Global Sensitivity Analysis: a novel generation of mighty estimators based on rank statistics

Fabrice Gamboa, Pierre Gremaud, Thierry Klein, Agnès Lagnoux

Global Sensitivity Analysis: a novel generation of mighty estimators based on rank statistics

Fabrice Gamboa[1], Pierre Gremaud[2], Thierry Klein[3], and Agnès Lagnoux[4]

[1]Institut de Mathématiques de Toulouse and ANITI; UMR5219. Université de Toulouse; CNRS. UT3, F-31062 Toulouse, France.
[2]Department of Mathematics. NC State University. Raleigh, North Carolina 27695, USA.
[3]Institut de Mathématiques de Toulouse; UMR5219. Université de Toulouse; ENAC - Ecole Nationale de l'Aviation Civile , Université de Toulouse, France
[4]Institut de Mathématiques de Toulouse; UMR5219. Université de Toulouse; CNRS. UT2J, F-31058 Toulouse, France.

November 6, 2020

Abstract

We propose a new statistical estimation framework for a large family of global sensitivity analysis indices. Our approach is based on rank statistics and uses an empirical correlation coefficient recently introduced by Chatterjee [9]. We show how to apply this approach to compute not only the Cramér-von-Mises indices, which are directly related to Chatterjee's notion of correlation, but also first-order Sobol indices, general metric space indices and higher-order moment indices. We establish consistency of the resulting estimators and demonstrate their numerical efficiency, especially for small sample sizes. In addition, we prove a central limit theorem for the estimators of the first-order Sobol indices.

Key words Global sensitivity analysis, Cramér-von-Mises distance, Pick-Freeze method, Chatterjee's coefficient of correlation, Sobol indices estimation.

AMS subject classification 62G05, 62G20, 62G30.

1 Introduction

The use of complex computer models for the analysis of applications from the sciences, engineering and other fields is by now routine. Often, the models are expensive to run in terms of computational time. It is thus crucial to understand, with just a few runs, the global influence of one or several inputs on the output of the system under study [33].

When these inputs are regarded as random elements, this problem is generally referred to as Global Sensitivity Analysis (GSA). We refer to [11, 32, 35] for an overview of the practical aspects of GSA.

A popular and highly useful tool to quantify input influence is the Sobol indices. These indices were first introduced in [36] and are well tailored to the case of scalar outputs. Thanks to the Hoeffding decomposition [21], the Sobol indices compare the conditional variance of the output knowing some of the input variables to the total variance of the output. Many different estimation procedures of the Sobol indices have been proposed and studied. Some are based on Monte-Carlo or quasi Monte-Carlo design of experiments (see [23, 27] and references therein for more details). In particular, an efficient estimation of the Sobol indices can be performed through the so-called Pick-Freeze method. For the description of this method and its theoretical study (consistency, Central Limit Theorem (CLT), concentration inequalities and Berry-Esseen bounds), we refer to [22, 16] and references therein. Some other estimation procedures are based on different designs of experiment using for example polynomial chaos (see [37] and the reference therein for more details).

Various generalizations of the Sobol indices have been developed. The issue of vectorial outputs, as in the case with time dependent or functional quantities of interest, is addressed in [1, 15, 24]. In particular, in [15], the authors recover the indices from [24] and show that they are a proper generalization of the classical Sobol indices in higher dimension. Moreover, they provide the theoretical study of their Pick-Freeze estimators and extend their definitions to the case of outputs valued in a separable Hilbert space.

Since Sobol indices are variance based, they only quantify the second-order influence of the inputs. Many authors proposed other criteria to compare the conditional distribution of the output knowing some of the inputs to the distribution of the output. In [27, 29, 28], the authors use higher moments to define new indices while, in [4, 5, 10], the use of divergences or distances between measures allows to define new indices. In [13], contrast functions are exploited to build indices that are goal oriented. Although these works define nice theoretical indices, the existence of a relevant statistical estimation procedure is still in most cases an open question. The case of vectorial-valued computer codes is considered in [17] where a sensitivity index based on the whole distribution is defined. Within this framework, the authors show that the Pick-Freeze estimation procedure provides an asymptotically Gaussian estimator of the index. The cost of an estimator naturally depends on the cost of each evaluation of the code and on the number of evaluations. The Pick-Freeze scheme requires $3N$ evaluations of the output code for the evaluation of a single index and leads to a convergence rate $\sqrt{N}$. Hence, if the number of input variables is p, the total number of calls of the code is $(p+3)N$ that grows linearly with p. This approach has been generalized in [14], where the authors considered computer codes valued on a compact Riemannian manifold. They use the Pick-Freeze scheme to provide a consistent estimator requiring $4N$ evaluations of the output code. The authors of [19] extend the previous indices to general metric spaces and propose U-statistics-based estimators improving the classical Pick-Freeze procedure.

We emphasize that the Pick-Freeze estimation procedure allows the estimation of several sensitivity indices: the classical Sobol indices for real-valued outputs, as well as their generalization for vectorial-valued codes, but also the indices based on higher moments [29] and the Cramér-von-Mises indices which take into account on the whole distribution [17, 14]. In addition, the Pick-Freeze estimators have desirable statistical properties such

as consistency, fixed rate of convergence and exponential inequalities. They have, however, two major drawbacks. First, they rely on a particular experimental design that may be unavailable in practice. Second, the number of model calls to estimate all first-order Sobol indices grows linearly with the number of input parameters. For example, if we consider $p = 99$ input parameters and only $n = 1000$ calls are allowed, then only a sample of size $n/(p+1) = 10$ is available to estimate each single first-order Sobol index.
In a recent work [9], Chatterjee studies the dependence between two variables by introducing an empirical correlation coefficient based on rank statistics, see Section 2.3 below for the precise definition. Further, the quantification of the dependence has also been investigated in the bivariate case (namely, in the copula setting), see [38, 12, 3]. The striking point of [9] is that this empirical correlation coefficient converges almost surely to the Cramér-von-Mises index introduced in [17] as the sample size goes to infinity. In this paper, we show how to embed Chatterjee's method in the GSA framework, thereby eliminating the two drawbacks of the classical Pick-Freeze estimation mentioned above. In addition, we generalize Chatterjee's approach to allow the estimation of a large class of GSA indices which includes the Sobol indices and the higher-order moment indices proposed by Owen [27, 29, 28]. Using a single sample of size n, it is now possible to estimate at the same time all the first-order Sobol indices, the Cramér-von-Mises indices, and other useful sensitivity indices. Furthermore, we show here that this new procedure provides estimators also converging at rate $\sqrt{n}$ by proving a CLT.
The paper is organized as follows. In Section 2, we recall the definition of the Cramér-von-Mises indices and their classical Pick-Freeze estimation. Further, we show how they can be also estimated using Chatterjee's method. In Section 3, we present the generalization of Chatterjee's method to estimate sensitivity indices together with the consistency of the estimation procedure. In addition, we recover the first-order Sobol indices and prove the asymptotic normality of their estimators. Section 4 considers other classical sensitivity indices while Section 5 is dedicated to a numerical comparison between the Pick-Freeze estimation procedure and Chatterjee's method. We first compare the numerical performances of both estimators on a linear model. Finally, we consider a real life application. As expected, Chatterjee's estimation method outperforms the classical Pick-Freeze procedure, even for small sample sizes (which are common in practice). Conclusions and perspectives are offered in Section 6.
After a first submission of this paper, we have been aware of the very nice work of Broto *et al* ([8]) concerning the statistical estimation of Shapley effect where the use of closest neighbors is also put in action to built consistent estimates.

2 Sensitivity analysis based on Cramér-von-Mises indices

2.1 Definition of Sobol and Cramér-von-Mises indices

The quantity of interest (QoI) Y is obtained from the numerical code and is regarded as a function f of the vector of the distributed input $(X_i)_{i=1,\cdots,p}$

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

where f is defined on the state space $E_1 \times \ldots \times E_p$, $X_i \in E_i$, $i = 1, \ldots, p$. Classically, the X_i's are assumed to be independent random variables and a sensitivity analysis is performed using the Hoeffding decomposition [2, 39] leading to the standard Sobol indices [35]. This assumption is made throughout the paper, unless explicitly stated otherwise. More precisely, assume f to be real-valued and square integrable and let $\mathbf{u}$ be a subset of $\{1, \ldots, p\}$ and $\sim\mathbf{u}$ its complementary set in $\{1, \ldots, p\}$. Setting $X_{\mathbf{u}} = (X_i, i \in \mathbf{u})$ and $X_{\sim\mathbf{u}} = (X_i, i \in \sim\mathbf{u})$, the corresponding Sobol indices take the form

$$S^{\mathbf{u}} = \frac{\mathrm{Var}\left(\mathbb{E}[Y|X_{\mathbf{u}}]\right)}{\mathrm{Var}(Y)} \quad \text{and} \quad S^{\sim\mathbf{u}} = \frac{\mathrm{Var}\left(\mathbb{E}[Y|X_{\sim\mathbf{u}}]\right)}{\mathrm{Var}(Y)}. \tag{2}$$

By definition, the Sobol indices quantify the fluctuations of the output Y around its mean. When the practitioner is not interested in the mean behavior of Y but rather in its median, in its tail, or even in its quantiles, the Sobol indices become less appropriate to quantify sensitivity. GSA must then be performed in a framework which takes into account more than one specific moment, such as the variance for Sobol indices. The Cramér-von-Mises indices introduced in [17] provide alternative indices based on the whole distribution. They are defined by

$$S^{\mathbf{u}}_{2,CVM} = \frac{\int_{\mathbb{R}} \mathbb{E}\left[\left(F(t) - F^{\mathbf{u}}(t)\right)^2\right] dF(t)}{\int_{\mathbb{R}} F(t)(1 - F(t))dF(t)} \tag{3}$$

where F is the cumulative distribution function of Y

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

and $F^{\mathbf{u}}$ is its Pick-Freeze version, namely the conditional distribution function of Y conditionally on $X_{\mathbf{u}}$:

$$F^{\mathbf{u}}(t) = \mathbb{P}\left(Y \leqslant t|X_{\mathbf{u}}\right) = \mathbb{E}\Big[\mathbb{1}_{\{Y \leqslant t\}}|X_{\mathbf{u}}\Big] \quad (t \in \mathbb{R}).$$

Such a definition stems from the Hoeffding decomposition of the collection of the indicator random variables $(\mathbb{1}_{\{Y \leqslant t\}})_{t \in \mathbb{R}}$. It is worth noting that this definition naturally extends to multivariate outputs.

2.2 Classical estimation of Cramér-von-Mises indices using the Pick-Freeze method

The estimation of the Cramér-von-Mises index (3) reduces to the estimation of both its numerator and its denominator. The numerator of $S^{\mathbf{u}}_{2,CVM}$ can be rewritten as

$$\begin{aligned}\int_{\mathbb{R}} \mathbb{E}\left[\left(F(t) - F^{\mathbf{u}}(t)\right)^2\right] dF(t) &= \mathbb{E}\left[\mathbb{E}\left[\left(F(Y') - F^{\mathbf{u}}(Y')\right)^2 |Y'\right]\right] \\ &= \mathbb{E}\left[\mathrm{Var}\left(\mathbb{E}\left[\mathbb{1}_{\{Y \leqslant Y'\}}|X_{\mathbf{u}}, Y'\right] |Y'\right)\right]\end{aligned}$$

where Y' is an independent copy of Y. A Monte-Carlo scheme can be used to estimate the Cramér-von-Mises indices. The corresponding Pick-Freeze approach from [16, 17, 22] relies on expressing the variances of the conditional expectations in terms of covariances

which are easily and well estimated by their empirical versions. To that end, we define, for any subset $\mathbf{u}$ of $\{1,\dots,p\}$

$$Y^{\mathbf{u}} := f(X^{\mathbf{u}}). \tag{4}$$

where $X^{\mathbf{u}}$ is such that $X^{\mathbf{u}}_{\mathbf{u}} = X_{\mathbf{u}}$ and $X^{\mathbf{u}}_i = X'_i$ if $i \in\sim \mathbf{u}$, X'_i being an independent copy of X_i. The estimation procedure relies on the following lemma which is still valid for any function $g \in L^2$ (not only $g(y) = \mathbb{1}_{\{y\leqslant t\}}$).

Lemma 2.1.

$$\mathrm{Var}(\mathbb{E}[\mathbb{1}_{\{Y\leqslant t\}}|X_{\boldsymbol{u}}]) = \mathrm{Cov}(\mathbb{1}_{\{Y\leqslant t\}}, \mathbb{1}_{\{Y^{\mathbf{u}}\leqslant t\}}). \tag{5}$$

Proof. Let $Z = \mathbb{1}_{\{Y\leqslant t\}}$ and $Z^{\mathbf{u}} = \mathbb{1}_{\{Y^{\mathbf{u}}\leqslant t\}}$. Since, Z and $Z^{\mathbf{u}}$ share the same distribution and are independent conditionally to $X_{\mathbf{u}}$, we have

$$\begin{aligned}
\mathrm{Var}(\mathbb{E}[Z|X_{\mathbf{u}}]) &= \mathbb{E}[\mathbb{E}[Z|X_{\mathbf{u}}]^2] - \mathbb{E}[\mathbb{E}[Z|X_{\mathbf{u}}]]^2 \\
&= \mathbb{E}[\mathbb{E}[Z|X_{\mathbf{u}}]\mathbb{E}[Z^{\mathbf{u}}|X_{\mathbf{u}}]] - \mathbb{E}[\mathbb{E}[Z|X_{\mathbf{u}}]]\mathbb{E}[\mathbb{E}[Z^{\mathbf{u}}|X_{\mathbf{u}}]] \\
&= \mathbb{E}[\mathbb{E}[ZZ^{\mathbf{u}}|X_{\mathbf{u}}]] - \mathbb{E}[Z]\mathbb{E}[Z^{\mathbf{u}}] \\
&= \mathbb{E}[ZZ^{\mathbf{u}}] - \mathbb{E}[Z]\mathbb{E}[Z^{\mathbf{u}}] \\
&= \mathrm{Cov}(Z, Z^{\mathbf{u}}).
\end{aligned}$$

□

Consequently, the Monte-Carlo estimation can be done as follows. An n-sample $(Y_1,\dots,Y_n)$ of the output Y and an n-sample $(Y^{\mathbf{u}}_1,\dots,Y^{\mathbf{u}}_n)$ of its Pick-Freeze version $Y^{\mathbf{u}}$ are required. In addition, in order to deal with the integral with respect to $dF(t)$ in (3), a third independent n sample $(W_1,\dots,W_n)$ of the output Y is necessary. Then the empirical estimator of $S^1_{2,CVM}$ is

$$\frac{\frac{1}{n}\sum_{k=1}^n \left(\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{Y_j\leqslant W_k\}}\mathbb{1}_{\{Y_j^{\mathbf{u}}\leqslant W_k\}} - \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{Y_j\leqslant W_k\}}\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{Y_j^{\mathbf{u}}\leqslant W_k\}}\right)}{\frac{1}{n}\sum_{k=1}^n \left(\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{Y_j\leqslant W_k\}} - \left(\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{Y_j\leqslant W_k\}}\right)^2\right)}. \tag{6}$$

As showed in [17], this estimator is consistent and asymptotically Gaussian (i.e. the rate of convergence is $\sqrt{n}$). The limiting variance can be computed explicitly, allowing the practitioner to build confidence intervals. In particular, if one wants to estimate all the first-order indices (that is the p first-order Sobol indices) and the p Cramér-von-Mises indices, $(p+2)n$ calls of the computer code are required. The number of calls grows linearly with respect to the number of input parameters. This is a practical issue for large input dimension domains. A second drawback of this estimation scheme comes from the need of the particular Pick-Freeze design that is not always available.

