Sobol-Hoeffding Decomposition with Application to Global Sensitivity Analysis

Olivier Le Maître[1]
with Colleague & Friend
Omar Knio

[1] LIMSI, CNRS
UPR-3251, Orsay, France
https://perso.limsi.fr/olm/

PhD course on UQ - DTU

Table of content

L_2 functions over unit-hypercubes

Let $L_2(\mathcal{U}^d)$ be the space of real-valued **squared-integrable functions** over the d-dimensional hypercube $\mathcal{U}$:

$$f : \boldsymbol{x} \in \mathcal{U}^d \mapsto f(\boldsymbol{x}) \in \mathbb{R}, \quad f \in L_2(\mathcal{U}^d) \Leftrightarrow \int_{\mathcal{U}^d} f(\boldsymbol{x})^2 d\boldsymbol{x} < \infty.$$

$L_2(\mathcal{U}^d)$ is equipped with the inner product $\langle \cdot, \cdot \rangle$,

$$\forall f, g \in L_2(\mathcal{U}^d), \quad \langle u, v \rangle := \int_{\mathcal{U}^d} f(\boldsymbol{x}) g(\boldsymbol{x}) d\boldsymbol{x},$$

and norm $\|\cdot\|_2$,

$$\forall f \in L_2(\mathcal{U}^d), \quad \|f\|_2 := \langle f, f \rangle^{1/2}.$$

Limsi

L_2 functions over unit-hypercubes

Let $L_2(\mathcal{U}^d)$ be the space of real-valued **squared-integrable functions** over the d-dimensional hypercube $\mathcal{U}$:

$$f : \boldsymbol{x} \in \mathcal{U}^d \mapsto f(\boldsymbol{x}) \in \mathbb{R}, \quad f \in L_2(\mathcal{U}^d) \Leftrightarrow \int_{\mathcal{U}^d} f(\boldsymbol{x})^2 d\boldsymbol{x} < \infty.$$

NB: all subsequent developments immediately extend to product-type situations, where

$$\boldsymbol{x} \in A = A_1 \times \cdots \times A_d \subseteq \mathbb{R}^d,$$

and weighted spaces $L_2(A, \rho)$,

$$\rho : \boldsymbol{x} \in A \mapsto \rho(\boldsymbol{x}) \geq 0, \quad \rho(\boldsymbol{x}) = \rho_1(x_1) \times \cdots \times \rho_d(x_d).$$

(*e.g.*: ρ is a pdf of a random vector $\boldsymbol{x}$ with mutually independent components.)

Ensemble notations

Let $\mathfrak{D} = \{1, 2, \ldots, d\}$.
Given $\mathfrak{i} \subseteq \mathfrak{D}$, we denote $\mathfrak{i}_\sim := \mathfrak{D} \setminus \mathfrak{i}$ **its complement set in** $\mathfrak{D}$, such that

$$\mathfrak{i} \cup \mathfrak{i}_\sim = \mathfrak{D}, \quad \mathfrak{i} \cap \mathfrak{i}_\sim = \emptyset.$$

For instance

- $\mathfrak{i} = \{1, 2\}$ and $\mathfrak{i}_\sim = \{3, \ldots, d\}$,
- $\mathfrak{i} = \mathfrak{D}$ and $\mathfrak{i}_\sim = \emptyset$.

Ensemble notations

Let $\mathfrak{D} = \{1, 2, \ldots, d\}$.
Given $\mathfrak{i} \subseteq \mathfrak{D}$, we denote $\mathfrak{i}_\sim := \mathfrak{D} \setminus \mathfrak{i}$ **its complement set in** $\mathfrak{D}$, such that

$$\mathfrak{i} \cup \mathfrak{i}_\sim = \mathfrak{D}, \quad \mathfrak{i} \cap \mathfrak{i}_\sim = \emptyset.$$

Given $\boldsymbol{x} = (x_1, \ldots, x_d)$, we denote $\boldsymbol{x}_\mathfrak{i}$ the vector having for components the $x_{i \in \mathfrak{i}}$, that is

$$\mathfrak{D} \supseteq \mathfrak{i} = \{i_1, \ldots, i_{|\mathfrak{i}|}\} \Rightarrow \boldsymbol{x}_\mathfrak{i} = (x_{i_1}, \ldots, x_{i_{|\mathfrak{i}|}}),$$

where $|\mathfrak{i}| := \mathrm{Card}(\mathfrak{i})$. For instance

$$\int_{\mathcal{U}^{|\mathfrak{i}|}} f(\boldsymbol{x}) d\boldsymbol{x}_\mathfrak{i} = \int_{\mathcal{U}^{|\mathfrak{i}|}} f(x_1, \ldots, x_d) \prod_{i \in \mathfrak{i}} dx_i,$$

and

$$\int_{\mathcal{U}^{d-|\mathfrak{i}|}} f(\boldsymbol{x}) d\boldsymbol{x}_{\mathfrak{i}_\sim} = \int_{\mathcal{U}^{d-|\mathfrak{i}|}} f(x_1, \ldots, x_d) \prod_{i \in \mathfrak{D}}^{i \notin \mathfrak{i}} dx_i,$$

