Derivative-based generalized sensitivity indices and Sobol' indices

Matieyendou Lamboni

Derivative-based generalized sensitivity indices and Sobol' indices

Matieyendou Lamboni[a,b]

[a] *University of Guyane, Department DFRST, 97346 Cayenne, French Guiana, France*
[b] *228-UMR Espace-Dev, University of Guyane, University of Réunion, IRD, University of Montpellier.*

Abstract

In uncertainty quantification, multivariate sensitivity analysis (MSA), including variance-based sensitivity analysis, and derivative global sensitivity measure (DGSM) are widely used for assessing the effects of input factors on the model outputs. While MSA allows for identifying the order and the strength of interactions among inputs, DGSM provides only a global effect of inputs by making use of model derivatives. It is interesting to combine the advantages of both approaches and to come up with generalized sensitivity indices (GSIs) from MSA based on model derivatives. First, we derive the mathematical expressions of the total effect and total-interaction effect functionals based on derivatives. Second, we construct minimum variance unbiased estimators (MVUEs) of the total-effect and total-interaction effect covariance matrices, and third, we provide the estimators of the total and total-interaction GSIs as well as their consistency and asymptotic normality. Finally, we demonstrate the applicability of these new results by means of simulations.

Keywords: Derivatives, Generalized sensitivity indices, Matrix norms, MVU estimators, U-statistics.

1. Introduction

Complex mathematical models (either multivariate or single response) are widely used as experimental tools for supporting decision making in natural or human-induced phenomena. They often include numerous uncertain input factors, and it is interesting to assess the effects of input factors on the whole model outputs prior to scenario-building, model-reducing, or model diagnostic activities.

Multivariate sensitivity analysis (MSA) ([1; 2; 3; 4; 5]), including variance-based sensitivity analysis (VbSA) ([6; 7; 8; 9]), is the standard way of assessing the impor-

tance of input factors on the model output(s) as well as interactions among input factors by making use of the model runs. The estimations of generalized sensitivity indices (GSIs) from MSA, including Sobol' indices, have been largely investigated ([10; 1; 11; 2; 4; 12; 13; 14; 15; 16; 17; 18; 19; 20; 21; 22]). Among the sample-based methods, the estimators of the first-order and total indices proposed in [1; 10; 12] allow for improving the estimates of sensitivity indices, as they are based on minimum variance unbiased estimators.

For response models with available gradients, derivative global sensitivity measure (DGSM) ([23; 24; 25]) is an appropriate way of assessing a global impact of input factors, as it is computationally more attractive than VbSA or MSA (sample-based methods). Given that the DGSM index and the total index (from VbSA) can provide different ranking of input factors, upper bounds and lower bounds of the total index and the total-interaction index based on derivatives were proposed ([26; 27; 28]). An upper bound of the total index (resp. total-interaction index), which is a (known) constant times the DGSM index (resp. cross-derivatives index), is used for the screening purpose. Indeed, while a small value of the upper bound of the Sobol total index means that the associated input does not really act in the model; a big value of the upper bound does not bring much information, regarding factors classification, and big values of upper bounds can happen especially in the case of the total-interaction index ([16; 27]).

In this paper, we combine the advantages of DGSM and MSA approaches to come up with a new way of computing the GSIs by making use of model derivatives. First, we derive the mathematical expressions of the total effect and total-interaction effect functionals using the model derivatives, cumulative distribution functions (CDFs), and probability density functions (PDFs). Second, we construct minimum variance unbiased estimators (MVUEs) of the total-effect and total-interaction effect covari-

ance matrices, and third, we provide estimators of the total and total-interaction GSIs as well as their consistency and asymptotic normality. Finally, we demonstrate the applicability of these new results by means of numerical tests.

The paper is organized as follows: in Section 2, we recall two definitions of GSIs using the sensitivity functionals and the Frobenius norm. The sensitivity functionals are the functionals that will lead to access the first-order, total and total-interaction GSIs. Section 3 provides the derivative-based GSIs by deriving the mathematical expressions of the sensitivity functionals using the model derivatives, CDFs, and PDFs. The derivation of the derivative-based total-interaction GSI motivates the proposition of a new Hoeffding decomposition in Section 2. While in Section 4, we propose and study estimators of the GSIs based on model derivatives, Section 5 deals with the computational issues. It provides an algorithm for computing the GSIs values and illustrates our approach on test cases. We conclude this work in Section 6.

Notation

This section defines the symbols that will be used throughout the paper. For integer $d \in \mathbb{N}^*$ and $j = 1, \ldots, d$, we use $\mu_j(x_j) = \rho_j(x_j)\, dx_j$ for an absolutely continuous probability measure on an open interval $\Omega_j \subseteq \mathbb{R}$ w.r.t. the Lebesgue measure, X_j for a random variable or factor from μ_j, x_j for a sample value of X_j, and $\mathbf{X} = \{X_j,\, j \in \{1, \ldots, d\}\}$ for a random vector. We use ρ_j (resp. F_j) for the positive and continuous probability density function (PDF) (resp. the cumulative distribution function: CDF) with $j = 1, \ldots, d$. We use $\mu(\mathbf{x})$, Ω, $\rho(\mathbf{x})$, $F(\mathbf{x})$ for the joint probability measure, the joint support, the joint PDF and the joint CDF of $\mathbf{X}$, respectively, and $\mathbb{E}$ and $\mathbb{V}$ for the expectation and variance taking w.r.t. $\mathbf{X}$. We use $\xrightarrow{D}$ and $\xrightarrow{P}$ for the convergence in distribution and in probability, respectively.

For integer $n \in \mathbb{N}^*$, the function $\mathbf{f} : \mathbb{R}^d \to \mathbb{R}^n$ denotes a deterministic function

that includes d input factors $\mathbf{X}$. We use $u \subseteq \{1, 2, \ldots, d\}$ for a non-empty subset of $\{1, 2, \ldots, d\}$, $\bar{u} = \{1, 2, \ldots, d\} \backslash u$ for the complement of u w.r.t. $\{1,\ 2,\ \ldots,\ \mathrm{d}\ \}$, and $|u|$ for its cardinality (i.e., the number of elements in u). For a given u, we use $\mathbf{X}_u = \{X_j,\, j \in u\}$ for a subset of input factors and $\mathbf{X}_{\sim u} = \{X_j,\, j \in \bar{u}\}$ for the vector containing all input factors, except $\mathbf{X}_u$. We have the following partition: $\mathbf{X} = (\mathbf{X}_u, \mathbf{X}_{\sim u})$.

For an $n \times n$ square matrix $\Sigma = (\sigma_{ij},\, i,\, j \in \{1,\, \ldots,\, n\})$, the trace (Tr), the Frobenius norm ($||\Sigma||_F$), and the vectorization (Vec(Σ)) of Σ are defined as follows:

$$\mathrm{Tr}(\Sigma) = \sum_{i=1}^{n} \sigma_{ii}\,,$$

$$||\Sigma||_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} |\sigma_{ij}|^2 = \mathrm{Tr}\left(\Sigma\Sigma^T\right)\,,$$

$$\mathrm{Vec}(\Sigma) = \left[\sigma_{i1}\,\forall\, i \in \{1,\, \ldots,\, n\};\, \sigma_{i2}\,\forall\, i \in \{1,\, \ldots,\, n\};\, \ldots;\, \sigma_{in}\,\forall\, i \in \{1,\, \ldots,\, n\}\right]^T\,.$$

By definition, the vectorization of Σ is a vector containing the first column of Σ, followed by the second column, and so on.

The function $\mathbf{f}(\cdot)$ may be subjected to the transformation of the form:

$$D_{\mathbf{w}}\mathbf{f}(\mathbf{X})\,,$$

where $D_{\mathbf{w}}$ is a weighting matrix. In the case of the model outputs with different units, some transformations may be used to obtain unit-less outputs. A classical way to accomplish this is to divide each output by its standard deviation, that is, $D_{\mathbf{w}}^{-1} = diag(\mathbf{w})$ is a diagonal matrix with $\mathbf{w}$ representing the vector of the standard deviations of the outputs. In the following text, we use $\mathbf{f}(\cdot)$ as either the original function or a given transformation of the latter.

In what follows, we consider only independent input factors (assumption A1) and measurable and differentiable functions $\mathbf{f} : \mathbb{R}^d \to \mathbb{R}^n$ having finite second moments, that is, $\mathbb{E}\left[||\mathbf{f}(\mathbf{X})||_{L^2}^2\right] < +\infty$.

2. Multivariate sensitivity analysis: two types of generalized sensitivity indices

This section gives two definitions of generalized sensitivity indices, including Sobol' indices, using the sensitivity functionals. We propose i) a new Hoeffding decomposition; ii) a link between the sensitivity functionals; and iii) the two main definitions of GSIs.

2.1. Hoeffding decomposition

Under the independence assumption A1, the multivariate Hoeffding decomposition ([29]; [30]) is given by

$$\begin{aligned} \mathbf{f}(\mathbf{X}) &= \mathbf{f}_\emptyset + \sum_{j=1}^{d} \mathbf{f}_j(X_j) + \sum_{j_1<j_2}^{d} \mathbf{f}_{j_1 j_2}(X_{j_1}, X_{j_2}) + \ldots + \mathbf{f}_{1\ldots d}(X_1, \ldots, X_d) \\ &= \mathbf{f}_\emptyset + \sum_{\substack{w,\, w\subseteq\{1,2,\ldots d\} \\ |w|>0}} \mathbf{f}_w(\mathbf{X}_w)\,, \end{aligned} \tag{2.1}$$

where $\mathbf{f}_\emptyset = \mathbb{E}\left[\mathbf{f}(\mathbf{X})\right]$ is the expectation of the model output, $\mathbf{f}_j(X_j) = \mathbb{E}\left[\mathbf{f}(\mathbf{X})|X_j\right] - \mathbf{f}_\emptyset$, and $\mathbf{f}_w(\mathbf{X}_w) = \mathbb{E}\left[\mathbf{f}(\mathbf{X})|\mathbf{X}_w\right] - \sum_{v,\, v\subset w} \mathbf{f}_v(\mathbf{X}_v)$ for a non-empty subset $w \subseteq \{1, 2, \ldots, d\}$.

The functional $\mathbb{E}\left[\mathbf{f}(\mathbf{X})|\mathbf{X}_u\right] - \mathbf{f}_\emptyset$ generalizes $\mathbf{f}_j(X_j) = \mathbb{E}\left[\mathbf{f}(\mathbf{X})|X_j\right] - \mathbf{f}_\emptyset$ from (2.1) to cope with any subset $\mathbf{X}_u$ of input factors. It allows for quantifying the single contribution of the input $\mathbf{X}_u$. We refer to the latter as the first-order functional, that is, $\mathbf{f}_u^{fo}(\mathbf{X}_u) = \mathbb{E}\left[\mathbf{f}(\mathbf{X})|\mathbf{X}_u\right] - \mathbf{f}_\emptyset$.

2.2. New Hoeffding decomposition: link between the total effect and total-interaction effect functionals

The total-effect and total-interaction effect functionals can be easily derived by re-organizing the Hoeffding decomposition. The new decomposition aims at managing the Hoeffding decomposition with only $2^{|u|}$ components, with $u \subseteq \{1, 2, \ldots, d\}$ and $|u|$ its cardinality. The idea consists in expanding the function $\mathbf{f}(\cdot)$ as a sum of collections of functions defined in Equation (2.1). For a given subset u, a collection (of functions) relies on one component of a partition of the set $\{w, w \subseteq \{1, 2, \ldots, d\}\}$ into $2^{|u|}$ subsets.

Definition 1. *Let u be a non-empty subset of $\{1, 2, \ldots, d\}$, and $\bar{u} = \{1, 2, \ldots, d\} \backslash u$ be a set containing all the elements of $\{1, 2, \ldots, d\}$, except those of u.*
For a given v with $v \subseteq u$, we define the set $\mathcal{A}_v$ as a set containing elements of the form $\{v, w\}$ with $w \subseteq \bar{u}$, that is,

$$\mathcal{A}_v = \{\{v, w\}, w \subseteq \bar{u}\} \, . \tag{2.2}$$

By definition, it is obvious that $v \in \mathcal{A}_v$ and $\mathcal{A}_v \subset \{w, w \subseteq \{1, 2, \ldots, d\}\}$. In particular, $\mathcal{A}_\emptyset = \{w, w \subseteq \bar{u}\}$ and $\mathcal{A}_u = \{\{u, w\}, w \subseteq \bar{u}\}$. While $\mathcal{A}_u$ contains all the super-sets of u (i.e., sets that contain u), $\mathcal{A}_v$ contains some super-sets of v but not all of them. Furthermore, $\forall\, A \in \mathcal{A}_v$, A satisfies the following properties:

- $\forall\, j \in v$, then $j \in A$;
- $\forall\, j \in u \backslash v$, then $j \notin A$.

Using the set $\mathcal{A}_v$, with $v \subseteq u$, Lemma 1 gives a partition of $\{w, w \subseteq \{1, 2, \ldots, d\}\}$.

