subset u$  and  $i \in u \setminus v$ , then

$$\begin{aligned}
\int \eta_v(x_v) \eta_u(x_u) p_X(x) d\nu(x) &= \int \eta_v(x_v) \eta_u(x_u) p_X(x) d\nu_i(x_i) d\nu_{u^c}(x_{u^c}) d\nu_{u \setminus i}(x_{u \setminus i}) \\
&= \int \eta_v(x_v) \left( \int \eta_u(x_u) p_X(x) d\nu_i(x_i) d\nu_{u^c}(x_{u^c}) \right) d\nu_{u \setminus i}(x_{u \setminus i}) \\
&= 0 \quad \text{by assumption}
\end{aligned}$$

Conversely, assume that  $\exists u, \exists i \in u$  such that  $\int \eta_u(x_u) p_X(x) d\nu_i(x_i) d\nu_{u^c}(x_{u^c}) \neq 0$ , then,

$$\eta_v = \int \eta_u(x_u) p_X(x) d\nu_i(x_i) d\nu_{u^c}(x_{u^c}) \quad \text{with } v = u \setminus i$$

and

$$\begin{aligned}
\int \eta_u(x_u) \eta_v(x_v) p_X(x) d\nu(x) &= \int \eta_u(x_u) \left( \int \eta_u(x_u) p_X(x) d\nu_i \, d\nu_{u^c} \right) p_X(x) d\nu_i \, d\nu_{u^c} \, d\nu_{u \setminus i} \\
&= \int \left( \int \eta_u(x_u) p_X(x) d\nu_i(x_i) \, d\nu_{u^c}(x_{u^c}) \right)^2 d\nu_{u \setminus i}(x_{u \setminus i}) \\
&> 0
\end{aligned}$$

There is a contradiction, so that  $\int \eta_u(x_u) p_X(x) d\nu_i(x_i) \, d\nu_{u^c}(x_{u^c}) = 0 \quad \forall i \in u, \forall u$ .  
The second expression can be rewritten as :

$$\int \eta_u(x_u) p_X(x) d\nu_i(x_i) \, d\nu_{u^c}(x_{u^c}) = \mathbb{E}(\eta_u / X_{u \setminus i}) \quad \forall i \in u, \forall u \in S$$

Then, by Theorem 1, the minimization problem  $(\mathcal{P})$  admits a unique solution. ■

## A.2 Generalized sensitivity indices

*Proof of Proposition 2*

Under (C.1) and (C.2), Theorem 1 states the existence and the uniqueness decomposition of  $\eta$ :

$$\eta(\mathbf{X}) = \sum_{u \in S} \eta_u(X_u),$$

with  $H_u^0 \perp H_v^0, \forall v \subset u$ . Therefore,

$$\mathbb{E}(\eta(X)) = \mathbb{E}\left(\sum_{u \in S} \eta_u(X_u)\right) = \eta_0$$

and

$$\begin{aligned}
V(Y) = V(\eta(X)) &= \mathbb{E}(\eta^2(X)) - \eta_0^2 \\
&= \sum_{u \neq \emptyset} \mathbb{E}(\eta_u^2(X_u)) + \sum_{u \neq v} \mathbb{E}(\eta_u(X_u) \eta_v(X_v)) \\
&= \sum_{u \neq \emptyset} V(\eta_u(X_u)) + \sum_{\substack{u \neq \emptyset \\ u \neq \mathcal{P}_p}} \sum_{\substack{v \neq \emptyset \\ u \not\supset v, v \not\subset u}} \mathbb{E}(\eta_u(X_u) \eta_v(X_v)) \\
&= \sum_{u \neq \emptyset} \left[ V(\eta_u(X_u)) + \sum_{\substack{v \\ u \cap v \neq u, v}} \text{Cov}(\eta_u(X_u), \eta_v(X_v)) \right]
\end{aligned}$$

Thus, (6) holds, and equalities (7) and (8) follow obviously. ■

## B Examples of distribution function

### B.1 Boundedness of the inputs density function

*Proof of Proposition 3*

Let  $u \subset \mathcal{P}_p$ , and  $0 < M_1 \leq p_X \leq M_2$ . As  $p_{X_u}$  and  $p_{X_{u^c}}$  are marginals, they are upper bounded by  $M_2$ .