Sobol-Hoeffding decomposition

Any $f \in L_2(\mathcal{U}^d)$ has a **unique hierarchical orthogonal decomposition** of the form

$$f(\boldsymbol{x}) = f(x_1, \ldots, x_d) = f_0 + \sum_{i=1}^{d} f_i(x_i) + \sum_{i=1}^{d} \sum_{j=i+1}^{d} f_{i,j}(x_i, x_j) + \sum_{i=1}^{d} \sum_{j=i+1}^{d} \sum_{k=j+1}^{d} f_{i,j,k}(x_i, x_j, x_k) + \cdots + f_{1,\ldots,d}(x_1, \ldots, x_d).$$

Hierarchical: 1st order functionals (f_i) $\rightarrow$ 2nd order functionals $(f_{i,j})$ $\rightarrow$ 3rd order functionals $(f_{i,j,l})$ $\rightarrow \cdots \rightarrow$ d-th order functional $(f_{1,\ldots,d})$.

Decomposition in a sum of 2^k functionals

Using ensemble notations:

$$f(\boldsymbol{x}) = \sum_{\mathfrak{i} \subseteq \mathfrak{D}} f_{\mathfrak{i}}(\boldsymbol{x}_{\mathfrak{i}}).$$

Limsi

Sobol-Hoeffding decomposition

Any $f \in L_2(\mathcal{U}^d)$ has a **unique hierarchical orthogonal decomposition** of the form

$$f(\boldsymbol{x}) = \sum_{\mathfrak{i} \subseteq \mathfrak{D}} f_{\mathfrak{i}}(\boldsymbol{x}_{\mathfrak{i}}).$$

Orthogonal: the functionals if the S-H decomposition verify the following orthogonality relations:

$$\int_{\mathcal{U}} f_{\mathfrak{i}}(\boldsymbol{x}_{\mathfrak{i}}) dx_j = 0, \qquad \forall \mathfrak{i} \subseteq \mathfrak{D},\ j \in \mathfrak{i},$$

$$\int_{\mathcal{U}^d} f_{\mathfrak{i}}(\boldsymbol{x}_{\mathfrak{i}}) f_{\mathfrak{j}}(\boldsymbol{x}_{\mathfrak{j}}) d\boldsymbol{x} = \langle f_{\mathfrak{i}}, f_{\mathfrak{j}} \rangle = 0, \qquad \forall \mathfrak{i}, \mathfrak{j} \subseteq \mathfrak{D},\ \mathfrak{i} \neq \mathfrak{j}.$$

It follows the hierarchical construction

$$f_{\emptyset} = \int_{\mathcal{U}^d} f(\boldsymbol{x}) d\boldsymbol{x} = \langle f \rangle_{\emptyset_{\sim} = \mathfrak{D}}$$

$$f_{\{i\}} = \int_{\mathcal{U}^{d-1}} f(\boldsymbol{x}) d\boldsymbol{x}_{\{i\}_{\sim}} - f_{\emptyset} = \langle f \rangle_{\mathfrak{D} \setminus \{i\}} - f_{\emptyset} \qquad i \in \mathfrak{D}$$

$$f_{\mathfrak{i}} = \int_{\mathcal{U}^{|\mathfrak{i}_{\sim}|}} f(\boldsymbol{x}) d\boldsymbol{x}_{\mathfrak{i}_{\sim}} - \sum_{\mathfrak{j} \subsetneq \mathfrak{i}} f_{\mathfrak{j}} = \langle f \rangle_{\mathfrak{i}_{\sim}} - \sum_{\mathfrak{j} \subsetneq \mathfrak{i}} f_{\mathfrak{j}} \qquad \mathfrak{i} \in \mathfrak{D},\ |\mathfrak{i}| \geq 2.$$

Table of content

Parametric sensitivity analysis

Consider $\boldsymbol{x}$ as a set of d independent random parameters uniformly distributed on $\mathcal{U}^d$, and $f(\boldsymbol{x})$ a model-output depending on these random parameters. It is assumes that f is a 2nd order random variable: $f \in L_2(\mathcal{U}^d)$. Thus, f has a unique S-H decomposition

$$f(\boldsymbol{x}) = \sum_{\mathfrak{i} \subseteq \mathfrak{D}} f_{\mathfrak{i}}(\boldsymbol{x}_{\mathfrak{i}}).$$

