

Sensitivity Analysis and Generalized Chaos Expansions. Lower Bounds for Sobol indices

O Roustant, F. Gamboa, B Iooss

► **To cite this version:**

O Roustant, F. Gamboa, B Iooss. Sensitivity Analysis and Generalized Chaos Expansions. Lower Bounds for Sobol indices. 2019. hal-02140127

HAL Id: hal-02140127

<https://hal.archives-ouvertes.fr/hal-02140127>

Submitted on 27 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sensitivity Analysis and Generalized Chaos Expansions. Lower Bounds for Sobol indices.

O. Roustant¹, F. Gamboa², and B. Iooss^{3,2}

¹Mines Saint-Étienne, Univ. Clermont Auvergne, CNRS, UMR 6158 LIMOS,
F-42023 Saint-Étienne, France

²Institut de Mathématiques de Toulouse, Université Paul Sabatier, 31062
Toulouse Cedex 9, France

³Electricité de France R&D, 6 quai Watier, Chatou, F-78401, France

Abstract

The so-called polynomial chaos expansion is widely used in computer experiments. For example, it is a powerful tool to estimate Sobol' sensitivity indices. In this paper, we consider generalized chaos expansions built on general tensor Hilbert basis. In this frame, we revisit the computation of the Sobol' indices and give general lower bounds for these indices. The case of the eigenfunctions system associated with a Poincaré differential operator leads to lower bounds involving the derivatives of the analyzed function and provides an efficient tool for variable screening. These lower bounds are put in action both on toy and real life models demonstrating their accuracy.

Contents

1	Introduction	2
2	Background on sensitivity analysis	4
3	Generalized chaos expansions	6

4	Poincaré differential operator expansions	9
5	Weight-free derivative global sensitivity measures	14
6	Examples on analytical functions	17
6.1	A polynomial function with interaction	17
6.2	A separable function	18
7	Applications	21
7.1	A simplified flood model	21
7.2	An aquatic prey-predator chain	24
7.3	Conclusion on the applications	26
8	Further works	28

1 Introduction

Computer models simulating physical phenomena and industrial systems are commonly used in engineering and safety studies. They often take as inputs a high number of numerical and physical variables. For the development and the analysis of such computer models, the global sensitivity analysis methodology is an invaluable tool that allows to rank the relative importance of each input of the system [20], [18]. Referring to a probabilistic modeling of the model input variables, it accounts for the whole input range of variation, and tries to explain output uncertainties on the basis of input uncertainties. Thanks to the so-called functional ANOVA (analysis of variance) decomposition [2], the Sobol' indices give, for a square integrable non-linear model and stochastically independent input variables, the parts of the output variance due to each input and to each interaction between inputs [32], [15]. In addition, the total Sobol' index provides the overall contribution of each input [16], including interactions with other inputs. More generally, we recall that a Sobol' index associated to a subset of variables I is the ratio of the ANOVA index (that is the L^2 norm of the contribution associated to I in the ANOVA decomposition), and the variance of the output (see Section 2 for the precise definition).

Many methods exist to accurately compute or statistically estimate the first-order Sobol' indices. For a general overview on these methods, we refer to [18] and references therein. One of the most popular and powerful method

is polynomial chaos (PC) expansion [13], [36]. It consists in approximating the response onto the specific basis made by the orthonormal polynomials built on the input distributions. Its strength stands on the fact that, once the expansion is computed, the Parseval formula gives directly all the ANOVA indices (in particular the total Sobol' indices) [36, 7]. Of course in practice the PC expansion is truncated. An obvious but important fact is that this truncated PC expansion provides a lower bound for the true ANOVA index. In this paper, we consider general tensor Hilbert basis called generalized chaos (GC). Further, we use the previous trick to produce general lower bounds (see Section 3). Then, a smart choice of the GC produces new interesting lower bounds involving the derivatives of the function of interest (see Section 4). More precisely, this special Hilbert basis is obtained by diagonalizing the Poincaré differential operators (PDO), associated with the input distributions (this operator is related to Poincaré inequality, see [3] or [5]). Notice that other special GC expansions based on the diagonalization of reproducing kernels has been recently studied and used for global sensitivity purposes in [29].

In general, the estimation of the total Sobol' indices (and other ANOVA indices) suffers from the curse of dimensionality (number of inputs) and can be too costly in terms of number of model evaluations [28]. Low-cost computations of upper and lower bounds for total Sobol' indices are then very useful. DGSM (Derivative-based Global Sensitivity Measures, see [34]), computed from some integral of the squared derivatives of the model output, may give such economical upper and lower bounds [22, 21]. Indeed, in many physical models the so-called adjoint method allows at weak extra cost the evaluation of the derivatives of the model (see for example the recent review [1]). Concerning the upper bounds, optimal and general (for any distribution type of the input) results are obtained in [30]. For lower bounds, only special cases (uniform, Normal and Gamma) have been investigated in [37, 23] (see [21] for a review). The bounds given in [23] are quite rough as they are smaller than the first-order Sobol' indices. In our work, we follow the tracks opened by [37] using PC expansions, but for both much more general distributions and expansions. Indeed, for a wide class of input distributions the PDO generalized chaos expansion leads naturally to quantities built on the derivatives.

Notice that the diagonalization of PDO used here, leads to orthogonal polynomial only for the Gaussian distribution (see [3] and [4]). Indeed, the PDO considered in this paper only involves the integration with respect to the input distribution of the squared derivatives (and not a reweighted input

distribution). Apart from this particular probability distribution, orthogonal polynomials cannot be interpreted, in general, as eigenfunctions of a PDO. Consequently, in general the Hilbert basis built by diagonalizing a PDO is not a polynomial basis. For example, for the uniform distribution, it is the Fourier basis.

The paper is structured as follows. Section 2 recalls the required mathematical tools for global sensitivity analysis (ANOVA decomposition and DGSM). Section 3 rephrases the ANOVA decomposition with Hilbert spaces, and introduces the generalized chaos expansion. Section 4 then focuses on PDO expansions, and their link to PC expansions. Section 5 gives an alternative proposition of orthonormal functions which lead to weight-free DGSM. Section 6 gives analytical examples. Section 7 illustrates on real life applications. Section 8 gives some perspectives for future works.

2 Background on sensitivity analysis

To begin with, let $X = (X_1, \dots, X_d)$ denotes the vector of independent input variables with distribution $\mu = \mu_1 \otimes \dots \otimes \mu_d$. Here the μ_i 's are continuous probability measures on \mathbb{R} . Let further h be a multivariate function of interest $h : \Delta \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$. We assume that $h(X) \in \mathcal{H} := L^2(\mu)$.

One of the main tool in global sensitivity analysis is the Sobol'-Hoeffding decomposition of h , (see [15, 11, 2, 32]). It provides a unique expansion of h as

$$h(X) = h_0 + \sum_{i=1}^d h_i(X_i) + \sum_{1 \leq i < j \leq d} h_{i,j}(X_i, X_j) + \dots + h_{1,\dots,d}(X_1, \dots, X_d)$$

with $\mathbb{E}[h_I(X_I)|X_J] = 0$ for all $I \subseteq \{1, \dots, d\}$ and all $J \subsetneq I$ (with the notation $X_I := (X_i : i \in I)$). Furthermore, $h_0 = \mathbb{E}[h(X)]$ and

$$h_I(X_I) = \mathbb{E}[h(X)|X_I] - \sum_{J \subsetneq I} h_J(X_J) = \sum_{J \subsetneq I} (-1)^{|I|-|J|} \mathbb{E}[h(X)|X_J].$$

Notice that the condition

$$\mathbb{E}[h_I(X_I)|X_J] = 0 \quad \text{for all } J \subsetneq I,$$

warrants both the uniqueness of the decomposition and the orthogonality of $h_I(X_I)$ to any square integrable random variables depending only on X_J with $J \cap I \subsetneq I$.

This last property leads to the so-called ANOVA decomposition for the variance of $h(X)$

$$D := \text{var}(h(X)) = \sum_{I \subseteq \{1, \dots, d\}} \text{var}(h_I(X_I)). \quad (1)$$

Notice further that the Sobol'-Hoeffding decomposition is a particular case of the multivariate decomposition built on a finite family of commuting projectors P_1, \dots, P_d and obtained by expanding the following product (see [24]),

$$\begin{aligned} I_d &= (P_1 + (I_d - P_1)) \dots (P_d + (I_d - P_d)) \\ &= \sum_{I \subseteq \{1, \dots, d\}} \underbrace{\prod_{j \notin I} P_j \prod_{k \in I} (I - P_k)}_{\Pi_I}. \end{aligned}$$

Obviously, Π_I is also a projector. In the Sobol'-Hoeffding decomposition the projection $P_j h$ is $\int h(x) d\mu_j(x_j)$.

In sensitivity analysis, one classically considers the Sobol' indices. These indices are defined, for $I \subseteq \{1, \dots, d\}$, as $S_I = D_I/D$ where $D_I := \text{var}(h_I(X_I))$. From (1) one directly obtains

$$D = \sum_I D_I, \quad 1 = \sum_I S_I.$$

Another interesting index is the total Sobol' one that includes all the contributions on the total variance of a variable group. In this paper, the total index associated to one variable is the object under study. For $I \subseteq \{1, \dots, d\}$, the total Sobol' index associated to I is defined as $S_I^{\text{tot}} := \frac{D_I^{\text{tot}}}{D}$ with

$$D_I^{\text{tot}} := \sum_{J \supseteq I} D_J.$$

To end this section, we recall the other popular global sensitivity index that will appear in our bounds. This is the so-called Derivative Global Sensitivity Measure (DGSM) introduced and studied in [33] and [22]. It is defined, for $I \subseteq \{1, \dots, d\}$, under smoothness and integrability assumptions on h as

$$\nu_I = \int \left(\frac{\partial^{|I|} h(x)}{\partial x_I} \right)^2 \mu(dx).$$

3 Generalized chaos expansions

In order to present the generalized chaos expansions, it is convenient to first rephrase the classical functional ANOVA decomposition presented in the previous section as a Hilbert space decomposition. The next proposition is devoted to this task. In particular, we emphasize that the operator giving one ANOVA term is an orthogonal projection. Then, we discuss the construction of Hilbert basis tailored to ANOVA decomposition. Part of the material is inspired from [38] and [2].