2.3 Chatterjee's method

In [9], Chatterjee considers a pair of real-valued random variables (V, Y) and an i.i.d. sample $(V_j, Y_j)_{1\leqslant j\leqslant n}$. In order to simplify the presentation, we assume that the laws of V and Y are both diffuse (ties are excluded). The pairs $(V_{(1)}, Y_{(1)}),\dots,(V_{(n)}, Y_{(n)})$ are rearranged in such a way that

$$V_{(1)} < \ldots < V_{(n)}.$$

Let r_j be the rank of $Y_{(j)}$, that is,

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

The new correlation coefficient defined by Chatterjee in [9] is

$$\xi_n(V, Y) := 1 - \frac{3 \sum_{j=1}^{n-1} |r_{j+1} - r_j|}{n^2 - 1}. \tag{7}$$

The author proves that $\xi_n(V, Y)$ converges almost surely to a deterministic limit $\xi(V, Y)$ which is equal to the Cramér-von-Mises sensitivity index $S^V_{2,CVM}$ with respect to V as soon as V is one of the random variables X_1, ..., X_p in the model (1) that are assumed to be real-valued. Further, he also proves a CLT when V and Y are independent.
Chatterjee also provides a rank statistics analogue to Lemma 2.1. More precisely, let $\pi(j)$ be the rank of V_j in the sample $(V_1, \dots, V_n)$ of V and define

$$N(j) = \begin{cases} \pi^{-1}(\pi(j) + 1) & \text{if } \pi(j) + 1 \leqslant n, \\ \pi^{-1}(1) & \text{if } \pi(j) = n. \end{cases} \tag{8}$$

Observe that $\xi_n(V, Y)$ can be rewritten as Q_n / S_n where

$$\begin{aligned} Q_n &= \frac{1}{n} \sum_{j=1}^{n} \left(\min\{F_n(Y_j), F_n(Y_{N(j)})\} - (1 - F_n(Y_j))^2 \right) \\ &= \frac{1}{n} \sum_{j=1}^{n} \left(\frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\{Y_k \leqslant Y_j\}} \mathbb{1}_{\{Y_k \leqslant Y_{N(j)}\}} - \left(\frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\{Y_j \leqslant Y_k\}} \right)^2 \right), \\ S_n &= \frac{1}{n} \sum_{j=1}^{n} F_n(Y_j)(1 - F_n(Y_j)), \end{aligned}$$

where F_n stands for the empirical distribution function of Y: $F_n(t) = \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\{Y_k \leqslant t\}}$. The analogue of the Pick-Freeze version Y^V with respect to V of Y becomes Y_N and Lemma 2.1 is replaced by the formula

$$\mathbb{E}[\mathbb{1}_{\{Y_j \geqslant t\}} \mathbb{1}_{\{Y_{N(j)} \geqslant t\}} | V_1, \dots, V_n] = G_{V_j}(t) G_{V_{N(j)}}(t) \tag{9}$$

for all $j = 1, \dots, n$ that is mentioned in the proof of Lemma 7.10 in [9, p.24], with G_V the conditional survival function: $G_V(t) = \mathbb{P}(Y \geqslant t | V)$.

Remark 2.2. In [9], the author considers also the random variables $V_{n,j}$ due to the fact that ties are possible. In our paper, we assume that the distributions of V and Y are diffuse rendering the introduction of the $V_{n,j}$'s unworthy since in this case, $V_{n,j} = V_{N(j)}$.

It is worth noticing that a unique n sample of input-output provides consistent estimations of the p first-order Cramér-von-Mises indices.

3 Generalization of Chatterjee's method

3.1 A universal estimation procedure of sensitivity indices

In this section, we propose a universal estimation procedure of expectations of the form

$$\mathbb{E}[\mathbb{E}[g(Y)|V]\mathbb{E}[h(Y)|V]],$$

for two integrable functions g and h. This result is a generalization of (9) and can be interpreted as an approximation of (5). To this end, we introduce the function Ψ_V defined by

$$\Psi_V(g) = \mathbb{E}[g(Y)|V] \tag{10}$$

for any integrable function g. Let $\mathcal{F}_n$ be the σ-algebra generated by $\{V_1, \ldots, V_n\}$. Note that in Section 2.3, we have considered $g(x) = g_t(x) = \mathbb{1}_{\{x \geqslant t\}}$ so that $\Psi_V(g) = \mathbb{P}(Y \geqslant t|V) = G_V(t)$.

Lemma 3.1. *Let g and h be two integrable functions such that gh is also integrable. Let $(V_j, Y_j)_{1 \leqslant j \leqslant n}$ be an n-sample of (V, Y). Consider a $\mathcal{F}_n$-measurable random permutation σ_n such that $\sigma_n(j) \neq j$, for all $j = 1, \ldots, n$. Then*

$$\mathbb{E}\left[g(Y_j)h(Y_{\sigma_n(j)})|V_1, \ldots, V_n\right] = \Psi_{V_j}(g)\Psi_{V_{\sigma_n(j)}}(h). \tag{11}$$

The previous lemma (the proof of which has been postponed to Appendix A) leads to a generalization of the first part of the numerator of ξ_n defined in (7). Following the same lines as in [9], one may prove that such a quantity converges almost surely as $n \to \infty$ under some mild conditions.

Proposition 3.2. *Let g and h be two bounded measurable functions. Consider a $\mathcal{F}_n$-measurable random permutation σ_n with no fix point (i.e. $\sigma_n(j) \neq j$), for all $j = 1, \ldots, n$. In addition, we assume that for any $j = 1, \ldots, n$, $V_{\sigma_n(j)} \to V_j$ as $n \to \infty$ with probability one. Then $\chi_n(V, Y; g, h)$ defined by*

$$\chi_n(V, Y; g, h) = \frac{1}{n}\sum_{j=1}^{n} g(Y_j)h(Y_{\sigma_n(j)}) \tag{12}$$

converges almost surely as $n \to \infty$ to

$$\chi(V, Y; g, h) = \mathbb{E}[\Psi_V(g)\Psi_V(h)], \tag{13}$$

where Ψ_V has been defined in (10).

The reader is referred to Appendix A for the detailed proof of Proposition 3.2.

3.2 Recovering the first-order Sobol indices

We can now leverage the above results and construct a new family of estimators for Sobol indices. More precisely, let us consider the model (1) and assume we want to estimate the first-order Sobol index S^1 defined in (2) with respect to $V = X_1$ assumed to be real-valued. We then define N as in (8) where π is the rank of X_1. Taking $g(x) = h(x) = x$ and $\sigma_n = N$, (11) provides the analogue to ξ_n to estimate the classical Sobol indices:

$$\xi_n^{\text{Sobol}}(X_1, Y) := \frac{\frac{1}{n}\sum_{j=1}^{n} Y_j Y_{N(j)} - \left(\frac{1}{n}\sum_{j=1}^{n} Y_j\right)^2}{\frac{1}{n}\sum_{j=1}^{n}(Y_j)^2 - \left(\frac{1}{n}\sum_{j=1}^{n} Y_j\right)^2}, \tag{14}$$

where the denominator is reduced to the empirical variance of Y. As the functions g and h are here unbounded, Proposition 3.2 does not apply and thus offers no asymptotic

information. However, the quantity of interest Y being generally bounded in practice, appropriately truncated versions of g and h could be considered.
This estimator can be compared to the classical Pick-Freeze estimator which is constructed as follows. For the estimation of S^1 for instance, an n-sample $(Y_1, \dots, Y_n)$ of the output Y and an n-sample $(Y_1^1, \dots, Y_n^1)$ of its Pick-Freeze version Y^1 are required. The natural estimator of S^1 is then given by

$$S_n^1 = \frac{\frac{1}{n}\sum_{j=1}^n Y_j Y_j^1 - \left(\frac{1}{n}\sum_{j=1}^n Y_j\right)\left(\frac{1}{n}\sum_{j=1}^n Y_j^1\right)}{\frac{1}{n}\sum_{j=1}^n (Y_j)^2 - \left(\frac{1}{n}\sum_{j=1}^n Y_j\right)^2}. \tag{15}$$

A slightly different estimator that uses all the information available is introduced in [22]:

$$T_n^1 = \frac{\frac{1}{n}\sum_{j=1}^n Y_j Y_j^1 - \left(\frac{1}{n}\sum_{j=1}^n \frac{Y_j+Y_j^1}{2}\right)^2}{\frac{1}{n}\sum_{j=1}^n \frac{(Y_j)^2+(Y_j^1)^2}{2} - \left(\frac{1}{n}\sum_{j=1}^n \frac{Y_j+Y_j^1}{2}\right)^2}. \tag{16}$$

As for the Cramér-von-Mises estimation scheme, such an estimation procedure has been proved to be consistent and asymptotically normal (i.e. the rate of convergence is $\sqrt{n}$) in [22, 16]. The limiting variance can be computed explicitly, allowing the practitioner to build confidence intervals. In addition, the sequence of estimators $(T_n^1)_n$ is asymptotically efficient to estimate S^1 from such a design of experiment (see, [39] for the definition of the asymptotic efficiency and [16] for the details of the result).

3.3 A CLT for the mighty estimator of the classical first-order Sobol indices

We establish a CLT for the estimator $\xi_n^{\text{Sobol}}(X_1, Y)$ of the first-order Sobol index

$$S^1 = \frac{\text{Var}(\mathbb{E}[Y|X_1])}{\text{Var}(Y)}.$$

with respect to X_1 (assumed to be real-valued) under some mild assumptions on the model f and the random input X_1 in (1). The proof of the theorem is given in Appendix B.

Theorem 3.3. *Assume that X_1 is uniformly distributed on $[0,1]$ and f in (1) is a twice differentiable function with respect to its first coordinate. Further, we suppose that f and its two first derivatives (with respect to its first coordinate) are bounded. Then*

$$\sqrt{n}\left(\xi_n^{Sobol}(X_1, Y) - S^1\right)$$

is asymptotically Gaussian with zero mean and explicit variance σ^2 given in Appendix B.5.

Remark 3.4. The boundedness of f implies that f has a fourth moment, that is the minimal assumption to get a CLT.

The assumption on the distribution of X_1 can be relaxed as stated in the following corollary.

Corollary 3.5. *Let F_{X_1} be the cumulative distribution function of X_1. Assume that $f \circ F_{X_1}^{-1}$ is a twice differentiable function such that $f \circ F_{X_1}^{-1}$ and its two first derivatives are bounded. Then the conclusion of Theorem 3.3 still holds.*

Observe that Theorem 3.3 and Corollary 3.5 naturally allow to build statistical tests for testing

$$H_0 : S^1 = 0 \quad \text{against} \quad H_1 : S^1 \neq 0.$$

One can note that Chatterjee [9] result allows to test the independence of the input X_1 with respect to the output Y which is a stronger assumption than $S^1 = 0$, this was for example studied in [34]. In addition our result allows to compute the power of the statistical test against any alternative of the kind $H_{1,x} : S^1 \neq x$ for any $x \leq 0$.

Remark 3.6. A careful reading of the different steps of the proof shows that Theorem 3.3 can be slightly extended to more general situations involving more than two successive order statistics and with more general second variable $(X_2, \ldots, X_p)$. See the forthcoming paper [18].

4 Recovering other classical indices

4.1 Sensitivity indices in general metric spaces

In this section, we consider a computer code of the form (1) valued in a general metric space $\mathcal{M}$ as presented in [19]. In this context, the authors of [19] consider a family of test functions parametrized by m elements of $\mathcal{M}$ ($m \in \mathbb{N}^*$). For any $a = (a_i)_{i=1,\ldots,m} \in \mathcal{M}^m$, the test functions

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

are assumed to be L^2-functions with respect to the product measure $\mathbb{P}^{\otimes m} \otimes \mathbb{P}$ on $\mathcal{M}^m \times \mathcal{M}$ where $\mathbb{P}$ is the distribution of the output, still denoted by Y. Then they define the general metric space sensitivity index with respect to X_1 by

$$S^1_{2,GMS} := \frac{\int_{\mathcal{M}^m} \mathbb{E}\left[\left(\mathbb{E}[T_a(Y)] - \mathbb{E}[T_a(Y)|X_1]\right)^2\right] d\mathbb{P}^{\otimes m}(a)}{\int_{\mathcal{M}^m} \mathrm{Var}(T_a(Y)) d\mathbb{P}^{\otimes m}(a)}. \tag{17}$$

This general class of indices encompasses the classical sensitivity indices, for instance, the Sobol indices and the Cramér-von-Mises indices. Naturally, a Monte-Carlo procedure based on the Pick-Freeze scheme can be performed to estimate $S^1_{2,GMS}$.