Further, the integrals of f with respect to $\mathfrak{i}_\sim$ are in this context conditional expectations,

$$\mathbb{E}\left[f|\boldsymbol{x}_{\mathfrak{i}}\right] = \int_{\mathcal{U}^{|\mathfrak{i}_\sim|}} f(\boldsymbol{x}) d\boldsymbol{x}_{\mathfrak{i}_\sim} = g(\boldsymbol{x}_{\mathfrak{i}}) \quad \forall \mathfrak{i} \subseteq \mathfrak{D},$$

so the S-H decomposition follows the hierarchical structure

$$\begin{aligned} f_\emptyset &= \mathbb{E}\left[f\right] & & \\ f_{\{i\}} &= \mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right] - \mathbb{E}\left[f\right] & & i \in \mathfrak{D} \\ f_{\mathfrak{i}} &= \mathbb{E}\left[f|\boldsymbol{x}_{\mathfrak{i}}\right] - \sum_{\mathfrak{j} \subsetneq \mathfrak{i}} f_{\mathfrak{j}} & & \mathfrak{i} \subseteq \mathfrak{D},\ |\mathfrak{i}| \geq 2. \end{aligned}$$

Limsi

Variance decomposition

Because of the orthogonality of the S-H decomposition the variance $\mathbb{V}[f]$ of the model-output can be decomposed as

$$\mathbb{V}[f] = \sum_{\mathfrak{i}\subseteq\mathfrak{D}}^{\mathfrak{i}\neq\emptyset} \mathbb{V}[f_{\mathfrak{i}}], \quad \mathbb{V}[f_{\mathfrak{i}}] = \langle f_{\mathfrak{i}}, f_{\mathfrak{i}} \rangle .$$

$\mathbb{V}[f_{\mathfrak{i}}]$ is interpreted as the contribution to the total variance $\mathbb{V}[f]$ of the interaction between parameters $x_{i\in\mathfrak{i}}$.

The S-H decomposition thus provide a rich mean of analyzing the respective contributions of individual or sets of parameters to model-output variability. However, as there are $2^d - 1$ contributions, so one needs more "abstract" characterizations.

Sensitivity indices

To facilitate the hierarchization of the respective influence of each parameter x_i, the partial variances $\mathbb{V}[f_{\mathfrak{i}}]$ are normalized by $\mathbb{V}[f]$ to obtain the **sensitivity indices**:

$$S_{\mathfrak{i}}(f) = \frac{\mathbb{V}[f_{\mathfrak{i}}]}{\mathbb{V}[f]} \leq 1, \quad \sum_{\mathfrak{i}\subseteq\mathfrak{D}}^{\mathfrak{i}\neq\emptyset} S_{\mathfrak{i}}(f) = 1.$$

The **order of the sensitivity indices** $S_{\mathfrak{i}}$ is equal to $|\mathfrak{i}| = \mathrm{Card}(\mathfrak{i})$.

1st order sensitivity indices. The d **first order indices** $S_{\{i\}\in\mathfrak{D}}$ characterize the **fraction of the variance due the parameter x_i only**, *i.e.* **without any interaction with others**. Therefore,

$$1 - \sum_{i=1}^{d} S_{\{i\}}(f) \geq 0,$$

measures globally the effect on the variability of all interactions between parameters. **If $\sum_{i=1}^{d} S_{\{i\}} = 1$, the model is said additive**, because its S-H decomposition is

$$f(x_i, \ldots, x_d) = f_0 + \sum_{i=1}^{d} f_i(x_i),$$

and the impact of the parameters can be studied separately.

Sensitivity indices

To facilitate the hierarchization of the respective influence of each parameter x_i, the partial variances $\mathbb{V}[f_{\mathfrak{i}}]$ are normalized by $\mathbb{V}[f]$ to obtain the **sensitivity indices**:

$$S_{\mathfrak{i}}(f) = \frac{\mathbb{V}[f_{\mathfrak{i}}]}{\mathbb{V}[f]} \leq 1, \quad \sum_{\mathfrak{i} \subseteq \mathfrak{D}}^{\mathfrak{i} \neq \emptyset} S_{\mathfrak{i}}(f) = 1.$$

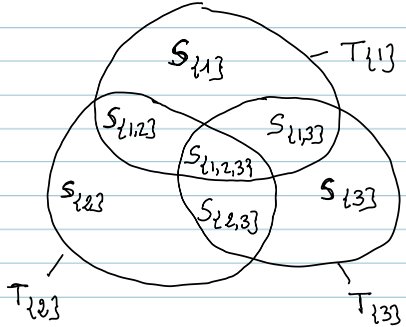
The **order of the sensitivity indices** $S_{\mathfrak{i}}$ is equal to $|\mathfrak{i}| = \mathrm{Card}(\mathfrak{i})$.