As a consequence,  $p_{X_u} p_{X_{u^c}} \leq \frac{M_2^2}{M_1} M_1 \leq \frac{M_2^2}{M_1} p_X$ , so that  $p_X \geq M_1 M_2^{-2} p_{X_u} p_{X_{u^c}}$ , with  $0 < M_1 M_2^{-2} < 1$ . ■

*Proof of Example 1*



- $\nu$  is a product of measure as  $\frac{d\nu}{d\nu_L} = \prod_{i=1}^p \nu_i(x_i)$ , with  $\nu_i \sim N(m_i, \sigma_i^2)$ . So

$$\nu = \bigotimes_{i=1}^p \nu_i$$

- $p_X$  is given by

$$\begin{aligned} p_X(\mathbf{x}) &= \frac{dP_X}{d\nu}(\mathbf{x}) = \frac{dP_X}{d\nu_L} \times \frac{d\nu_L}{d\nu}(\mathbf{x}) \\ &= \alpha + (1 - \alpha) \left| \frac{\Sigma}{\Omega} \right|^{1/2} \exp -\frac{1}{2} {}^t(\mathbf{x} - m)(\Omega^{-1} - \Sigma^{-1})(\mathbf{x} - m) \end{aligned} \quad (33)$$

First, we have  $p_X(\mathbf{x}) \geq \alpha > 0$ .

Further, the sufficient and necessary condition to have  $p_X \leq M_2 < \infty$  is to get  $(\Omega^{-1} - \Sigma^{-1})$  positive definite. Indeed, if  $(\Omega^{-1} - \Sigma^{-1})$  admits a negative eigenvalue,  $p_X$  can not be bounded. Thus,  $0 < \alpha \leq p_X \leq M_2$  iff  $(\Omega^{-1} - \Sigma^{-1})$  is positive definite.

■

## B.2 Examples of distribution of two inputs

*Proof of Proposition 4*

Condition (C.5) is immediate with Equation 9. Let prove that (C.5) is equivalent to (C.6).

If (C.6) holds, then  $c(u, v) \geq M$ . Conversely, we assume that  $0 < M < 1$ , and

$$\tilde{C}(u, v) = \frac{C(u, v) - Muv}{1 - M}$$

It is enough to show it is a copula : Obviously,  $\tilde{C}(0, u) = C(u, 0) = 0$  and  $\tilde{C}(1, u) = \tilde{C}(u, 1) = u \quad \forall u \in [0, 1]$ . By second order derivation, it comes that  $\tilde{c}(u, v) = \frac{c(u, v) - M}{1 - M}$ , so  $\tilde{c}(u, v) \geq 0$  by hypothesis (C.5).

■

## C Estimation

### C.1 Model of $p = 2$ input variables

*Proof of Proposition 7*

- We first show first that  $(\mathcal{S})$  admits an unique solution.

Under (C.1) and (C.2), by Theorem 1, the decomposition of  $\eta(\mathbf{X})$  is unique and

$$\eta(X_1, X_2) = \eta_0 + \eta_1(X_1) + \eta_2(X_2) + \eta_{12}(X_1, X_2)$$

with

$$\begin{cases} \eta_0 \in H_\emptyset \\ \eta_i \in H_i^0 \perp H_\emptyset, & i = 1, 2 \\ \eta_{12} \in H_{12}^0 \perp H_i^0, & i = 1, 2, H_{12}^0 \perp H_\emptyset \end{cases}$$

Thus,

$$\begin{pmatrix} Id & 0 & 0 & 0 \\ 0 & Id & P_{H_1^0} & 0 \\ 0 & P_{H_2^0} & Id & 0 \\ 0 & 0 & 0 & Id \end{pmatrix} \begin{pmatrix} \eta_0 \\ \eta_1 \\ \eta_2 \\ \eta_{12} \end{pmatrix} = \begin{pmatrix} P_{H_0}(\eta) \\ P_{H_1^0}(\eta) \\ P_{H_2^0}(\eta) \\ P_{H_{12}^0}(\eta) \end{pmatrix} \quad (34)$$

So  $(\eta_0, \eta_1, \eta_2, \eta_{12})$  is solution of  $(\mathcal{S})$ .