Proposition 1 (Hilbert space decomposition for ANOVA). *For all subset I of $\{1, \dots, d\}$, the map $\Pi_I : h \in \mathcal{H} \mapsto h_I$ is an orthogonal projection. The image spaces $\mathcal{H}_I = \Pi_I(\mathcal{H}) = \{h \in \mathcal{H}, h = h_I\}$, called ANOVA spaces, are Hilbert spaces that form an orthogonal decomposition of \mathcal{H} :*

$$\mathcal{H} = \bigoplus_{I \subseteq \{1, \dots, d\}}^{\perp} \mathcal{H}_I \quad (2)$$

Proof. First, Π_I is a projector since applying twice the ANOVA decomposition leaves it unchanged. Now, let $g, h \in \mathcal{H}$. We have:

$$\langle \Pi_I g, h \rangle = \mathbb{E}(g_I(X_I)h(X)) = \sum_{J \subseteq \{1, \dots, d\}} \mathbb{E}(g_I(X_I)h_J(X_J))$$

where we wrote the ANOVA decomposition of h . Now, if $J \neq I$, then $I \cap J \subsetneq I$ or $I \cap J \subsetneq J$, thus $\mathbb{E}(g_I(X_I)h_J(X_J)) = 0$ by the uniqueness property of ANOVA decomposition. Hence,

$$\langle \Pi_I g, h \rangle = \mathbb{E}(g_I(X_I)h_I(X_I)) = \langle g, \Pi_I h \rangle,$$

which proves that the projector Π_I is self-adjoint, and thus orthogonal.

Consequently, Π_I is continuous and \mathcal{H}_I is a Hilbert space as a closed subspace of \mathcal{H} . The direct sum (2) results from the existence and uniqueness of ANOVA decomposition. As shown above, the uniqueness property implies that $\mathcal{H}_I \perp \mathcal{H}_J$ if $I \neq J$. \square

Corollary 1 (Hilbert space decomposition for total effects). *Let I be a subset of $\{1, \dots, d\}$. Then the map $\Pi_I^{tot} : h \in \mathcal{H} \mapsto h_I^{tot} = \sum_{J \supseteq I} h_J$ is an orthogonal projection. The image space $\mathcal{H}_I^{tot} = \Pi_I^{tot}(\mathcal{H}) = \{h \in \mathcal{H}, h = h_I^{tot}\}$ is the Hilbert space*

$$\mathcal{H}_I^{tot} = \bigoplus_{J \supseteq I}^{\perp} \mathcal{H}_J. \quad (3)$$

Proof. Observe that $\Pi_I^{\text{tot}} = \sum_{J \supseteq I} \Pi_J$. As the Π_J are commuting orthogonal projections, Π_I^{tot} is an orthogonal projection. The remainder is straightforward. \square

We now exhibit Hilbert bases of \mathcal{H} that are adapted to the ANOVA decomposition, in the sense that each element belongs to one ANOVA space \mathcal{H}_I . This provides Hilbert bases for all \mathcal{H}_I and $\mathcal{H}_I^{\text{tot}}$.

Definition 1 (Generalized chaos). *For $i = 1, \dots, d$, let $(e_{i,n})_{n \in \mathbb{N}}$ be a Hilbert basis of $L^2(\mu_i)$, with $e_{i,0} = 1$. For a multi-index $\underline{\ell} = (\ell_1, \dots, \ell_d) \in \mathbb{N}^d$, the generalized chaos of order $\underline{\ell}$ is defined as the following $L^2(\mu)$ function:*

$$e_{\underline{\ell}}(x) := \left(\bigotimes_{i=1, \dots, d} e_{i, \ell_i} \right) (x) = e_{1, \ell_1}(x_1) \times \dots \times e_{d, \ell_d}(x_d).$$

The so-called polynomial chaos introduced by [39], built with the orthogonal polynomials associated to the Gaussian distribution (Hermite polynomials (H_n)), is a special case of the previous definition (with $e_{i,n} = H_n$). Similarly, this is also the case for the generalized polynomial chaos corresponding to orthogonal polynomials associated to other probability distributions. For history on polynomial chaos and generalized polynomial chaos, we refer to the introduction of [12]. Other examples of generalized chaos in the context of sensitivity analysis are the Fourier bases, investigated in [8], and the Haar systems originally used by Sobol' [31].

Proposition 2.

1. *The whole set of generalized chaos $\mathcal{T} := (e_{\underline{\ell}})_{\underline{\ell} \in \mathbb{N}^d}$ is a Hilbert basis of \mathcal{H} , and each $e_{\underline{\ell}}$ belongs to (exactly) one \mathcal{H}_I , where I is the set containing the indices of active variables: $I = \{i \in \{1, \dots, d\} : \ell_i \geq 1\}$.*
2. *For all $I \subseteq \{1, \dots, d\}$,*
 - *The subset of basis functions that involve exactly the variables in I , $\mathcal{T}_I := \{e_{\underline{\ell}}, \text{ with } \ell_i \geq 1 \text{ if } i \in I \text{ and } \ell_i = 0 \text{ if } i \notin I\}$ is a Hilbert basis of \mathcal{H}_I .*
 - *The subset of basis functions that involve at least the variables in I , $\mathcal{T}_I^{\text{tot}} := \{e_{\underline{\ell}}, \text{ with } \ell_i \geq 1 \text{ if } i \in I\}$ is a Hilbert basis of $\mathcal{H}_I^{\text{tot}}$.*

Notice that in the definition of \mathcal{T}_I and $\mathcal{T}_I^{\text{tot}}$, the index n_i is non zero, which means that x_i is active.

Proof. The fact that \mathcal{T} is a Hilbert basis of \mathcal{H} is well known. Let us see that $e_{\underline{\ell}}$ belongs to \mathcal{H}_I , with $I = \{i \in \{1, \dots, d\} : \ell_i \geq 1\}$. For that, we need to check that the ANOVA decomposition of $e_{\underline{\ell}}$ consists of only one non-zero term corresponding to the subset I and equal to $e_{\underline{\ell}}$. As $e_{\underline{\ell}}$ is a function of x_I , it remains to check the non-overlapping condition. Let J be a strict subset of I (possibly empty). Then,

$$\mathbb{E} \left[\prod_{i \in I} e_{i, \ell_i}(X_I) | X_J \right] = \prod_{j \in J} e_{j, \ell_j}(X_j) \prod_{i \in I \setminus J} \mathbb{E} [e_{i, \ell_i}(X_I)]$$

Let us choose $i \in I \setminus J$. Then, $\ell_i \geq 1$, implying that $\mathbb{E} [e_{i, \ell_i}(X_I)] = 0$ (as e_{i, ℓ_i} is orthogonal to $e_{i, 0} = 1$). Finally $e_{\underline{\ell}}$ belongs to \mathcal{H}_I . Now let us fix a subset I of $\{1, \dots, d\}$, and consider for instance \mathcal{T}_I (the proof is similar for $\mathcal{T}_I^{\text{tot}}$). Clearly, as a subset of \mathcal{T} , the set \mathcal{T}_I is a collection of orthonormal functions. Furthermore, by the proof above, each $e_{\underline{\ell}}$ of \mathcal{T}_I belongs to \mathcal{H}_I . To see that \mathcal{T}_I is dense in \mathcal{H}_I , let us choose $h \in \mathcal{H}_I$. Since \mathcal{T} is a Hilbert basis of \mathcal{H} , then h can be written as

$$h = \sum_{\underline{\ell} \in \mathbb{N}^d} c_{\underline{\ell}} e_{\underline{\ell}} = \sum_{e_{\underline{\ell}} \in \mathcal{T}_I} c_{\underline{\ell}} e_{\underline{\ell}} + \sum_{e_{\underline{\ell}} \notin \mathcal{T}_I} c_{\underline{\ell}} e_{\underline{\ell}}$$

where $(c_{\underline{\ell}})_{\underline{\ell} \in \mathbb{N}^d}$ is a squared integrable sequence of real numbers. Recall that each $e_{\underline{\ell}}$ belongs to \mathcal{H}_J , with $J = \{i \in \{1, \dots, d\} \text{ s.t. } \ell_i \geq 1\}$. Thus, if $e_{\underline{\ell}} \notin \mathcal{T}_I$, then $J \neq I$. Hence, $e_{\underline{\ell}} \in \mathcal{H}_I^\perp$ (as $\mathcal{H}_J \perp \mathcal{H}_I$). Since $h \in \mathcal{H}_I$, it implies that $\sum_{e_{\underline{\ell}} \notin \mathcal{T}_I} c_{\underline{\ell}} e_{\underline{\ell}} = 0$. \square

The previous results imply that the variance D_I (resp. D_I^{tot}) of the output explained by a set I (resp. supersets of I) of input variables, is equal to the squared norm of the orthogonal projection onto \mathcal{H}_I (resp. $\mathcal{H}_I^{\text{tot}}$). Hence, lower bounds can be obtained by projecting onto smaller subspaces.