Total sensitivity indices. The first order SI $S_{\{i\}}$ measures the variability due to parameter x_i alone. The **total SI $T_{\{i\}}$ measures the variability due to the parameter x_i, including all its interactions** with other parameters:

$$T_{\{i\}} := \sum_{\mathfrak{i} \ni i} S_{\mathfrak{i}} \geq S_{\{i\}}.$$

Important point: for x_i to be deemed non-important or non-influent on the model-output, $S_{\{i\}}$ **and** $T_{\{i\}}$ have to be negligible.
Observe that $\sum_{i \in \mathfrak{D}} T_{\{i\}} \geq 1$, the excess from 1 **characterizes the presence of interactions** in the model-output.

Sensitivity indices



Case of $d=3$.

$$\sum S_i = 1$$

$$\sum S_{\{i\}} \leq 1$$

$$\sum T_{\{i\}} \geq 1$$

Sensitivity indices

In many uncertainty problem, the **set of uncertain parameters can be naturally grouped into subsets** depending on the process each parameter accounts for. For instance, boundary conditions BC, material property φ, external forcing F, and $\mathfrak{D}$ is the union of these distinct subsets:

$$\mathfrak{D} = \mathfrak{D}_{BC} \cup \mathfrak{D}_{\varphi} \cup \mathfrak{D}_{F}.$$

The notion of **first order and total sensitivity indices can be extended to characterize the influence of the subsets of parameters**. For instance,

$$S_{\mathfrak{D}_{\varphi}} = \sum_{\mathfrak{i} \subseteq \mathfrak{D}_{\varphi}} S_{\mathfrak{i}},$$

measures the fraction of variance induced by the material uncertainty alone, while

$$T_{\mathfrak{D}_{F}} = \sum_{\mathfrak{i} \cap \mathfrak{D}_{F} \neq \emptyset} S_{\mathfrak{i}}.$$

measures the fraction of variance due to the external forcing uncertainty and all its interactions.

Let (ξ_1, ξ_2) be two independent centered, normalized random variables

$$\xi_i \sim \mathrm{N}(0,1), \quad i = 1, 2.$$

Consider the model-output $f : (\xi_1, \xi_2) \in \mathbb{R}^2 \mapsto \mathbb{R}$ given by

$$f(\xi_1, \xi_2) = (\mu_1 + \sigma_1 \xi_1) + (\mu_2 + \sigma_2 \xi_2) \, .$$

1. Determine the S-H decomposition of f
2. Compute the 1st order and total sensitivity indices of f
3. Comment
4. Repeat for $f(\xi_1, \xi_2) = (\mu_1 + \sigma_1 \xi_1)\,(\mu_2 + \sigma_2 \xi_2)$.

$$f(\xi_1, \xi_2) = \mu_1 + \mu_2 + \sigma_1\xi_1 + \sigma_2\xi_2$$

- $\mathbb{E}[f] = (\mu_1 + \mu_2)$
- $\mathbb{E}[f|\xi_1] = (\mu_1 + \mu_2) + \sigma_1\xi_1$
  $$\Rightarrow f_1(\xi_1) = \mathbb{E}[f|\xi_1] - \mathbb{E}[f] = \sigma_1\xi_1$$
- $\mathbb{E}[f|\xi_2] = (\mu_1 + \mu_2) + \sigma_2\xi_2$
  $$\Rightarrow f_2(\xi_2) = \mathbb{E}[f|\xi_2] - \mathbb{E}[f] = \sigma_2\xi_2$$
- $\mathbb{E}[f|\xi_1, \xi_2] = f(\xi_1, \xi_2)$
  $$\Rightarrow f_{1,2}(\xi_1, \xi_1) = \mathbb{E}[f|\xi_1, \xi_2] - \mathbb{E}[f] - f_1(\xi_1) - f_2(\xi_2) = 0$$
- Then, $\mathbb{V}[f] = \sigma_1^2 + \sigma_2^2$, so
  $$S_1 = T_1 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \text{ and } S_2 = T_2 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$
- Comment: obvious case, as *f* is a linear (additive) model.

$$f(\xi_1,\xi_2) = \mu_1\mu_2 + \mu_2\sigma_1\xi_1 + \mu_1\sigma_2\xi_2 + \sigma_1\sigma_2\xi_1\xi_2.$$