Estimation procedure based on U-statistics In [19], the authors propose a more efficient estimation procedure based on U-statistics (see [19, Equation (13)]). More precisely, for any $1 \leqslant i \leqslant m+2$, let $\mathbf{y}_i = (y_i, y_i^1)$ and define

$$\begin{aligned} \Phi_1(\mathbf{y}_1, \ldots, \mathbf{y}_{m+1}) &:= T_{y_1,\ldots,y_m}(y_{m+1}) T_{y_1,\ldots,y_m}(y^1_{m+1}) \\ \Phi_2(\mathbf{y}_1, \ldots, \mathbf{y}_{m+2}) &:= T_{y_1,\ldots,y_m}(y_{m+1}) T_{y_1,\ldots,y_m}(y^1_{m+2}) \\ \Phi_3(\mathbf{y}_1, \ldots, \mathbf{y}_{m+1}) &:= T_{y_1,\ldots,y_m}(y_{m+1})^2 \\ \Phi_4(\mathbf{y}_1, \ldots, \mathbf{y}_{m+2}) &:= T_{y_1,\ldots,y_m}(y_{m+1}) T_{y_1,\ldots,y_m}(y_{m+2}). \end{aligned}$$

In addition, set

$$m(1) = m(3) = m + 1 \quad \text{and} \quad m(2) = m(4) = m + 2 \tag{18}$$

and define for $j = 1, \ldots, 4$,

$$I(\Phi_j) := \int_{\mathcal{M}^{m(j)}} \Phi_j(\mathbf{y}_1, \ldots, \mathbf{y}_{m(j)}) d\mathbb{P}_{\mathbf{Y}}^{\otimes m(j)}(\mathbf{y}_1 \ldots, \mathbf{y}_{m(j)}), \tag{19}$$

where $\mathbb{P}_{\mathbf{Y}}$ stands for the law of $\mathbf{Y} = (Y, Y^1)^\top$. Finally, we introduce the application Ψ from $\mathbb{R}^4$ to $\mathbb{R}$ defined by

$$\begin{array}{rccc} \psi: & \mathbb{R}^4 & \to & \mathbb{R} \\ & (x, y, z, t) & \mapsto & \frac{x-y}{z-t}. \end{array} \tag{20}$$

Then one can express $S^1_{2,GMS}$ in the following way

$$S^1_{2,GMS} = \psi\left(I(\Phi_1), I(\Phi_2), I(\Phi_3), I(\Phi_4)\right). \tag{21}$$

Following the framework of Hoeffding [21], we replace the functions Φ_1, Φ_2, Φ_3, and Φ_4 by their symmetrized version Φ^s_1, Φ^s_2, Φ^s_3, and Φ^s_4:

$$\Phi^s_j(\mathbf{y}_1, \ldots, \mathbf{y}_{m(j)}) = \frac{1}{(m(j))!} \sum_{\tau \in \mathcal{S}_{m(j)}} \Phi_j(\mathbf{y}_{\tau(1)}, \ldots, \mathbf{y}_{\tau(m(j))})$$

for $j = 1, \ldots, 4$ where $\mathcal{S}_k$ is the symmetric group of order k. For $j = 1, \ldots 4$, the integrals $I(\Phi^s_j)$ are naturally estimated by U-statistics of order $m(j)$. More precisely, we consider an n i.i.d. sample $(\mathbf{Y}_1, \ldots, \mathbf{Y}_n)$ with distribution $\mathbb{P}_{\mathbf{Y}}$ and, for $j = 1, \ldots, 4$, we define

$$U_{j,n} := \binom{n}{m(j)}^{-1} \sum_{1 \leqslant i_1 < \cdots < i_{m(j)} \leqslant n} \Phi^s_j\left(\mathbf{Y}_{i_1}, \ldots, \mathbf{Y}_{i_{m(j)}}\right). \tag{22}$$

[21, Theorem 7.1] ensures that $U_{j,n}$ converges in probability to $I(\Phi_j)$ for any $j = 1, \ldots, 4$. Moreover, one may also prove that the convergence holds almost surely proceeding as in the proof of [17, Lemma 6.1]. Then we estimate $S^1_{2,GMS}$ by

$$S^1_{2,GMS,n} := \frac{U_{1,n} - U_{2,n}}{U_{3,n} - U_{4,n}} = \psi(U_{1,n}, U_{2,n}, U_{3,n}, U_{4,n}). \tag{23}$$

A novel estimation procedure In light of Section 3.1, we introduce a novel estimation $\xi^{\mathrm{GMS}}_n(X_1, Y)$ of $S^1_{2,GMS}$ in (17) as follows. The Pick-Freeze scheme is replaced by the use of the $Y_{N(i)}$'s where N is the permutation defined in (8) and the integration with respect to $\mathbb{P}^{\otimes m}$ is handled using a unique n-sample of Y. More precisely, the empirical estimator $\xi^{\mathrm{GMS}}_n(X_1, Y)$ of $S^1_{2,GMS}$ is given by the ratio between

$$\begin{aligned} &\frac{1}{n^m} \sum_{1 \leqslant i_1, \ldots, i_m \leqslant n} \left[\frac{1}{n} \sum_{j=1}^n T_{Y_{i_1}, \cdots, Y_{i_m}}(Y_j) T_{Y_{i_1}, \cdots, Y_{i_m}}(Y_{N_i(j)})\right] \\ &- \frac{1}{n^m} \sum_{1 \leqslant i_1, \ldots, i_m \leqslant n} \left[\frac{1}{n} \sum_{j=1}^n T_{Y_{i_1}, \cdots, Y_{i_m}}(Y_j)\right]^2 \end{aligned}$$

and

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

4.2 Owen higher-order moment indices

Following [28, 29], we consider extensions to Sobol indices obtained by replacing their numerator by higher-order moments. More precisely, for any integer $q \geqslant 2$, we set

$$H_q^1 := \mathbb{E}\left[(\mathbb{E}[Y|X_1] - \mathbb{E}[Y])^q\right]. \tag{24}$$

See [17] for known properties of H_q^1.
In order to construct a Pick-Freeze estimator for H_q^1, we refer the reader to [17]. More precisely, we first observe that

$$H_q^1 = \mathbb{E}\left[\prod_{m=1}^{q}\left((Y^1)^m - \mathbb{E}[Y]\right)\right] = \sum_{l=0}^{q}\binom{q}{l}(-1)^{q-l}\mathbb{E}\left[Y\right]^{q-l}\mathbb{E}\left[\prod_{m=1}^{l}(Y^1)^m\right]$$

with the usual convention $\prod_{m=1}^{0}(Y^1)^m = 1$. Here, $Y_1^1 = Y$ and, for $i = 2, \ldots, q$, Y_i^1 is constructed independently (similarly to Y^1 in (5)). Now we construct a Monte-Carlo scheme and consider the following Pick-Freeze design constituted by a n-sample $\left(Y_{i,j}^1\right)_{(i,j)\in I_q\times I_n}$ of $\left(Y_1^1, \ldots, Y_q^1\right)$ where, for any positive integer k, I_k stands for the set $\{1, \ldots, k\}$. The resulting Monte-Carlo estimator is then

$$H_{q,n}^1 = \sum_{l=0}^{q}\binom{q}{l}(-1)^{q-l}\left(\overline{P}_1^1\right)^{q-l}\overline{P}_l^1$$

where for any positive integer n, $j \in I_n$ and $l \in I_q$, we have set

$$P_{l,j}^1 = \binom{q}{l}^{-1}\sum_{k_1<\ldots<k_l\in I_q}\left(\prod_{i=1}^{l}Y_{k_i,j}^1\right) \quad \text{and} \quad \overline{P}_l^1 = \frac{1}{N}\sum_{j=1}^{N}P_{l,j}^1.$$

This setting generalizes the estimation procedure from [16] and uses more information by considering the means over the set of indices $k_1, \ldots, k_l \in I_d$, $k_n \neq k_m$.

A novel estimation procedure We generalize the procedure proposed by Chatterjee in order to estimate higher-order moment indices. To that end, we introduce, for all $m \in \{1, \ldots, q-1\}$ and $j \in \{1, \ldots, n\}$,

$$N_m(j) = \begin{cases}\pi^{-1}(\pi(j)+m) & \text{if } \pi(j)+m \leqslant n,\\ \pi^{-1}(\pi(j)+m-n) & \text{if } \pi(j)+m > n.\end{cases} \tag{25}$$

Note that $N_1 = N$. It remains to update Lemma 3.1 as follows.

Lemma 4.1. *Let $(g_m)_{m=0,\ldots,q-1}$ be a family of measurable functions in $L^1(\mathbb{R})$. Let $(V_j, Y_j)_{1\leqslant j\leqslant n}$ be an n-sample of (V, Y). Then*

$$\mathbb{E}\left[\prod_{m=0}^{q-1}g_m(Y_{N_m(i)})|V_1, \ldots, V_n\right] = \prod_{m=0}^{q-1}\psi_{V_{N_m(i)}}(g_m), \tag{26}$$

where by convention $N_0(j) = j$ for all $j = 1, \ldots, n$.

Finally, Lemma 4.1 (with $g_m(y) = y$, for all $y \in \mathbb{R}$ and $m = 0, \ldots, q-1$ together with (24)) allows us to propose a more efficient estimation procedure of the higher-order moment index H_q^1 introduced by Owen.

Remark 4.2. While the collection of all indices $\left(H_q^1\right)_q$ is more informative than the classical Sobol indices, it also has several drawbacks. First, these indices are moment-based and, as is well known, they are not stable when the moment order increases. Second, they may be negative when q is odd. To overcome this fact, one could introduce $\mathbb{E}\left[|\mathbb{E}[Y|X_1] - \mathbb{E}[Y]|^q\right]$ but the Pick-Freeze estimation procedure is then lost. Third, the Pick-Freeze estimation procedure is computationally expensive and may be unstable: it requires a $q \times n$-sample of the output Y. In order to properly assess the influence of an input on the law of the output, we need to estimate the first $K-1$ indices H_q^1: H_2^1, $\ldots$, H_K^1. Hence, we need to run the code $K \times n$ times. These indices are thus not attractive in practice.

5 Numerical experiments

5.1 Numerical comparison on the Sobol g-function: conventional Pick-Freeze estimators vs Chatterjee's estimators

In this section, we compare the performances of both estimation procedures on an analytic function: the so-called Sobol g-function, that is defined by

$$g(X_1, \ldots, X_p) = \prod_{i=1}^{p} \frac{|4X_i - 2| + a_i}{1 + a_i}, \tag{27}$$

where $(a_i)_{i \in \mathbb{N}}$ is a sequence of real numbers and the X_i's are i.i.d. random variables uniformly distributed on $[0, 1]$. In this setting, one may easily compute the exact expression of the first-order Sobol indices:

$$S^i = \frac{1/(3(1 + a_i^2))}{[\prod_{i=1}^{p} 1/(3(1 + a_i^2))] - 1}.$$

As expected, the lower the coefficient a_i, the more significant the variable X_i. In the sequel, we simply fix $a_i = i$.
Due to its complexity (non-linear and non-monotonic correlations) and the analytical expression of the Sobol indices, the Sobol g-function is a classical test example commonly used in GSA (see e.g. [32]).

Convergence as the sample size increases In Figure 1, we compare the estimations of the six first-order Sobol indices given by both methods ($p = 6$). In the Pick-Freeze estimations given by (16), several sizes of sample N have been considered: $N = 100$, 500, 1000, 5000, 10000, 50000, 100000, and 500000. The Pick-Freeze procedure requires $(p + 1) = 7$ samples of size N. To have a fair comparison, the sample sizes considered in the estimation of ξ_n^{Sobol} are $n = (p + 1)N = 7N$. We observe that both methods converge and give precise results for large sample sizes.

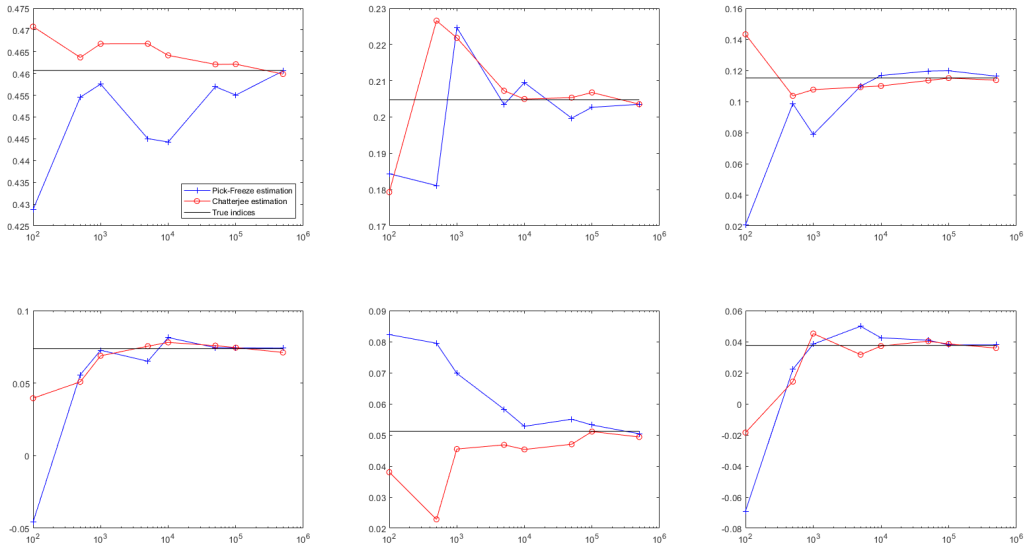

Figure 1: The Sobol g-function model (27). Convergence of both methods when N increases. The sixth first-order Sobol indices have been represented from left to right and up to bottom. Several sample sizes have been considered: $N = 100$, 500, 1000, 5000, 10000, 50000, 100000, and 500000 for the Pick-Freeze estimation procedure and correspondingly $(p+1)N$ for the estimation procedure proposed in [9]. The x-axis is in logarithmic scale.

Comparison of the mean square errors We now compare the efficiency of both methods at a fixed sample size. In that view, we assume that only $n = 700$ calls of the computer code f are allowed to estimate the six first-order Sobol indices. We repeat the estimation procedure 500 times. The boxplot of the mean square errors for the estimation of the first-order Sobol index S^1 with respect to X_1 has been represented in Figure 2. We observe that, for a fixed sample size $n = 700$ (corresponding to a Pick-Freeze sample size $N = 100$), Chatterjee's estimation procedure performs much better than the Pick-Freeze method with significantly lower mean errors. The same behavior can be observed for all the first Sobol indices as can be seen in Table 1 that provides some characteristics of the mean squares errors.