Corollary 2. *Let I be a subset of $\{1, \dots, d\}$ and let $h \in \mathcal{H}$. Then:*

- *For all subset G of \mathcal{H}_I , $D_I = \|\Pi_I(h)\|^2 \geq \|\Pi_G(h)\|^2$, with equality iff h has the form $h = g + f$ with $g \in G$ and $f \in \mathcal{H}_I^\perp$*
- *For all subset G of $\mathcal{H}_I^{\text{tot}}$, $D_I^{\text{tot}} = \|\Pi_I^{\text{tot}}(h)\|^2 \geq \|\Pi_G(h)\|^2$, with equality iff h has the form $h = g + f$ with $g \in G$ and $f \in (\mathcal{H}_I^{\text{tot}})^\perp$*

In practice, the subset G on which to project may be finite dimensional. For instance, it can be chosen by picking a finite number of orthonormal functions from the Hilbert basis obtained in Proposition 2. We illustrate this on the common case where I correspond to a single variable. Without loss of generality, we assume that $I = \{1\}$.

Corollary 3. *Let ϕ_1, \dots, ϕ_N be orthonormal functions in $\mathcal{H}_1^{\text{tot}}$. Then:*

$$D_1^{\text{tot}}(h) \geq \sum_{n=1}^N \left(\int h(x) \phi_n(x) \mu(dx) \right)^2$$

with equality iff h has the form $h(x) = \sum_{n=1}^N \alpha_n \phi_n(x) + g(x_2, \dots, x_N)$, where $g \in L^2(\otimes_{i=2, \dots, d} \mu_i)$. Furthermore, if all the ϕ_j 's belong to \mathcal{H}_1 , then the lower bound holds for D_1 .

Proof. This is a direct application of Corollary 2 with $G = \text{span}\{\phi_1, \dots, \phi_m\}$. The equality case is obtained by remarking that $(\mathcal{H}_1^{\text{tot}})^\perp$ is formed by functions of \mathcal{H} that do not involve x_1 : $(\mathcal{H}_1^{\text{tot}})^\perp = \bigoplus_{J \subseteq \{2, \dots, d\}} \mathcal{H}_J$. \square

4 Poincaré differential operator expansions

Generalized chaos expansions are defined from d Hilbert bases associated to probability measures on the real line μ_i ($i = 1, \dots, d$). Here, each μ_i is assumed to be absolutely continuous with respect to the Lebesgue measure. In this section, we exhibit a class of Hilbert basis which is well tailored to perform sensitivity analysis based on derivatives. They consist of eigenfunctions of an elliptic differential operator (DO). More precisely, we choose the DO associated to a 1-dimensional Poincaré inequality (assuming it holds)

$$\text{var}_{\mu_1}(h) \leq C \int_{\mathbb{R}} h'(x)^2 \mu_1(dx), \quad (4)$$

as it was successfully used to obtain accurate bounds for DGSM [30].

Before defining the so-called PDO expansions, we first recall the spectral theorem related to Poincaré inequalities. In what follows, for any positive integer ℓ , we denote by $H^\ell(\mu_1)$ the Sobolev space of order ℓ :

$$H^\ell(\mu_1) := \{h \in L^2(\mu_1) \text{ such that for all } k \leq \ell, h^{(k)} \in L^2(\mu_1)\} \quad (5)$$

Proposition 3 (Spectral theorem for Poincaré inequalities, [3, 30]). *Let $\mu_1(dt) = \rho(t)dt$ be a continuous measure on a bounded interval $I = (a, b)$ of \mathbb{R} , where $\rho(t) = e^{-V(t)}$. Assume that V is continuous and piecewise C^1 on $\bar{I} = [a, b]$. Then consider the differential operator*

$$Lh = h'' - V'h' \tag{6}$$

defined on $\mathcal{H}' = \{h \in H^2(\mu_1) \text{ s.t. } h'(a) = h'(b) = 0\}$. Then L admits a spectral decomposition. That is, there exists an increasing sequence $(\lambda_n)_{n \geq 0}$ of non-negative values that tends to infinity, and a set of orthonormal functions e_n which form a Hilbert basis of $L^2(\mu_1)$ such that $Le_n = -\lambda_n e_n$. Furthermore, all the eigenvalues λ_n are simple. The first eigenvalue is $\lambda_0 = 0$, and the corresponding eigenspace consists of constant functions (we can choose $e_0 = 1$). The first positive eigenvalue λ_1 is called spectral gap, and equal to the inverse of the Poincaré constant $C_P(\mu_1)$, i.e. the smallest constant satisfying Inequality (4).

Remark 1. *The assumptions of Proposition 3 guarantee that L admits a spectral decomposition, and correspond to a continuous probability distribution defined on a compact support, whose density is continuous and does not vanish. However, the spectral decomposition can exist for more general cases. For instance, it exists for the Normal distribution on \mathbb{R} : the corresponding eigenfunctions consist of Hermite polynomials and eigenvalues to non-negative integers. On the other hand, the spectral decomposition does not exist for the Laplace (double-exponential) distribution on the whole \mathbb{R} .*

The key property in our context is given by the equation

$$\langle h', e'_n \rangle = \lambda_n \langle h, e_n \rangle, \tag{7}$$

corresponding to the weak formulation of the spectral problem $Le_n = -\lambda_n e_n$ associated to the Poincaré inequality, and holding for all $n \geq 0$, and all $h \in H^1(\mu_1)$. It implies that geometric quantities involved in PDO expansions can be rewritten with derivatives. In particular, for a centered function h , we have:

$$\|h\|^2 = \sum_{n=1}^{\infty} \langle h, e_n \rangle^2 = \sum_{n=1}^{\infty} \frac{1}{\lambda_n^2} \langle h', e'_n \rangle^2.$$

Let us come back to the d -dimensional situation, where $\mu = \bigotimes_{i=1,\dots,d} \mu_i$. For each measure μ_i , we make the assumptions of Proposition 3 (see also Remark 1 for alternative conditions). We denote by L_i the corresponding operator and $\lambda_{i,n}, e_{i,n}$ ($n \geq 0$) its eigenvalues and eigenfunctions. We define $H^1(\mu)$ similarly to $H^1(\mu_1)$ (Equation 5). We can now define the PDO expansion and then state the main result.

Definition 2 (PDO expansions). *We call Poincaré differential operator (PDO) expansion the generalized chaos expansion corresponding to the Hilbert bases formed by the eigenfunctions of L_1, \dots, L_d .*

Proposition 4 (Poincaré-based lower bounds). *For all h in $H^1(\mu)$, we have*

$$D_1^{\text{tot}}(h) = \sum_{\ell_1 \geq 1, \ell_2, \dots, \ell_d} \langle h, e_{1,\ell_1} \dots e_{d,\ell_d} \rangle^2 \quad (8)$$

$$= \sum_{\ell_1 \geq 1, \ell_2, \dots, \ell_d} \frac{1}{\lambda_{1,\ell_1}^2} \left\langle \frac{\partial h}{\partial x_1}, e'_{1,\ell_1} e_{2,\ell_2} \dots e_{d,\ell_d} \right\rangle^2. \quad (9)$$

In particular, limiting ourselves to the first eigenfunction in all dimensions, and to first and second order tensors involving x_1 , we obtain the lower bound

$$D_1^{\text{tot}}(h) \geq \langle h, e_{1,1} \rangle^2 + \sum_{i=2}^d \langle h, e_{1,1} e_{i,1} \rangle^2 \quad (10)$$

$$= C_P(\mu_1)^2 \left(\left\langle \frac{\partial h}{\partial x_1}, e'_{1,1} \right\rangle^2 + \sum_{i=2}^d \left\langle \frac{\partial h}{\partial x_1}, e'_{1,1} e_{i,1} \right\rangle^2 \right). \quad (11)$$

Proof. By Proposition 2, the subset of (e_ℓ) corresponding to $\ell_1 \geq 1$ is a Hilbert basis of $\mathcal{H}_1^{\text{tot}}$. This gives (8). Now, for $\ell_1 \geq 1$:

$$\langle h, e_{1,\ell_1} \dots e_{d,\ell_d} \rangle = \frac{1}{\lambda_{1,\ell_1}} \left\langle \frac{\partial h}{\partial x_1}, e'_{1,\ell_1} e_{2,\ell_2} \dots e_{d,\ell_d} \right\rangle$$

This is obtained by applying Eq. (7) to $x_1 \mapsto h(x)$ and integrating with

respect to x_2, \dots, x_d :

$$\begin{aligned}
\langle h, e_{1,\ell_1} \dots e_{d,\ell_d} \rangle &= \int \langle h(\bullet, x_2, \dots, x_d), e_{1,\ell_1} \rangle_{L^2(\mu_1)} \prod_{i=2}^d e_{i,\ell_i} \mu_i(dx_i) \\
&= \frac{1}{\lambda_{1,\ell_1}} \int \left\langle \frac{\partial h(\bullet, x_2, \dots, x_d)}{\partial x_1}, e'_{1,\ell_1} \right\rangle_{L^2(\mu_1)} \prod_{i=2}^d e_{i,\ell_i} \mu_i(dx_i) \\
&= \frac{1}{\lambda_{1,\ell_1}} \left\langle \frac{\partial h}{\partial x_1}, e'_{1,\ell_1} e_{2,\ell_2} \dots e_{d,\ell_d} \right\rangle
\end{aligned}$$

This gives (9). The remainder is straightforward, knowing that $C_P(\mu_1) = 1/\lambda_{1,1}$. \square