- $\mathbb{E}[f] = \mu_1\mu_2$
- $\mathbb{E}[f|\xi_1] = \mu_1\mu_2 + \mu_2\sigma_1\xi_1$
  $$\Rightarrow f_1(\xi_1) = \mathbb{E}[f|\xi_1] - \mathbb{E}[f] = \mu_2\sigma_1\xi_1$$
- $\mathbb{E}[f|\xi_2] = \mu_1\mu_2 + \mu_1\sigma_2\xi_2$
  $$\Rightarrow f_2(\xi_2) = \mathbb{E}[f|\xi_2] - \mathbb{E}[f] = \mu_1\sigma_2\xi_2$$
- $\mathbb{E}[f|\xi_1,\xi_2] = f(\xi_1,\xi_2)$
  $$\Rightarrow f_{1,2}(\xi_1,\xi_1) = \mathbb{E}[f|\xi_1,\xi_2] - \mathbb{E}[f] - f_1(\xi_1) - f_2(\xi_2) = \sigma_1\sigma_2\xi_1\xi_2$$
- Then, $\mathbb{V}[f] = \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2$, so
  $$S_1 = \frac{\mu_2^2\sigma_1^2}{\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2} \text{ and } S_2 = \frac{\mu_1^2\sigma_2^2}{\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2},$$
  $$T_1 = \frac{\mu_2^2\sigma_1^2 + \sigma_1\sigma_2}{\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2} \text{ and } T_2 = \frac{\mu_1^2\sigma_2^2 + \sigma_1\sigma_2}{\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2},$$
- Comment: fraction of variance due to interactions is
  $$\sigma_1^2\sigma_2^2/(\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2)$$

$$f(\xi_1, \xi_2) = \mu_1\mu_2 + \mu_2\sigma_1\xi_1 + \mu_1\sigma_2\xi_2 + \sigma_1\sigma_2\xi_1\xi_2.$$

- Then, $\mathbb{V}[f] = \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2$, so

$$S_1 = \frac{\mu_2^2\sigma_1^2}{\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2} \text{ and } S_2 = \frac{\mu_1^2\sigma_2^2}{\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2},$$

$$T_1 = \frac{\mu_2^2\sigma_1^2 + \sigma_1\sigma_2}{\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2} \text{ and } T_2 = \frac{\mu_1^2\sigma_2^2 + \sigma_1\sigma_2}{\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2},$$

- Comment: fraction of variance due to interactions is

$$\sigma_1^2\sigma_2^2/(\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2)$$

- Example: $(\mu_1, \sigma_1) = (1, 3)$ and $(\mu_2, \sigma_2) = (2, 2)$, so

$$S_1 = 9/19, \quad S_2 = 1/19, \quad S_{1,2} = 9/19.$$

  One can draw the conclusions:
  - ξ_1 is the most influential variable as $S_1 > S_2$ and $T_1 > T_2$.
  - interactions are important as $1 - S_1 - S_2 = 9/19 \approx 0.5$, especially for ξ_2 for which $(T_2 - S_2)/T_2 = 9/10 \approx 1$.

Table of content

1-st order sensitivity indices by Monte-Carlo sampling.

The S_i can be computed by MC sampling as follow. Recall that

$$S_{\{i\}}(f) = \frac{\mathbb{V}\left[f_{\{i\}}\right]}{\mathbb{V}[f]} = \frac{\mathbb{V}[\mathbb{E}[f|x_i])]}{\mathbb{V}[f]} = \frac{\mathbb{E}\left[\mathbb{E}[f|x_i]^2\right] - \mathbb{E}[\mathbb{E}[f|x_i]]^2}{\mathbb{V}[f]}.$$

Observe: $\mathbb{E}[\mathbb{E}[f|x_i]] = \mathbb{E}[f]$.

$\mathbb{E}[f]$ and $\mathbb{V}[f]$ can be **estimated using MC sampling**.
Let $\chi_M = \{\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(M)}\}$ be a set of **independent samples drawn uniformly** in $\mathcal{U}^d$, the **mean and variance estimators** are:

$$\widehat{\mathbb{E}[f]} = \frac{1}{M}\sum_{l=1}^{M} f(\boldsymbol{x}^{(l)}), \quad \widehat{\mathbb{V}[f]} = \frac{1}{M-1}\sum_{l=1}^{M} f(\boldsymbol{x}^{(l)})^2 - \widehat{\mathbb{E}[f]}^2.$$

It now remains to compute the variance of conditional expectations, $\mathbb{V}\left[\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right]\right]$.

Any idea?

Monte-Carlo estimator for variance of conditional expectation.