Lemma 1. *Let u be a non-empty subset of $\{1, 2, \ldots, d\}$, v_1, v_2 be two subsets of u (i.e., $v_1 \subseteq u$, $v_2 \subseteq u$).*

(i) If $v_1 \neq v_2$ then we have

$$\mathcal{A}_{v_1} \bigcap \mathcal{A}_{v_2} = \emptyset \, . \tag{2.3}$$

(ii) The partition of the set $\{w,\, w \subseteq \{1,\, 2,\, \ldots,\, d\}\}$ *is given by*

$$\{w,\, w \subseteq \{1,\, 2,\, \ldots,\, d\}\} = \bigcup_{v,\, v \subseteq u} \mathcal{A}_v\,. \tag{2.4}$$

(iii) If we use $\mathcal{B}_{v_1}$ *for the set containing all the super-sets of* v_1*, we have*

$$\mathcal{B}_{v_1} = \bigcup_{v,\, v_1 \subseteq v \subseteq u} \mathcal{A}_v\,. \tag{2.5}$$

Proof. *See AppendixA.*

□

Lemma 1 gives an interesting tool for managing and controlling 2^d elements with only $2^{|u|}$ elements. For instance, when $u = \{j\}$, the set $\{w,\, w \subseteq \{1,\, 2,\, \ldots,\, d\}\}$ is completely controlled by $\mathcal{A}_u$ and $\mathcal{A}_\emptyset$, and this result was obtained in [26; 13]. When $u = \{j_1,\, j_2\}$, $\{w,\, w \subseteq \{1,\, 2,\, \ldots,\, d\}\}$ can be managed with the following four sets $\mathcal{A}_\emptyset$, $\mathcal{A}_{\{j_1\}}$, $\mathcal{A}_{\{j_2\}}$, $\mathcal{A}_u$. The set $\mathcal{A}_u$ with $|u| = 1$ (resp. $|u| = 2$) is particularly interesting in SA, as it can lead to assess the total effect (resp. the total-interaction effect of the second order) of input factor(s) ([12; 10; 1; 27; 16; 17]). Indeed, $\mathcal{A}_{\{j\}}$ is a set containing all the super-sets of $\{j\}$, and it is sufficient to assess the total effect of X_j. In the same sense, while $\mathcal{A}_u$ can lead to assess the total-interaction effect of the $|u|^{\text{th}}$ order, $\mathcal{B}_{v_1}$ allows for quantifying the total-interaction effect of the $|v_1|^{\text{th}}$ order, with $v_1 \subseteq u$. If $|u| = 1$, the total-interaction effect of X_u comes down to the total effect of X_u.

Now, if we define $\mathbf{f}_{\mathcal{A}_v}(\mathbf{X}) = \sum_{w,\, w \in \mathcal{A}_v} \mathbf{f}_w(\mathbf{X}_w)$ with $\mathbf{f}_w(\mathbf{X}_w)$ defined in Equation (2.1), we can give a new decomposition of $\mathbf{f}$ (see Proposition 1).

Proposition 1. *Let* u *be a non-empty subset of* $\{1,\, 2,\, \ldots,\, d\}$*,* $v_1 \subseteq u$ *be a subset of* u*. If assumption A1 (independence of input factors) holds, then*

(i) a new Hoeffding decomposition of $\mathbf{f}$ *is given by*

$$\mathbf{f}(\mathbf{X}) \;=\; \sum_{v,\, v \subseteq u} \mathbf{f}_{\mathcal{A}_v}(\mathbf{X}_v, \mathbf{X}_{\sim v})\,, \tag{2.6}$$

where the components $\mathbf{f}_{\mathcal{A}_v}(\mathbf{x})$, $v \subseteq u$ *are mutually orthogonal.*

(*ii*) *The functional*

$$\begin{aligned} \mathbf{f}^{sup}_{v_1}(\mathbf{X}) &= \mathbf{f}_{\mathcal{B}_{v_1}}(\mathbf{X}) \\ &= \sum_{v,\, v_1 \subseteq v \subseteq u} \mathbf{f}_{\mathcal{A}_v}(\mathbf{X})\,, \end{aligned} \tag{2.7}$$

allows for quantifying the total-interaction effect of $\mathbf{X}_{v_1}$.

(*iii*) *The functional*

$$\mathbf{f}^{tot}_{u}(\mathbf{X}) = \sum_{\substack{v,\, v \subseteq u \\ |v|>0}} \mathbf{f}_{\mathcal{A}_v}(\mathbf{X}_v, \mathbf{X}_{\sim v})\,, \tag{2.8}$$

allows for quantifying the total effect of $\mathbf{X}_u$.

Proof. *See AppendixB.*

□

In what follows, we refer to $\mathbf{f}^{sup}_{u}(\mathbf{X})$ as the total-interaction effect functional (TIEF) of $\mathbf{X}_u$ and $\mathbf{f}^{tot}_{u}$ as the total effect functional (TEF) of $\mathbf{X}_u$. Lemma 2 provides a link between TIEF and TEF.

Lemma 2. *Let* $u \subseteq \{1, 2, \ldots, d\}$ *be a non-empty subset. If assumption A1 holds, then we have*

$$\mathbf{f}^{tot}_{u}(\mathbf{X}_u, \mathbf{X}_{\sim u}) = \sum_{v,\, v \subset u} (-1)^{|u|-|v|+1} \mathbf{f}^{sup}_{u \setminus v}(\mathbf{X})\,, \tag{2.9}$$

and

$$\mathbf{f}^{sup}_{u}(\mathbf{X}) = \sum_{v,\, v \subset u} (-1)^{|u|-|v|+1} \mathbf{f}^{tot}_{u \setminus v}(\mathbf{X})\,. \tag{2.10}$$

Proof. *See AppendixC.*

□

2.3. Definition of generalized sensitivity indices

When we use the variance as a measure of the variability of the model outputs, a definition of the sensitivity indices for the multivariate-response models should be

based on the following covariance matrices.

The first-order covariance matrix of $\mathbf{X}_u$ is given by

$$\Sigma_u = \mathbb{V}\left[\mathbf{f}_u^{fo}(\mathbf{X})\right] . \tag{2.11}$$

Further, the total-effect covariance matrix of $\mathbf{X}_u$ is given by

$$\Sigma_u^{tot} = \mathbb{V}\left[\mathbf{f}_u^{tot}(\mathbf{X})\right] . \tag{2.12}$$

Likewise, the total-interaction covariance matrix of $\mathbf{X}_u$ is given by

$$\Sigma_u^{sup} = \mathbb{V}\left[\mathbf{f}_u^{sup}(\mathbf{X})\right] . \tag{2.13}$$

For the single-response models ($n = 1$), the prioritization of input factors based on the covariance matrices is straightforward, as the covariance matrices are scalars. In the case of the multivariate-response models with $n > 1$, Lamboni [1] proposed to apply matrix norms on the covariance matrices in order to prioritize input factors. In this paper, we consider two types of generalized sensitivity indices from [1].

Definition 2. *Let Σ, Σ_u, Σ_u^{sup}, and Σ_u^{tot} be the covariance matrices of the model outputs, the first-order, the total-interaction effect and the total-effect functionals, respectively.*

The first-type GSIs are defined below ([3; 2; 4]).
The first-order GSI of $\mathbf{X}_u$ is defined as follows:

$$GSI_u^F = \frac{\left|\left|\Sigma_u^{1/2}\right|\right|_F^2}{||\Sigma^{1/2}||_F^2} . \tag{2.14}$$

Further, the total GSI of $\mathbf{X}_u$ is given by

$$GSI_{T_u}^F = \frac{\left|\left|\left(\Sigma_u^{tot}\right)^{1/2}\right|\right|_F^2}{||\Sigma^{1/2}||_F^2} , \tag{2.15}$$

and the total-interaction GSI of $\mathbf{X}_u$ *is given by*

$$GSI^F_{sup,u} = \frac{\left|\left|\left(\Sigma^{sup}_u\right)^{1/2}\right|\right|^2_F}{||\Sigma^{1/2}||^2_F}\,. \tag{2.16}$$

Likewise, the second-type GSIs are defined as follows ([1]):

$$GSI^{l_2}_u = \frac{||\Sigma_u||_F}{n\,||\Sigma^{1/2}||^2_F}\,; \tag{2.17}$$

$$GSI^{l_2}_{T_u} = \frac{||\Sigma^{tot}_u||_F}{n\,||\Sigma^{1/2}||^2_F}\,; \tag{2.18}$$

and

$$GSI^{l_2}_{sup,u} = \frac{||\Sigma^{sup}_u||_F}{n\,||\Sigma^{1/2}||^2_F}\,. \tag{2.19}$$

Remark 1. *The first-type GSIs such as* GSI^F_u, $GSI^F_{T_u}$, $GSI^F_{sup,u}$ *are equivalent to the classical definition, that is,*

$$GSI^F_u = \frac{Tr(\Sigma_u)}{Tr(\Sigma)}, \quad GSI^F_{T_u} = \frac{Tr(\Sigma^{tot}_u)}{Tr(\Sigma)}, \quad GSI^F_{sup,u} = \frac{Tr(\Sigma^{sup}_u)}{Tr(\Sigma)}\,.$$

In the case of single response models ($n = 1$)*, the two types GSIs come down to Sobol' indices. Thus, both types of GSIs extend Sobol' indices to cope with multivariate response models. The* l_2*-based definition of GSIs (second-type GSIs) explicitly includes the off-diagonal elements of the covariance matrices. Therefore, it accounts for the correlations among the component of sensitivity functionals.*

3. Generalized sensitivity and Sobol' indices using model derivatives

Section 2 provides the definitions of generalized sensitivity indices using the main sensitivity functionals. To propose the GSIs and Sobol' indices based on model derivatives, we are going to express the TIEF and TEF as functions of the model derivatives.

In what follows, we assume that the function $\mathbf{f} : \mathbb{R}^d \to \mathbb{R}^n$ is a measurable and differentiable function with respect to each input (assumption A3). Namely, we use $\mathbf{X}$ for d input factors and $\mathbf{x}$ for a value of $\mathbf{X}$; u for a non-empty subset of $\{1, \ldots, d\}$; $(\mathbf{x}_u, \mathbf{x}_{\sim u})$, $(\mathbf{y}_u, \mathbf{x}_{\sim u})$ and $(\mathbf{z}_u, \mathbf{x}_{\sim u})$ for three sample values of $\mathbf{X}$. The usual total differential of $\mathbf{f}$ ($d\mathbf{f}$) with higher-order terms is given as follows ([31]):

$$
\begin{aligned}
d\mathbf{f} &= \sum_{j=1}^{d} \frac{\partial \mathbf{f}}{\partial x_j}(\mathbf{x})\, dx_j + \sum_{1\le i<j\le d}^{d} \frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}_{\{i,j\}}}(\mathbf{x})\, d\mathbf{x}_{\{i,j\}} + \ldots + \frac{\partial^d \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x})\, d\mathbf{x} \\
&= \sum_{\substack{v,\, v\subseteq\{1,\ldots,d\} \\ |v|>0}} \frac{\partial^{|v|}\mathbf{f}}{\partial \mathbf{x}_v}(\mathbf{x})\, d\mathbf{x}_v\,, \qquad (3.20)
\end{aligned}
$$

where $\frac{\partial^{|v|}\mathbf{f}}{\partial \mathbf{x}_v}(\mathbf{x})$ stands for the $|v|^{\text{th}}$ cross-partial derivatives of each component of $\mathbf{f}$ with respect to each x_j, with $j \in v$. By integrating Equation (3.20), we obtain the increment of $\mathbf{f}$ ([31]), that is,

$$
\mathbf{f}(\mathbf{z}_u, \mathbf{x}_{\sim u}) - \mathbf{f}(\mathbf{y}_u, \mathbf{x}_{\sim u}) = \sum_{j\in u} \int_{y_j}^{z_j} \frac{\partial \mathbf{f}}{\partial x_j}(\mathbf{x})\, dx_j + \sum_{\substack{v,\, v\subseteq u \\ |v|>1}} \int_{\mathbf{y}_v}^{\mathbf{z}_v} \frac{\partial^{|v|}\mathbf{f}}{\partial \mathbf{x}_v}(\mathbf{x})\, d\mathbf{x}_v \qquad (3.21)
$$

When we multiply Equation (3.21) by the probability density function $\rho_u(\mathbf{y}_u)$ and integrate it over the joint support Ω_u, we obtain the total-effect functional given by (see Equation (B.1))

$$
\begin{aligned}
\mathbf{f}_u^{tot}(\mathbf{z}_u, \mathbf{x}_{\sim u}) &= \mathbf{f}(\mathbf{z}_u, \mathbf{x}_{\sim u}) - \int_{\Omega_u} \mathbf{f}(\mathbf{y}_u, \mathbf{x}_{\sim u})\, d\mu(\mathbf{y}_u) \\
&= \sum_{j\in u} \int_{\Omega_j} \int_{y_j}^{z_j} \frac{\partial \mathbf{f}}{\partial x_j}(\mathbf{x}) \rho_j(y_j)\, dx_j dy_j \\
&\quad + \sum_{\substack{v,\, v\subseteq u \\ |v|>1}} \int_{\Omega_v} \int_{\mathbf{y}_v}^{\mathbf{z}_v} \frac{\partial^{|v|}\mathbf{f}}{\partial \mathbf{x}_v}(\mathbf{x}) \rho_v(\mathbf{y}_v)\, d\mathbf{x}_v d\mathbf{y}_v\,. \qquad (3.22)
\end{aligned}
$$

Under assumption A1 (independent input factor), $\rho_v(\mathbf{y}_v) = \prod_{j\in v} \rho_j(y_j)$, and the right-hand terms of Equation (3.22) become an iteration of one-dimensional integral. For $j \in u$, consider the measurable function $t_j : \Omega_j \to [0,\, 1]$ given by $t_j(x_j) = \frac{x_j - y_j}{z_j - y_j}$. Using (3.22), a change of variables gives

$$\mathbf{f}_u^{tot}(\mathbf{z}_u, \mathbf{x}_{\sim u}) = \sum_{j \in u} \int_{\Omega_j} \int_0^1 \frac{\partial \mathbf{f}}{\partial x_j} \left(t_j z_j + y_j(1-t_j), \mathbf{x}_{\sim j}\right) \times (z_j - y_j) \times \rho_j(y_j)\, dt_j dy_j + \sum_{\substack{v,\, v \subseteq u \\ |v|>1}} \int_{\Omega_v} \int_{[0,1]^{|v|}} \frac{\partial^{|v|}\mathbf{f}}{\partial \mathbf{x}_v} \left([t_j z_j + y_j(1-t_j)\, \forall\, j \in v],\, \mathbf{x}_{\sim v}\right) \times \prod_{j \in v} (z_j - y_j) \times \rho_j(y_j)\, d\mathbf{t}_v d\mathbf{y}_v \,. \tag{3.23}$$

Now, we can express the TIEF and TEF as functions of the model derivatives, the cumulative distribution functions (CDF), and the probability density functions (PDF). Theorem 1 provides these results.