Case of uniform distributions: Fourier expansion. Let us assume that μ_1 is uniform on $[-1/2, 1/2]$. Then, the differential operator L is the usual Laplacian, and its eigenfunctions correspond to Fourier basis. More precisely, using the Neumann boundary conditions $h'(a) = h'(b) = 0$, one can check that the eigenvalues are $\lambda_\ell = \ell^2 \pi^2$, ($\ell = 0, 1, \dots$), and a set of orthonormal eigenfunctions is given by $e_0 = 1$ and

$$e_\ell(x_1) = \sqrt{2} \cos(\pi \ell (x_1 + 1/2))$$

for $\ell > 0$. Denote by $|\underline{\ell}|_0$ the number of non-zero coefficients of the multi-index $\underline{\ell} = (\ell_1, \dots, \ell_d)$. When the other μ_i 's are also uniform on $[-1/2, 1/2]$, we obtain a multivariate Parseval formula for D_1^{tot} :

$$\begin{aligned}
D_1^{\text{tot}}(h) &= \sum_{\ell_1 \geq 1, \ell_2, \dots, \ell_d} 2^{|\underline{\ell}|_0} \langle h, \prod_{i=1}^d \cos(\pi \ell_i (x_i + 1/2)) \rangle^2 \\
&= \sum_{\ell_1 \geq 1, \ell_2, \dots, \ell_d} 2^{|\underline{\ell}|_0} \frac{1}{\pi^2 \ell_1^2} \left\langle \frac{\partial h}{\partial x_1}, \sin(\pi \ell_1 (x_1 + 1/2)) \prod_{i=2}^d \cos(\pi \ell_i (x_i + 1/2)) \right\rangle^2
\end{aligned}$$

Limiting for instance the sum to first terms, we obtain the lower bounds

$$\begin{aligned}
D_1^{\text{tot}}(h) &\geq 2 \langle h, \sin(\pi x_1) \rangle^2 + 4 \sum_{i=2}^d \langle h, \sin(\pi x_1) \sin(\pi x_i) \rangle^2 \quad (12) \\
&\geq \frac{2}{\pi^2} \left(\left\langle \frac{\partial h}{\partial x_1}, \cos(\pi x_1) \right\rangle^2 + 2 \sum_{i=2}^d \left\langle \frac{\partial h}{\partial x_1}, \cos(\pi x_1) \sin(\pi x_i) \right\rangle^2 \right) \quad (13)
\end{aligned}$$

Extension of PDO expansions to weighted Poincaré inequalities.

PDO expansions correspond to diffusion operators associated to Poincaré inequalities. They can be extended to weighted Poincaré inequalities

$$\text{var}_{\mu_1}(h) \leq C \int_{\mathbb{R}} h'(x)^2 w(x) \mu_1(dx), \quad (14)$$

defined for some suitable positive weight w . Such inequalities have recently been used in sensitivity analysis [35]. They are also useful when a probability distribution does not admit a Poincaré inequality such as the Cauchy distribution [5]. The weighted Poincaré inequality (14) corresponds to the differential operator

$$Lh = wh'' + (w' - wV')h'. \quad (15)$$

Similarly to (7), rewriting geometrical quantities with derivatives can be done with the formula:

$$\langle h', e'_n \rangle_w = \lambda_n \langle h, e_n \rangle, \quad (16)$$

where $\langle \cdot, \cdot \rangle_w$ is the weighted dot product $\langle f, g \rangle_w := \int f(x)g(x)w(x)\mu(dx)$. Proposition 4 can be adapted accordingly.

When PDO expansions coincide with PC expansions. There are exactly three cases where PDO expansions coincide with PC expansions, even when considering their extension to weighted Poincaré inequalities. Indeed, it can be shown that orthogonal polynomials are eigenfunctions of diffusion operators only for the Normal, Gamma and Beta distributions, corresponding respectively to Hermite, Laguerre and Jacobi orthogonal polynomials ([3], § 2.7). These differential operators correspond to weighted Poincaré inequalities with weight $w(x) = x$ for the Gamma distribution $d\mu_1(x) \propto x^{\alpha-1}e^{-\alpha x}$ on \mathbb{R}^+ , and weight $w(x) = 1 - x^2$ for the Beta distribution $d\mu_1(x) \propto (1 - x)^{\alpha-1}(1 + x)^{\beta-1}$ on $[-1, 1]$. Notice that in [35], w is chosen such that the eigenfunction associated to λ_1 is a first-order polynomial. Except for the three cases mentioned above, the other eigenfunctions cannot be all polynomials.

5 Weight-free derivative global sensitivity measures

The lower bounds of total indices obtained with generalized chaos expansions may involve weighted DGSM. For instance, in PDO expansions, weights involve the eigenfunction derivatives (Equation (11)). The presence of weight can be a drawback when the integral has to be estimated with a small sample size, as it can increase the variance of the Monte Carlo estimator. In this section, we show how to choose the two first orthonormal functions of GC expansions in order to obtain weight-free DGSM. Interestingly, this is related to Fisher information and Cramér-Rao bounds.

Proposition 5 (Lower bounds with weight-free DGSM, for pdf vanishing at the boundaries). *Assume that $\frac{\partial h(x)}{\partial x_1}$ is in $L^2(\mu)$, and that the probability distributions μ_i are absolutely continuous on their support (a_i, b_i) with $-\infty \leq a_i < b_i \leq +\infty$. For each i , denote by p_i the corresponding probability density function. Assume that p_i belongs to $H^1(\mu_i)$, do not vanish on (a_i, b_i) but vanishes at the boundaries: $p_i(a_i) = p_i(b_i) = 0$. Finally, assume that p'_i is not identically zero, and that p'_i/p_i is in $L^2(\mu_i)$. Define $Z_i(x_i) = (\ln p_i)'(x_i)$ and $I_i = \text{var}(Z_i(X_i))$. Then, we have the inequality:*

$$D_1^{\text{tot}} \geq I_1^{-1} c_1^2 + I_1^{-1} \sum_{j=2}^d I_j^{-1} c_{1,j}^2 \quad (17)$$

with

$$\begin{aligned} c_1 &= \int h(x) Z_1(x_1) \mu(dx) = - \int \frac{\partial h(x)}{\partial x_1} \mu(dx) \\ c_{1,j} &= \int h(x) Z_1(x_1) Z_j(x_j) \mu(dx) = - \int \frac{\partial h(x)}{\partial x_1} Z_j(x_j) \mu(dx) \end{aligned}$$

Furthermore, if all the cross derivatives $\frac{\partial^2 h(x)}{\partial x_1 \partial x_j}$ are in $L^2(\mu)$, then

$$c_{1,j} = \int \frac{\partial^2 h(x)}{\partial x_1 \partial x_j} \mu(dx)$$

The cases of equality correspond to functions h of the form

$$h(x) = \alpha_1 Z_1(x_1) + \sum_{j=2}^d \alpha_j Z_1(x_1) Z_j(x_j) + h(x_2, \dots, x_d). \quad (18)$$

Proof. For $i = 1, \dots, d$, let $e_{i,1}(x_i) := I_i^{-1/2} Z_i(x_i)$. Then define

$$\phi_1(x) = e_{1,1}(x_1), \quad \text{and for } j = 2, \dots, d: \quad \phi_j(x) = e_{1,1}(x_1)e_{j,1}(x_j).$$

By definition, the norm of each $e_{i,1}$ is equal to 1. Furthermore, Z_i is centered, since

$$\mathbb{E}[Z_i] = \int_{a_i}^{b_i} p_i'(x_i) dx_i = [p_i(x_i)]_{a_i}^{b_i} = 0.$$

This implies that $e_{i,1}$ is orthogonal to $e_{i,0} = 1$. By Proposition 2, the ϕ_i 's are then orthonormal functions of $\mathcal{H}_1^{\text{tot}}$. The inequality is then given by Corollary 3, with first expressions of c_1 and $c_{1,j}$. The other ones are obtained by integrating by part, using that the values at the boundaries of the p_j 's are zero. \square

The proposition can be adapted when the probability density functions do not vanish at the boundaries of their support, by modifying the definition of the Z_j 's. Notice that the expressions of c_1 and $c_{1,j}$ that involve derivatives then contain corrective terms, and are of limited practical interest. For instance, denoting $[h]_{a_1}^{b_1} = h(b_1) - h(a_1)$ and $h_0 = \int h(x)\mu(dx)$, we have:

$$c_1 = \left[\left(\int h(x_1, x_{-1}) \mu_{-1}(dx_{-1}) - h_0 \right) p_1(x_1) \right]_{a_1}^{b_1} - \int \frac{\partial h(x)}{\partial x_1} \mu(dx).$$

Nevertheless, the first expressions of c_1 and $c_{1,j}$ remain valid and, by analogy to Proposition 5, have a close connection to derivative-based lower bounds.

Proposition 6 ([Lower bounds with weight-free DGSM, general case]). *Assume that $\frac{\partial h(x)}{\partial x_1}$ is in $L^2(\mu)$, and that the probability distributions μ_i are absolutely continuous on their support (a_i, b_i) with $-\infty \leq a_i < b_i \leq +\infty$. For each i , denote by p_i the corresponding probability density function. Assume that p_i belongs to $H^1(\mu_i)$ and do not vanish on (a_i, b_i) . Finally, assume that p_i' is not identically zero, and that p_i'/p_i is in $L^2(\mu_i)$. Define $Z_i(x_i) = (\ln p_i)'(x_i) - [p_i(x_i)]_{a_i}^{b_i}$ and $I_i = \text{var}(Z_i(X_i))$. Then Inequality (17) holds with $c_1 = \int h(x) Z_1(x_1) \mu(dx)$ and $c_{1,j} = \int h(x) Z_1(x_1) Z_j(x_j) \mu(dx)$. The equality case is the same as in Proposition 5, and given by (18).*

Remark 1. *The expressions of Z_i and I_i in Proposition 6 correspond respectively to the score and to the Fisher information at $\theta = 0$ of a parametric family of probability distributions obtained by translation $p_{i,\theta_i}(x_i) = p_i(x_i + \theta_i)$. In this framework, the lower bound (17) corresponds to the Cramér-Rao lower bound.*

Dist. name	Support	p	Z	I
Normal	\mathbb{R}	$\frac{1}{s\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(x-m)^2}{s^2}\right)$	$-(X-m)/s^2$	$1/s^2$
Laplace	\mathbb{R}	$\frac{1}{2s} \exp\left(\frac{ x-m }{s}\right)$	$-\text{sgn}(X-m)/s$	$1/s^2$
Cauchy	\mathbb{R}	$\frac{1}{\pi} \frac{s}{(x-x_0)^2+s^2}$	$\frac{-2(x-x_0)}{(x-x_0)^2+s^2}$	$1/(2s^2)$

Table 1: Useful quantities for derivative-based lower bounds. For readability, we have removed the subscript j for p, Z, I . The parameter s is a scale parameter, and can be different from the standard deviation.