Observe:

$$\int_{\mathcal{U}^{d+|\mathrm{i}_\sim|}} f(\boldsymbol{x}_\mathrm{i}, \boldsymbol{x}_{\mathrm{i}_\sim}) f(\boldsymbol{x}_\mathrm{i}, \boldsymbol{x}'_{\mathrm{i}_\sim}) \mathrm{d}\boldsymbol{x}_\mathrm{i} d\boldsymbol{x}_{\mathrm{i}_\sim} d\boldsymbol{x}_{\mathrm{i}_\sim'}$$
$$= \int_{\mathcal{U}^{|\mathrm{i}|}} d\boldsymbol{x}_\mathrm{i} \int_{\mathcal{U}^{|\mathrm{i}_\sim|}} f(\boldsymbol{x}_\mathrm{i}, \boldsymbol{x}_{\mathrm{i}_\sim}) d\boldsymbol{x}_{\mathrm{i}_\sim} \int_{\mathcal{U}^{|\mathrm{i}_\sim|}} f(\boldsymbol{x}_\mathrm{i}, \boldsymbol{x}'_{\mathrm{i}_\sim}) d\boldsymbol{x}'_{\mathrm{i}_\sim}$$
$$= \int_{\mathcal{U}^{|\mathrm{i}|}} d\boldsymbol{x}_\mathrm{i} \left[\int_{\mathcal{U}^{|\mathrm{i}_\sim|}} f(\boldsymbol{x}_\mathrm{i}, \boldsymbol{x}_{\mathrm{i}_\sim}) d\boldsymbol{x}_{\mathrm{i}_\sim} \right]^2 .$$

Monte-Carlo estimator for variance of conditional expectation.

Observe:

$$\begin{aligned}
&\int_{\mathcal{U}^{d+|\mathrm{i}_\sim|}} f(\boldsymbol{x}_{\mathrm{i}}, \boldsymbol{x}_{\mathrm{i}_\sim}) f(\boldsymbol{x}_{\mathrm{i}}, \boldsymbol{x}'_{\mathrm{i}_\sim}) \mathrm{d}\boldsymbol{x}_{\mathrm{i}}\, d\boldsymbol{x}_{\mathrm{i}_\sim}\, d\boldsymbol{x}_{\mathrm{i}_\sim'} \\
&= \int_{\mathcal{U}^{|\mathrm{i}|}} d\boldsymbol{x}_{\mathrm{i}} \int_{\mathcal{U}^{|\mathrm{i}_\sim|}} f(\boldsymbol{x}_{\mathrm{i}}, \boldsymbol{x}_{\mathrm{i}_\sim}) d\boldsymbol{x}_{\mathrm{i}_\sim} \int_{\mathcal{U}^{|\mathrm{i}_\sim|}} f(\boldsymbol{x}_{\mathrm{i}}, \boldsymbol{x}'_{\mathrm{i}_\sim}) d\boldsymbol{x}'_{\mathrm{i}_\sim} \\
&= \int_{\mathcal{U}^{|\mathrm{i}|}} d\boldsymbol{x}_{\mathrm{i}} \left[\int_{\mathcal{U}^{|\mathrm{i}_\sim|}} f(\boldsymbol{x}_{\mathrm{i}}, \boldsymbol{x}_{\mathrm{i}_\sim}) d\boldsymbol{x}_{\mathrm{i}_\sim} \right]^2 .
\end{aligned}$$

$$\mathbb{V}\left[\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right]\right] = \mathbb{E}\left[\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right]^2\right] - \mathbb{E}\left[\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right]\right]^2 = \mathbb{E}\left[\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right]^2\right] - \mathbb{E}\left[f\right]^2$$

Monte-Carlo estimator for variance of conditional expectation.

Observe:

$$\int_{\mathcal{U}^{d+|\mathrm{i}_\sim|}} f(\boldsymbol{x}_\mathrm{i}, \boldsymbol{x}_{\mathrm{i}_\sim}) f(\boldsymbol{x}_\mathrm{i}, \boldsymbol{x}'_{\mathrm{i}_\sim}) \mathrm{d}\boldsymbol{x}_\mathrm{i} \, d\boldsymbol{x}_{\mathrm{i}_\sim} \, d\boldsymbol{x}_{\mathrm{i}_\sim'}$$
$$= \int_{\mathcal{U}^{|\mathrm{i}|}} d\boldsymbol{x}_\mathrm{i} \int_{\mathcal{U}^{|\mathrm{i}_\sim|}} f(\boldsymbol{x}_\mathrm{i}, \boldsymbol{x}_{\mathrm{i}_\sim}) d\boldsymbol{x}_{\mathrm{i}_\sim} \int_{\mathcal{U}^{|\mathrm{i}_\sim|}} f(\boldsymbol{x}_\mathrm{i}, \boldsymbol{x}'_{\mathrm{i}_\sim}) d\boldsymbol{x}'_{\mathrm{i}_\sim}$$
$$= \int_{\mathcal{U}^{|\mathrm{i}|}} d\boldsymbol{x}_\mathrm{i} \left[\int_{\mathcal{U}^{|\mathrm{i}_\sim|}} f(\boldsymbol{x}_\mathrm{i}, \boldsymbol{x}_{\mathrm{i}_\sim}) d\boldsymbol{x}_{\mathrm{i}_\sim} \right]^2 .$$