Now, assume there exists an another solution of the system, say  $(\tilde{\eta}_0, \dots, \tilde{\eta}_{\mathcal{P}_p}) \in H_0 \times \dots \times H_{\mathcal{P}_p}^0$ , then

$$\begin{aligned} \begin{cases} \eta_0 - \tilde{\eta}_0 = 0 \\ \eta_1 - \tilde{\eta}_1 + P_{H_1^0}(\eta_2 - \tilde{\eta}_2) = 0 \\ P_{H_2^0}(\eta_1 - \tilde{\eta}_1) + \eta_2 - \tilde{\eta}_2 = 0 \\ \eta_{12} - \tilde{\eta}_{12} = 0 \end{cases} &\Rightarrow \begin{cases} \eta_0 = \tilde{\eta}_0 \\ P_{H_1^0}(\eta_1 - \tilde{\eta}_1 + \eta_2 - \tilde{\eta}_2) = 0 \\ P_{H_2^0}(\eta_1 - \tilde{\eta}_1 + \eta_2 - \tilde{\eta}_2) = 0 \\ \eta_{12} = \tilde{\eta}_{12} \end{cases} \\ &\Rightarrow \begin{cases} \eta_0 = \tilde{\eta}_0 \\ \eta_1 - \tilde{\eta}_1 + \eta_2 - \tilde{\eta}_2 \in H_1^{0\perp} \cap H_2^{0\perp} \\ \eta_{12} = \tilde{\eta}_{12} \end{cases} \end{aligned}$$

As  $\eta_1 - \tilde{\eta}_1 \in H_1^0$  and  $\eta_2 - \tilde{\eta}_2 \in H_2^0$ , it follows that

$$\begin{aligned} \begin{cases} \langle \eta_1 - \tilde{\eta}_1, \eta_1 - \tilde{\eta}_1 + \eta_2 - \tilde{\eta}_2 \rangle = 0 \\ \langle \eta_2 - \tilde{\eta}_2, \eta_1 - \tilde{\eta}_1 + \eta_2 - \tilde{\eta}_2 \rangle = 0 \end{cases} &\Rightarrow \begin{cases} \|\eta_1 - \tilde{\eta}_1\|^2 + \langle \eta_1 - \tilde{\eta}_1, \eta_2 - \tilde{\eta}_2 \rangle = 0 \\ \|\eta_2 - \tilde{\eta}_2\|^2 + \langle \eta_1 - \tilde{\eta}_1, \eta_2 - \tilde{\eta}_2 \rangle = 0 \end{cases} \\ &\Rightarrow \|\eta_1 - \tilde{\eta}_1 + \eta_2 - \tilde{\eta}_2\|^2 = 0 \\ &\Rightarrow \eta_1 - \tilde{\eta}_1 + \eta_2 - \tilde{\eta}_2 = 0 \end{aligned}$$

As 0 can be uniquely decomposed into  $H^0$  as a sum of zero, then,

$$\eta_1 - \tilde{\eta}_1 = \eta_2 - \tilde{\eta}_2 = 0$$

- Let now compute

$$\begin{pmatrix} P_{H_0}(\eta) \\ P_{H_1^0}(\eta) \\ P_{H_2^0}(\eta) \\ P_{H_{12}^0}(\eta) \end{pmatrix}$$

First of all, it is obvious that the constant term  $\eta_0 = \mathbb{E}(\eta)$  and that  $\eta_{12}$  is obtained by subtracting  $\eta$  with all other terms of the right of the decomposition.

Now, let us use the projector's property of embedded spaces. Indeed, as  $H_i^0 \subset H_i$ ,  $\forall i = 1, 2$ , it comes

$$P_{H_i^0}(\eta) = P_{H_i^0}(P_{H_i}(\eta)) = P_{H_i^0}(\underbrace{\mathbb{E}(\eta/X_i)}_{\varphi(X_i)})$$

$\varphi$  is a function of  $X_i$ , so it can be decomposed into the following expression :

$$\varphi(X_i) = \varphi_0 + \varphi_i(X_i), \quad \varphi_0 \in H_0, \quad \varphi_i \in H_i^0$$

with  $\varphi_0 = \mathbb{E}(\varphi) = \mathbb{E}(\eta)$ .

Hence, one can easily deduce  $P_{H_i^0}(\eta)$ ,  $i = 1, 2$ , as the term  $\varphi_i = \mathbb{E}(\eta/X_i) - \mathbb{E}(\eta)$

We obtain

$$\begin{pmatrix} P_{H_0}(\eta) \\ P_{H_1^0}(\eta) \\ P_{H_2^0}(\eta) \\ P_{H_{12}^0}(\eta) \end{pmatrix} = \begin{pmatrix} \mathbb{E}(\eta) \\ \mathbb{E}(\eta/X_1) - \mathbb{E}(\eta) \\ \mathbb{E}(\eta/X_2) - \mathbb{E}(\eta) \\ \eta - \mathbb{E}(\eta/X_1) - \mathbb{E}(\eta/X_2) + \mathbb{E}(\eta) \end{pmatrix} \quad (35)$$

■

## Numerical procedure

Gauss-Seidel algorithm requires the estimation of conditional expectation at each iteration. To do this, we use the local polynomial estimation with a leave-one-out technique. To considerably reduce time cost in application, the Sherman-Morrison formula is exploited.