Theorem 1. *Let* $\mathbf{X}$ *be* d *input factors,* $\mathbf{X}^{(1)}$ *be an i.i.d copy of* $\mathbf{X}$*,* F_j *(resp.* ρ_j*) be the CDF (resp. PDF) of* X_j *with* $j = 1, \ldots, d$*, and* u *be a non-empty subset of* $\{1, 2, \ldots, d\}$*. If assumptions*

A1: the input factors $\mathbf{X}$ *are independent,*

A2: the model outputs $\mathbf{f}(\mathbf{X})$ *have finite second moments,*

A3: each component of $\mathbf{f}$ *is measurable and differentiable with respect to* $\mathbf{x}_u$*,*

A4: for all $v \subseteq u$*,* $\frac{\partial^{|v|}\mathbf{f}}{\partial \mathbf{x}_v}(\mathbf{X})$ *are measurable and have finite second moments,*

A5: the PDF ρ_j *is continuous on the support* Ω_j *and* $\rho_j(x_j) \in]0, +\infty[\, \forall\, x_j \in \Omega_j$

hold, then

(i) we have the following expressions of the TEF and TIEF, respectively.

$$\mathbf{f}_u^{tot}(\mathbf{X}) = \sum_{\substack{v,\, v \subseteq u \\ |v|>0}} \mathbb{E}_{\mathbf{X}_u^{(1)}} \left[\frac{\partial^{|v|}\mathbf{f}}{\partial \mathbf{x}_v} \left(\mathbf{X}_v^{(1)}, \mathbf{X}_{\sim v}\right) \prod_{j \in v} \frac{F_j(X_j^{(1)}) \left(1 - \mathbb{1}_{[X_j^{(1)} = X_j]}\right) - \mathbb{1}_{[X_j^{(1)} > X_j]}}{\rho_j(X_j^{(1)})} \right], \tag{3.24}$$

and

$$\mathbf{f}_u^{sup}(\mathbf{X}) = (-1)^{|u|+1} \mathbb{E}_{\mathbf{X}_u^{(1)}} \left[\frac{\partial^{|u|}\mathbf{f}}{\partial \mathbf{x}_u} \left(\mathbf{X}_u^{(1)}, \mathbf{X}_{\sim u}\right) \prod_{j \in u} \frac{F_j(X_j^{(1)}) \left(1 - \mathbb{1}_{[X_j^{(1)} = X_j]}\right) - \mathbb{1}_{[X_j^{(1)} > X_j]}}{\rho_j(X_j^{(1)})} \right], \tag{3.25}$$

where $\mathbb{E}_{\mathbf{X}_u^{(1)}}$ *means that the expectation is taken with respect to* $\mathbf{X}_u^{(1)}$, $\prod_{j\in v}$ *is the product of a sequence depending on* j *with* $j \in v$, *and* $\mathbb{1}_{[X_j^{(1)}=X_j]} = 1$ *if* $X_j^{(1)} = X_j$ *and* 0 *otherwise.*

(ii) We have the following expansion of the function $\mathbf{f}$.

$$\mathbf{f}(\mathbf{X}) = \mathbf{f}_\emptyset + \sum_{\substack{v,\, v\subseteq\{1,\ldots,d\}\\ |v|>0}} \mathbb{E}_{\mathbf{X}^{(1)}}\left[\frac{\partial^{|v|}\mathbf{f}}{\partial \mathbf{x}_v}\left(\mathbf{X}_v^{(1)}, \mathbf{X}_{\sim v}\right)\prod_{j\in v}\frac{F_j(X_j^{(1)})\left(1-\mathbb{1}_{[X_j^{(1)}=X_j]}\right)-\mathbb{1}_{[X_j^{(1)}>X_j]}}{\rho_j(X_j^{(1)})}\right], \tag{3.26}$$

with $\mathbf{f}_\emptyset = \mathbb{E}\left[\mathbf{f}(\mathbf{X})\right]$.

(iii) The total-interaction effect covariance matrix becomes

$$\Sigma_u^{sup} = \mathbb{V}\left[\mathbb{E}\left(\frac{\partial^{|u|}\mathbf{f}}{\partial \mathbf{x}_u}\left(\mathbf{X}_u^{(1)}, \mathbf{X}_{\sim u}\right)\prod_{j\in u}\frac{F_j(X_j^{(1)})\left(1-\mathbb{1}_{[X_j^{(1)}=X_j]}\right)-\mathbb{1}_{[X_j^{(1)}>X_j]}}{\rho_j(X_j^{(1)})}\,|\,\mathbf{X}\right)\right]. \tag{3.27}$$

Likewise, the total effect covariance matrix is given by

$$\Sigma_u^{tot} = \mathbb{V}\left[\mathbb{E}\left(\sum_{\substack{v,\, v\subseteq u\\ |v|>0}}\frac{\partial^{|v|}\mathbf{f}}{\partial \mathbf{x}_v}\left(\mathbf{X}_v^{(1)}, \mathbf{X}_{\sim v}\right)\prod_{j\in v}\frac{F_j(X_j^{(1)})\left(1-\mathbb{1}_{[X_j^{(1)}=X_j]}\right)-\mathbb{1}_{[X_j^{(1)}>X_j]}}{\rho_j(X_j^{(1)})}\,|\,\mathbf{X}\right)\right]. \tag{3.28}$$

Proof. *See AppendixD.*

□

From Theorem 1, it comes out that the total effect and total-interaction effect covariance matrices are the first-order covariance matrices of the functions

$$\sum_{\substack{v,\, v\subseteq u\\ |v|>0}}\frac{\partial^{|v|}\mathbf{f}}{\partial \mathbf{x}_v}\left(\mathbf{X}_v^{(1)}, \mathbf{X}_{\sim v}\right)\prod_{j\in v}\frac{F_j(X_j^{(1)})\left(1-\mathbb{1}_{[X_j^{(1)}=X_j]}\right)-\mathbb{1}_{[X_j^{(1)}>X_j]}}{\rho_j(X_j^{(1)})},$$

and

$$\frac{\partial^{|u|}\mathbf{f}}{\partial \mathbf{x}_u}\left(\mathbf{X}_u^{(1)}, \mathbf{X}_{\sim u}\right)\prod_{j\in u}\frac{F_j(X_j^{(1)})\left(1-\mathbb{1}_{[X_j^{(1)}=X_j]}\right)-\mathbb{1}_{[X_j^{(1)}>X_j]}}{\rho_j(X_j^{(1)})},$$

respectively. Thus, the estimator of the first-order covariance matrices from [1] can be adapted to obtain the estimators of GSIs using model derivatives.

Remark 2. *The derivative-based expressions of the TEF and TIEF in Equations (3.24-3.25) are still suitable for functions that are continuous on the joint support* Ω *and differentiable almost everywhere.*

It is obvious that the TEF and TIEF are centered, that is,

$$\mathbb{E}\left[\mathbf{f}_u^{tot}(\mathbf{X})\right] = \mathbb{E}\left[\mathbf{f}_u^{sup}(\mathbf{X})\right] = \mathbf{0}\,.$$

4. Estimators of generalized sensitivity indices using model derivatives

The theory of U-statistics allows for easily deriving the properties of estimators ([32; 33; 34; 29; 35; 36]). Lamboni [10; 1; 12] introduced the theory of U-statistics in sensitivity analysis by deriving minimum variance unbiased estimators of (the non-normalized) GSIs and Sobol' indices as well as the consistency and the asymptotic normality of such estimators. The main idea consists in i) constructing a kernel (i.e., a random function which expectation is exactly our parameter of interest); ii) proposing the estimator of that parameter using the kernel; iii) deriving the statistical performance of the proposed estimator. In this section, we follow these main steps to construct estimators of the total-effect, the total-interaction effect covariance matrices and GSIs by making use of model derivatives.

4.1. Kernel functions for the total-effect and the total-interaction effect covariance matrices

Let $u \subseteq \{1, \ldots, d\}$ be a non-empty subset, $\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}\right)$ be 2 i.i.d copies of $\mathbf{X}$, $\left(\mathbf{X}_u^{(3)}, \mathbf{X}_u^{(4)}\right)$ be 2 i.i.d copies of $\mathbf{X}_u$. For all $v \subseteq u$ with $|v| > 0$, we define the weight-derivative function ($\mathbf{d}(\cdot)$) as follows:

$$\mathbf{d}\left(\mathbf{X}^{(1)}, \mathbf{X}_v^{(3)}\right) = \frac{\partial^{|v|}\mathbf{f}}{\partial \mathbf{x}_v}\left(\mathbf{X}_v^{(3)}, \mathbf{X}_{\sim v}^{(1)}\right) \prod_{j \in v} \frac{F_j\left(X_j^{(3)}\right)\left(1 - \mathbb{1}_{\left[X_j^{(3)} = X_j^{(1)}\right]}\right) - \mathbb{1}_{\left[X_j^{(3)} > X_j^{(1)}\right]}}{\rho_j\left(X_j^{(3)}\right)}\,. \tag{4.29}$$

To estimate the total-effect and the total-interaction covariance matrices using model derivatives, we consider two functions with two types of inputs $\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}\right)$ and $\left(\mathbf{X}_u^{(3)}, \mathbf{X}_u^{(4)}\right)$. For the total-interaction covariance matrix, we consider the function $\mathbf{K}^{sup}(\cdot)$ given by

$$\begin{aligned}
&\mathbf{K}^{sup}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \mathbf{X}_u^{(3)}, \mathbf{X}_u^{(4)}\right) = \\
&\frac{1}{4}\left(\left[\mathbf{d}(\mathbf{X}^{(1)}, \mathbf{X}_u^{(3)}) - \mathbf{d}(\mathbf{X}^{(2)}, \mathbf{X}_u^{(3)})\right] \times \left[\mathbf{d}(\mathbf{X}^{(1)}, \mathbf{X}_u^{(4)}) - \mathbf{d}(\mathbf{X}^{(2)}, \mathbf{X}_u^{(4)})\right]^T\right) \\
&+\frac{1}{4}\left(\left[\mathbf{d}(\mathbf{X}^{(1)}, \mathbf{X}_u^{(4)}) - \mathbf{d}(\mathbf{X}^{(2)}, \mathbf{X}_u^{(4)})\right] \times \left[\mathbf{d}(\mathbf{X}^{(1)}, \mathbf{X}_u^{(3)}) - \mathbf{d}(\mathbf{X}^{(2)}, \mathbf{X}_u^{(3)})\right]^T\right).
\end{aligned} \tag{4.30}$$

Likewise, we consider the function $\mathbf{K}^{tot}(\cdot)$ for the total-effect covariance matrix with

$$\begin{aligned}
&\mathbf{K}^{tot}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \mathbf{X}_u^{(3)}, \mathbf{X}_u^{(4)}\right) = \\
&\frac{1}{4}\left(\sum_{\substack{v,\, v\subseteq u \\ |v|>0}} \sum_{\substack{v_1,\, v_1\subseteq u \\ |v_1|>0}} \left[\mathbf{d}(\mathbf{X}^{(1)}, \mathbf{X}_v^{(3)}) - \mathbf{d}(\mathbf{X}^{(2)}, \mathbf{X}_v^{(3)})\right] \times \left[\mathbf{d}(\mathbf{X}^{(1)}, \mathbf{X}_{v_1}^{(4)}) - \mathbf{d}(\mathbf{X}^{(2)}, \mathbf{X}_{v_1}^{(4)})\right]^T\right) \\
&+\frac{1}{4}\left(\sum_{\substack{v,\, v\subseteq u \\ |v|>0}} \sum_{\substack{v_1,\, v_1\subseteq u \\ |v_1|>0}} \left[\mathbf{d}(\mathbf{X}^{(1)}, \mathbf{X}_{v_1}^{(4)}) - \mathbf{d}(\mathbf{X}^{(2)}, \mathbf{X}_{v_1}^{(4)})\right] \times \left[\mathbf{d}(\mathbf{X}^{(1)}, \mathbf{X}_v^{(3)}) - \mathbf{d}(\mathbf{X}^{(2)}, \mathbf{X}_v^{(3)})\right]^T\right).
\end{aligned} \tag{4.31}$$

The functions $\mathbf{K}^{sup}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \mathbf{X}_u^{(3)}, \mathbf{X}_u^{(4)}\right)$ and $\mathbf{K}^{tot}\left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \mathbf{X}_u^{(3)}, \mathbf{X}_u^{(4)}\right)$ are symmetric under independent permutations of their first arguments $(\mathbf{X}^{(1)}, \mathbf{X}^{(2)})$ and second arguments $(\mathbf{X}_u^{(3)}, \mathbf{X}_u^{(4)})$. Indeed, the values of such functions do not change if we permute the position of $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ in one hand, and the position of $\mathbf{X}_u^{(3)}$ and $\mathbf{X}_u^{(4)}$ in the other hand. Theorem 2 gives other properties of such functions.