Examples. First consider the case of normal distributions $\mu_i \sim \mathcal{N}(m_i, v_i)$ ($i = 1, \dots, d$). Applying Inequality (17) gives

$$D_1^{\text{tot}} \geq v_1 \left(\int \frac{\partial h(x)}{\partial x_1} \mu(dx) \right)^2 + v_1 \sum_{j=2}^d v_j \left(\int \frac{\partial^2 h(x)}{\partial x_1 \partial x_j} \mu(dx) \right)^2. \quad (19)$$

Here, the inequality is equivalent to Inequality (11) obtained with the Poincaré differential operator of Section 4, since Z_i is a first-order polynomial, and thus equal to the first eigenvector of L (Hermite polynomial). The case of equality corresponds to functions of the form

$$h(x) = \alpha_1(x_1 - m_1) + \sum_{j=2}^m \alpha_j(x_1 - m_1)(x_j - m_j) + g(x_2, \dots, x_d).$$

Other inequalities can be established for standard probability distributions. Table 1 summarizes the results for some of them. Notice that the equality case does not always correspond to polynomials (see the form of Z). Interestingly, an inequality is obtained for the Cauchy distribution, whereas the theory of Section 4 does not apply as this distribution does not admit a Poincaré constant. On the other hand, some probability distributions for which Section 4 is applicable, do not satisfy the assumptions of Proposition 6, such as the uniform (p'_i is identically zero) or the triangular distributions (p'_i/p_i does not belong to $L^2(\mu_i)$).

Link to other works. Here, we briefly compare our lower bounds to those presented in the recent review [21].

For the uniform distribution on $[0, 1]$, we can obtain both a better upper bound and a description of the equality case. For that, we apply Corollary 3 to the orthonormal function obtained from x_1^m , i.e. $\phi(x_1) = (x_1^m - m_1)/s_1$ with $m_1 = 1/(m + 1)$ and $s_1^2 = \left(\frac{m}{m+1}\right)^2 \frac{1}{2m+1}$. Then after some algebra and an integration by part, we obtain

$$D_1^{\text{tot}} \geq \frac{2m+1}{m^2} \left(\int (h(1, x_{-1}) - h(x)) dx - w_1^{(m+1)} \right)^2$$

where $w_1^{(m+1)} = \int \frac{\partial h(x)}{\partial x_1} x_1^{m+1} dx$. This improves on the lower bound found in [21], Theorem 2, which has the same form, but with the smaller multiplicative constant $\frac{2m+1}{(m+1)^2}$. Furthermore, the lower bound above is attained when h has the form $h(x) = \alpha_1 x_1^m + g(x_2, \dots, x_d)$. However, notice that these two lower bounds are only a lower bound for $D_1 \leq D_1^{\text{tot}}$, and can be improved by considering additional orthonormal functions belonging to $\mathcal{H}_1^{\text{tot}} \setminus \mathcal{H}_1$.

For normal distributions, Inequality (19) improves the lower bound given by [23], i.e.

$$D_1^{\text{tot}} \geq v_1 \left(\int \frac{\partial h(x)}{\partial x_1} \mu(dx) \right)^2.$$

Here also, this latter lower bound is only a lower bound of $D_1 \leq D_1^{\text{tot}}$ since it corresponds to the case in Corollary 3 where the ϕ_j 's (here $\phi_1(x) = Z_1(x_1)/I_1^{1/2}$) only depend on x_1 .

6 Examples on analytical functions

This section briefly illustrates PDO expansions for the uniform distribution on benchmark functions from sensitivity analysis. We assess the accuracy of the lower bounds of total indices, when only the two first eigenvalues are used.

6.1 A polynomial function with interaction

Example 1. *Let us consider $g(x_1, x_2) = x_1 + ax_1x_2$, and let μ be the uniform distribution on $[-1/2, 1/2]^2$. The inequalities obtained by truncating the PDO*

expansions to the first eigenvalue are:

$$\begin{aligned} D_1 &= \frac{1}{12} \approx 0.0833 \geq 0.0821 \approx \frac{8}{\pi^4} \\ D_1^{\text{tot}} &= \frac{1}{12} + \frac{a^2}{144} \approx 0.0833 + 0.0069 a^2 \\ &\geq 0.0821 + 0.0067 a^2 \approx \frac{8}{\pi^4} + \frac{64}{\pi^8} a^2 \end{aligned}$$

We can see that for a polynomial function of degree 1 with respect to x_1 , the lower bound obtained by restricting the PDO expansion to the first eigenvalue is very accurate. Hence, we do not lose a lot of information by ignoring that the function is a polynomial. This is an ideal situation for polynomial chaos.

Let us give some computing details on the previous inequalities. It is easy to check that the two terms x_1 , ax_1x_2 correspond to the main effect and second order interaction respectively. The partial variances are given by $D_1 = 1/12$ and $D_{1,2} = a^2/144$. Hence, $D_1^{\text{tot}} = 1/12 + a^2/144$. Restricting the PDO expansion to the first term, a lower bound is given by Inequality (13):

$$D_1^{\text{tot}} \geq \frac{2}{\pi^2} \left(\left\langle \frac{\partial g}{\partial x_1}, \cos(\pi x_1) \right\rangle^2 + 2 \left\langle \frac{\partial g}{\partial x_1}, \cos(\pi x_1) \sin(\pi x_2) \right\rangle^2 \right).$$

The two terms of the lower bound above correspond to a lower bound of D_1 and $D_{1,2}$ respectively. A direct computation gives:

$$\begin{aligned} \text{LB}_1 &:= \frac{2}{\pi^2} \left\langle \frac{\partial g}{\partial x_1}, \cos(\pi x_1) \right\rangle^2 = \frac{2}{\pi^2} \left(\frac{2}{\pi} \right)^2 = \frac{8}{\pi^4} \\ \text{LB}_{1,2} &:= \frac{2}{\pi^2} 2 \left\langle \frac{\partial g}{\partial x_1}, \cos(\pi x_1) \sin(\pi x_2) \right\rangle^2 = \frac{2}{\pi^2} \cdot 2 \cdot \left(\frac{4a}{\pi^3} \right)^2 = \frac{64a^2}{\pi^8} \end{aligned}$$

The result follows.

6.2 A separable function

Example 2. Consider the g -Sobol' function on $[-1/2, 1/2]$ defined by

$$g(x) = \prod_{i=1}^d (1 + h_i(x_i))$$

with $h_i(x_i) = (4|x_i| - 1)/(1 + a_i)$ ($i = 1, \dots, d$), and let μ be the uniform distribution on $[-1/2, 1/2]^d$. The inequalities obtained by truncating the PDO expansions to the first two eigenvalues are:

$$D_i = \frac{1}{3} \frac{1}{(1 + a_i)^2} \geq \frac{32}{\pi^4} \frac{1}{(1 + a_i)^2} := LB_i \quad (20)$$

$$D_i^{\text{tot}} = D_i \prod_{j \neq i}^d (1 + D_j) \geq LB_i \cdot \sum_{j \neq i}^d LB_j \quad (21)$$

Notice that $32/\pi^4 \approx 0.328$ is very close to $1/3$. Hence, the lower bound for D_i is very accurate. Obviously, a very sharp inequality $D_i^{\text{tot}} \geq LB_i \prod_{i \neq j}^d (1 + LB_j)$ could have been deduced, but this is unrealistic in practice, since the separable form of the function is unknown. The lower bound (21) for D_i^{tot} is actually a very good approximation of the variance explained by second-order interactions involving x_i , equal to $D_i \cdot \sum_{j \neq i}^d D_j$. Hence, Inequality (21) will be less fine in presence of higher order interactions, (tuned by the values of the a_j 's). Then, more than two eigenvalues in PDO expansions must be considered.

Let us give some computing details on the previous inequalities. Without loss of generality, we write the proof for $i = 1$. Let us first recall the computation of Sobol' indices for the g-Sobol' function. As all the h_i are centered, the Sobol'-Hoeffding decomposition is given by $g_I(x_I) = \prod_{i \in I} h_i(x_i)$. In particular $D_1 = \int h_1^2 d\mu_1 = \frac{1}{3} \frac{1}{(1+a_1)^2}$. Furthermore, the variance of a second order interaction is, for $i \neq 1$:

$$D_{1,i} = E(h_1(x_1)^2 h_i(x_i)^2) = D_1 D_i, \quad (22)$$

and variance explained by second-order interactions containing x_1 is equal to

$$\sum_{i=2}^d D_{1,i} = D_1 \sum_{i=2}^d D_i.$$

Finally the total effect is the variance of $\sum_{I \supseteq \{1\}} \prod_{i \in I} h_i$, equal to

$$D_1^{\text{tot}} = \sum_{I \supseteq \{1\}} \prod_{i \in I} D_i = D_1 \prod_{i=2}^d (1 + D_i).$$

Let us now consider lower bounds. To obtain accurate lower bounds, we need to consider the first two non-zero eigenvalues. Indeed, the first non-zero eigenvector is even and all the dot products are 0. By using Equation (9) and the results about uniform distributions presented in Section 4, we obtain:

$$D_1^{\text{tot}}(g) \geq \frac{1}{\lambda_2^2} \left(\left\langle \frac{\partial g}{\partial x_1}, e'_{1,2} \right\rangle^2 + \sum_{i=2}^d \left\langle \frac{\partial g}{\partial x_1}, e'_{1,2} e_{i,2} \right\rangle^2 \right), \quad (23)$$

with $e_{i,2} = \sqrt{2} \cos(2\pi x_i)$ (we omit the '-' sign) and $\lambda_2 = 4\pi^2$. We could have also used (8), but using derivatives simplifies the computations here.