We are going to review these methods for estimating  $\mathbb{E}(Y/X = x)$ , when  $Y$  and  $X$  are supposed to be real random variable.

The local polynomial estimation [29] consists in approximating  $m(x) = \mathbb{E}(Y/X = x)$  by a  $q^{th}$ -order polynomial fitted by a weight least squared estimation.

An explicit solution of  $\hat{m}(x)$  is given by :

$$\hat{m}(x) = {}^t(1 \ 0 \cdots 0)[{}^t\mathbb{X}(x)D(x)\mathbb{X}(x)]^{-1} \cdot {}^t\mathbb{X}(x)D(x)\mathbb{Y} \quad (36)$$

with

$$\mathbb{X}(x) = \begin{pmatrix} 1 & X_1 - x & \cdots & (X_1 - x)^q \\ \vdots & \vdots & & \vdots \\ 1 & X_n - x & \cdots & (X_n - x)^q \end{pmatrix} \quad D(x) = \begin{pmatrix} K\left(\frac{X_1 - x}{h}\right) & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & K\left(\frac{X_n - x}{h}\right) \end{pmatrix}$$

$$\mathbb{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$$

The leave-one-out technique on local estimation consists in estimating  $m$  in every observation point  $X_1, \dots, X_n$ , i.e. computing Equation (36) when the  $k^{th}$  line of matrices has been removed for estimating  $\hat{m}(X_k)$ . It means that we would need to inverse  ${}^t\mathbb{X}_{-k}(x)D_{-k}(x)\mathbb{X}_{-k}(x)$   $n$  times, which is very expensive. To avoid these expensive computations, Sherman and Morrison [30] proposed a formula :

**Lemma 2.** If  $A$  is a square invertible matrix, and  $u, v$  are vectors such that  $1 + {}^tvA^{-1}u \neq 0$ , then

$$(A + u{}^tv)^{-1} = A^{-1} - \frac{A^{-1}u{}^tvA^{-1}}{1 + {}^tvA^{-1}u} \quad (37)$$

In our problem, set  $S_n(x) = {}^t\mathbb{X}(x)D(x)\mathbb{X}(x)$ .  $S_n(x)$  can be rewritten as :

$$S_n(x) = \sum_{i=1}^n \Phi_i(x){}^t\Phi_i(x) = \underbrace{\sum_{i \neq k} \Phi_i(x){}^t\Phi_i(x)}_{S_{-k}(x)} + \Phi_k(x){}^t\Phi_k(x) \quad (38)$$

where

$$\Phi_i(x) = {}^t(K(\frac{X_i - x}{h})\mathbb{X}_i(x)), \quad \mathbb{X}_i(x) = (1 \cdots (X_i - x)^q), \quad \forall i = 1, \dots, n$$

Thus,  $S_{-k}(x)$ , corresponding to  ${}^t\mathbb{X}(x)D(x)\mathbb{X}(x)$  when the  $k^{th}$  line has been removed, is of the form :

$$S_{-k}(x) = S_n(x) - \Phi_k(x){}^t\Phi_k(x), \quad \forall k = 1, \dots, n$$

The Sherman-Morrison formula gives :

$$S_{-k}^{-1}(x) = S_n^{-1}(x) + \frac{S_n^{-1}(x)\Phi_k(x)^t\Phi_k(x)S_n^{-1}(x)}{1 - {}^t\Phi_k(x)S_n^{-1}(x)\Phi_k(x)}, \quad \forall k = 1, \dots, n \quad (39)$$

As  $\hat{m}(X_k) = {}^t(1 \ 0 \cdots 0)S_{-k}^{-1}(X_k) \cdot {}^t\mathbb{X}_{-k}(X_k)D_{-k}(X_k)\mathbb{Y}$ ,  $\forall k$ , it is faster to estimate  $S_n^{-1}(x)$  and  $\Phi_k(x)$  at each point of the design.

## Acknowledgement

This work has been supported by French National Research Agency (ANR) through COSINUS program (project COSTA-BRAVA number ANR-09-COSI-015).