Theorem 2. *If assumptions A1-A5 hold, then we have*

$$\mathbb{E}\left[\mathbf{K}^{sup}\left(\mathbf{X}^{(1)},\,\mathbf{X}^{(2)},\mathbf{X}_u^{(3)},\,\mathbf{X}_u^{(4)}\right)\right] \;=\; \Sigma_u^{sup}\,, \tag{4.32}$$

$$\mathbb{E}\left[\mathbf{K}^{tot}\left(\mathbf{X}^{(1)},\,\mathbf{X}^{(2)},\mathbf{X}_u^{(3)},\,\mathbf{X}_u^{(4)}\right)\right] \;=\; \Sigma_u^{tot}\,. \tag{4.33}$$

Proof. *See AppendixE.*

□

Theorem 2 shows that the functions $\mathbf{K}(\cdot)$ (resp. $\mathbf{K}^{tot}(\cdot)$) are unbiased estimators of Σ_u^{sup} (resp. Σ_u^{tot}). Both functions are called kernels of degree (2, 2) in the theory of U-statistics of two samples. Theorems 3-4 propose the estimators of Σ_u^{sup} and Σ_u^{tot}.

Theorem 3. *Let* $\mathcal{X} = \left(\mathbf{X}^{(1)},\,\mathbf{X}^{(2)}\right)$, $\mathcal{Y} = \left(\mathbf{X}_u^{(3)},\,\mathbf{X}_u^{(4)}\right)$, *and* $\mathcal{X}_i$, $\mathcal{Y}_i$, $i = 1, 2, \ldots, m$, *be two independent samples of size* m *from* $\mathcal{X}$ *and* $\mathcal{Y}$, *respectively. If assumptions A1-A5 hold, then*

(*i*) *the minimum variance unbiased estimator of* Σ_u^{sup} *for a given* m *and degree* (2, 2) *is given by*

$$\begin{aligned}\widehat{\Sigma_u^{sup}} \;=\; \frac{1}{4m}\sum_{i=1}^{m}\Bigg(&\left[\mathbf{d}(\mathbf{X}_i^{(1)},\mathbf{X}_{i,u}^{(3)}) - \mathbf{d}(\mathbf{X}_i^{(2)},\mathbf{X}_{i,u}^{(3)})\right]\times\left[\mathbf{d}(\mathbf{X}_i^{(1)},\mathbf{X}_{i,u}^{(4)}) - \mathbf{d}(\mathbf{X}_i^{(2)},\mathbf{X}_{i,u}^{(4)})\right]^T\\ &+\left[\mathbf{d}(\mathbf{X}_i^{(1)},\mathbf{X}_{i,u}^{(4)}) - \mathbf{d}(\mathbf{X}_i^{(2)},\mathbf{X}_{i,u}^{(4)})\right]\times\left[\mathbf{d}(\mathbf{X}_i^{(1)},\mathbf{X}_{i,u}^{(3)}) - \mathbf{d}(\mathbf{X}_i^{(2)},\mathbf{X}_{i,u}^{(3)})\right]^T\Bigg)\,,\end{aligned} \tag{4.34}$$

and we have

$$\mathbb{E}\left(\widehat{\Sigma_u^{sup}}\right) = \Sigma_u^{sup}\,. \tag{4.35}$$

(*ii*) *If* $m \to +\infty$, $\widehat{\Sigma_u^{sup}}$ *is consistent, that is,*

$$\widehat{\Sigma_u^{sup}} \xrightarrow{\mathcal{P}} \Sigma_u^{sup}\,. \tag{4.36}$$

(*iii*) *If* $m \to +\infty$, $Vec\left[\widehat{\Sigma_u^{sup}}\right]$ *follows a normal distribution, that is,*

$$\sqrt{m}\left(Vec\left[\widehat{\Sigma_u^{sup}}\right] - Vec\left[\Sigma_u^{sup}\right]\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\mathbf{0},\,\mathbb{V}\left(Vec[\mathbf{K}^{sup}(\mathcal{X},\,\mathcal{Y})]\right)\right)\,. \tag{4.37}$$

Proof. *See AppendixF.*

□

Theorem 4. *Let* $\mathcal{X} = \left(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}\right)$, $\mathcal{Y} = \left(\mathbf{X}_u^{(3)}, \mathbf{X}_u^{(4)}\right)$, *and* $\mathcal{X}_i$, $\mathcal{Y}_i$, $i = 1, 2, \ldots, m$, *be two independent samples of size* m *from* $\mathcal{X}$ *and* $\mathcal{Y}$, *respectively. If assumptions A1-A5 hold, then*

(*i*) *the minimum variance unbiased estimator of* Σ_u^{tot} *for a given* m *and degree* $(2, 2)$ *is given by*

$$\begin{aligned}\widehat{\Sigma_u^{tot}} &= \frac{1}{4m}\sum_{i=1}^{m}\sum_{\substack{v,\, v\subseteq u\\ |v|>0}}\sum_{\substack{v_1,\, v_1\subseteq u\\ |v_1|>0}} \\ &\times\left(\left[\mathbf{d}(\mathbf{X}_i^{(1)}, \mathbf{X}_{i,v}^{(3)}) - \mathbf{d}(\mathbf{X}_i^{(2)}, \mathbf{X}_{i,v}^{(3)})\right] \times \left[\mathbf{d}(\mathbf{X}_i^{(1)}, \mathbf{X}_{i,v_1}^{(4)}) - \mathbf{d}(\mathbf{X}_i^{(2)}, \mathbf{X}_{i,v_1}^{(4)})\right]^T\right. \\ &\left.+\left[\mathbf{d}(\mathbf{X}_i^{(1)}, \mathbf{X}_{i,v_1}^{(4)}) - \mathbf{d}(\mathbf{X}_i^{(2)}, \mathbf{X}_{i,v_1}^{(4)})\right] \times \left[\mathbf{d}(\mathbf{X}_i^{(1)}, \mathbf{X}_{i,v}^{(3)}) - \mathbf{d}(\mathbf{X}_i^{(2)}, \mathbf{X}_{i,v}^{(3)})\right]^T\right),\end{aligned} \tag{4.38}$$

and we have

$$\mathbb{E}\left(\widehat{\Sigma_u^{tot}}\right) = \Sigma_u^{tot}. \tag{4.39}$$

(*ii*) *If* $m \to +\infty$, $\widehat{\Sigma_u^{tot}}$ *is consistent, that is,*

$$\widehat{\Sigma_u^{tot}} \xrightarrow{\mathcal{P}} \Sigma_u^{tot}. \tag{4.40}$$

(*iii*) *If* $m \to +\infty$, $Vec\left[\widehat{\Sigma_u^{tot}}\right]$ *follows a normal distribution, that is,*

$$\sqrt{m}\left(Vec\left[\widehat{\Sigma_u^{tot}}\right] - Vec\left[\Sigma_u^{tot}\right]\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\mathbf{0}, \mathbb{V}(Vec[\mathbf{K}^{tot}(\mathcal{X}, \mathcal{Y})])\right). \tag{4.41}$$

Proof. *The proof is similar to the proof of Theorem 3.*

□

Remark 3. *The property of minimum variance is valid for a class of estimators that are based on the weight-derivative function and make use of the same information such as the degree* $(2, 2)$ *and the same sample size* (m).

Theorems 3-4 give the formulas for computing the TIEF and TEF covariance matrices. For computing GSIs using matrix norms, we use the formulas proposed in Corollaries 1-2.

Corollary 1. *Under assumptions A1-A5,*

(*i*) *the minimum variance unbiased estimator of* $Tr(\Sigma_u^{sup})$ *is given by*

$$\widehat{Tr(\Sigma_u^{sup})} = \frac{1}{2m}\sum_{i=1}^{m} Tr\Big(\Big[\mathbf{d}(\mathbf{X}_i^{(1)},\mathbf{X}_{i,u}^{(3)}) - \mathbf{d}(\mathbf{X}_i^{(2)},\mathbf{X}_{i,u}^{(3)})\Big] \times \Big[\mathbf{d}(\mathbf{X}_i^{(1)},\mathbf{X}_{i,u}^{(4)}) - \mathbf{d}(\mathbf{X}_i^{(2)},\mathbf{X}_{i,u}^{(4)})\Big]^T\Big) , \tag{4.42}$$

and we have

$$\mathbb{E}\left[\widehat{Tr(\Sigma_u^{sup})}\right] = Tr(\Sigma_u^{sup}) ; \tag{4.43}$$

$$\widehat{Tr(\Sigma_u^{sup})} \xrightarrow{\mathcal{P}} Tr(\Sigma_u^{sup}) ; \tag{4.44}$$

$$\sqrt{m}\left(\widehat{Tr(\Sigma_u^{sup})} - Tr(\Sigma_u^{sup})\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\mathbf{0},\, \mathbb{V}(Tr[\mathbf{K}^{sup}(\mathcal{X},\,\mathcal{Y})])\right) . \tag{4.45}$$

(*ii*) *A consistent estimator of* $||\Sigma_u^{sup}||_F$ *is given by*

$$\widehat{||\Sigma_u^{sup}||}_F = \left|\left|\widehat{\Sigma_u^{sup}}\right|\right|_F , \tag{4.46}$$

and we have

$$\widehat{||\Sigma_u^{sup}||}_F \xrightarrow{\mathcal{P}} ||\Sigma_u^{sup}||_F . \tag{4.47}$$

Proof. *The proof is obvious bearing in mind Theorem 3, the Slutsky theorem, and the linearity of the trace and expectation.*

□

Likewise, we use the following formulas for computing the non-normalized total GSIs of $\mathbf{X_u}$.

Corollary 2. *Under assumptions A1-A5,*

(*i*) *the minimum variance unbiased estimator of* $Tr(\Sigma_u^{tot})$ *is given by*

$$\widehat{Tr(\Sigma_u^{tot})} = \frac{1}{2m}\sum_{i=1}^{m}\sum_{\substack{v,\, v\subseteq u\\ |v|>0}}\sum_{\substack{v_1,\, v_1\subseteq u\\ |v_1|>0}} Tr\Big(\Big[\mathbf{d}(\mathbf{X}_i^{(1)},\mathbf{X}_{i,v}^{(3)}) - \mathbf{d}(\mathbf{X}_i^{(2)},\mathbf{X}_{i,v}^{(3)})\Big] \times \Big[\mathbf{d}(\mathbf{X}_i^{(1)},\mathbf{X}_{i,v_1}^{(4)}) - \mathbf{d}(\mathbf{X}_i^{(2)},\mathbf{X}_{i,v_1}^{(4)})\Big]^T\Big) , \tag{4.48}$$

and we have

$$\mathbb{E}\left[\widehat{Tr\left(\Sigma_u^{tot}\right)}\right] \;=\; Tr\left(\Sigma_u^{tot}\right)\;; \tag{4.49}$$

$$\widehat{Tr\left(\Sigma_u^{tot}\right)} \xrightarrow{\mathcal{P}} Tr\left(\Sigma_u^{tot}\right)\;; \tag{4.50}$$

$$\sqrt{m}\left(\widehat{Tr\left(\Sigma_u^{tot}\right)} - Tr\left(\Sigma_u^{tot}\right)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\mathbf{0},\, \mathbb{V}(Tr[\mathbf{K}^{tot}(\mathcal{X},\mathcal{Y})])\right)\;. \tag{4.51}$$

(ii) A consistent estimator of $||\Sigma_u^{tot}||_F$ *is given by*

$$\widehat{||\Sigma_u^{tot}||}_F = \left|\left|\widehat{\Sigma_u^{tot}}\right|\right|_F\;, \tag{4.52}$$

and we have

$$\widehat{||\Sigma_u^{tot}||}_F \xrightarrow{\mathcal{P}} \left|\left|\Sigma_u^{tot}\right|\right|_F\;. \tag{4.53}$$

Proof. *The proof is obvious bearing in mind Theorem 4, the Slutsky theorem, and the linearity of the trace and expectation.*

□

Now, we can derive the estimators of the normalized GSIs proposed in Definition 2. Theorem 5 gives the estimators of both types of GSIs, and it provides the performance of these estimators.