	Pick-Freeze			Chatterjee		
	Mean	Median	Stdev	Mean	Median	Stdev
mse S^1	0.0095548	0.0039458	0.0145033	0.0010218	0.0004498	0.0013999
mse S^2	0.0105727	0.0046104	0.0148873	0.0017314	0.0006870	0.0027436
mse S^3	0.0101785	0.0041789	0.0143846	0.0016667	0.0006409	0.0024392
mse S^4	0.0105463	0.0047284	0.0178064	0.0018522	0.0008126	0.0025296
mse S^5	0.0097979	0.0042995	0.0135533	0.0016285	0.0006855	0.0024264
mse S^6	0.0096109	0.0046822	0.0134822	0.0015590	0.0007080	0.0021333

Table 1: The Sobol g-function model (27). Some characteristics of the mean square errors for the estimation of the six first-order Sobol indices with a fixed sample size and 500 replications. In Chatterjee's methodology, the sample size considered is $n = 700$ while in the Pick-Freeze estimation procedure, it is $N = 100$.

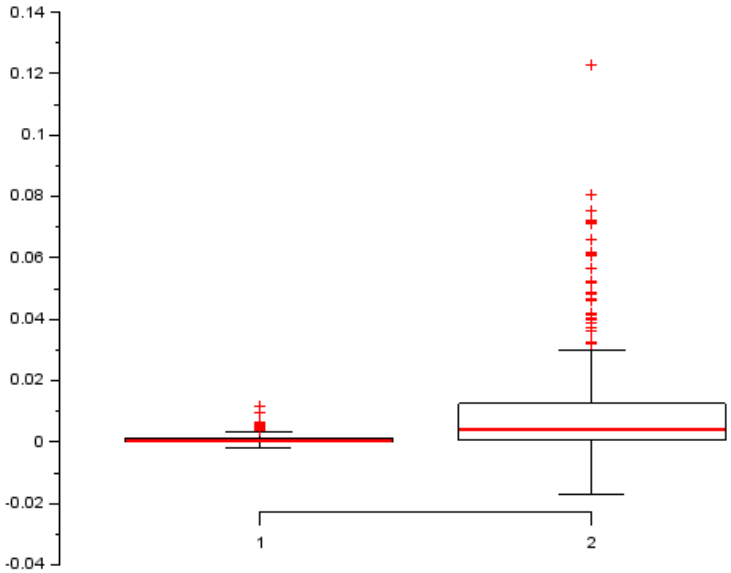

Figure 2: The Sobol g-function model (27). Boxplot of the mean square errors of the estimation of S^1 with a fixed sample size and 500 replications. The results of Chatterjee's methodology with $n = 700$ are provided in the left panel. The results of the Pick-Freeze estimation procedure with $N = 100$ are provided in the right panel.

Performances for small sample sizes or for large number of input variables As expected, we can observe in Table 2 that Chatterjee's procedure proceeds much better than the Pick-Freeze methodology for small sample sizes. Similarly, if the number of input variables increases drastically, we can observe the same behavior as can be seen in Figure 3. In that case, we consider the model (27) for several values of p: 6, 10, 15, 20, 30, 40, and 50.

	Pick-Freeze			Chatterjee		
	$N = 10$	$N = 50$	$N = 100$	$n = 70$	$n = 350$	$n = 700$
mse S^1	0.1128686	0.0172275	0.0095548	0.0116790	0.0022941	0.0010218
mse S^2	0.1509575	0.0223196	0.0105727	0.0177522	0.0033719	0.0017314
mse S^3	0.1469124	0.0220015	0.0101785	0.0175517	0.0032474	0.0016667
mse S^4	0.1591130	0.0196357	0.0105463	0.0159360	0.0033948	0.0018522
mse S^5	0.1646339	0.0240353	0.0097979	0.0158563	0.0032230	0.0016285
mse S^6	0.1466408	0.0217638	0.0096109	0.0166701	0.0029653	0.0015590

Table 2: The Sobol g-function model (27). Mean squares errors of the estimation of the six first-order Sobol indices with small sample sizes and with both methods.

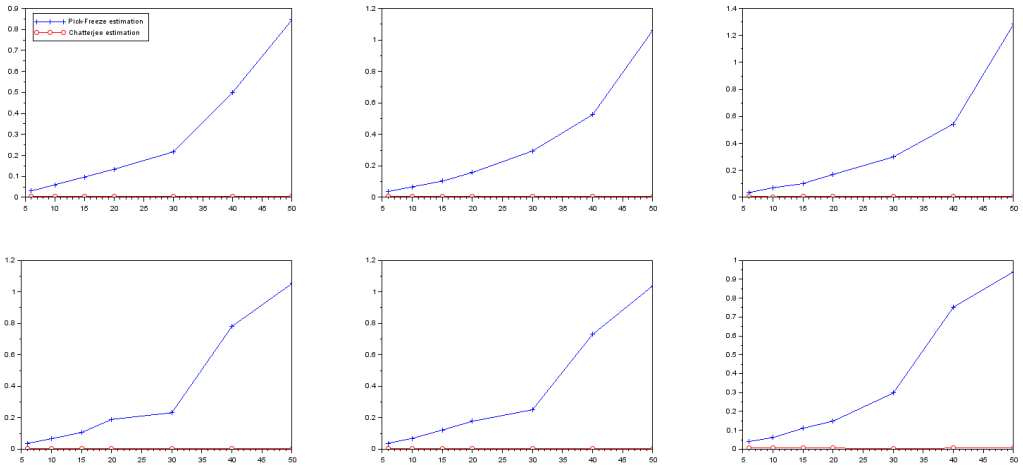

Figure 3: The Sobol g-function model (27). Mean square errors of the estimation of the six first-order Sobol indices with respect to the number of input variables with a fixed sample size and 500 replications. We consider the sample sizes $n = 200$ in Chatterjee's methodology and $N = n/(p + 1)$ in the Pick-Freeze procedure. The number of input variables considered are $p = 6, 10, 15, 20, 30, 40,$ and 50.

5.2 An application in biology

Here, we illustrate the nature and the performance of the Cramér-von-Mises indices and their corresponding Chatterjee estimators as a screening mechanism for high-dimensional problems. To do so, we consider the neurovascular coupling model from [20]. Mathemat-

ically, this corresponds to the following differential-algebraic equation (DAE) system

$$\frac{dW}{dt} = G(W, Z, X), \tag{28}$$

$$0 = H(W, Z, X), \tag{29}$$

where $W = (W_1, \ldots, W_N)$ and $Z = (Z_1, \ldots, Z_M)$ correspond respectively to the differential and algebraic state variables of the models. The variables $X = (X_1, \ldots, X_p)$ correspond to the uncertain parameters of the model. Our quantity of interest corresponds to the time average over $[0, T]$ of W^* (which is one of the differential state variables W_1, ..., W_N), i.e.

$$Y = \frac{1}{T} \int_0^T W^*(t)\, dt. \tag{30}$$

As above, we regard Y as a function of the unknown parameters, i.e., $Y = f(X_1, \ldots, X_p)$. In our implementation, the values of W^* are obtained by solving the above DAE system (Equations (28) and (29)) by the MATLAB routine ode15s (it can be checked that (28) and (29) form an index one system). Further, in the current example, $N = 67$ and $p = 160$ and the distributions of most of the X_i's are uniform and allowed to vary $\pm 10\%$ from nominal values (see [20] for additional details).
We compare the results from the Chatterjee estimators as described above to those resulting from the linear regression

$$f(X_1, \ldots, X_{160}) \approx \lambda_0 + \sum_{j=1}^{160} \lambda_j X_j.$$

As shown in [20], the above approximation performs well for the considered QoI. We assign to each variable $X_1, \ldots, X_{160}$ a relative importance L_j where

$$L_j = \frac{|\lambda_j|}{\sum_{\ell=1}^{160} |\lambda_\ell|}, \qquad j = 1, \ldots, 160.$$

Figure 4 displays the results. Both screening approaches identify the same to three influential parameters. More parameters are identified as being non-influential through the linear regression approach than using the Cramér-von-Mises indices.

6 Conclusion

In this paper, we explain how to use the estimator proposed by Chatterjee in [9] to provide a very nice and mighty procedure to estimate both all the order one Sobol indices and the so-called Cramér-von-Mises indices [17] at a small cost (only n calls of the computer code). We also extend Chatterjee's method to estimate more general quantities. Furthermore, we show a CLT for our estimations of Sobol indices. As examples, we consider two indices already introduced in sensitivity analysis: the indices adapted to output valued in general metric spaces defined in [19] and the higher-moment indices [28, 29]. A general CLT will be established soon in [18].

Acknowledgment. We warmly thank Robin Morillo for the numerical study provided in Section 5.2. Moreover, we deeply thank the anonymous referee of the early version of our paper who pushed us to prove the CLT.

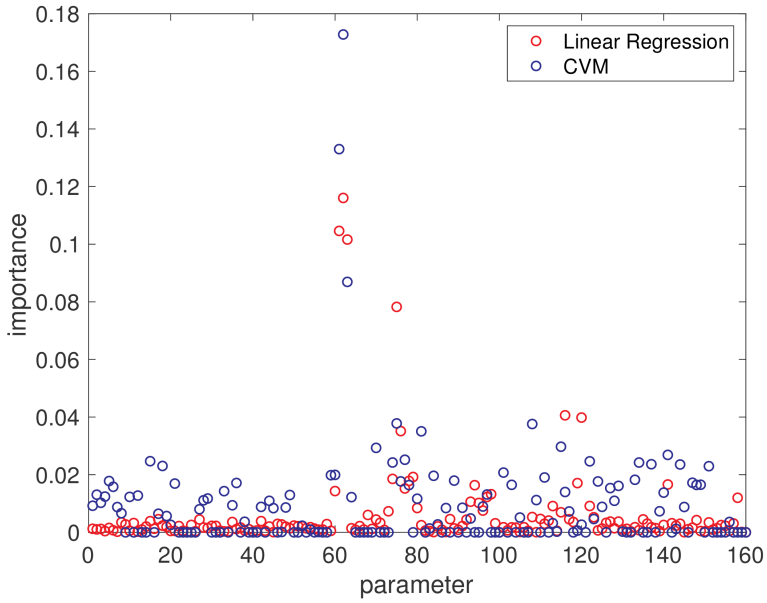

Figure 4: Chatterjee estimators corresponding to the Cramér-von-Mises indices as a screening mechanics for the DAE system given by (28) and (29).

Support from the ANR-3IA Artificial and Natural Intelligence Toulouse Institute is gratefully acknowledged. This work was also supported by the National Science Foundation under grant and DMS-1745654.

A Proof of the consistency

Proof of Lemma 3.1. Using the measurability of σ_n and by independence, we have

$$\begin{aligned}
\mathbb{E}\left[g(Y_j)h(Y_{\sigma_n(j)})|\mathcal{F}_n\right] &= \mathbb{E}\Big[g(Y_j)\sum_{\substack{l=1,\\ l\neq j}}^{n} h(Y_l)\mathbb{1}_{\{\sigma_n(j)=l\}}|\mathcal{F}_n\Big] \\
&= \sum_{\substack{l=1,\\ l\neq j}}^{n} \mathbb{1}_{\{\sigma_n(j)=l\}}\mathbb{E}\Big[g(Y_j)h(Y_l)|\mathcal{F}_n\Big] \\
&= \sum_{\substack{l=1,\\ l\neq j}}^{n} \mathbb{1}_{\{\sigma_n(j)=l\}}\mathbb{E}\Big[g(Y_j)|\mathcal{F}_n\Big]\mathbb{E}\Big[h(Y_l)|\mathcal{F}_n\Big] \\
&= \mathbb{E}\Big[g(Y_j)|V_j\Big]\sum_{\substack{l=1,\\ l\neq j}}^{n} \mathbb{1}_{\{\sigma_n(j)=l\}}\mathbb{E}\Big[h(Y_l)|V_l\Big] \\
&= \Psi_{V_j}(g)\sum_{\substack{l=1,\\ l\neq j}}^{n} \mathbb{1}_{\{\sigma_n(j)=l\}}\Psi_{V_l}(h) = \Psi_{V_j}(g)\Psi_{V_{\sigma_n(j)}}(h).
\end{aligned}$$

□

Proof of Proposition 3.2. We follow the steps of the proof of Corollary 7.12 in [9]. Our

proof is significantly simpler since σ_n is assumed to have no fix points and V is continuous so that the are no ties in the sample. To simplify the notation, we denote $\chi_n(V, Y; g, h)$ and $\chi(V, Y; g, h)$ by χ_n and χ respectively.
We first prove that, for any measurable function φ,

$$\varphi(V_1) - \varphi(V_{\sigma_n(1)}) \to 0 \tag{31}$$

in probability as $n \to \infty$. Let $\varepsilon > 0$. By the special case of Lusin's theorem (see [9, Lemma 7.5]), there exists a compactly supported continuous function $\tilde{\varphi}\colon \mathbb{R} \to \mathbb{R}$ such that $\mathbb{P}(\{x;\, \varphi(x) \neq \tilde{\varphi}(x)\}) < \varepsilon$, where $\mathbb{P}$ stands for the distribution of V. Then for any $\delta > 0$,

$$\begin{aligned}\mathbb{P}\Big(\big|\varphi(V_1) - \varphi(V_{\sigma_n(1)})\big| > \delta \Big) \leqslant \mathbb{P}\left(\big|\tilde{\varphi}(V_1) - \tilde{\varphi}(V_{\sigma_n(1)})\big| > \delta \right) \\ + \mathbb{P}\left(\varphi(V_1) \neq \tilde{\varphi}(V_1)) + \mathbb{P}(\varphi(V_{\sigma_n(1)}) \neq \tilde{\varphi}(V_{\sigma_n(1)})\right).\end{aligned} \tag{32}$$

By continuity of $\tilde{\varphi}$ and since $V_{\sigma_n(1)} \to V_1$ as $n \to \infty$ with probability one, the first term in the right hand side of (32) converges to 0 as $n \to \infty$. By construction of g, the second term is lower than ε. Turning to the third one, we have thus

$$\begin{aligned}\mathbb{E}[\varphi(V_{\sigma_n(1)})] &= \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[\varphi(V_{\sigma_n(j)})] = \frac{1}{n}\sum_{j=1}^{n}\sum_{\substack{l=1\\ l\neq j}}^{n} \mathbb{E}[\varphi(V_l)\mathbb{1}_{\{\sigma_n(j)=l\}}] \\ &= \frac{1}{n}\sum_{l=1}^{n}\sum_{\substack{j=1\\ j\neq l}}^{n} \mathbb{E}[\varphi(V_l)\mathbb{1}_{\{\sigma_n(j)=l\}}] = \frac{1}{n}\sum_{l=1}^{n} \mathbb{E}[\varphi(V_l)\sum_{\substack{j=1\\ j\neq l}}^{n} \mathbb{1}_{\{\sigma_n(j)=l\}}] \\ &= \frac{1}{n}\sum_{l=1}^{n} \mathbb{E}[\varphi(V_l)] = \mathbb{E}[\varphi(V_1)]\end{aligned}$$

where we have used the fact that by assumption σ_n has no fix point and the V_i's have no ties. This yields

$$\mathbb{P}(\varphi(V_{\sigma_n(1)}) \neq \tilde{\varphi}(V_{\sigma_n(1)})) = \mathbb{P}(\varphi(V_1) \neq \tilde{\varphi}(V_1)) < \varepsilon,$$

and, since ε and δ are arbitrary, (31) is therefore proved.
Now, since $x \mapsto \Psi_x$ is a measurable function and applying (31), we have

$$\begin{cases} \Psi_{V_1}(g) - \Psi_{V_{\sigma_n(1)}}(g) & \to 0, \\ \Psi_{V_1}(h) - \Psi_{V_{\sigma_n(1)}}(h) & \to 0, \end{cases} \quad \text{in probability as } n \to \infty. \tag{33}$$

Lemma 3.1 and the dominated convergence theorem lead to

$$\begin{aligned}\mathbb{E}[\chi_n] &= \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[g(Y_j)h(Y_{\sigma_n(j)})] = \mathbb{E}[g(Y_1)h(Y_{\sigma_n(1)})] \\ &= \mathbb{E}[\Psi_{V_1}(g)\Psi_{V_{\sigma_n(1)}}(h)] \to \mathbb{E}[\Psi_V(g)\Psi_V(h)] = \chi(V, Y; g, h)\end{aligned} \tag{34}$$

where we have taken into account the fact that $\Psi_V(g)$ and $\Psi_V(h)$ are bounded (due to the boundedness of g and h) and used (33).
The last step of the proof consists in comparing χ_n with $\mathbb{E}[\chi_n]$ using Mc Diarmid's concentration inequality [25]. To be self-contained, we recall this result.