The first term gives a lower bound for D_1 . We have:

$$\frac{\partial g}{\partial x_1}(x) = \frac{4}{1+a_1} \text{sgn}(x_1) \prod_{i \geq 2} (1 + h_i(x_i)).$$

Due to the tensor form of the g-Sobol' function partial derivative, the dot product is expressed as a product of one-dimensional dot-products. Furthermore, as all the h_i 's are centered, the dot-products in dimensions $2, \dots, d$ are equal to 1. Finally,

$$\begin{aligned} \left\langle \frac{\partial g}{\partial x_1}, e'_{1,2} \right\rangle &= \langle h'_1, e'_{1,2} \rangle_1 = \frac{4}{1+a_1} \int_{-1/2}^{1/2} \text{sgn}(x_1) e'_{1,2}(x_1) dx_1 \\ &= \frac{4}{1+a_1} \sqrt{2} \cdot 2 \int_0^{1/2} 2\pi \sin(2\pi x_1) dx_1 = \frac{16\sqrt{2}}{1+a_1}. \end{aligned}$$

This gives the announced lower bound for the main effect (Equation (21)):

$$\text{LB}_1 = \frac{1}{\lambda_2^2} \left\langle \frac{\partial g}{\partial x_1}, e'_{1,2} \right\rangle^2 = \frac{1}{(4\pi^2)^2} \left(\frac{16\sqrt{2}}{1+a_1} \right)^2 = \frac{32}{\pi^4} \frac{1}{(1+a_1)^2}. \quad (24)$$

Now, let us compute the second term in (23). Notice that it is a lower bound for the variance explained by second-order interactions involving x_1 , as computed in (22). As above, exploiting the tensor form, we have:

$$\left\langle \frac{\partial g}{\partial x_1}, e'_{1,2} e_{i,2} \right\rangle = \langle h'_1, e'_{1,2} \rangle_1 \cdot \langle (1 + h_i), e_{i,2} \rangle_i.$$

The first term has already been computed above. For the second one, we use the property of eigenvectors (7):

$$\langle (1 + h_i), e_{i,2} \rangle_i = \frac{1}{\lambda_2} \langle h'_i, e'_{i,2} \rangle_i$$

and we recognize the quantity computed above where we replace 1 by i , equal to $\frac{1}{\lambda_2} \frac{16\sqrt{2}}{1 + a_i} = \sqrt{\text{LB}_i}$. Finally, plugging this result in (23) together with (24) gives the announced lower bound (21).

7 Applications

In this section, two numerical models representing real physical phenomena are used in order to illustrate the usefulness of the lower bounds of total Sobol' indices provided by PDO expansions. More precisely, we restrict ourselves to the simplest lower bound provided by considering only the first eigenfunctions in all dimensions, given by the two equivalent Equations (10) and (11). The first equation gives a derivative-free lower bound of the total index, here called PDO lower bound. The second one gives a derivative-based version, here called PDO-der lower bound.

Whereas the PDO and PDO-der lower bounds are theoretically equal, their estimated values will differ. Estimations of integrals and square products have been performed via crude Monte Carlo samples. We have centered the function f . It does not change the value of sensitivity indices but reduces the estimation error. The use of Monte Carlo samples allows to provide confidence intervals on the estimates by the way of a bootstrap resampling technique. Boxplots will be used to graphically represent these estimation uncertainties. Finally, the computation of eigenvalues, eigenfunctions and eigenfunction derivatives has been done with the numerical method presented in [30].

7.1 A simplified flood model

Our first model simulates flooding events by comparing the height of a river to the height of a dyke. It involves the characteristics of the river stretch, as already studied in [25, 30]. The model has 8 input random variables (r.v.), each one follows a specific probability distribution (truncated Gumbel,

truncated normal, triangular or uniform). When the height of a river is over the height of the dyke, flooding occurs. The model output is the cost (in million euros) of the damage on the dyke which writes:

$$Y = \mathbb{1}_{S>0} + \left[0.2 + 0.8 \left(1 - \exp^{-\frac{1000}{S^4}} \right) \right] \mathbb{1}_{S \leq 0} + \frac{1}{20} (H_d \mathbb{1}_{H_d > 8} + 8 \mathbb{1}_{H_d \leq 8}), \quad (25)$$

where $\mathbb{1}_A(x)$ is the indicator function which is equal to 1 for $x \in A$ and 0 otherwise, H_d is the height of the dyke (uniform r.v.) and S is the maximal annual overflow (in meters) based on a crude simplification of the 1D hydrodynamical equations of Saint-Venant under the assumptions of uniform and constant flowrate and large rectangular section. S is calculated as

$$S = \left(\frac{Q}{BK_s \sqrt{\frac{Z_m - Z_v}{L}}} \right)^{0.6} + Z_v - H_d - C_b, \quad (26)$$

with Q the maximal annual flowrate (truncated Gumbel r.v.), K_s the Strickler coefficient (truncated Gaussian r.v.), Z_m and Z_v the upstream and downstream riverbed levels (triangular r.v.), L and B the length and width of the water section (triangular r.v.) and C_b the bank level (triangular r.v.). For this model, first-order and total Sobol' indices have been estimated in [25] with high precision (large sample size) via a Monte-Carlo based algorithms.

Fig. 1 shows the PDO lower bounds. By looking at the values of first-order and total Sobol' indices (horizontal straight lines), we notice that rather large interaction effects are present between four inputs of the model (Q , K_s , Z_v and H_d). First, the bounds estimated with the sample size $n = 100$ have large uncertainties. It shows that this sample size is too small for this complex model (it includes non-linear and interaction effects). Secondly, concerning the estimation of the bounds, the convergence is reached, with very small uncertainties on the estimates from $n = 10\,000$. From this sample size, we can visually check (e.g. looking at the third quartile) that estimated lower bounds are smaller than the corresponding true Sobol' indices. Moreover, for smaller sample sizes as $n = 1\,000$, results for all the inputs show sufficient accuracies (easy discrimination between the bounds). Finally, except for K_s and H_d , the bounds are informative because:

- The PDO bounds are very close to the theoretical values of total Sobol' indices, which is remarkable as only the first eigenvalue was used.

- The PDO lower bounds for total indices are larger than their respective first-order Sobol' indices.

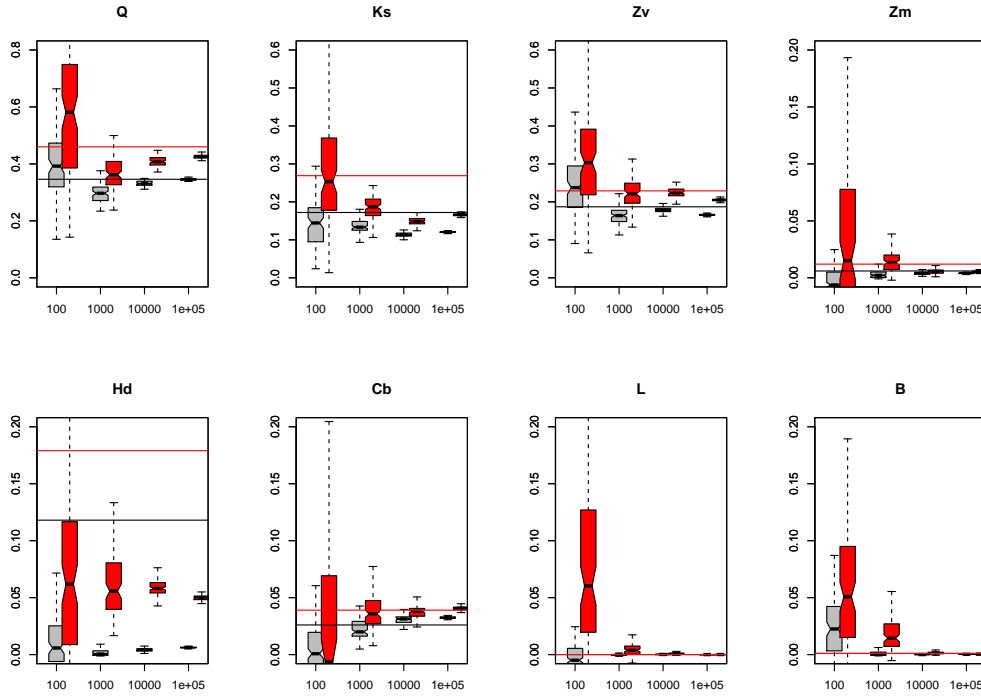


Figure 1: PDO bounds for the 8 inputs of the flood model application for four different sample sizes n (10^2 , 10^3 , 10^4 and 10^5). Red (resp. gray) boxplots are lower bounds of total (resp. first-order) Sobol' indices. Horizontal lines indicate the 'true' values of the Sobol' indices.

Fig. 2 shows that the PDO-der lower bounds give significantly better results than the PDO bounds, especially for small sample sizes. In particular, when the Sobol' indices are close to zero, the bounds perfectly match their respective Sobol' indices from $n = 100$. This result clearly favors the use of derivative-based lower bounds for the screening step when model derivatives can be computed.