Theorem 5. *If assumptions A1-A5 hold, then*

(i) the first-type and total GSI of $\mathbf{X_u}$ *is given by*

$$\widehat{GSI_{T_u}^F} = \frac{\widehat{Tr\left(\Sigma_u^{tot}\right)}}{Tr\left(\widehat{\Sigma}\right)}\,, \tag{4.54}$$

where $\widehat{\Sigma}$ *is the classical estimator of the variance-covariance matrix of the model outputs, that is,* $\Sigma = \mathbb{V}[\mathbf{f}(\mathbf{X})]$*, and it is estimated using* M *model runs.*

Moreover, if $m \to +\infty$ *and* $m/M \to 0$ *then we have*

$$\widehat{GSI_{T_u}^F} \xrightarrow{\mathcal{P}} GSI_{T_u}^F\,, \tag{4.55}$$

and

$$\sqrt{m}\left(\widehat{GSI_{T_u}^F} - GSI_{T_u}^F\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\mathbf{0},\, \frac{\mathbb{V}(Tr[\mathbf{K}^{tot}(\mathcal{X},\mathcal{Y})])}{(Tr\,[\Sigma])^2}\right)\;. \tag{4.56}$$

(*ii*) *The first-type and total-interaction GSI of* $\mathbf{X_u}$ *is given by*

$$\widehat{GSI^F_{sup,u}} = \frac{\widehat{Tr(\Sigma^{sup}_u)}}{Tr\left(\widehat{\Sigma}\right)}, \tag{4.57}$$

and if $m \to +\infty$ *and* $m/M \to 0$ *we have*

$$\widehat{GSI^F_{sup,u}} \xrightarrow{\mathcal{P}} GSI^F_{sup,u}. \tag{4.58}$$

$$\sqrt{m}\left(\widehat{GSI^F_{sup,u}} - GSI^F_{sup,u}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\mathbf{0},\, \frac{\mathbb{V}(Tr[\mathbf{K}^{sup}(\mathcal{X},\,\mathcal{Y})])}{(Tr[\Sigma])^2}\right). \tag{4.59}$$

(*iii*) *The second-type (total and total-interaction) GSIs of* $\mathbf{X_u}$ *are given by*

$$\widehat{GSI^{l_2}_{T_u}} = \frac{\widehat{||\Sigma^{tot}_u||_F}}{n\,Tr\left(\widehat{\Sigma}\right)}, \tag{4.60}$$

with

$$\widehat{GSI^{l_2}_{T_u}} \xrightarrow{\mathcal{P}} GSI^{l_2}_{T_u}; \tag{4.61}$$

and

$$\widehat{GSI^{l_2}_{sup,u}} = \frac{\widehat{||\Sigma^{sup}_u||_F}}{n\,Tr\left(\widehat{\Sigma}\right)}, \tag{4.62}$$

with

$$\widehat{GSI^{l_2}_{sup,u}} \xrightarrow{\mathcal{P}} GSI^{l_2}_{sup,u}. \tag{4.63}$$

Proof. *See AppendixG.*

□

5. Computational issues

This section illustrates the computations of GSIs and Sobol' indices using derivatives of different functions. First, we provide an algorithm for computing the GSIs using the proposed estimators, and second, we apply such algorithm on different functions.

5.1. Algorithm, design scheme and computational cost

For a given sample size m, the following steps allow for computing the total-effect and total-interaction effect covariance matrices.

Algorithm 1. *Estimations of the d total-effect covariance matrices.*

- *Sample 4 input values (matrices) of type $m \times d$ ($\mathcal{X}_1, \ldots, \mathcal{X}_4$)*
- *For each input factor X_j, replace the j^{th} column of $\mathcal{X}_1$ and $\mathcal{X}_3$ with the j^{th} column of $\mathcal{X}_2$ (i.e, X_{2j}) to obtain 2 new matrices ($\mathcal{X}_{1,2j}$, $\mathcal{X}_{3,2j}$)*
- *Compute the model derivative with respect to X_j for the two input values $\mathcal{X}_{1,2j}$, $\mathcal{X}_{3,2j}$ to obtain $\frac{\partial \mathbf{f}}{\partial \mathbf{x}_j}(\mathcal{X}_{1,2j})$ and $\frac{\partial \mathbf{f}}{\partial \mathbf{x}_j}(\mathcal{X}_{3,2j})$*
- *Run the weight-derivative function from Equation (4.29) for $(\mathcal{X}_1\, X_{2j})$ (resp. $(\mathcal{X}_3\, X_{2j})$) as follows:*

$$\mathbf{d}\,(\mathcal{X}_1\, X_{2j}) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}_j}(\mathcal{X}_{1,2j})\, \frac{F_j(X_{2j})\left(1 - \mathbb{1}_{[X_{2j}=X_{1j}]}\right) - \mathbb{1}_{[X_{2j}>X_{1j}]}}{\rho_j(X_{2j})}\,,$$

  with X_{1j} the j^{th} column of $\mathcal{X}_1$

- *Repeat the loop for one more time by replacing $\mathcal{X}_2$ with $\mathcal{X}_4$*
- *Use the estimators in (4.38), (4.54), and (4.60) to obtain the estimates of the total-effect covariance matrices and GSIs.*

Algorithm 1 is still suitable for computing the total-interaction effect covariance matrices (4.34) and GSIs (4.57, 4.62) of the subset of input factors $\mathbf{X}_v$, with $|v| > 1$. We should modify the steps ii)-iv) of Algorithm 1 as follows: replace the columns of $\mathcal{X}_1$ and $\mathcal{X}_3$ that belong to v with the same columns of $\mathcal{X}_2$ to obtain 2 new matrices, and compute the weight-derivative function as defined in (4.29).

The total number of the evaluations of the model derivatives to obtain the values of the d total indices (i.e., computational cost of this algorithm) is $4 \times m \times d$. For the second-order total-interaction indices, up to $4 \times m \times d(d-1)$ evaluations of the model derivatives are necessary for obtaining the indices estimates, as a null partial derivative yields to null cross-partial derivatives. These computational costs depend on the crude estimators used in this paper.

5.2. Test cases

In this section, we performed some numerical tests to assess the effectiveness of our estimations. To illustrate our approach, we considered two types of functions as follows: functions with a small number of inputs ($d = 3$), and functions with a medium number of inputs ($d = 6$, $d = 10$). We computed the model derivatives using the finite difference method, and we computed the GSIs for two values of the sample size (m), that is, $m = 1000$, 5000 using Sobol's sequence or Quasi-Monte Carlo ([37]). For each sample size (m), we replicated the process of computing the indices $R = 50$ times by changing the seed randomly when sampling input values.

5.2.1. Multivariate Ishigami's function ($d = 3$, $n = 3$)

The multivariate Ishigami function includes three independent input factors following a uniform distribution on $[-\pi, \pi]$, and it provides three outputs ([1]). It is defined as follows:

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \sin(x_1) + 7\sin^2(x_2) + 0.1\, x_3^4 \sin(x_1) \\ \sin(x_1) + 5.896\sin^2(x_2) + 0.1\, x_3^4 \sin(x_1) \\ \sin(x_1) + 6.494\sin^2(x_2) + 0.125\, x_3^4 \sin(x_1) \end{bmatrix}. \tag{5.64}$$

The true and estimated values of the GSIs for this function are listed in Table 1.

5.2.2. Block-additive function ($d = 6$, $n = 1$)

The block-additive function includes six independent inputs following a uniform distribution on $[-1, 1]$ ([27]). It is defined as follows:

$$\mathbf{f}(\mathbf{x}) = \cos(-0.8 - 1.1x_1 + 1.1x_5 + x_3) + \sin(0.5 + 0.9x_4 + x_2 - 1.1x_6)\,. \tag{5.65}$$

The true values of Sobol' indices and the estimates are listed in Table 2.

	First-type GSIs			Second-type GSIs		
	$GSI^F_{T.j}$	$\widehat{GSI^F_{T.j}}$		$GSI^{l2}_{T.j}$	$\widehat{GSI^{l2}_{T.j}}$	
		$m = 1000$	5000		$m = 1000$	5000
X_1	0.628	0.639 (0.14)	0.630 (0.06)	0.209	0.213 (0.04)	0.210 (0.02)
X_2	0.372	0.340 (0.35)	0.375 (0.11)	0.125	0.135 (0.09)	0.125 (0.04)
X_3	0.284	0.272 (0.13)	0.277 (0.04)	0.093	0.091 (0.04)	0.092 (0.01)
	$GSI^F_{sup,u}$	$\widehat{GSI^F_{sup,u}}$		$GSI^{l2}_{sup,u}$	$\widehat{GSI^{l2}_{sup,u}}$	
$X_1 : X_2$	0	e-23 (e-22)	3e-24 (e-23)	0	3e-23 (e-22)	9e-24 (e-23)
$X_1 : X_3$	0.284	0.336 (0.74)	0.273 (0.21)	0.093	0.155 (0.22)	0.101 (0.05)
$X_2 : X_3$	0	6e-25 (6e-24)	-6e-25 (3e-24)	0	e-24 (2e-24)	8e-25 (e-24)

Table 1: True and estimated values of two types of generalized sensitivity indices for the multivariate Ishigami function. The estimated GSIs (average over 50 replications) are followed by their standard deviations (in bracket). While the top part of this table focuses on the total GSI, the bottom part deals with the total-interaction GSIs. For concise reporting of small indices, we use e-a for 10^{-a}.

5.2.3. Multivariate Sobol's function ($d = 10$, $n = 4$)

The multivariate Sobol function includes 10 independent input factors following a uniform distribution on [0, 1] ([1]). It is defined as follows:

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \prod_{j=1}^{d=10} \frac{|4\,x_j - 2| + \mathcal{A}[1,j]}{1 + \mathcal{A}[1,j]} \\ \prod_{j=1}^{d=10} \frac{|4\,x_j - 2| + \mathcal{A}[2,j]}{1 + \mathcal{A}[2,j]} \\ \prod_{j=1}^{d=10} \frac{|4\,x_j - 2| + \mathcal{A}[3,j]}{1 + \mathcal{A}[3,j]} \\ \prod_{j=1}^{d=10} \frac{|4\,x_j - 2| + \mathcal{A}[4,j]}{1 + \mathcal{A}[4,j]} \end{bmatrix}. \tag{5.66}$$

According to the values of $\mathcal{A}$ (matrix of type $4 \times d$), this function has different properties. If

$$\mathcal{A} = \begin{bmatrix} 0 & 0 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 & 6.52 \\ 0 & 1 & 4.5 & 9 & 99 & 99 & 99 & 99 & 99 & 99 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 & 50 \end{bmatrix},$$

the values of the GSIs are listed in Table 3.

	First-type GSIs			Second-type GSIs		
	$GSI^F_{T.j}$	$\widehat{GSI^F_{T.j}}$		$GSI^{l2}_{T.j}$	$\widehat{GSI^{l2}_{T.j}}$	
		$m = 1000$	5000		$m = 1000$	5000
X_1	0.231	0.229 (0.02)	0.231 (0.01)	0.231	0.229 (0.02)	0.231 (0.01)
X_2	0.214	0.216 (0.02)	0.213 (0.01)	0.214	0.216 (0.02)	0.213 (0.01)
X_3	0.196	0.196 (0.01)	0.196 (0.004)	0.196	0.196 (0.01)	0.196 (0.004)
X_4	0.176	0.176 (0.01)	0.175 (0.005)	0.176	0.176 (0.01)	0.175 (0.005)
X_5	0.231	0.230 (0.02)	0.231 (0.007)	0.231	0.230 (0.02)	0.231 (0.007)
X_6	0.256	0.254 (0.01)	0.257 (0.006)	0.256	0.254 (0.01)	0.257 (0.006)
	$GSI^F_{sup,u}$	$\widehat{GSI^F_{sup,u}}$		$GSI^{l2}_{sup,u}$	$\widehat{GSI^{l2}_{sup,u}}$	
$X_1 : X_2$	0	2e-20 (e-19)	-9e-22 (6e-21)	0	2e-20 (e-19)	1e-21 (6e-21)
$X_1 : X_3$	0.067	0.069 (0.01)	0.067 (0.006)	0.067	0.069 (0.01)	0.067 (0.006)
$X_1 : X_4$	0	3e-22 (e-21)	-9e-22 (5e-21)	0	8e-22 (e-21)	1e-22 (5e-21)
$X_1 : X_5$	0.078	0.080 (0.01)	0.079 (0.006)	0.078	0.080 (0.01)	0.079 (0.006)
$X_1 : X_6$	0	-4e-21 (3e-20)	-3e-22 (e-21)	0	5e-21 (3e-20)	6e-22 (e-21)
$X_2 : X_3$	0	e-22 (2e-21)	7e-22 (3e-21)	0	9e-22 (2e-21)	9e-22 (3e-21)
$X_2 : X_4$	0.040	0.041 (0.007)	0.040 (0.004)	0.040	0.041 (0.007)	0.040 (0.004)
$X_2 : X_5$	0	6e-23 (e-21)	e-22 (3e-21)	0	5e-22 (e-21)	e-21 (3e-21)
$X_2 : X_6$	0.053	0.054 (0.014)	0.053 (0.005)	0.053	0.054 (0.014)	0.053 (0.005)
$X_3 : X_4$	0	2e-22 (e-21)	-4e-20 (3e-19)	0	4e-22 (e-21)	5e-20 (3e-19)
$X_3 : X_5$	0.067	0.066 (0.010)	0.068 (0.006)	0.067	0.066 (0.010)	0.068 (0.006)
$X_3 : X_6$	0	2e-20 (e-19)	7e-22 (5e-21)	0	2e-20 (e-19)	1e-21 (4e-21)
$X_4 : X_5$	0	-4e-22 (2e-21)	2e-22 (2e-21)	0	9e-22 (2e-21)	1e-21 (2e-21)
$X_4 : X_6$	0.046	0.045 (0.01)	0.045 (0.004)	0.046	0.045 (0.01)	0.045 (0.004)
$X_5 : X_6$	0	-5e-23 (e-21)	e-22 (2e-21)	0	4e-22 (e-21)	8e-22 (2e-21)

Table 2: True and estimated values of Sobol' indices for the block-additive function. The estimated indices (average over 50 replications) are followed by their standard deviations (in bracket). While the top part of this table focuses on the total indices, the bottom part deals with the total-interaction indices. For concise reporting of small indices, we use e-a for 10^{-a}.