Theorem A.1 (Mac Diarmid's bounded difference concentration inequality [25]). *Let $W = (W_1, \dots, W_n)$ be a family of independent variables with W_i taking its values in a set A_k. Consider a real-valued function φ defined on $\Pi_{k=1}^n A_k$ satisfying*

$$|\varphi(w) - \varphi(w')| \leqslant c_k \tag{35}$$

as soon as the vectors w and w' differ only on the k-th coordinate. Then we have, for any $t > 0$,

$$\mathbb{P}\left(|\varphi(W) - \mathbb{E}[\varphi(W)]| \geqslant t\right) \leqslant 2\exp\left(-\frac{2t^2}{\sum_{k=1}^n c_k}\right).$$

Sharper constants can be obtained in Mc Diarmid's inequality by using the inequalites from [6, 7]. As we are interested in asymptotic results the accuracy of the constant has no impact on the result.
Following the same lines as in the proof of [9, Lemma 7.11], Theorem A.1 then implies

$$\mathbb{P}(|\chi_n - \mathbb{E}[\chi_n]| \geqslant t) \leqslant 2\exp\{-2n^2t^2/C^2\}, \tag{36}$$

where C is a universal constant and we conclude the proof by combining (34) and (36). □

B Proof of the asymtotic normality

Framework and goal We consider the model defined in (1) that can be rewritten as

$$Y = f(X, W) \tag{37}$$

where $X = X_1$ and $W = (X_2, \dots, X_p)$ are two independent inputs of the numerical code f that is assumed to be bounded.
The random variables X and W are defined on a product space $\Omega = \Omega_X \times \Omega_W$; so that for any $\omega \in \Omega$, there exists $\omega_X \in \Omega_X$ and $\omega_W \in \Omega_W$ and we have $(X, W)(\omega) = (X(\omega_X), W(\omega_W))$. Further, we consider π_W the projection on Ω_W and the product measure $\mathbb{P} = \mathbb{P}_X \otimes \mathbb{P}_W = \mathcal{L}_X \otimes \mathcal{L}_W$, where $\mathcal{L}_X$ is the distribution of X and $\mathcal{L}_W$ is the distribution of W. Naturally, $\mathbb{P}_W = \mathbb{P} \circ \pi_W^{-1}$.

We aim to prove a CLT for the estimator $\xi_n^{\text{Sobol}}(X, Y)$ of the classical first-order Sobol index with respect to X:

$$S^X = \frac{\text{Var}(\mathbb{E}[Y|X])}{\text{Var}(Y)}.$$

This estimator has been defined in (14) and is also equal to

$$\xi_n^{\text{Sobol}}(X, Y) = \frac{\frac{1}{n}\sum_{j=1}^{n-1} Y_{\sigma_n(j)} Y_{\sigma_n(j+1)} - \left(\frac{1}{n}\sum_{j=1}^n Y_j\right)^2}{\frac{1}{n}\sum_{j=1}^n (Y_j)^2 - \left(\frac{1}{n}\sum_{j=1}^n Y_j\right)^2},$$

where the denominator is reduced to the empirical variance of Y.

Notation It is convenient to have short expressions for terms that converge in probability to zero. We follow [39]. The notation $o_{\mathbb{P}}(1)$ (respectively $O_{\mathbb{P}}(1)$) stands for a sequence of random variables that converges to zero in probability (resp. is bounded in probability) as $n \to \infty$. More generally, for a sequence of random variables R_n,

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

For deterministic sequences X_n and R_n, the stochastic notation reduce to the usual o and O.

B.1 Proof of Theorem 3.3

The proof will proceed as follows. First, we prove a CLT for

$$\left(\frac{1}{n} \sum_{j=1}^{n-1} Y_{\sigma_n(j)} Y_{\sigma_n(j+1)}, \frac{1}{n} \sum_{j=1}^{n} Y_j, \frac{1}{n} \sum_{j=1}^{n} Y_j^2 \right) = \left(\frac{1}{n} \sum_{j=1}^{n-1} Y_{\sigma_n(j)} Y_{\sigma_n(j+1)}, \frac{1}{n} \sum_{j=1}^{n} Y_{\sigma_n(j)}, \frac{1}{n} \sum_{j=1}^{n} Y_{\sigma_n(j)}^2 \right).$$

Since f is bounded, notice that it amounts to prove a CLT for

$$\left(\frac{1}{n} \sum_{j=1}^{n-1} Y_{\sigma_n(j)} Y_{\sigma_n(j+1)}, \frac{1}{n} \sum_{j=1}^{n-1} Y_{\sigma_n(j)}, \frac{1}{n} \sum_{j=1}^{n-1} Y_{\sigma_n(j)}^2 \right),$$

where as mentioned in Section 3.3, $\sigma_n = N$ defined in (8). Secondly, we use the so-called delta method [39, Theorem 3.1] to conclude to Theorem 3.3.

It is worth noticing that the permutation on the W's do not affect the result as seen in the sequel. For $j = 1, \dots n-1$, introducing

$$\Delta_{n,j} := f\left(X_{\sigma_n(j)}, W_j\right) - f\left(\frac{j}{n+1}, W_j\right)$$

and

$$\boxed{W_{n,j} := \left(\frac{j}{n+1}, W_j\right)} \tag{38}$$

leads to

$$Y_{\sigma_n(j)} = f\left(X_{\sigma_n(j)}, W_{\sigma_n(j)}\right) \overset{\mathcal{L}}{=} f\left(X_{\sigma_n(j)}, W_j\right) = \Delta_{n,j} + f\left(W_{n,j}\right)$$

and

$$\begin{aligned} Y_{\sigma_n(j)} Y_{\sigma_n(j+1)} &= f\left(X_{\sigma_n(j)}, W_{\sigma_n(j)}\right) f\left(X_{\sigma_n(j+1)}, W_{\sigma_n(j+1)}\right) \\ &\overset{\mathcal{L}}{=} f\left(X_{\sigma_n(j)}, W_j\right) f\left(X_{\sigma_n(j+1)}, W_{j+1}\right) \\ &= \left(f\left(W_{n,j}\right) + \Delta_{n,j}\right)\left(f\left(W_{n,j+1}\right) + \Delta_{n,j+1}\right) \\ &= f\left(W_{n,j}\right) f\left(W_{n,j+1}\right) + \Delta_{n,j} f\left(W_{n,j+1}\right) + \Delta_{n,j+1} f\left(W_{n,j}\right) + \Delta_{n,j}\Delta_{n,j+1}. \end{aligned}$$

Thus we are led to establish a CLT for

$$Z_n = \frac{1}{n}\sum_{j=1}^{n-1} \begin{pmatrix} f(W_{n,j})f(W_{n,j+1}) + \Delta_{n,j} f\left(W_{n,j+1}\right) + \Delta_{n,j+1} f\left(W_{n,j}\right) + \Delta_{n,j}\Delta_{n,j+1} \\ f(W_{n,j}) + \Delta_{n,j} \\ \left(f(W_{n,j}) + \Delta_{n,j}\right)^2 \end{pmatrix}. \quad (39)$$

Let us discard the negligible terms in the CLT for Z_n. In that view, noticing that

$$\mathbb{E}\left[X_{\sigma_n(j)}\right] = \frac{j}{n+1}$$

and

$$\mathrm{Var}(X_{\sigma_n(j)}) = \frac{j(n-j+1)}{(n+1)^2(n+2)} = \mathbb{E}\left[\left(X_{\sigma_n(j)} - \frac{j}{n+1}\right)^2\right] \leqslant \frac{4}{n+2},$$

we first establish

$$X_{\sigma_n(j)} - \frac{j}{n+1} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right). \quad (40)$$

As explained bellow (40) will imply

$$\frac{1}{n}\sum_{j=1}^{n-1} \Delta_{n,j}^2 = O_{\mathbb{P}}\left(\frac{1}{n}\right) \quad \text{and} \quad \frac{1}{n}\sum_{j=1}^{n-1} \Delta_{n,j}\Delta_{n,j+1} = O_{\mathbb{P}}\left(\frac{1}{n}\right). \quad (41)$$

First of all, we expand $\Delta_{n,j}$ using the Taylor-Lagrange formula, for any $j = 1, \dots n-1$ and we obtain

$$\Delta_{n,j} = \left(X_{\sigma_n(j)} - \frac{j}{n+1}\right) f_x\left(W_{n,j}\right) + \frac{1}{2}\left(X_{\sigma_n(j)} - \frac{j}{n+1}\right)^2 f_{xx}\left(\delta_{n,j}, W_{\sigma_n(j)}\right), \quad (42)$$

where $\delta_{n,j}$ (resp. $\delta_{n,j+1}$) lies in the unordered segment $(X_{\sigma_n(j)}, j/(n+1))$ (resp. $(X_{\sigma_n(j+1)}, (j+1)/(n+1))$) and where f_x and f_{xx} are the first and second derivatives of f with respect to the first coordinate. This leads to expansions for $\Delta_{n,j}^2$ and $\Delta_{n,j}\Delta_{n,j+1}$:

$$\begin{aligned}
\Delta_{n,j}^2 &= \left(X_{\sigma_n(j)} - \frac{j}{n+1}\right)^2 \left(f_x\left(W_{n,j}\right) + \frac{1}{2}\left(X_{\sigma_n(j)} - \frac{j}{n+1}\right) f_{xx}\left(\delta_{n,j}, W_{\sigma_n(j)}\right)\right)^2 \\
\Delta_{n,j}\Delta_{n,j+1} &= \left(X_{\sigma_n(j)} - \frac{j}{n+1}\right)\left(X_{\sigma_n(j+1)} - \frac{j+1}{n+1}\right) \\
&\quad \times \left(f_x\left(W_{n,j}\right) + \frac{1}{2}\left(X_{\sigma_n(j)} - \frac{j}{n+1}\right) f_{xx}\left(\delta_{n,j}, W_{\sigma_n(j)}\right)\right) \\
&\quad \times \left(f_x\left(W_{n,j+1}\right) + \frac{1}{2}\left(X_{\sigma_n(j+1)} - \frac{j+1}{n+1}\right) f_{xx}\left(\delta_{n,j+1}, W_{\sigma_n(j+1)}\right)\right).
\end{aligned}$$

Finally, using the boundedness of f, f_x, and f_{xx}, together with (40), (41) follows. Remark that the proof of (41) yields also

$$\frac{1}{n}\sum_{j=1}^{n-1} \Delta_{n,j} = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right), \quad (43)$$

from which it is clear that this term will contribute in the CLT on Z_n. Then (41) entails that the asymptotic study reduces to that of the empirical mean of

$$Z_{n,j} = B_{n,j} + C_{n,j}$$

where

$$B_{n,j} := \begin{pmatrix} f\left(W_{n,j}\right) f\left(W_{n,j+1}\right) \\ f(W_{n,j}) \\ f(W_{n,j})^2 \end{pmatrix} \text{ and } C_{n,j} := \begin{pmatrix} \Delta_{n,j} f\left(W_{n,j+1}\right) + \Delta_{n,j+1} f\left(W_{n,j}\right) \\ \Delta_{n,j} \\ 2\Delta_{n,j} f(W_{n,j}) \end{pmatrix}. \quad (44)$$

First, we consider $B_{n,j}$ in (44) and we establish the following result, the proof of which has been postponed to Appendix B.3.

Lemma B.1. *As $n \to \infty$, the random vector B_n given by*

$$\frac{1}{n}\sum_{j=1}^{n-1} B_{n,j} = \frac{1}{n}\sum_{j=1}^{n-1} \left(f\left(W_{n,j}\right) f\left(W_{n,j+1}\right), f\left(W_{n,j}\right), f\left(W_{n,j}\right)^2\right)^\top$$

satisfies a CLT. More precisely,

$$\sqrt{n}\Big(B_n - m_B\Big) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_3(0, \Sigma_B),$$

where

$$m_B := \Big(\mathbb{E}[YY'], \mathbb{E}[Y], \mathbb{E}[Y^2]\Big)^\top, \quad (45)$$

$Y' = f(X, W')$, W' is an independent copy of W, and Σ_B has an explicit expression given in Appendix B.3.

Remark that Y' is the so-called Pick-Freeze version of Y with respect to X. Secondly, we establish a conditional CLT for the empirical mean of the $C_{n,j}$'s defined in (44). The reader is referred to Appendix B.4 for the proof of this result.

Lemma B.2. *There exists a measurable set $\Pi \in \Omega_W$ having $\mathbb{P}_W$-probability one such that, for any $\omega_W \in \Pi$, we have*

$$\sqrt{n} C_n(\cdot, \omega_W) \xrightarrow[n\to\infty]{\mathcal{L}_X} \mathcal{N}_3(0, \Sigma_C).$$

Moreover, Σ_C does not depend on ω_W and has an explicit expression given Appendix B.4.