## References

- [1] W. Hoeffding : *A class of statistics with asymptotically normal distribution*, The annals of mathematical statistics, 1948.
- [2] A.B. Owen : *Lattice sampling revisited: Monte Carlo variance of means over randomized orthogonal arrays*, The Annals of Statistics, 1994.
- [3] I.M. Sobol : *Sensitivity estimates for nonlinear mathematical models*, Wiley Ed., vol.1, no.4, 1993.
- [4] I.M. Sobol : *Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates*, Mathematics and Computers in Simulations, Vol.55, 271-280, 2001.
- [5] A.B. Owen : *Multidimensional variation for quasi-Monte Carlo*, 2005.
- [6] R.I. Cukier, J.H. Schaibly, K.E. Shuler : *Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. III. Analysis of the approximations*, The Journal of Chemical Physics, 1975.
- [7] R.I. Cukier, H.B. Levine, K.E. Shuler : *Nonlinear sensitivity analysis of multiparameter model systems*, Journal of computational physics, 1978.
- [8] S. Tarantola, D. Gatelli, T. Mara : *Random Balance Designs for the Estimation of First Order Global Sensitivity Indices*, Reliability Engineering and System Safety, 2006.
- [9] J.Y. Tissot, C. Prieur : *A bias correction method for the estimation of sensitivity indices based on random balance designs*, 2011, <http://hal.archives-ouvertes.fr/hal-00507526/fr/>
- [10] S. Da Veiga : *Analyse d'incertitudes et de sensibilité Application aux modèles de cinétique chimique*, Ph.D.thesis, p.133-134, <http://www.gdr-mascotnum.fr/media/thesedaveiga.pdf>, 2007.
- [11] J. Jacques, C. Lavergne, N; Devictor : *Sensitivity analysis in presence of model uncertainty and correlated inputs*, Reliability Engineering and System Safety, 2006.
- [12] T. Mara, S. Tarantola : *variance-based sensitivity analysis of computer models with dependent inputs*, Reliability Engineering and System Safety, 2011.
- [13] C. Xu, G. Gertner : *Uncertainty and sensitivity analysis for models with correlated parameters*, Reliability Engineering and System Safety, 2007.
- [14] G. Li, H. Rabitz : *Global sensitivity analysis with independent and/or correlated inputs*, Journal of Physical Chemistry A, 2010.
- [15] C. Xu, G. Gertner : *Extending a global sensitivity analysis technique to models with correlated parameters*, Computational Statistics and Data Analysis, 2007.
- [16] E. Borgonovo : *A new uncertainty importance measure*, Reliability Engineering and System Safety, 2006.
- [17] E. Borgonovo, W. Castaings, S. Tarantola : *Moment Independent Importance Measures: New Results and Analytical Test Cases*, Risk Analysis, 2011.
- [18] G. Hooker : *Generalized functional ANOVA diagnostics for high-dimensional functions of dependent variables*, Journal of Computational and Graphical Statistics, Vol.16, No.3, 2007.
- [19] C.J. Stone : *The use of Polynomial Splines and their tensor products in multivariate function estimation*, The Annals of Statistics, Vol.22, No.1, 1994.

- [20] A.W. Van Der Vaart : *Asymptotic Statistics*, Cambridge University Press.
- [21] R.B. Nelsen : *An introduction to copulas*, Springer, 2006.
- [22] A. Sklar : *Random Variables, Joint Distribution Functions, and Copulas*, Kybernetika, Vol. 9, 1971.
- [23] D. Morgenstern : *Einfache Beispiele Zweidimensionaler Verteilungen* Mitteilungsblatt für Mathematische Statistik, 8, 234235, 1956.
- [24] Z. Ouyang, H. Liao X. Yang : *Modeling dependence based on mixture copulas and its application in risk management*, Appl. Math. J. Chinese Univ., 2009.
- [25] R. Kress : *Numerical analysis*, Springer Ed., 1998.
- [26] J. De Pillis : *Gauss-Seidel convergence for operators on Hilbert space* , SIAM Journal of Numerical Analysis, 1973.
- [27] J.L. Jensen : *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*, Acta Math., 1906.
- [28] S. Da Veiga, F. Wahl, F. Gamboa : *Local Polynomial Estimation for Sensitivity Analysis on Models With Correlated Inputs*, 2008.
- [29] J. Fan, I. Gijbels : *Local polynomial modelling and its application*, Chapman & Hall Ed., 1996.
- [30] J. Sherman, W.J. Morrison : *Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix*, The Annals of Mathematical Statistics, vol.21, No.1, 1950.
- [31] A. Saltelli, K.Chan, E.M.Scott : *Sensitivity Analysis*, Wiley Ed., 2000.
- [32] D.G. Luenberger : *Optimization by vector space methods*, Wiley Ed.