5.3. Numerical results and discussion

Tables 1-3 report the estimated values of sensitivity indices for the three functions. From Tables 1-3, it comes out that our estimators give accurate estimates of both GSIs and Sobol' indices in average, and these estimates allow for identifying the most influential input factors. Regarding the precision of our estimates (standard deviations), we obtain better estimates of indices in the case of the block-additive function and the multivariate Sobol function when the sample size is 1000. For the multivariate Ishigami function, we have the worst precision for input X_2, showing a big impact of the sample size on the estimates and/or a numerical instability of

	First-type GSIs			Second-type GSIs		
	$GSI^F_{T.j}$	$\widehat{GSI^F_{T.j}}$		$GSI^{l2}_{T.j}$	$\widehat{GSI^{l2}_{T.j}}$	
		$m = 1000$	5000		$m = 1000$	5000
X_1	0.605	0.619 (0.19)	0.617 (0.13)	0.147	0.152 (0.05)	0.150 (0.03)
X_2	0.406	0.395 (0.14)	0.409 (0.04)	0.103	0.100 (0.03)	0.101 (0.01)
X_3	0.034	0.034 (0.008)	0.034 (0.003)	0.008	0.008 (0.002)	0.008 (0.001)
X_4	0.021	0.022 (0.006)	0.020 (0.002)	0.005	0.005 (0.002)	0.004 (0.001)
X_5	0.014	0.014 (0.003)	0.015 (0.001)	0.003	0.003 (0.001)	0.004 (0.001)
X_6	0.013	0.012 (0.003)	0.013 (0.002)	0.003	0.003 (0.001)	0.003 (0.0004)
X_7	0.011	0.011 (0.003)	0.011 (0.001)	0.003	0.003 (0.001)	0.003(0.0003)
X_8	0.011	0.011 (0.003)	0.011 (0.001)	0.002	0.003 (0.001)	0.003 (0.0003)
X_9	0.010	0.010 (0.004)	0.010 (0.002)	0.002	0.002 (0.001)	0.002 (0.0004)
X_{10}	0.009	0.010 (0.002)	0.009 (0.001)	0.002	0.002 (0.001)	0.002 (0.0002)

Table 3: True and estimated values of two types of the total GSIs for the multivariate Sobol function of type A. The estimated GSIs (average over 50 replications) are followed by their standard deviations (in bracket).

our estimators. This numerical instability decreases when increasing the sample size ($m = 5000$), and we obtain reasonable precision for all input factors.

It happens to obtain negative estimates of some null total-interaction indices (Tables 1-2), although the negative estimates are too small (around -10^{-20}). This result seems to suggest the difficulty of our estimator of the total-interaction index to better estimate null values of indices according to the precision required. The second-type GSIs, which are also equivalent to Sobol' indices in the case of single response mpdels, avoid obtaining negatives estimates of indices even for small or null values of indices by definition. Moreover, for the three functions considered in this paper, the second-type GSIs give the same ranking of input factors compared to the first-type GSIs.

6. Conclusion

In this paper, we propose and study a novel way of computing two types of generalized sensitivity indices, including Sobol' indices, using model derivatives, the cumulative distribution function, and the probability density function. First, we

derive the mathematical expressions of the total-effect and total-interaction effect functionals and covariance matrices based on the model derivatives. Second, we construct minimum variance unbiased estimators of the total-effect and total-interaction effect covariance matrices, and third, we provide estimators of the total and total-interaction GSIs as well as their consistency and asymptotic normality. Finally, we provide an algorithm for computing these covariance matrices and the GSIs, including the Sobol indices.

The numerical tests confirmed the accuracy of our estimates in general, except for some null indices. While a medium value of the sample size ($m = 1000$) gives interesting results for the Sobol and block additive functions, our estimators exhibit a numerical instability in the case of Ishigami's function. Our estimators of the first-type GSIs and Sobol' indoces show some difficulties for estimating null indices by providing negative estimates (around -10^{-20}). The second-type GSIs avoid obtaining negative estimates by definition, and they provide the same ranking of input factors compared to the first-type GSIs for the models considered in this paper. In the next future, it is interesting to i) improve the robustness of our estimators, ii) investigate a dimension-free algorithm for computing our indices using derivatives, iii) compare our estimates with those obtained by using derivatives-free algorithms (i.e., that do not use derivatives), iv) extend our approach to cope with dependent input factors.

Acknowledgments

We would like to thank the two referees for their comments and suggestions that have helped improving this paper. We also thank the Editor in Chief for considering our manuscript.

References

[1] M. Lamboni, Multivariate sensitivity analysis: minimum variance unbiased estimators of the first-order and total-effect covariance matrices, Reliability Engineering & System Safety (2018) – doi:https://doi.org/10.1016/j.ress.2018.06.004.

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

[3] M. Lamboni, D. Makowski, S. Lehuger, B. Gabrielle, H. Monod, Multivariate global sensitivity analysis for dynamic crop models, Fields Crop Reasearch 113 (2009) 312–320.

[4] F. Gamboa, A. Janon, T. Klein, A. Lagnoux, Sensitivity indices for multivariate outputs, Comptes Rendus de l'Académie des Sciences (2014) In press.

[5] S. Xiao, Z. Lu, L. Xu, Multivariate sensitivity analysis based on the direction of eigen space through principal component analysis, Reliability Engineering & System Safety 165 (2017) 1 – 10.

[6] I. M. Sobol, Sensitivity analysis for non-linear mathematical models, Mathematical Modelling and Computational Experiments 1 (1993) 407–414.

[7] A. Saltelli, K. Chan, E. Scott, Variance-Based Methods, Probability and Statistics, John Wiley and Sons, 2000.

[8] R. Ghanem, D. Higdon, H. Owhadi, Handbook of Uncertainty Quantification, Springer International Publishing, 2017.

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

[10] M. Lamboni, Uncertainty quantification: a minimun variance unbiased (joint) estimator of the non-normalized sobol indices, Stat Papers (2018) – doi:https://doi.org/10.1007/s00362-018-1010-4.

[11] O. Garcia-Cabrejo, A. Valocchi, Global sensitivity analysis for multivariate output using polynomial chaos expansion, Reliability Engineering & System Safety 126 (2014) 25 – 36.

[12] M. Lamboni, Global sensitivity analysis: a generalized, unbiased and optimal estimator of total-effect variance, Statistical Papers (2016) 1–26doi:10.1007/s00362-016-0768-5.
URL http://dx.doi.org/10.1007/s00362-016-0768-5

[13] M. Lamboni, Global sensitivity analysis: an efficient numerical method for approximating the total sensitivity index, International Journal for Uncertainty Quantification 6 (1) (2016) 1–17.

[14] E. Plischke, E. Borgonovo, C. L. Smith, Global sensitivity measures from given data, European Journal of Operational Research 226 (3) (2013) 536 – 550.

[15] E. Borgonovo, S. Tarantola, E. Plischke, M. D. Morris, Transformations and invariance in the sensitivity analysis of computer experiments, Journal of the Royal Statistical Society Series B 76 (5) (2014) 925–947.

[16] J. Fruth, O. Roustant, S. Kuhnt, Total interaction index: A variance-based sensitivity index for second-order interaction screening, Journal of Statistical Planning and Inference 147 (2014) 212 – 223.

[17] R. Liu, A. B. Owen, Estimating mean dimensionality of analysis of variance decompositions, Journal of the American Statistical Association 101 (474) (2006) 712–721.

[18] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, S. Tarantola, Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, Computer Physics Communications 181 (2) (2010) 259–270.

[19] A. Saltelli, Making best use of model evaluations to compute sensitivity indices, Computer Physics Communications 145 (2002) 280–297.

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

[21] A. B. Owen, Better estimation of small Sobol' sensitivity indices, ACM Trans. Model. Comput. Simul. 23 (2013) 111–1117.

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

[23] M. Morris, Factorial sampling plans for preliminary computational experiments, Technometrics 33 (1991) 161–174.

[24] I. M. Sobol, S. Kucherenko, Derivative based global sensitivity measures and the link with global sensitivity indices, Mathematics and Computers in Simulation 79 (2009) 3009–3017.

[25] S. Kucherenko, M. Rodriguez-Fernandez, C. Pantelides, N. Shah, Monte Carlo evaluation of derivative-based global sensitivity measures, Reliability Engineering and System Safety 94 (2009) 1135–1148.

[26] M. Lamboni, B. Iooss, A.-L. Popelin, F. Gamboa, Derivative-based global sensitivity measures: General links with Sobol' indices and numerical tests, Mathematics and Computers in Simulation 87 (0) (2013) 45 – 54.

[27] O. Roustant, J. Fruth, B. Iooss, S. Kuhnt, Crossed-derivative based sensitivity measures for interaction screening, Mathematics and Computers in Simulation 105 (2014) 105 – 118.

[28] S. Kucherenko, S. Song, Derivative-based global sensitivity measures and their link with sobol' sensitivity indices. in: Cools R., Nuyens D. (eds) Monte Carlo and Quasi-Monte Carlo methods, Springer Proceedings in Mathematics & Statistics 163. doi:https://doi.org/10.1007/978-3-319-33507-0_23.

[29] W. Hoeffding, A class of statistics with asymptotically normal distribution, Annals of Mathematical Statistics 19 (1948) 293–325.

[30] B. Efron, C. Stein, The jacknife estimate of variance, The Annals of Statistics 9 (1981) 586–596.

[31] R. Courant, Differential and integral calculus, Vol. II, Blackie and Son Limited, Great Britain, 1936.

[32] E. L. Lehmann, Consistency and unbiasedness of certain nonparametric tests, Annals of Mathematical Statistics 22 (1951) 165–179.

[33] T. S. Ferguson, A Course in Large Sample Theory, Chapman-Hall, New York, 1996.

[34] E. L. Lehmann, Elements of Large Sample Theory, Springer, 1999.

[35] W. Hoeffding, A non-parametric test for independence, Annals of Mathematical Statistics 19 (1948) 546–557.

[36] N. Sugiura, Multisample and multivariate nonparametric tests based on U-statistics and their asymptotic efficiencies, Osaka J. Math. 2 (2) (1965) 385–426. URL `http://projecteuclid.org/euclid.ojm/1200691466`

[37] C. Dutang, P. Savicky, randtoolbox: Generating and Testing Random Numbers, R package version 1.13 (2013).

AppendixA. Proof of Lemma 1

For point (i), as $v_1 \neq v_2$, there exists at least one element $j \in v_2$ and $j \notin v_1$. Bearing in mind the definition and properties of $\mathcal{A}_v$ (see Definition 1), we have: $\forall\, A_2 \in \mathcal{A}_{v_2}$, $j \in A_2$. As $j \in (u \backslash v_1)$, we can write $j \notin A_1$, $\forall\, A_1 \in \mathcal{A}_{v_1}$. Therefore, $\forall\, A_1 \in \mathcal{A}_{v_1}$, $A_1 \notin \mathcal{A}_{v_2}$.

For point (ii), it is obvious that $\cup_{v \subseteq u} \mathcal{A}_v \subseteq \{w,\, w \subseteq \{1,\, 2,\, \ldots,\, d\}\}$ by definition. To show that $\{w,\, w \subseteq \{1,\, 2,\, \ldots,\, d\}\} \subseteq \cup_{v \subseteq u} \mathcal{A}_v$, let us consider a set $w_0 \in \{w,\, w \subseteq \{1,\, 2,\, \ldots,\, d\}\}$. If $w_0 \subseteq u$ then there exists $v \subseteq u$ such that $v = w_0$, and therefore $w_0 \in \mathcal{A}_v$. If $w_0 \cap u = \emptyset$ then $w_0 \in \mathcal{A}_\emptyset$ ($v = \emptyset$). If $w_0 = w_1 \cup w_2$ with $w_1 \subseteq u$ and $w_2 \cap u = \emptyset$ then $w_0 \in \mathcal{A}_{w_1} \bigcup \mathcal{A}_\emptyset$.

Point (iii) is obvious using point (i).

AppendixB. Proof of Proposition 1

Point (i) is a consequence of Lemma 1 and the fact that the terms in Equation (2.1) are mutually orthogonal and centered.

Point (ii) is obvious using the set $\mathcal{B}_{v_1}$ from Lemma 1.

For Point (iii), the functional $\mathbf{f}_u^{tot}(\mathbf{X}) = \sum_{v,\, v \subseteq u \atop |v|>0} \mathbf{f}_{\mathcal{A}_v}(\mathbf{X}_v, \mathbf{X}_{\sim v})$ contains all information brought by $\mathbf{X}_u$ to the model, and it is sufficient to assess the overall contribution of input $\mathbf{X}_u$. We also have ([1])

$$\mathbf{f}_u^{tot}(\mathbf{X}) \;=\; \mathbf{f}(\mathbf{X}) - \mathbb{E}\left[\mathbf{f}(\mathbf{X}) \,|\, \mathbf{X}_{\sim u}\right]. \tag{B.1}$$

AppendixC. Proof of Lemma 2

Bearing in mind (2.7), we have

$$\begin{aligned}\sum_{v,\, v\subset u}(-1)^{|u|-|v|+1}\mathbf{f}^{sup}_{u\backslash v}(\mathbf{X}) &= \sum_{v,\, v\subset u}(-1)^{|u|-|v|+1}\sum_{w,\,(u\backslash v)\subseteq w\subseteq u}\mathbf{f}_{\mathcal{A}_w}(\mathbf{X})\\ &= \sum_{v,\, v\subset u}\sum_{w,\,(u\backslash v)\subseteq w\subseteq u}(-1)^{|u|-|v|+1}\mathbf{f}_{\mathcal{A}_w}(\mathbf{X}) \qquad \text{(C.1)}\end{aligned}$$

For $v_1 \subseteq u$ and $|v_1| > 0$, the sum $\sum_{w,\,(u\backslash v)\subseteq w\subseteq u}$ in Equation (C.1) contains v_1 for $\{v,\, (u\backslash v) \subseteq v_1\} \equiv \{v = u\backslash t,\ t \subseteq v_1 \text{ and } t \neq \emptyset\}$. Thus, Equation (C.1) becomes

$$\begin{aligned}\sum_{v,\, v\subset u}(-1)^{|u|-|v|+1}\mathbf{f}^{sup}_{u\backslash v}(\mathbf{X}) &= \sum_{\substack{v_1,\, v_1\subseteq u\\ |v_1|>0}}\sum_{\substack{t,\, t\subseteq v_1\\ |t|>0}}(-1)^{|t|+1}\mathbf{f}_{\mathcal{A}_{v_1}}(\mathbf{X})\\ &= \sum_{\substack{v_1,\, v_1\subseteq u\\ |v_1|>0}}\mathbf{f}_{\mathcal{A}_{v_1}}(\mathbf{X})\,,\end{aligned}$$

as $\sum_{|t|=1}^{|v_1|}(-1)^{|t|+1}\binom{|v_1|}{|t|} = -\sum_{|t|=1}^{|v_1|}(-1)^{|t|}\binom{|v_1|}{|t|} = 1$. The first result holds using Equation (2.8).