Considering the characteristic function of the vector $\sqrt{n}(B_n - \mathbb{E}[B_n], C_n)$, one may write

$$\mathbb{E}\left[e^{i\left(\sqrt{n}\langle s, (B_n - \mathbb{E}[B_n])\rangle + \sqrt{n}\langle t, C_n\rangle\right)}\right] = \mathbb{E}\left[e^{i\sqrt{n}\langle s, (B_n - \mathbb{E}[B_n])\rangle} \mathbb{E}\left[e^{i\sqrt{n}\langle t, C_n\rangle}\Big|\mathcal{F}_W\right]\right]$$

for any s and $t \in \mathbb{R}^3$. On the one hand, $\mathbb{E}\left[e^{i\sqrt{n}\langle t, C_n\rangle}\Big|\mathcal{F}_W\right]$ converges almost surely to $\exp\{-t^\top \Sigma_C t/2\}$ which is not random. On the other hand, $\sqrt{n}\langle s, (B_n - \mathbb{E}[B_n])\rangle]$ converges in distribution to a Gaussian random variable denoted by B_s. By Slutsky's lemma,

$$\left(\sqrt{n}\langle s, (B_n - \mathbb{E}[B_n])\rangle], \mathbb{E}\left[e^{i\sqrt{n}\langle t, C_n\rangle}\Big|\mathcal{F}_W\right]\right)$$

converges in distribution to $(B_s, \exp\{-t^\top \Sigma_C t/2\})$. We consider the application $h\colon (u,v) \in \mathbb{R}\times D(0,1) \mapsto e^{iu}v \in \mathbb{C}$ where $D(0,1)$ is the unit disc in $\mathbb{C}$. The continuity and the boundedness of h lead to the convergence in distribution of $e^{i\sqrt{n}\langle s,(B_n-\mathbb{E}[B_n])\rangle}\left[e^{i\sqrt{n}\langle t,C_n\rangle}\middle|\mathcal{F}_W\right]$ and we conclude to the asymptotic normality of $\sqrt{n}(B_n-\mathbb{E}[B_n], C_n)$ to a six-dimensional Gaussian random vector with zero mean and variance-covariance matrix $\begin{pmatrix}\Sigma_B & 0\\ 0 & \Sigma_C\end{pmatrix}$. It remains to apply the so-called delta method [39, Theorem 3.1] and Slutsky's lemma to get the required result. The details of the computation of the asymptotic variance σ^2 can be found in Appendix B.5.

B.2 Technical results

B.2.1 Convergence of random measures

In the sequel, we will denote by $\mathcal{L}_Z$ the law of a random vector Z.

Lemma B.3. *There exists a measurable set $\Pi \subset \Omega_W$ with $\mathbb{P}_W$-probability one such that for any $\omega_W \in \Pi$,*

$$\pi_n(\omega_W) := \frac{1}{n}\sum_{j=1}^{n-2}\delta_{\left(\frac{j}{n+1},\frac{j+1}{n+1},\frac{j+2}{n+1},W_j(\omega_W),W_{j+1}(\omega_W)\right)} \Rightarrow \pi := \mathcal{L}_{(X,X,X)}\otimes\mathcal{L}_W\otimes\mathcal{L}_W,$$

as $n\to\infty$ where as before X is uniformly distributed on $[0,1]$ and $\Rightarrow$ stands for the weak convergence of measures.

Proof of Lemma B.3. Let $\omega_W \in \Omega_W$. Let us consider the continuous and bounded functions defined on $\mathbb{R}^5$ by

$$g_{s,s',s'',t,t'}(x,x',x'',w,w') = \exp\{i(sx+s'x'+s''x''+tw+t'w')\},$$

for any s, s', s'', t, and t' real numbers. To prove the weak convergence of the measures $(\pi_n(\omega_W))_n$, we show that $\pi_n(\omega_W)(g_{s,s',s'',t,t'})$ converges almost surely for any s, s', s'', t, and $t' \in \mathbb{Q}$ as $n\to\infty$. Finally, we will conclude by density of rational numbers in $\mathbb{R}$.
Let $(s,s',s'',t,t') \in \mathbb{Q}^5$ be fixed. To ease the reading, we use the shorthand notation g for $g_{s,s',s'',t,t'}$ and we omit the notation ω_W as classically done in probability.

One has

$$\pi_n(g) = \int g d\pi_n = \frac{1}{n}\sum_{j=1}^{n-2} e^{i\left(s\frac{j}{n+1}+s'\frac{j+1}{n+1}+s''\frac{j+2}{n+1}+tW_j+t'W_{j+1}\right)}.$$

Obviously, by the independence of the sequence W_n and the convergence theorem of Riemann sums,

$$\begin{aligned}\mathbb{E}\left[\pi_n(g)\right] &= \mathbb{E}\left[e^{itW}\right]\mathbb{E}\left[e^{it'W}\right]\frac{1}{n}\sum_{j=1}^{n-2} e^{i\left(s\frac{j}{n+1}+s'\frac{j+1}{n+1}+s''\frac{j+2}{n+1}\right)}\\ &\underset{n\to\infty}{\to} \mathbb{E}\left[e^{itW}\right]\mathbb{E}\left[e^{it'W}\right]\int_0^1 e^{i(s+s'+s'')x}dx.\end{aligned}$$

Observe that the almost sure convergence of π_n is equivalent to the almost sure convergence of its real part and that of its imaginary part. Setting

$$U_{n,j} = s\frac{j}{n+1} + s'\frac{j+1}{n+1} + s''\frac{j+2}{n+1} + tW_j + t'W_{j+1},$$

we have

$$\Re(\pi_n(g)) = \frac{1}{n}\sum_{j=1}^{n-2} \cos\left(U_{n,j}\right).$$

In order to apply the Borel-Cantelli lemma, we need to control the fourth moment

$$\mathbb{E}\left[\left(\Re(\pi_n(g)) - \mathbb{E}[\Re(\pi_n(g))]\right)^4\right] = \frac{1}{n^4}\mathbb{E}\left[\left(\sum_{j=1}^{n-2}\cos\left(U_{n,j}\right) - \mathbb{E}[\cos\left(U_{n,j}\right)]\right)^4\right].$$

The random variables $\cos\left(U_{n,j}\right) - \mathbb{E}\left[\cos\left(U_{n,j}\right)\right]$ are real-valued, centered, and bounded so that we can apply inequality (2.14) page 37 in [30]. Then we obtain

$$\mathbb{E}\left[\left(\sum_{j=1}^{n-2}\cos\left(U_{n,j}\right) - \mathbb{E}\left[\cos\left(U_{n,j}\right)\right]\right)^4\right] \leqslant 224n^2\left(\Lambda_2(\alpha^{-1})\right)^2 \tag{46}$$

where

$$\Lambda_2(\alpha^{-1}) = \sup_{0\leqslant m<n}(m+1)(\alpha_m)^{\frac{1}{2}},$$

where $(\alpha_m)_m$ is the sequence f the strong mixing coefficients of the sequence $(U_{n,j})$. Now since the random variable Z_j^n only depends on (W_j, W_{j+1}), α_m equal zero as soon as $m \geqslant 2$. Hence, there exists a positive constant K such that

$$\frac{1}{n^4}\mathbb{E}\left[\left(\sum_{j=1}^{n-2}\cos\left(U_{n,j}\right) - \mathbb{E}\left[\cos\left(U_{n,j}\right)\right]\right)^4\right] \leqslant \frac{K}{n^2}.$$

It follows by Borel-Cantelli lemma that the real part of $\pi_n(g)$ converges almost surely. Since the imaginary part can be treated using the exact same steps, the proof of Lemma B.3 is almost complete. Hence, there exists a Borel set $N_{s,s',s'',t,t'}$ with $\mathbb{P}(N_{s,s',s'',t,t'}) = 1$ so that the previous convergence holds on $\Omega_W \backslash N_{s,s',s'',t,t'}$. It remains to define $\Pi := \Omega_W \backslash \cup_{(s,s',s'',t,t')\in\mathbb{Q}^5} N_{s,s',s'',t,t'}$. Obviously, one has $\mathbb{P}(\Pi) = 1$ and the almost sure convergence holds on Π for all functions $g_{s,s',s'',t,t'}$ with $(s, s', s'', t, t') \in \mathbb{Q}^5$.

Finally, the result holding for any five-uplet $(s, s', s'', t, t') \in \mathbb{Q}^5$, we conclude to the required result by density of rational numbers in $\mathbb{R}$. □

The obvious following corollary is a direct consequence of Lemma B.3.

Corollary B.4. *We use the notation of Lemma B.3. For any $\omega_W \in \Pi$, as $n \to \infty$,*

$$\eta_n := \frac{1}{n}\sum_{j=1}^{n-1} \delta_{\left(\frac{j}{n+1}, \frac{j+1}{n+1}\right)} \Rightarrow \eta := \mathcal{L}_{(X,X)},$$

$$\kappa_n := \frac{1}{n}\sum_{j=1}^{n-1} \delta_{\left(\frac{j}{n+1}, \frac{j+1}{n+1}, \frac{j+2}{n+1}\right)} \Rightarrow \kappa := \mathcal{L}_{(X,X,X)},$$

$$\mu_n(\omega_W) := \frac{1}{n}\sum_{j=1}^{n} \delta_{\left(\frac{j}{n+1}, W_j(\omega_W)\right)} \Rightarrow \mu := \mathcal{L}_X \otimes \mathcal{L}_W,$$

$$\nu_n(\omega_W) := \frac{1}{n}\sum_{j=1}^{n-1} \delta_{\left(\frac{j}{n+1}, \frac{j+1}{n+1}, W_j(\omega_W), W_{j+1}(\omega_W)\right)} \Rightarrow \nu := \mathcal{L}_{(X,X)} \otimes \mathcal{L}_W \otimes \mathcal{L}_W.$$

B.2.2 Generalized L-Statistics

Lemma B.5. *Let $(E_i)_{i\geqslant 1}$ be a sequence of i.i.d. random variables with standard exponential distribution and let ψ be a bounded measurable function on $[0,1]$. We assume that the set of discontinuity points of ψ has null Lebesgue measure. Then, the sequence*

$$\left(n^{-1/2}\sum_{j=1}^{n-1} \psi(j/n)(E_j - 1)\right)_{n\in\mathbb{N}^*}$$

converges in distribution towards a centered Gaussian law. The asymptotic variance is $\sigma_\psi^2 := \int_{[0,1]} \psi^2(x)dx$.

Proof of Lemma B.5. For $k \in \mathbb{N}^*$, let cum_k denotes the cumulant of order k of

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n-1} \psi(j/n)(E_j - 1).$$

Obviously, $\mathrm{cum}_1 = 0$ and, for $k \geqslant 2$,

$$\mathrm{cum}_k = \frac{1}{n\sqrt{n^{k-2}}}\sum_{j=1}^{n-1} \left(\psi(j/n)\right)^k.$$

So that, $\lim_{n\to\infty} \mathrm{cum}_2 = \int \psi^2(x)dx$ while, for $k \geqslant 3$, $\lim_{n\to\infty} \mathrm{cum}_k = 0$. □

Remark B.6. The previous lemma obviously extends to the case of a continuous function $\Psi = (\psi_i)$ valued in $\mathbb{R}^d$ ($d \geqslant 1$). In this case, the asymptotic covariance matrix Σ_Ψ is the Gram matrix $\left(\int_{[0,1]} \psi_i(x)\psi_j(x)dx;\ 1 \leqslant i,j \leqslant d\right)$. Indeed, the previous lemma holds for any linear combination of such random vector sequence. A direct computation of the asymptotic variance leads to the quadratic form built on Σ_Ψ.

The next lemma is a generalization of the CLT for a L-statistics (see, e.g., [39, Chapter 22]).

Lemma B.7. *Let $(U, \mathbb{B}(U))$ be a Polish space where $\mathbb{B}(U)$ denotes the Borel σ algebra of U. We consider a sequence $(\chi_j)_{1\leqslant j\leqslant n,\, n\in\mathbb{N}^*}$ valued in U and Q a probability measure on $U \times [0,1]$. We assume that the sequence of empirical measures $\left(\frac{1}{n}\sum_{j=1}^{n-1} \delta_{\chi_j, j/n}\right)_{n\in\mathbb{N}^*}$ converges in distribution to Q.*

Let ψ be a bounded measurable real function on $U \times [0,1]$. We assume that the set of discontinuity points of ψ has null Q-probability. Then,

$$D_n := \frac{1}{\sqrt{n}} \sum_{j=1}^{n-1} \psi\left(\chi_j, j/n\right) \left(X_{\sigma_n(j)} - \frac{j}{n+1} \right) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}\left(0, s_\psi^2\right).$$

where the asymptotic variance s_ψ^2 is given in (48)*.*

Proof of Lemma B.7. Recall that the sequence (E_i) has been defined in Lemma B.5. We have

$$\begin{aligned} X_{\sigma_n(j)} - \frac{j}{n+1} &\overset{\mathcal{L}}{=} \frac{\sum_{i=1}^{j} E_i}{\sum_{i=1}^{n+1} E_i} - \frac{j}{n+1} \\ &= \frac{1}{\frac{1}{n+1}\sum_{i=1}^{n+1} E_i} \left(\frac{1}{n+1} \sum_{i=1}^{j} E_i - \frac{j}{(n+1)^2} \sum_{i=1}^{n+1} E_i \right) \\ &= \frac{1}{\frac{1}{n+1}\sum_{i=1}^{n+1} E_i} \left(\frac{1}{n+1} \sum_{i=1}^{j} (E_i - 1) - \frac{j}{(n+1)^2} \sum_{i=1}^{n+1} (E_i - 1) \right), \end{aligned}$$

so that,

$$\begin{aligned} D_n \overset{\mathcal{L}}{=} \frac{1}{\sqrt{n}(n+1)} & \frac{1}{\frac{1}{n+1}\sum_{i=1}^{n+1} E_i} \\ & \sum_{j=1}^{n-1} \psi\left(\chi_j, j/n\right) \left(\sum_{i=1}^{j} (E_i - 1) - \frac{j}{n+1} \sum_{i=1}^{n+1} (E_i - 1) \right). \end{aligned}$$

Using the assumption on the empirical measure, we get

$$\frac{1}{n} \sum_{j=1}^{n} \psi\left(\chi_j, j/n\right) \frac{j}{n+1} \to I := \int_{U\times[0,1]} x\psi(\chi, x) dQ(\chi, x).$$