$$\mathbb{V}\left[\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right]\right] = \mathbb{E}\left[\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right]^2\right] - \mathbb{E}\left[\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right]\right]^2 = \mathbb{E}\left[\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right]^2\right] - \mathbb{E}\left[f\right]^2$$
$$= \lim_{M\to\infty} \frac{1}{M} \sum_{l=1}^{M} f\left(\boldsymbol{x}^{(l)}_{\{i\}_\sim}, \boldsymbol{x}^{(l)}_{\{i\}}\right) f\left(\tilde{\boldsymbol{x}}^{(l)}_{\{i\}_\sim}, \boldsymbol{x}^{(l)}_{\{i\}}\right) - \mathbb{E}\left[f\right]^2$$

$\left(\text{independent samples } \boldsymbol{x}^{(l)}_{\{i\}_\sim}, \tilde{\boldsymbol{x}}^{(l)}_{\{i\}_\sim} \text{ and } \boldsymbol{x}^{(k)}_{\{i\}}\right)$

Monte-Carlo estimators for 1st order SI $S_{\{i\}}$.

1. Draw 2 independent sample sets, with size M, χ_M and $\tilde{\chi}_M$
2. Compute estimators $\widehat{\mathbb{E}[f]}$ and $\widehat{\mathbb{V}[f]}$ from χ_M (or $\tilde{\chi}_M$) [M model evaluations]
3. For $i = 1, 2, \dots, d$:
   - Estimate variance of conditional expectation through

$$\mathbb{V}\left[\widehat{\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right]}\right] = \frac{1}{M}\sum_{l=1}^{M} f\left(\boldsymbol{x}^{(l)}_{\{i\}_\sim}, \boldsymbol{x}^{(l)}_{\{i\}}\right) f\left(\tilde{\boldsymbol{x}}^{(l)}_{\{i\}_\sim}, \boldsymbol{x}^{(l)}_{\{i\}}\right) - \widehat{\mathbb{E}[f]}^2$$

[M new model evaluations at $(\tilde{\boldsymbol{x}}^{(l)}_{\{i\}_\sim}, \boldsymbol{x}^{(l)}_{\{i\}})$]

   - Estimator of the 1-st order SI:

$$\widehat{S_{\{i\}}(f)} = \frac{\mathbb{V}\left[\widehat{\mathbb{E}\left[f|\xi_{\{i\}}\right]}\right]}{\widehat{\mathbb{V}[f]}}.$$

Requires $(d+1) \times M$ model evaluations.

Monte-Carlo estimation

$$\mathcal{X}_M : \begin{bmatrix} x_1^{(1)} & \dots & x_i^{(1)} & \dots & x_d^{(1)} \\ \vdots & & & & \vdots \\ x_1^{(l)} & \dots & x_i^{(l)} & \dots & x_d^{(l)} \\ \vdots & & & & \vdots \\ x_1^{(M)} & \dots & x_i^{(M)} & \dots & x_d^{(M)} \end{bmatrix} \qquad \tilde{\mathcal{X}}_M : \begin{bmatrix} \tilde{x}_1^{(1)} & \dots & \tilde{x}_i^{(1)} & \dots & \tilde{x}_d^{(1)} \\ \vdots & & & & \vdots \\ \tilde{x}_1^{(l)} & \dots & \tilde{x}_i^{(l)} & \dots & \tilde{x}_d^{(l)} \\ \vdots & & & & \vdots \\ \tilde{x}_1^{(M)} & \dots & \tilde{x}_i^{(M)} & \dots & \tilde{x}_d^{(M)} \end{bmatrix}.$$

$$\tilde{\mathcal{X}}_M^{\{i\}} : \begin{bmatrix} \tilde{x}_1^{(1)} & \dots & x_i^{(1)} & \dots & \tilde{x}_d^{(1)} \\ \vdots & & \vdots & & \vdots \\ \tilde{x}_1^{(l)} & \dots & x_i^{(l)} & \dots & \tilde{x}_d^{(l)} \\ \vdots & & \vdots & & \vdots \\ \tilde{x}_1^{(M)} & \dots & x_i^{(M)} & \dots & \tilde{x}_d^{(M)} \end{bmatrix}.$$

$$\mathbb{V}\left[\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right]\right] \approx \frac{1}{M}\sum_{l=1}^{M} f(\boldsymbol{x}_{\{i\}_\sim}, \boldsymbol{x}_{\{i\}_\sim}) f(\tilde{\boldsymbol{x}}_{\{i\}_\sim}, \boldsymbol{x}_{\{i\}_\sim}) - \left(\frac{1}{M}\sum_{l=1}^{M} f(\boldsymbol{x}_{\{i\}_\sim}, \boldsymbol{x}_{\{i\}_\sim})\right)^2.$$

Limsi

Total sensitivity indices by Monte-Carlo sampling.