The last result in (2.10) is obtained using Equation (2.9). Indeed,

$$\begin{aligned}\sum_{v,\, v\subset u}(-1)^{|u|-|v|+1}\mathbf{f}^{tot}_{u\backslash v}(\mathbf{X}) &= \sum_{v,\, v\subset u}(-1)^{|u|-|v|+1}\sum_{v_1,\, v_1\subset(u\backslash v)}(-1)^{|u|-|v|-|v_1|+1}\mathbf{f}^{sup}_{(u\backslash v)\backslash v_1}(\mathbf{X})\\ &= \sum_{v,\, v\subset u}\sum_{v_1,\, v_1\subset(u\backslash v)}(-1)^{-|v_1|}\mathbf{f}^{sup}_{u\backslash(v\cup v_1)}(\mathbf{X})\,.\end{aligned}$$

For $w = \{v_1 \cup v\}$, we have $|w| = |v_1| + |v|$ as $v_1 \cap v = \emptyset$. We can see that

$\{(v, v_1),\, v \subset u,\, v_1 \subset (u \backslash v)\} \equiv \{(w, v_1),\, w \subset u,\, v_1 \subseteq w\}$. Thus, we have

$$\begin{aligned}
\sum_{v,\, v \subset u} (-1)^{|v|+|u|-1} \mathbf{f}_{u \backslash v}^{tot}(\mathbf{X}) &= \sum_{w,\, w \subset u} \sum_{v_1,\, v_1 \subseteq w} (-1)^{-|v_1|} \mathbf{f}_{u \backslash w}^{sup}(\mathbf{X}) \\
&= \sum_{w,\, w \subset u} \sum_{|v_1|=0}^{|w|} \binom{|w|}{|v_1|} (-1)^{-|v_1|} \mathbf{f}_{u \backslash w}^{sup}(\mathbf{X}) \\
&= \mathbf{f}_u^{sup}(\mathbf{X})\,,
\end{aligned}$$

as $\sum_{|v_1|=0}^{|w|} \binom{|w|}{|v_1|} (-1)^{-|v_1|} = 0$ if $|w| > 0$ and 1 otherwise.

AppendixD. Proof of Theorem 1

Let start with Point (i). Replacing $\mathbf{z}_u$ with $\mathbf{x}_u$ in Equation (3.23) yields to

$$\begin{aligned}
\mathbf{f}_u^{tot}(\mathbf{x}) &= \sum_{j \in u} \int_{\Omega_j} \int_0^1 \frac{\partial \mathbf{f}}{\partial x_j}(t_j x_j + y_j(1-t_j), \mathbf{x}_{\sim j}) \times (x_j - y_j) \times \rho_j(y_j)\, dt_j dy_j \\
&\quad + \sum_{\substack{v,\, v \subseteq u \\ |v|>1}} \int_{\Omega_v} \int_{[0,1]^{|v|}} \frac{\partial^{|v|} \mathbf{f}}{\partial \mathbf{x}_v} \left([t_j x_j + y_j(1-t_j)\, \forall\, j \in v],\, \mathbf{x}_{\sim v}\right) \\
&\quad \times \prod_{j \in v} (x_j - y_j) \times \rho_j(y_j)\, d\mathbf{t}_v d\mathbf{y}_v\,. \qquad \text{(D.1)}
\end{aligned}$$

For $j \in u$, consider the measurable function $r_j(y_j) = t_j x_j + y_j(1-t_j)$. We can see that $y_j = \frac{r_j - t_j x_j}{1-t_j}$, $dy_j = \frac{dr_j}{1-t_j}$, and $x_j - y_j = \frac{x_j - r_j}{1-t_j}$. By making a change of variables and bearing in mind Fubini-Lebesgue's theorem, we have

$$\begin{aligned}
\mathbf{f}_u^{tot}(\mathbf{x}) &= \sum_{j \in u} \int_{\Omega_j} \frac{\partial \mathbf{f}}{\partial x_j}(r_j, \mathbf{x}_{\sim j}) \left[\int_0^1 \frac{x_j - r_j}{(1-t_j)^2} \rho_j \left(\frac{r_j - t_j x_j}{1-t_j} \right) dt_j \right] dr_j \\
&\quad + \sum_{\substack{v,\, v \subseteq u \\ |v|>1}} \int_{\Omega_v} \frac{\partial^{|v|} \mathbf{f}}{\partial \mathbf{x}_v} \left([r_j\, \forall\, j \in v],\, \mathbf{x}_{\sim v}\right) \\
&\quad \times \left[\int_{[0,1]^{|v|}} \prod_{j \in v} \frac{x_j - r_j}{(1-t_j)^2} \rho_j \left(\frac{r_j - t_j x_j}{1-t_j} \right) d\mathbf{t}_v \right] d\mathbf{r}_v\,.
\end{aligned}$$

As the integral

$$\begin{aligned}\int_0^1 \frac{x_j - r_j}{(1-t_j)^2}\rho_j\left(\frac{r_j - t_j x_j}{1-t_j}\right) dt_j &= \left[-F_j\left(\frac{r_j - t_j x_j}{1-t_j}\right)\right]_0^1 \\ &= F_j(r_j) - \lim_{t_j \to 1} F_j\left(\frac{r_j - t_j x_j}{1-t_j}\right) \\ &= F_j(r_j) - \mathbb{1}_{[r_j > x_j]} - F_j(r_j)\mathbb{1}_{[r_j = x_j]}\,,\end{aligned}$$

the first result holds, that is,

$$\begin{aligned}\mathbf{f}_u^{tot}(\mathbf{x}) &= \sum_{j\in u}\int_{\Omega_j}\frac{\partial \mathbf{f}}{\partial x_j}(r_j, \mathbf{x}_{\sim j})\left[F_j(r_j)(1-\mathbb{1}_{[r_j=x_j]}) - \mathbb{1}_{[r_j>x_j]}\right] dr_j \\ &\quad + \sum_{\substack{v,\, v\subseteq u \\ |v|>1}}\int_{\Omega_v}\frac{\partial^{|v|}\mathbf{f}}{\partial \mathbf{x}_v}\left([r_j\,\forall\, j\in v],\, \mathbf{x}_{\sim v}\right)\prod_{j\in v}\left[F_j(r_j)(1-\mathbb{1}_{[r_j=x_j]}) - \mathbb{1}_{[r_j>x_j]}\right] d\mathbf{r}_v\,,\end{aligned}$$

or equivalently

$$\begin{aligned}\mathbf{f}_u^{tot}(\mathbf{X}) &= \sum_{j\in u}\mathbb{E}_{X_j^{(1)}}\left[\frac{\partial \mathbf{f}}{\partial X_j}(X_j^{(1)}, \mathbf{X}_{\sim j})\frac{F_j(X_j^{(1)})(1-\mathbb{1}_{[X_j^{(1)}=X_j]}) - \mathbb{1}_{[X_j^{(1)}>X_j]}}{\rho_j(X_j^{(1)})}\right] \\ &\quad + \sum_{\substack{v,\, v\subseteq u \\ |v|>1}}\mathbb{E}_{\mathbf{X}_v^{(1)}}\left[\frac{\partial^{|v|}\mathbf{f}}{\partial \mathbf{x}_v}\left(\mathbf{X}_v^{(1)}, \mathbf{X}_{\sim v}\right)\prod_{j\in v}\frac{F_j(X_j^{(1)})(1-\mathbb{1}_{[X_j^{(1)}=X_j]}) - \mathbb{1}_{[X_j^{(1)}>X_j]}}{\rho_j(X_j^{(1)})}\right].\end{aligned}$$

To derive the expression of TIEF based on model derivatives, we combine Equa-

tion (2.10) and Equation (3.24). Indeed,

$$\begin{aligned}
\mathbf{f}_u^{sup}(\mathbf{X}) &= \sum_{v,\, v\subset u} (-1)^{|u|-|v|+1} \sum_{\substack{v_1,\, v_1\subseteq (u\setminus v)\\ |v_1|>0}} \mathbb{E}_{\mathbf{X}_{v_1}^{(1)}} \left[\frac{\partial^{|v_1|}\mathbf{f}}{\partial \mathbf{x}_{v_1}} \left(\mathbf{X}_{v_1}^{(1)}, \mathbf{X}_{\sim v_1}\right) \right. \\
&\qquad \left. \times \prod_{j\in v_1} \frac{F_j(X_j^{(1)})(1-\mathbb{1}_{[X_j^{(1)}=X_j]}) - \mathbb{1}_{[X_j^{(1)}>X_j]}}{\rho_j(X_j^{(1)})} \right] \\
&= \sum_{v,\, v\subset u} \sum_{\substack{v_1,\, v_1\subseteq (u\setminus v)\\ |v_1|>0}} (-1)^{|u|-|v|+1} \mathbb{E}_{\mathbf{X}_{v_1}^{(1)}} \left[\frac{\partial^{|v_1|}\mathbf{f}}{\partial \mathbf{x}_{v_1}} \left(\mathbf{X}_{v_1}^{(1)}, \mathbf{X}_{\sim v_1}\right) \right. \\
&\qquad \left. \times \prod_{j\in v_1} \frac{F_j(X_j^{(1)})(1-\mathbb{1}_{[X_j^{(1)}=X_j]}) - \mathbb{1}_{[X_j^{(1)}>X_j]}}{\rho_j(X_j^{(1)})} \right] .
\end{aligned}$$

For a given $v_1 \subseteq u$ with $v_1 \neq \emptyset$, we can see that $\{v,\, v_1 \subseteq (u\setminus v)\} \equiv \{v = u\setminus t,\, v_1 \subseteq t \subseteq u\}$.

Thus, we have

$$\begin{aligned}
\mathbf{f}_u^{sup}(\mathbf{X}) &= \sum_{\substack{v_1,\, v_1\subseteq u\\ |v_1|>0}} \sum_{t,\, v_1\subseteq t\subseteq u} (-1)^{|t|+1} \mathbb{E}_{\mathbf{X}_{v_1}^{(1)}} \left[\frac{\partial^{|v_1|}\mathbf{f}}{\partial \mathbf{x}_{v_1}} \left(\mathbf{X}_{v_1}^{(1)}, \mathbf{X}_{\sim v_1}\right) \right. \\
&\qquad \left. \times \prod_{j\in v_1} \frac{F_j(X_j^{(1)})(1-\mathbb{1}_{[X_j^{(1)}=X_j]}) - \mathbb{1}_{[X_j^{(1)}>X_j]}}{\rho_j(X_j^{(1)})} \right] \\
&= (-1)^{|u|+1} \mathbb{E}_{\mathbf{X}_u^{(1)}} \left[\frac{\partial^{|v_1|}\mathbf{f}}{\partial \mathbf{x}_u} \left(\mathbf{X}_u^{(1)}, \mathbf{X}_{\sim u}\right) \prod_{j\in u} \frac{F_j(X_j^{(1)})(1-\mathbb{1}_{[X_j^{(1)}=X_j]}) - \mathbb{1}_{[X_j^{(1)}>X_j]}}{\rho_j(X_j^{(1)})} \right] ,
\end{aligned}$$

due to the fact that

$$\begin{aligned}
\sum_{t,\, v_1\subseteq t\subseteq u} (-1)^{|t|+1} &= \sum_{|t|=|v_1|}^{|u|} \binom{|u|-|v_1|}{|t|-|v_1|} (-1)^{|t|+1} = \sum_{i=0}^{|u|-|v_1|} \binom{|u|-|v_1|}{i} (-1)^{i+|v_1|+1} \\
&= (-1)^{|u|+1} \mathbb{1}_{[v_1=u]} .
\end{aligned}$$

Point (ii) is a particular case of Equation (3.24). Indeed, if we choose $u = \{1, \ldots, d\}$, we have $\mathbf{f}_u^{tot}(\mathbf{X}) = \mathbf{f}(\mathbf{X}) - \mathbb{E}[\mathbf{f}(\mathbf{X})]$ (see Equation (B.1)).

Point (iii) is obvious.