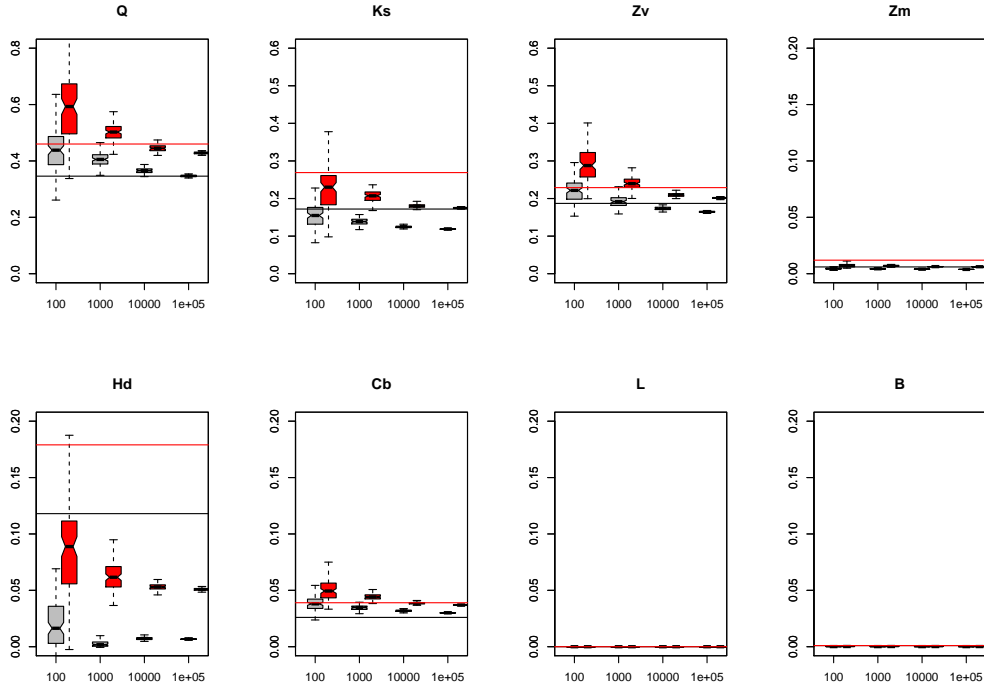


Figure 2: PDO-der bounds for the 8 inputs of the flood model application. The legend details are the same as Figure 1.

7.2 An aquatic prey-predator chain

This application is related to the modeling of an aquatic ecosystem called MELODY (MESocosm structure and functioning for representing LOTic DYnamic ecosystems). This model simulates the functioning of aquatic mesocosms as well as the impact of toxic substances on the dynamics of their populations. Inside this model, the Periphyton-Grazers sub-model is representative of processes involved in dynamics of primary producers and primary consumers, i.e. photosynthesis, excretion, respiration, egestion, mortality, sloughing and predation [6]. It contains a total number of $d = 20$ uncertain input variables. In order to conduct sensitivity analysis, [6] has defined that each of these input variables are random following a uniform distribution law, defined by their minimal and maximal values.

The PDO-der upper bound of total Sobol' indices [34] was then applied in [19] on one model output (the periphyton biomass) at only one reference

time, day 60 of simulations, which corresponds to the period of maximum periphyton biomass and a growth phase for grazers, according to experimental data. A design of experiments of size $n = 100$ was then provided, and simulated with MELODY. A model output vector of size 100 is obtained, as well as the derivatives of the output with respect to each input at each point of the design (matrix of size 100×20). In this section, we analyze the same data that has been studied in [19].

Fig. 3 shows the PDO lower bounds, as well as the first-order Sobol' indices estimates (via the local polynomials sample based technique [10]). Good results are obtained on the first-order lower bounds which have reduced estimation uncertainties and are always smaller than the estimated first-order Sobol' indices. Less accurate estimates are obtained for the lower bounds of total indices. They remain informative because they are clearly larger than the first-order Sobol' indices. This last result proves that large interactions between inputs dominate in this prey-predator model, which confirms the first analysis of [19] (the sum of all the first-order Sobol' indices is much smaller than one). The new results of Fig. 3 prove the strong influence of some inputs which have large total lower bounds. For example, 5 inputs have total lower bound median values larger than 20%: Maximum photosynthesis rate (n°1), Maximum consumption rate (n°2), Rate of change per 10°C (n°9), Grazers preference for periphyton (n°11) and Intrinsic mortality rate (n°16). This result cannot be found from the first-order Sobol' indices which are rather small (except for the Maximum photosynthesis rate).

Fig. 4 shows the PDO-der lower bounds, as well as the PDO-der upper bounds of the total Sobol' indices (see [19]) whose confidence intervals are also obtained by bootstrap. In this figure lower and upper bounds of total Sobol' indices have been truncated to one in order to only consider realistic values. Indeed, values larger than one are theoretically impossible but can sometimes be found due to numerical estimation errors. First, some partial checks can be done by looking at the median of the estimated values, e.g. by observing that the lower bounds are smaller than the upper bounds for each input. Second, several PDO-der lower bounds estimates are much less accurate than the (derivative-free) PDO lower bounds, especially when their values are large, for example the inputs n°1 and n°11. Even in this case of large values, informative results can be deduced by taking their median values: the total Sobol' indices of the input n°1 (resp. n°11) approximately lie in $[0.85, 1]$ (resp. $[0.4, 1]$). From these total lower bounds, a coarse importance hierarchy can then be proposed between the most influential inputs.

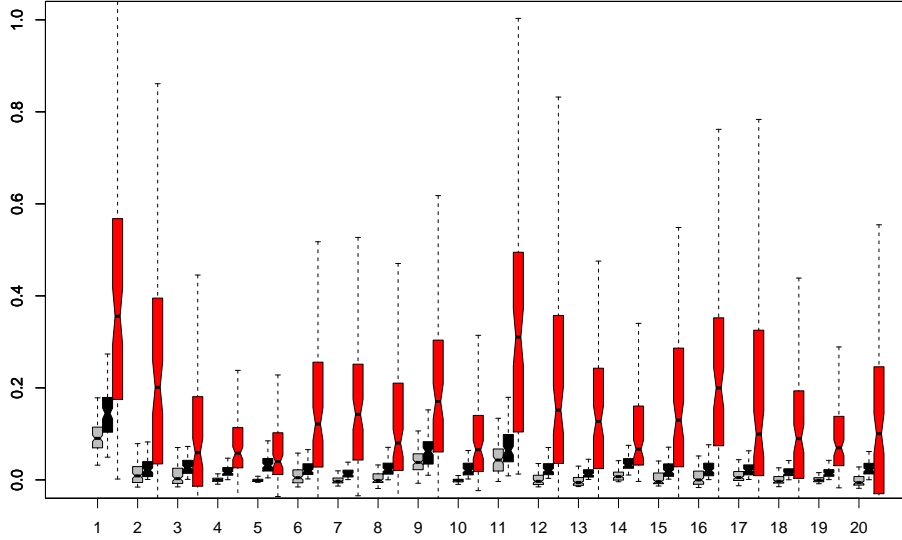


Figure 3: PDO bounds for the 20 inputs of the prey-predator model. Gray, black and red boxplots are respectively the lower bounds of the first-order Sobol' indices, the estimates of the first-order Sobol' indices and the lower bounds of the total Sobol' indices.

Finally, we observe the excellent results for non influential inputs which have all their PDO-der lower and upper bounds close to zero (inputs 2, 5, 7, 8, 10, 13, 15, 18, 19, 20). This is not the case with the PDO lower bounds (see Fig. 3) which are more difficult to exploit. A convenient usage would be to estimate both derivative-free and derivative-based lower bounds, and to keep the smallest value. Indeed, the PDO bound is more accurate when the Sobol' index is much larger than zero, whereas the PDO-der bound is much smaller when the Sobol' index is close to zero.

7.3 Conclusion on the applications

On the two previous applications, we have tested the simplest PDO and PDO-der lower bounds, obtained by keeping only the first eigenvalue in all dimensions, for real-world models involving non-linear and interaction effects. Several conclusions can be made:

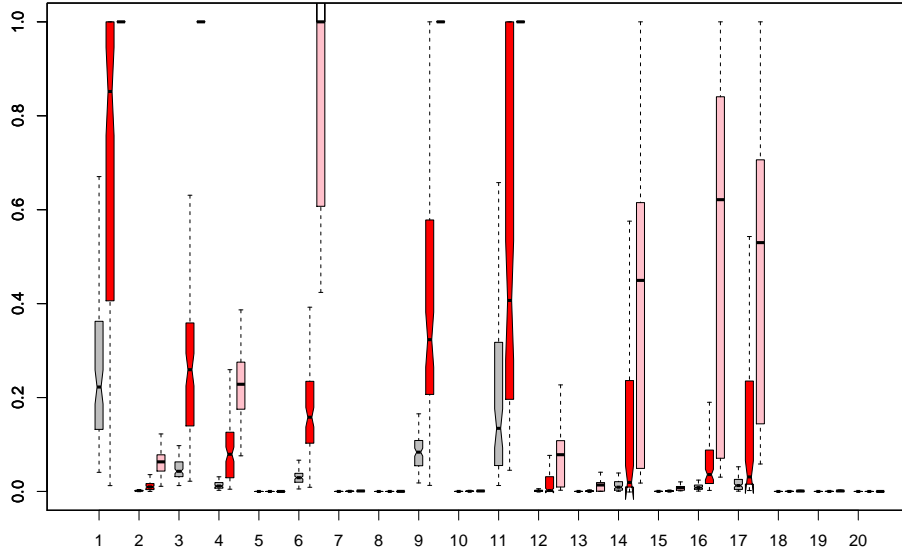


Figure 4: PDO-der bounds for the 20 inputs of the prey-predator model. As in Figure 3, gray and red boxplots are respectively lower bounds of the first-order and total Sobol' indices. The additional pink boxplots correspond to the upper bound of the total Sobol' indices.

- Lower bounds can be easily computed for any probability distribution of the inputs;
- The estimation error can be large for small sample sizes. Estimating some bootstrap confidence intervals is essential to evaluate the quality of the estimates;
- The lower bounds of the total Sobol' indices are most of the times informative, i.e. larger than the (estimated) first order Sobol' indices;
- Using derivatives (then DGSM) is sometimes preferable to obtain lower bounds, especially for the screening step (identification of non influential inputs with negligible total Sobol' indices). With DGSM, excellent results are obtained for screening, even for small sample size cases.