Further, by the weak law of large numbers, $(1/(n+1)) \sum_{i=1}^{n+1} E_i$ converges in probability to $\mathbb{E}[E_1] = 1$. Hence, by Slutsky's lemma, we are led to consider the random vector

$$V_n := \frac{1}{\sqrt{n}} \begin{pmatrix} \frac{1}{n+1} \sum_{j=1}^{n-1} \psi\left(\chi_j, j/n\right) \sum_{i=1}^{j} (E_i - 1) \\ \sum_{i=1}^{n+1} (E_i - 1) \end{pmatrix}.$$

Notice that the first coordinate of V_n can be rewritten as (up to the normalizing factor $n^{-1/2}$)

$$\sum_{i=1}^{n-1} \left(\frac{1}{n+1} \sum_{j=1}^{n-1} \psi\left(\chi_j, j/n\right) \mathbb{1}_{i \leqslant j} \right) (E_i - 1).$$

For $t \in [0,1]$, let $\phi(t) := \int_{U\times[t,1]} \psi(\chi, x) dQ(\chi, x)$. We will show below that

$$\lim_n \sup_{t\in[0,1]} \left| \left(\frac{1}{n+1} \sum_{j=1}^{n-1} \psi\left(\chi_j, j/n\right) \mathbb{1}_{i \leqslant j} \right) - \phi(t) \right| = 0. \tag{47}$$

Let assume for a while that this result holds. Then, in our study, we may replace V_n by

$$\widehat{V}_n := \frac{1}{\sqrt{n}} \begin{pmatrix} \frac{1}{n+1} \sum_{i=1}^{n-1} \phi(i/n)(E_i - 1) \\ \sum_{i=1}^{n+1} (E_i - 1) \end{pmatrix}$$

since (47) implies that $\lim_{n\to\infty} \mathbb{E}\|V_n - \widehat{V}_n\|^2 = 0$. Using Remark B.6, we obtain that the sequence $(\widehat{V}_n)_{n\in\mathbb{N}^*}$ converges in distribution to a centered Gaussian vector with covariance matrix

$$\begin{pmatrix} \int_0^1 \phi^2(t)dt & \int_0^1 \phi(t)dt \\ \int_0^1 \phi(t)dt & 1 \end{pmatrix}.$$

Finally, using the so-called delta method [39, Theorem 3.1], $(D_n)_{n\in\mathbb{N}^*}$ converges in distribution to a centered Gaussian variable with variance

$$s_\psi^2 = \int_0^1 (\phi(t) - I)^2 dt. \tag{48}$$

It remains to show that (47) holds. First let assume that $\psi \geqslant 0$. Set, for $j = 1, \dots n$, $\phi_n(j/n) := (1/(n+1)) \sum_{j=1}^{n-1} \psi(\chi_j, j/n)$ and consider the piece-wise linear extension ϕ_n defined on $[0,1]$. The second Dini's theorem [31] allows to conclude that the sequence of functions $(\phi_n)_{n\in\mathbb{N}^*}$ converges uniformly to ϕ yielding the result. In the general case, we may mimic this reasoning on $\psi^+ = \sup(\psi, 0)$ and $\psi^- = \sup(-\psi, 0)$ and so conclude. □

Notice that, using the definitions of ϕ and I and applying Fubini's theorem, s_ψ^2 can be explicited as follows:

$$\begin{aligned}
s_\psi^2 &= \int_0^1 (\phi(t) - I)^2 dt \\
&= \int_0^1 \left(\int_{U\times[0,1]} \psi(\chi, x)(\mathbb{1}_{t\leqslant x} - x) dQ(\chi, x) \right)^2 dt \\
&= \int_0^1 \iint_{(U\times[0,1])^2} \psi(\chi_1, x_1)\psi(\chi_2, x_2)(\mathbb{1}_{t\leqslant x_1} - x_1)(\mathbb{1}_{t\leqslant x_2} - x_2) dQ(\chi_1, x_1) dQ(\chi_2, x_2) dt \\
&= \iint_{(U\times[0,1])^2} \psi(\chi_1, x_1)\psi(\chi_2, x_2) \int_0^1 (\mathbb{1}_{t\leqslant x_1} - x_1)(\mathbb{1}_{t\leqslant x_2} - x_2) dt dQ(\chi_1, x_1) dQ(\chi_2, x_2) \\
&= \iint_{(U\times[0,1])^2} \psi(\chi_1, x_1)\psi(\chi_2, x_2)(x_1 \wedge x_2 - x_1 x_2) dQ(\chi_1, x_1) dQ(\chi_2, x_2).
\end{aligned} \tag{49}$$

B.3 Proof of Lemma B.1

One has

$$\mathbb{E}[B_n] = \frac{1}{n} \sum_{j=1}^{n-1} \left(\mathbb{E}\left[f\left(W_{n,j}\right) f\left(W_{n,j+1}\right)\right], \mathbb{E}\left[f\left(W_{n,j}\right)\right], \mathbb{E}\left[f\left(W_{n,j}\right)^2\right] \right)^\top,$$

the first coordinate of which converges as $n \to \infty$ to

$$\begin{aligned}
\int_0^1 \mathbb{E}\left(f\left(x, W\right) f\left(x', W'\right)\right] \mathrm{d}\eta(x, x') &= \mathbb{E}\left[\mathbb{E}\left[f\left(X, W\right) f\left(X, W'\right) | X\right]\right] \\
&= \mathbb{E}\left[f\left(X, W\right) f\left(X, W'\right)\right] = \mathbb{E}\left[YY'\right].
\end{aligned}$$

The two other coordinates can be handled similarly leading to

$$\mathbb{E}[B_n] \underset{n\to\infty}{\to} \left(\mathbb{E}[YY'], \mathbb{E}[Y], \mathbb{E}[Y^2]\right)^\top = m_B.$$

We apply the CLT for dependent variables proved in [26] to $\widetilde{B}_{n,j}^1$, the centered version of the random variables $f\left(W_{n,j}\right) f\left(W_{n,j+1}\right)/\sqrt{n}$ with $m = 1$, α=0, and because f is

bounded (so is $\widetilde{B}^1_{n,j}$). Assumptions (1) and (2) in [26] obviously hold, the assumption (3) is naturally fulfilled and assumption (4) is a mere consequence of Chebyshev's inequality and the boundedness of f. Now, it remains to check that assumption (5) holds. We have

$$\begin{aligned}
&\sum_{i,j=1}^{n-1} \mathrm{Cov}(\widetilde{B}^1_{n,i}, \widetilde{B}^1_{n,j}) \\
&\quad = \frac{1}{n} \sum_{i,j=1}^{n-1} \mathrm{Cov}\left(f\left(W_{n,i}\right) f\left(W_{n,i+1}\right), f\left(W_{n,j}\right) f\left(W_{n,j+1}\right)\right) \\
&\quad = \frac{1}{n} \sum_{j=1}^{n-1} \mathrm{Var}\left(f\left(W_{n,j}\right) f\left(W_{n,j+1}\right)\right) \\
&\qquad\qquad + \frac{1}{n} \sum_{j=1}^{n-2} \mathrm{Cov}\left(f\left(W_{n,j}\right) f\left(W_{n,j+1}\right), f\left(W_{n,j+1}\right) f\left(W_{n,j+2}\right)\right).
\end{aligned}$$

On the one hand, by Corollary B.4,

$$\begin{aligned}
\frac{1}{n} \sum_{j=1}^{n-1} \mathrm{Var}\left(f\left(W_{n,j}\right) f\left(W_{n,j+1}\right)\right) &\underset{n\to\infty}{\to} \int \mathrm{Var}\left(f\left(x, W\right) f\left(x', W'\right)\right) \mathrm{d}\eta(x, x') \\
&= \int_0^1 \mathrm{Var}\left(f\left(x, W\right) f\left(x, W'\right)\right) \mathrm{d}x \\
&= \mathbb{E}\left[\mathrm{Var}\left(f\left(X, W\right) f\left(X, W'\right) | X\right)\right] \\
&= \mathbb{E}\left[\mathrm{Var}\left(YY' | X\right)\right],
\end{aligned}$$

where W' is an independent copies of W, $Y = f(X, W)$ and $Y' = f(X, W')$. On the other hand, by Corollary B.4,

$$\begin{aligned}
&\frac{1}{n} \sum_{j=1}^{n-2} \mathrm{Cov}\left(f\left(W_{n,j}\right) f\left(W_{n,j+1}\right), f\left(W_{n,j+1}\right) f\left(W_{n,j+2}\right)\right) \\
&\underset{n\to\infty}{\to} \mathbb{E}\left[\mathrm{Cov}\left(f\left(X, W\right) f\left(X, W'\right), f\left(X, W'\right) f\left(X, W''\right) | X\right)\right] = \mathbb{E}\left[\mathrm{Cov}\left(YY', YY'' | X\right)\right],
\end{aligned}$$

where W' and W'' are two independent copies of W. Further, $Y = f(X, W)$, $Y' = f(X, W')$, and $Y'' = f(X, W'')$. Actually, notice that all linear combination of the coordinates of

$$\left(f(W_{n,j})f(W_{n,j+1}), f(W_{n,j}), f(W_{n,j})^2\right)^\top \tag{50}$$

is a one-dependent random variable. In addition, following the same lines as above, one may check that any linear combination still satisfies the assumptions of [26]. Hence, any linear combination of the coordinates of B_n satisfies a CLT so that Lemma B.1 is proved, up to the computation of the asymptotic variance-covariance matrix Σ_B done in the paragraph that follows.

Computation of the asymptotic covariance matrix Σ_B

We consider a linear combination of the random vector in (50) given by

$$uf(W_{n,j})f(W_{n,j+1}) + vf(W_{n,j}) + wf(W_{n,j})^2,$$

where $(u, v, w) \in \mathbb{R}^3$. This one-dimensional random vector is one-dependent and its centered version normalized by $\sqrt{n}$, denoted by $\widetilde{B}_{n,j}$, satisfies the assumptions of [26]. To calculate the asymptotic variance-covariance matrix Σ_B, we compute explicitly the limit of

$$\sum_{i,j=1}^{n-1} \mathrm{Cov}(\widetilde{B}_{n,i}, \widetilde{B}_{n,j}),$$

as $n \to \infty$ using Corollary B.4. It remains to take $(u, 0, 0)$, $(0, v, 0)$ and $(0, 0, w)$ to get the diagonal terms of the asymptotic variance-covariance matrix and to solve a three-dimensional system of equations to get the remaining terms. Finally, as computed previously and using notation of Corollary B.4, the first diagonal term of Σ_B is :

$$\begin{aligned}
\Sigma_B^{1,1} &= \int \mathrm{Var}\left(f\left(x, W\right) f\left(x', W'\right)\right) \mathrm{d}\eta(x, x') \\
&\quad + \int \mathrm{Cov}\left(f\left(x, W\right) f\left(x', W'\right), f\left(x', W'\right) f\left(x'', W''\right)\right) \mathrm{d}\kappa(x, x', x'') \\
&= \int_0^1 \mathrm{Var}\left(f\left(x, W\right) f\left(x, W'\right)\right) \mathrm{d}x \\
&\quad + \int_0^1 \mathrm{Cov}\left(f\left(x, W\right) f\left(x, W'\right), f\left(x, W'\right) f\left(x, W''\right)\right) \mathrm{d}x \\
&= \mathbb{E}\left[\mathrm{Var}\left(f\left(X, W\right) f\left(X, W'\right) | X\right)\right] + \mathbb{E}\left[\mathrm{Cov}\left(f\left(X, W\right) f\left(X, W'\right), f\left(X, W'\right) f\left(X, W''\right) | X\right)\right] \\
&= \mathbb{E}\left[\mathrm{Var}\left(YY' | X\right)\right] + \mathbb{E}\left[\mathrm{Cov}\left(YY', YY'' | X\right)\right],
\end{aligned}$$

where we remind that $Y = f(X, W)$, $Y' = f(X, W')$, and $Y'' = f(X, W'')$ with W' and W'' independent copies of W. The other terms are

$$\begin{aligned}
\Sigma_B^{2,2} &= \int_0^1 \mathrm{Var}\left(f\left(x, W\right)\right) \mathrm{d}x = \mathbb{E}\left[\mathrm{Var}\left(f\left(X, W\right) | X\right)\right] = \mathbb{E}\left[\mathrm{Var}(Y | X)\right], \\
\Sigma_B^{3,3} &= \int_0^1 \mathrm{Var}\left(f\left(x, W\right)^2\right) \mathrm{d}x = \mathbb{E}\left[\mathrm{Var}\left(Y^2 | X\right)\right], \\
\Sigma_B^{1,2} &= \Sigma_B^{2,1} = 2\int_0^1 \mathrm{Cov}\left(f\left(x, W\right) f\left(x, W'\right), f\left(x, W\right)\right) \mathrm{d}x = 2\mathbb{E}\left[\mathrm{Cov}\left(YY', Y | X\right)\right], \\
\Sigma_B^{1,3} &= \Sigma_B^{3,1} = 2\int_0^1 \mathrm{Cov}\left(f\left(x, W\right) f\left(x, W'\right), f\left(x, W\right)^2\right) \mathrm{d}x = 2\mathbb{E}\left[\mathrm{Cov}\left(YY', Y^2 | X\right)\right], \\
\Sigma_B^{2,3} &= \Sigma_B^{3,2} = \int_0^1 \mathrm{Cov}\left(f\left(x, W\right), f\left(x, W\right)^2\right) \mathrm{d}x = \mathbb{E}\left[\mathrm{Cov}(Y, Y^2 | X)\right].
\end{aligned}$$