AppendixE. Proof of Theorem 2

First, we establish some useful equalities for the proofs, and second, we derive the two results of Theorem 2. We know that $\mathbf{d}(\mathbf{X}^{(1)}, \mathbf{X}_v^{(3)})$ is centered, that is, $\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(1)}, \mathbf{X}_v^{(3)})\right] = \mathbf{0}$ (see Remark 2).
Bearing in mind that $\mathbf{X}^{(j_1)}$ and $\mathbf{X}_u^{(j_2)}$ are independent for $j_1 = 1, 2$ and $j_2 = 3, 4$, $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ (resp. $\mathbf{X}_u^{(3)}$ and $\mathbf{X}_u^{(4)}$) have the same distribution, we have the following equalities.

$$\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_v^{(3)}) \,|\, \mathbf{X}^{(j_1)}\right] = \mathbb{E}_{\mathbf{X}_v^{(3)}}\left[\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_v^{(3)})\right] = \mathbb{E}_{\mathbf{X}_v^{(4)}}\left[\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_v^{(4)})\right] , \quad \text{(E.1)}$$

where $\mathbb{E}_{\mathbf{X}_v^{(j_2)}}\left[\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_v^{(j_2)})\right]$ means that the expectation is taken with respect to $\mathbf{X}_u^{(j_2)}$ with $j_1 \in \{1, 2\}$ and $j_2 \in \{3, 4\}$.

According to Equation (3.25) and using Equation (E.1), the total-interaction covariance matrix is given by

$$\Sigma_u^{sup} = \mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_u^{(3)})\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_u^{(4)})^T\right] , \quad \text{(E.2)}$$

with $j_1 \in \{1, 2\}$. Indeed,

$$\begin{aligned}
\Sigma_u^{sup} &= \mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_u^{(3)})\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_u^{(4)})^T\right] \\
&= \mathbb{E}\left(\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_u^{(3)})\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_u^{(4)})^T \,|\, \mathbf{X}^{(j_1)}\right]\right) \\
&= \mathbb{E}\left(\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_u^{(3)}) \,|\, \mathbf{X}^{(j_1)}\right]\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_u^{(4)})^T \,|\, \mathbf{X}^{(j_1)}\right]\right) \\
&= \mathbb{E}\left(\mathbb{E}_{\mathbf{X}_u^{(3)}}\left[\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_u^{(3)})\right]\mathbb{E}_{\mathbf{X}_u^{(4)}}\left[\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_u^{(4)})^T\right]\right) \\
&= \mathbb{E}\left(\mathbb{E}_{\mathbf{X}_u^{(3)}}\left[\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_u^{(3)})\right]\mathbb{E}_{\mathbf{X}_u^{(3)}}\left[\mathbf{d}(\mathbf{X}^{(j_1)}, \mathbf{X}_u^{(3)})^T\right]\right) \\
&= \mathbb{V}\left[\mathbf{f}_u^{sup}(\mathbf{X}^{(j_1)})\right]
\end{aligned}$$

To prove Equation (4.32), it is sufficient to show that

$$\begin{aligned} A &= \frac{1}{4}\mathbb{E}\left[\left[\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_u^{(3)})-\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_u^{(3)})\right]\times\left[\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_u^{(4)})-\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_u^{(4)})\right]^T\right] \\ &= \frac{1}{2}\Sigma_u^{(sup)}, \end{aligned}$$

bearing in mind the symmetry of Equation (4.30). Indeed,

$$\begin{aligned} A &= \frac{1}{4}\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_u^{(3)})\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_u^{(4)})^T\right]-\frac{1}{4}\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_u^{(3)})\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_u^{(4)})^T\right] \\ &\quad -\frac{1}{4}\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_u^{(3)})\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_u^{(4)})^T\right]+\frac{1}{4}\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_u^{(3)})\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_u^{(4)})^T\right] \\ &= \frac{1}{2}\Sigma_u^{sup}, \end{aligned}$$

as $\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_u^{(3)})\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_u^{(4)})^T\right]=\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_u^{(3)})\right]\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_u^{(4)})^T\right]=\mathbf{0}$ and $\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_u^{(3)})\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_u^{(4)})^T\right]=0$.

The proof of the second results (4.33) is similar to the proof of Equation (4.32). To show (4.33), it is sufficient to prove that

$$\begin{aligned} B &= \frac{1}{4}\mathbb{E}\left(\sum_{\substack{v\subseteq u\\|v|=1}}^{|u|}\sum_{\substack{v_1\subseteq u\\|v_1|=1}}^{|u|}\left[\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_v^{(3)})-\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_v^{(3)})\right]\left[\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_{v_1}^{(4)})-\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_{v_1}^{(4)})\right]^T\right) \\ &= \frac{1}{2}\Sigma_u^{tot}. \end{aligned}$$

Indeed, as $\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_v^{(3)})\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_{v_1}^{(4)})^T\right]=\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_v^{(3)})\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_{v_1}^{(4)})^T\right]=0$, $\mathbf{f}_u^{tot}(\mathbf{X}^{(1)})=\sum_{\substack{v\subseteq u\\|v|=1}}^{|u|}\mathbb{E}_{\mathbf{X}_v^{(3)}}\left[\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_v^{(3)})\right]$, and $\mathbf{f}_u^{tot}(\mathbf{X}^{(2)})=\sum_{\substack{v_1\subseteq u\\|v_1|=1}}^{|u|}\mathbb{E}_{\mathbf{X}_{v_1}^{(3)}}\left[\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_{v_1}^{(3)})\right]$,

we have

$$
\begin{aligned}
B &= \frac{1}{4}\sum_{\substack{v\subseteq u\\|v|=1}}^{|u|}\sum_{\substack{v_1\subseteq u\\|v_1|=1}}^{|u|}\mathbb{E}\left(\left[\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_v^{(3)})-\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_v^{(3)})\right]\left[\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_{v_1}^{(4)})-\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_{v_1}^{(4)})\right]^T\right)\\
&= \frac{1}{4}\sum_{\substack{v\subseteq u\\|v|=1}}^{|u|}\sum_{\substack{v_1\subseteq u\\|v_1|=1}}^{|u|}\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_v^{(3)})\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_{v_1}^{(4)})^T\right]\\
&\quad+\frac{1}{4}\sum_{\substack{v\subseteq u\\|v|=1}}^{|u|}\sum_{\substack{v_1\subseteq u\\|v_1|=1}}^{|u|}\mathbb{E}\left[\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_v^{(3)}\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_{v_1}^{(4)})^T\right]\\
&= \frac{1}{4}\sum_{\substack{v\subseteq u\\|v|=1}}^{|u|}\sum_{\substack{v_1\subseteq u\\|v_1|=1}}^{|u|}\mathbb{E}\left[\mathbb{E}_{\mathbf{X}_v^{(3)}}\left(\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_v^{(3)})\right)\mathbb{E}_{\mathbf{X}_{v_1}^{(4)}}\left(\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_{v_1}^{(4)})^T\right)\right]\\
&\quad+\frac{1}{4}\sum_{\substack{v\subseteq u\\|v|=1}}^{|u|}\sum_{\substack{v_1\subseteq u\\|v_1|=1}}^{|u|}\mathbb{E}\left[\mathbb{E}_{\mathbf{X}_v^{(3)}}\left(\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_v^{(3)})\right)\mathbb{E}_{\mathbf{X}_{v_1}^{(4)}}\left(\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_{v_1}^{(4)})^T\right)\right]\\
&= \frac{1}{4}\mathbb{E}\left[\sum_{\substack{v\subseteq u\\|v|=1}}^{|u|}\mathbb{E}_{\mathbf{X}_v^{(3)}}\left(\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_v^{(3)})\right)\sum_{\substack{v_1\subseteq u\\|v_1|=1}}^{|u|}\mathbb{E}_{\mathbf{X}_{v_1}^{(3)}}\left(\mathbf{d}(\mathbf{X}^{(1)},\mathbf{X}_{v_1}^{(3)})^T\right)\right]\\
&\quad+\frac{1}{4}\mathbb{E}\left[\sum_{\substack{v\subseteq u\\|v|=1}}^{|u|}\mathbb{E}_{\mathbf{X}_v^{(3)}}\left(\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_v^{(3)})\right)\sum_{\substack{v_1\subseteq u\\|v_1|=1}}^{|u|}\mathbb{E}_{\mathbf{X}_{v_1}^{(3)}}\left(\mathbf{d}(\mathbf{X}^{(2)},\mathbf{X}_{v_1}^{(3)})^T\right)\right]\\
&= \frac{1}{2}\Sigma_u^{tot}.
\end{aligned}
$$

AppendixF. Proof of Theorem 3

Knowing that $\Sigma_u^{sup}=\mathbb{E}\left[\mathbf{K}^{sup}(\mathcal{X},\mathcal{Y})\right]$, the corresponding U-statistic of two sample associated with the *kernel* $\mathbf{K}^{sup}(\mathcal{X},\mathcal{Y})$ is $\widehat{\Sigma_u^{sup}}$.
It is obvious that Vec $\left[\widehat{\Sigma_u^{sup}}\right]$ is also an unbiased estimator of Vec$[\Sigma_u^{sup}]$, and it is symmetric w.r.t the two types of inputs. Therefore, it follows from the theory of U-statistics that Vec $\left[\widehat{\Sigma_u^{sup}}\right]$ has a minimum variance for the class of unbiased estimators that make use of the kernel $\mathbf{K}^{sup}(\mathcal{X},\mathcal{Y})$ and have finite 4^{th} moment (see [1; 10] for more details). We also have the following results:

(a) unbiasedness

$$\mathbb{E}\left[\text{Vec}\left[\widehat{\Sigma_u^{sup}}\right]\right] = \text{Vec}[\Sigma_u^{sup}]\,;$$

(b) consistency

$$\text{Vec}\left[\widehat{\Sigma_u^{sup}}\right] \xrightarrow{\mathcal{P}} \text{Vec}[\Sigma_u^{sup}]\,;$$

(c) asymptotic normality

$$\sqrt{m}\left(\text{Vec}\left[\widehat{\Sigma_u^{sup}}\right] - \text{Vec}[\Sigma_u^{sup}]\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\mathbf{0}, \mathbb{V}[\text{Vec}(\mathbf{K}^{sup})]\right)\,.$$

Points (a), (b), and (d) are equivalent to Points (i), (ii), and (iii) of Theorem 3.

AppendixG. Proof of Theorem 5

The results about the consistency are obvious bearing in mind Corollaries 1-2 and the Slutsky theorem.

For the asymptotic normality, first, bearing in mind Slutsky's theorem and using Corollary 1, we can write

$$\sqrt{m}\left(\frac{\widehat{\text{Tr}\left(\Sigma_u^{sup}\right)}}{\text{Tr}\left(\widehat{\Sigma}\right)} - \frac{\text{Tr}\left(\Sigma_u^{sup}\right)}{\text{Tr}\left(\widehat{\Sigma}\right)}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\mathbb{V}(\text{Tr}[\mathbf{K}^{sup}(\mathcal{X},\mathcal{Y})])}{[\text{Tr}(\Sigma)]^2}\right)\,.$$

Second, we need to show that $\sqrt{m}\left(\frac{\widehat{\text{Tr}\left(\Sigma_u^{sup}\right)}}{\text{Tr}\left(\widehat{\Sigma}\right)} - \frac{\text{Tr}\left(\Sigma_u^{sup}\right)}{\text{Tr}\left(\widehat{\Sigma}\right)}\right)$ and $\sqrt{m}\left(\frac{\widehat{\text{Tr}\left(\Sigma_u^{sup}\right)}}{\text{Tr}\left(\widehat{\Sigma}\right)} - \frac{\text{Tr}\left(\Sigma_u^{sup}\right)}{\text{Tr}(\Sigma)}\right)$ are asymptotically equivalent and therefore have the same asymptotic distribution. Indeed,

$$\sqrt{m}\left(\frac{\widehat{\text{Tr}\left(\Sigma_u^{sup}\right)}}{\text{Tr}\left(\widehat{\Sigma}\right)} - \frac{\text{Tr}\left(\Sigma_u^{sup}\right)}{\text{Tr}\left(\widehat{\Sigma}\right)}\right) - \sqrt{m}\left(\frac{\widehat{\text{Tr}\left(\Sigma_u^{sup}\right)}}{\text{Tr}\left(\widehat{\Sigma}\right)} - \frac{\text{Tr}\left(\Sigma_u^{sup}\right)}{\text{Tr}(\Sigma)}\right) = \sqrt{m}\text{Tr}\left(\Sigma_u^{sup}\right)\left(\frac{1}{\text{Tr}(\Sigma)} - \frac{1}{\text{Tr}\left(\widehat{\Sigma}\right)}\right) \to 0\,,$$

because, the probability

$$\begin{aligned}\mathbb{P}\left[\sqrt{m}\left|\frac{1}{\widehat{\mathrm{Tr}\left(\Sigma\right)}}-\frac{1}{\mathrm{Tr}\left(\Sigma\right)}\right|\geq\epsilon\right] &\leq \frac{m}{\epsilon^2}\mathbb{E}\left[\left(\frac{1}{\widehat{\mathrm{Tr}\left(\Sigma\right)}}-\frac{1}{\mathrm{Tr}\left(\Sigma\right)}\right)^2\right]\\ &\leq \frac{m}{\epsilon^2}\frac{C}{M}\xrightarrow{m,M\to+\infty}0\,,\end{aligned}$$

as $\sqrt{M}\left(\frac{1}{\widehat{\mathrm{Tr}(\Sigma)}}-\frac{1}{\mathrm{Tr}(\Sigma)}\right)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,C\right)$, which is obtained by applying the delta method to $\sqrt{M}\left(\widehat{\mathrm{Tr}\left(\Sigma\right)}-\mathrm{Tr}\left(\Sigma\right)\right)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,C_0\right)$.