8 Further works

In this paper, we revisit the so-called chaos expansion method for the evaluation of Sobol' indices. We summarize in a compact way the role played by the functional basis and the associated projection operators for evaluating by below these indices through a truncated Parseval formula. Generalized chaos basis built on the Poincaré differential operator associated to the input distribution leads to very interesting new lower bounds for the total Sobol' index in terms of DGSM. This bound appears to be sharp both on toy and real life models, allowing a fast screening of the model input based on the energy of the function derivatives. This opens some challenging problems in mathematical statistics. First, the bounds obtained by the brute force truncation method could certainly be merely improved considering accurate model selection methods as adaptive thresholding or l^1 regularization. Second, the statistical estimation of the lower bound is a non linear semi-parametric problem. By non linear, we mean that the quantity to be estimated depends in a non linear way (here quadratic), of the infinite dimensional parameter (the function of interest). The estimation of a quadratic functional have been addressed in [26, 27, 14, 9]. It involves U -statistics theory, and offers an excellent source of inspiration for further works in mathematical statistics having concrete computational applications. For example, the unbiased estimation of such quantity for small sample appears to be an interesting challenging issue. As ending remark, notice that the use of PDO also opens challenging questions concerning the construction of such operators (and eigenbasis). First, one may be interested to build a PDO that provides a lower bound involving weighted DGSM. Secondly, one may wish to consider the case of heavy tail input distributions (as the Cauchy one for example).

Software and acknowledgement

The implementations are partially based on the R package `sensitivity` [17]. The whole code should be included in a future version of that package.

Part of this research was conducted within the frame of the Chair in Applied Mathematics OQUAIDO, gathering partners in technological research (BRGM, CEA, IFPEN, IRSN, Safran, Storengy) and academia (CNRS, Ecole Centrale de Lyon, Mines Saint-Etienne, University of Grenoble, University of Nice, University of Toulouse) around advanced methods for Computer

Experiments. The authors thank the participants for fruitful discussions. In particular we are grateful to A. Joulin for the insightful idea of using the Poincaré differential operator for computing lower bounds. Support from the ANR-3IA Artificial and Natural Intelligence Toulouse Institute is gratefully acknowledged.

References

- [1] G. Allaire. A review of adjoint methods for sensitivity analysis, uncertainty quantification and optimization in numerical codes. Ingénieurs de l'Automobile, 836:33–36, 2015.
- [2] A. Antoniadis. Analysis of variance on function spaces. Statistics: A Journal of Theoretical and Applied Statistics, 15(1):59–71, 1984.
- [3] D. Bakry, I. Gentil, and M. Ledoux. Analysis and geometry of Markov diffusion operators, volume 348 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Cham, 2014.
- [4] D. Bakry and O. Mazet. Characterization of markov semigroups on associated to some families of orthogonal polynomials. In Séminaire de Probabilités XXXVII, pages 60–80. Springer, 2003.
- [5] M. Bonnefont, A. Joulin, and Y. Ma. A note on spectral gap and weighted Poincaré inequalities for some one-dimensional diffusions. ESAIM: Probability and Statistics, 20:18–29, 2016.
- [6] C. Ciric, P. Ciffroy, and S. Charles. Use of sensitivity analysis to identify influential and non-influential parameters within an aquatic ecosystem model. Ecological Modelling, 246:119–130, 2012.
- [7] T. Crestaux, O. L. Maître, and J.-M. Martinez. Polynomial chaos expansions for uncertainties quantification and sensitivity analysis. Reliability Engineering and System Safety, 94:1161–1172, 2009.
- [8] R. Cukier, H. Levine, and K. Shuler. Nonlinear sensitivity analysis of multiparameter model systems. Journal of computational physics, 26(1):1–42, 1978.

- [9] S. Da Veiga and F. Gamboa. Efficient estimation of sensitivity indices. Journal of Nonparametric Statistics, 25(3):573–595, 2013.
- [10] S. Da Veiga, F. Wahl, and F. Gamboa. Local polynomial estimation for sensitivity analysis on models with correlated inputs. Technometrics, 51(4):452–463, 2009.
- [11] B. Efron and C. Stein. The jackknife estimate of variance. The Annals of Statistics, 9(3):586–596, 1981.
- [12] O. G. Ernst, A. Mugler, H.-J. Starkloff, and E. Ullmann. On the convergence of generalized polynomial chaos expansions. ESAIM: Mathematical Modelling and Numerical Analysis, 46(2):317–339, 2012.
- [13] R. Ghanem and P. Spanos. Stochastic finite elements - A spectral approach. Berlin: Springer, 1991.
- [14] E. Giné and R. Nickl. A simple adaptive estimator of the integrated square of a density. Bernoulli, 14(1):47–61, 2008.
- [15] W. Hoeffding. A class of statistics with asymptotically normal distribution. Ann. Math. Statist., 19(3):293–325, 09 1948.
- [16] T. Homma and A. Saltelli. Importance measures in global sensitivity analysis of non linear models. Reliability Engineering and System Safety, 52:1–17, 1996.
- [17] B. Iooss, A. Janon, and G. Pujol. sensitivity: Global Sensitivity Analysis of Model Outputs, 2018. R package version 1.15.2.
- [18] B. Iooss and P. Lemaitre. A review on global sensitivity analysis methods. In C. Meloni and G. Dellino, editors, Uncertainty management in Simulation-Optimization of Complex Systems: Algorithms and Applications, pages 101–122. Springer, 2015.
- [19] B. Iooss, A.-L. Popelin, G. Blatman, C. Ciric, F. Gamboa, S. Lacaze, and M. Lamboni. Some new insights in derivative-based global sensitivity measures. In Proceedings of the PSAM11 ESREL 2012 Conference, pages 1094–1104, Helsinki, Finland, June 2012.

- [20] B. Iooss and A. Saltelli. Introduction: Sensitivity analysis. In R. Ghanem, D. Higdon, and H. Owhadi, editors, Springer Handbook on Uncertainty Quantification, pages 1103–1122. Springer, 2017.
- [21] S. Kucherenko and B. Iooss. Derivative-based global sensitivity measures. In R. Ghanem, D. Higdon, and H. Owhadi, editors, Springer Handbook on Uncertainty Quantification, pages 1241–1263. Springer, 2017.
- [22] S. Kucherenko, M. Rodriguez-Fernandez, C. Pantelides, and N. Shah. Monte carlo evaluation of derivative-based global sensitivity measures. Reliability Engineering and System Safety, 94:1135–1148, 2009.
- [23] S. Kucherenko and S. Song. Derivative-based global sensitivity measures and their link with Sobol’ sensitivity indices. In R. Cools and D. Nuyens, editors, Monte Carlo and Quasi-Monte Carlo Methods, pages 455–469, Cham, 2016. Springer International Publishing.
- [24] F. Kuo, I. Sloan, G. Wasilkowski, and H. Woźniakowski. On decompositions of multivariate functions. Mathematics of Computation, 79(270):953–966, 2010.
- [25] M. Lamboni, B. Iooss, A.-L. Popelin, and F. Gamboa. Derivative-based global sensitivity measures: General links with Sobol’ indices and numerical tests. Mathematics and Computers in Simulation, 87:45–54, 2013.
- [26] B. Laurent. Efficient estimation of integral functionals of a density. The Annals of Statistics, 24(2):659–681, 1996.
- [27] B. Laurent and P. Massart. Adaptive estimation of a quadratic functional by model selection. The Annals of Statistics, 28(5):1302–1338, 2000.
- [28] C. Prieur and S. Tarantola. Variance-based sensitivity analysis: Theory and estimation algorithms. In R. Ghanem, D. Higdon, and H. Owhadi, editors, Springer Handbook on Uncertainty Quantification, pages 1217–1239. Springer, 2017.

- [29] L. Pronzato. Sensitivity analysis via Karhunen-Loève expansion of a random field model: Estimation of Sobol’ indices and experimental design. Reliability Engineering and System Safety, 187:93–109, 2019.
- [30] O. Roustant, F. Barthe, and B. Iooss. Poincaré inequalities on intervals - application to sensitivity analysis. Electron. J. Statist., 11(2):3081–3119, 2017.
- [31] I. Sobol’. Multidimensional quadrature formulas and Haar functions. Izdat’ Nauka’, Moscow, 1969.
- [32] I. Sobol’. Sensitivity estimates for non linear mathematical models. Mathematical Modelling and Computational Experiments, 1:407–414, 1993.
- [33] I. Sobol’ and A. Gershman. On an alternative global sensitivity estimator. In Proceedings of SAMO 1995, pages 40–42, Belgirate, Italy, 1995.
- [34] I. Sobol’ and S. Kucherenko. Derivative based global sensitivity measures and their links with global sensitivity indices. Mathematics and Computers in Simulation, 79:3009–3017, 2009.
- [35] S. Song, T. Zhou, L. Wang, S. Kucherenko, and Z. Lu. Derivative-based new upper bound of Sobol’ sensitivity measure. Reliability Engineering & System Safety, 187:142 – 148, 2019.
- [36] B. Sudret. Global sensitivity analysis using polynomial chaos expansion. Reliability Engineering and System Safety, 93:964–979, 2008.
- [37] B. Sudret and C. V. Mai. Computing derivative-based global sensitivity measures using polynomial chaos expansions. Reliability Engineering & System Safety, 134:241–250, 2015.
- [38] J.-Y. Tissot. Sur la décomposition ANOVA et l’estimation des indices de Sobol’. Application à un modèle d’écosystème marin. PhD thesis, Grenoble University, 2012.
- [39] N. Wiener. The homogeneous chaos. American Journal of Mathematics, 60(4):897–936, 1938.