B.4 Proof of Lemma B.2

Let $\omega_W \in \Pi$ as defined in Lemma B.3. The aim is to establish a CLT for $\sqrt{n}C_{n,j}(\cdot, \omega_W)$. To ease the reading, we omit the notation $(\cdot, \omega_W)$ as classically done in probability. First, dealing with the first coordinate of $C_{n,j}$ defined in (44), one has

$$\begin{aligned}
&f\left(W_{n,j+1}\right) \Delta_{n,j} + f\left(W_{n,j}\right) \Delta_{n,j+1} \\
&\qquad = \left(f\left(W_{n,j}\right) + f\left(W_{n,j+1}\right)\right) \Delta_{n,j} + f\left(W_{n,j}\right) \left(\Delta_{n,j+1} - \Delta_{n,j}\right) \\
&\qquad = \left(X_{\sigma_n(j)} - \frac{j}{n+1}\right) \left(f\left(W_{n,j}\right) + f\left(W_{n,j+1}\right)\right) f_x\left(W_{n,j}\right) \\
&\qquad\quad + \frac{1}{2}\left(X_{\sigma_n(j)} - \frac{j}{n+1}\right)^2 \left(f\left(W_{n,j}\right) + f\left(W_{n,j+1}\right)\right) f_{xx}\left(\delta_{n,j}, W_j\right) \\
&\qquad\quad + f\left(W_{n,j}\right) \left(\Delta_{n,j+1} - \Delta_{n,j}\right)
\end{aligned}$$

using the expansion of $\Delta_{n,j}$ given in (42). By Lemma 40 and using the boundedness of f and f_{xx}, we get that

$$\frac{1}{n}\sum_{j=1}^{n-1}\left(X_{\sigma_n(j)}-\frac{j}{n+1}\right)^2\left(f\left(W_{n,j}\right)+f\left(W_{n,j+1}\right)\right)f_{xx}\left(\delta_{n,j},W_j\right)$$

is $O_{\mathbb{P}}\left(1/n\right)$. The same asymptotic behavior is observed for the telescopic sum

$$\frac{1}{n}\sum_{j=1}^{n-1}f\left(W_{n,j}\right)\left(\Delta_{n,j+1}-\Delta_{n,j}\right).$$

So that, using also the expansion of $\Delta_{n,j}$ given in (42), Lemma 40, and the boundedness of f and f_{xx}, the study of C_n reduces to that of the random vector

$$\frac{1}{n}\sum_{j=1}^{n-1}\left(X_{\sigma_n(j)}-\frac{j}{n+1}\right)f_x\left(W_{n,j}\right)\begin{pmatrix}f\left(W_{n,j}\right)+f\left(W_{n,j+1}\right)\\1\\2f\left(W_{n,j+1}\right)\end{pmatrix} \tag{51}$$

by the independence between σ_n and $W_1,\ldots,W_n$. In that view, let us consider the following linear combination

$$u(f(W_{n,j})+f(W_{n,j+1}))+v+w2f(W_{n,j+1}),$$

where $(u,v,w)\in\mathbb{R}^3$ and the empirical mean

$$\frac{1}{n}\sum_{j=1}^{n-1}\left(X_{\sigma_n(j)}-\frac{j}{n+1}\right)f_x\left(W_{n,j}\right)\\ \times\left(u(f(W_{n,j})+f(W_{n,j+1}))+v+w2f(W_{n,j+1})\right). \tag{52}$$

Now it remains to apply Lemma B.7[1] with $\chi_j=(W_j,W_{j+1})$ and $\psi=\psi_{uvw}$ with

$$\psi_{uvw}\left(\chi_j,\frac{j}{n+1},\frac{j+1}{n+1}\right)=f_x\left(W_{n,j}\right)\left(u(f(W_{n,j})+f(W_{n,j+1}))+v+w2f(W_{n,j+1})\right), \tag{53}$$

noticing that, as $n\to\infty$, $(1/n)\sum_{j=1}^{n-1}\delta_{\chi_j,j/(n+1),(j+1)/(n+1)}$ converges in distribution to $Q=\nu=\mathcal{L}_{(X,X)}\otimes\mathcal{L}_W\otimes\mathcal{L}_W$ by Corollary B.4. Thus we deduce that the empirical mean in (52) converges in distribution for any 3-uplet (u,v,w). Since any linear combination of the components of the random vector defined in (51) satisfies a CLT, so does the random vector itself. The proof of Lemma B.2 is now complete, up to the computation of the asymptotic variance-covariance matrix Σ_C done in the paragraph that follows.

Computation of the asymptotic covariance matrix Σ_C

We use the explicited expression (49) of the asymptotic variance σ_ψ^2 of Lemma B.7 (actually its slightly generalized version) with $Q=\nu=\mathcal{L}_{(X,X)}\otimes\mathcal{L}_W\otimes\mathcal{L}_W$ and with ψ given by (53). Then taking the values $(u,0,0)$, $(0,v,0)$ and $(0,0,w)$ leads to the diagonal terms of

[1] A slightly generalization of Lemma B.7 is required to handle the pair $(j/(n+1),(j+1)/(n+1))$ rather than the quantity j/n. Its proof comes directly following the same lines as in the proof of Lemma B.7

the asymptotic variance-covariance matrix Σ_C while solving a three-dimensional system of equations provides the remaining terms. For instance, reminding that $\chi_j = (W_j, W_{j+1})$ and $W_{n,j} = (W_j, j/(n+1))$ and

$$\psi_{100}\left(\chi_j, \frac{j}{n+1}, \frac{j+1}{n+1}\right) = f_x\left(W_{n,j}\right)\left(f(W_{n,j}) + f(W_{n,j+1})\right)$$

(namely, ψ with $(u, v, w) = (u, 0, 0)$), we have

$$\begin{aligned}\Sigma_C^{1,1} =& \iint \psi_1(\chi_1, x_1, x_1')\psi_1(\chi_2, x_2, x_2')x_1 \wedge x_2 \wedge x_1' \wedge x_2' \mathrm{d}\nu(\chi_1, x_1, x_1')\mathrm{d}\nu(\chi_2, x_2, x_2') \\ & - \left(\int \psi_1(\chi, x, x')x \mathrm{d}\nu(\chi, x, x')\right)^2 \\ =& \mathbb{E}[(Y_1 + Y_1')(Y_2 + Y_2')f_x(X_1, W_1)f_x(X_2, W_2)(X_1 \wedge X_2)] - \mathbb{E}[(Y + Y')f_x(X, W)X]^2.\end{aligned}$$

Finally, the other diagonal terms of Σ_C are:

$$\begin{aligned}\Sigma_C^{2,2} =& \mathbb{E}[f_x(X_1, W_1)f_x(X_2, W_2)(X_1 \wedge X_2)] - \mathbb{E}[f_x(X, W)X]^2 \\ \Sigma_C^{3,3} =& 4\mathbb{E}[Y_1'Y_2'f_x(X_1, W_1)f_x(X_2, W_2)(X_1 \wedge X_2)] - 4\mathbb{E}[Y'f_x(X, W)X]^2\end{aligned}$$

while the remaining terms are

$$\begin{aligned}\Sigma_C^{1,2} = \Sigma_C^{2,1} =& \mathbb{E}[(Y_1 + Y_1')f_x(X_1, W_2)f_x(X_2, W_2)(X_1 \wedge X_2)] \\ & - \mathbb{E}[(Y + Y')f_x(X, W)X]\mathbb{E}[f_x(X, W)X] \\ \Sigma_C^{1,3} = \Sigma_C^{3,1} =& 2\mathbb{E}[(Y_1 + Y_1')f_x(X_1, W_1)Y_2'f_x(X_2, W_2)(X_1 \wedge X_2)] \\ & - 2\mathbb{E}[(Y + Y')f_x(X, W)X]\mathbb{E}[Y'f_x(X, W)X] \\ \Sigma_C^{2,3} = \Sigma_C^{3,2} =& 2\mathbb{E}[f_x(X_1, W_1)Y_2'f_x(X_2, W_2)(X_1 \wedge X_2)] - 2\mathbb{E}[f_x(X, W)X]\mathbb{E}[Y'f_x(X, W)X].\end{aligned}$$

B.5 Asymptotic variance σ^2 of Theorem 3.3

We have proved yet that

$$\sqrt{n}\left(\begin{pmatrix} B_n \\ C_n \end{pmatrix} - \begin{pmatrix} m_B \\ 0 \end{pmatrix}\right) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_6\left(0, \begin{pmatrix} \Sigma_B & 0 \\ 0 & \Sigma_C \end{pmatrix}\right),$$

where the explicit expressions of m_B, Σ_B and Σ_C are given in (45) of Lemma B.1, Appendices B.3 and B.4 respectively. Applying the so-called delta method [39, Theorem 3.1] to the linear function $f(x, y) = x + y$, we conclude that

$$\sqrt{n}(Z_n - m_B) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_3\left(0, \Sigma_B + \Sigma_C\right),$$

Further, we notice that

$$\xi_n^{\mathrm{Sobol}}(X, Y) \overset{\mathcal{L}}{=} \Psi(Z_n)$$

with $\Psi(x, y, z) = (x - y^2)/(z - y^2)$. The so-called delta method [39, Theorem 3.1] then gives

$$\sqrt{N}\left(\xi_n^{\mathrm{Sobol}}(X, Y) - S^X\right) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_1(0, \sigma^2)$$

where $S^X = \mathrm{Var}(\mathbb{E}[Y|X])/\mathrm{Var}(Y)$ is the first-order Sobol index with respect to X and $\sigma^2 = g^T(\Sigma_B + \Sigma_C)g$ with

$$g = \nabla\Psi(m_B).$$

By assumption $\mathrm{Var}(Y) \neq 0$, Ψ is differentiable at m and we will see in the sequel that $g^T(\Sigma_B + \Sigma_C)g \neq 0$, so that the application of the delta method is justified. By differentiation, we get that, for any x, y, and z so that $z \neq y^2$:

$$\nabla\Psi(x, y, z) = \left(\frac{1}{z - y^2}, -2y\frac{z - x}{(z - y^2)^2}, -\frac{x - y^2}{(z - y^2)^2}\right)^T$$

so that

$$\begin{aligned} g = \nabla\Psi(m_B) &= \left(\frac{1}{\mathrm{Var}(Y)}, 2\mathbb{E}[Y]\frac{\mathbb{E}[YY'] - E[Y^2]}{\mathrm{Var}(Y)^2}, -\frac{S^X}{\mathrm{Var}(Y)}\right)^\top, \\ &= \frac{1}{\mathrm{Var}(Y)}\left(1, 2\mathbb{E}[Y](S^X - 1), -S^X\right)^\top. \end{aligned} \tag{54}$$

Hence the asymptotic variance σ^2 in Theorem 3.3 is finally given by

$$\sigma^2 = g^T\left(\Sigma_B + \Sigma_C\right)g$$

where g has been defined above in (54), Σ_B and Σ_C have been defined in Appendices B.3 and B.4 respectively. The matrix Σ_B rewrites as

$$\Sigma_B = \begin{pmatrix} v_{01} + c_{01,02} & 2c_{01,03} & 2c_{01,00} \\ 2c_{01,03} & \mathrm{Var}(Y)(1 - S^X) & 2c_{03,00} \\ 2c_{01,00} & 2c_{03,00} & v_{00} \end{pmatrix}$$

where

$$\begin{aligned} v_{ij} &= \mathbb{E}[\mathrm{Var}(A_iA_j|X)] \\ c_{ij,kl} &= \mathbb{E}[\mathrm{Cov}(A_iA_j, A_kA_l|X)] \end{aligned}$$

and $A_0 = Y$, $A_1 = Y'$, $A_2 = Y''$, and $A_3 = 1$ (Y and Y'' have been defined just before (50)). The matrix Σ_C rewrites as

$$\Sigma_C = \begin{pmatrix} s^2_{\psi 100} & s^2_{\psi 110} & s^2_{\psi 101} \\ s^2_{\psi 110} & s^2_{\psi 010} & s^2_{\psi 011} \\ s^2_{\psi 101} & s^2_{\psi 011} & s^2_{\psi 001} \end{pmatrix}$$

where s^2_ψ and ψ_{uvw} have been defined in (48) and (53) respectively. The proof of Theorem 3.3 is now complete.

References

[1] A. Alexanderian, P. Gremaud, and R. Smith. Variance-based sensitivity analysis for time-dependent processes. *Reliability Eng. Sys. Safety*, 196:106722, 2020.

[2] A. Antoniadis. Analysis of variance on function spaces. *Statistics: A Journal of Theoretical and Applied Statistics*, 15(1):59–71, 1984.

[3] M. Azadkia and S. Chatterjee. A simple measure of conditional dependence. *arXiv preprint arXiv:1910.12327*, 2019.

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

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

[6] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.

[7] S. Boucheron, G. Lugosi, P. Massart, et al. On concentration of self-bounding functions. *Electronic Journal of Probability*, 14:1884–1899, 2009.

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

[9] S. Chatterjee. A new coefficient of correlation. *Journal of the American Statistical Association*, (just-accepted):1–26, 2020.

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

[11] E. De Rocquigny, N. Devictor, and S. Tarantola. *Uncertainty in industrial practice*. Wiley Online Library, 2008.

[12] H. Dette, K. F. Siburg, and P. A. Stoimenov. A copula-based non-parametric measure of regression dependence. *Scandinavian Journal of Statistics*, 40(1):21–41, 2013.

[13] J.-C. Fort, T. Klein, and N. Rachdi. New sensitivity analysis subordinated to a contrast. *Communications in Statistics - Theory and Methods*, 2015 to appear.

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

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

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

[17] F. Gamboa, T. Klein, and A. Lagnoux. Sensitivity analysis based on Cramér von Mises distance. *SIAM/ASA Journal on Uncertainty Quantification*, 6(2):522–548, Apr. 2018.

[18] F. Gamboa, T. Klein, and A. Lagnoux. A central limit theorem for generalized L-statistics. Preprint, 2021.

[19] F. Gamboa, T. Klein, A. Lagnoux, and L. Moreno. Sensitivity analysis in general metric spaces. Preprint, Feb. 2019.

[20] J. Hart, P. Gremaud, and T. David. Global sensitivity analysis of high dimensional neuroscience models: an example of neurovascular coupling. *Bull Math Biol*, 2019.

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

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

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

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

[25] C. McDiarmid. On the method of bounded differences. *Surveys in combinatorics*, 141(1):148–188, 1989.

[26] S. Orey et al. A central limit theorem for m-dependent random variables. *Duke Mathematical Journal*, 25(4):543–546, 1958.

[27] A. B. Owen. Better estimation of small Sobol'sensitivity indices. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 23(2):11, 2013.

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

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

[30] E. Rio. *Théorie asymptotique des processus aléatoires faiblement dépendants*, volume 31. Springer Science & Business Media, 1999.

[31] W. Rudin. Real and complex analysis. 1987. *Cited on*, 156, 1987.

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

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

[34] H. Shi, M. Drton, and F. Han. On the power of chatterjee rank correlation, 2020.

[35] I. M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Math. Modeling Comput. Experiment*, 1(4):407–414 (1995), 1993.

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

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

[38] W. Trutschnig. On a strong metric on the space of copulas and its induced dependence measure. *Journal of mathematical analysis and applications*, 384(2):690–705, 2011.

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