The $T_{\{i\}}$ can be computed by MC sampling as follow. Recall that

$$T_{\{i\}}(f) = \sum_{\mathfrak{i} \ni \{i\}} S_{\mathfrak{i}}(f) = 1 - \sum_{\mathfrak{i} \subseteq \mathfrak{D} \setminus \{i\}} S_{\mathfrak{i}}(f) = 1 - \frac{\mathbb{V}\left[\mathbb{E}\left[f | \boldsymbol{x}_{\{i\}_\sim}\right]\right]}{\mathbb{V}[f]}.$$

As for $\mathbb{V}\left[\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right]\right]$, we can derive the following Monte-Carlo estimator for $\mathbb{V}\left[\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}_\sim}\right]\right]$, using the two independent sample sets χ_M and $\tilde{\chi}_M$

$$\mathbb{V}\left[\widehat{\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}_\sim}\right]}\right] = \frac{1}{M} \sum_{l=1}^{M} f\left(\boldsymbol{x}^{(l)}_{\{i\}_\sim}, \boldsymbol{x}^{(l)}_{\{i\}}\right) f\left(\boldsymbol{x}^{(l)}_{\{i\}_\sim}, \tilde{\boldsymbol{x}}^{(l)}_{\{i\}}\right) - \mathbb{E}[f]^2,$$

so finally

$$\widehat{T_{\{i\}}(f)} = 1 - \frac{1}{\widehat{\mathbb{V}[f]}} \left(\frac{1}{M} \sum_{l=1}^{M} f\left(\boldsymbol{x}^{(l)}_{\{i\}_\sim}, \boldsymbol{x}^{(l)}_{\{i\}}\right) f\left(\boldsymbol{x}^{(l)}_{\{i\}_\sim}, \tilde{\boldsymbol{x}}^{(l)}_{\{i\}}\right) - \mathbb{E}[f]^2 \right)$$

The MC estimation of the $T_{\{i\}}$ needs $d \times M$ additional model evaluations.

$$\mathcal{X}_M : \begin{bmatrix} x_1^{(1)} & \dots & x_i^{(1)} & \dots & x_d^{(1)} \\ \vdots & & & & \vdots \\ x_1^{(l)} & \dots & x_i^{(l)} & \dots & x_d^{(l)} \\ \vdots & & & & \vdots \\ x_1^{(M)} & \dots & x_i^{(M)} & \dots & x_d^{(M)} \end{bmatrix} \qquad \tilde{\mathcal{X}}_M : \begin{bmatrix} \tilde{x}_1^{(1)} & \dots & \tilde{x}_i^{(1)} & \dots & \tilde{x}_d^{(1)} \\ \vdots & & & & \vdots \\ \tilde{x}_1^{(l)} & \dots & \tilde{x}_i^{(l)} & \dots & \tilde{x}_d^{(l)} \\ \vdots & & & & \vdots \\ \tilde{x}_1^{(M)} & \dots & \tilde{x}_i^{(M)} & \dots & \tilde{x}_d^{(M)} \end{bmatrix}.$$

$$\tilde{\mathcal{X}}_M^{\{i\}} : \begin{bmatrix} x_1^{(1)} & \dots & \tilde{x}_i^{(1)} & \dots & x_d^{(1)} \\ \vdots & & \vdots & & \vdots \\ x_1^{(l)} & \dots & \tilde{x}_i^{(l)} & \dots & x_d^{(l)} \\ \vdots & & \vdots & & \vdots \\ x_1^{(M)} & \dots & \tilde{x}_i^{(M)} & \dots & x_d^{(M)} \end{bmatrix}.$$

$$\mathbb{V}\left[\mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right]\right] \approx \frac{1}{M}\sum_{l=1}^{M} f(\boldsymbol{x}_{\{i\}_\sim}, \boldsymbol{x}_{\{i\}_\sim}) f(\boldsymbol{x}_{\{i\}_\sim}, \tilde{\boldsymbol{x}}_{\{i\}_\sim}) - \left(\frac{1}{M}\sum_{l=1}^{M} f(\boldsymbol{x}_{\{i\}_\sim}, \boldsymbol{x}_{\{i\}_\sim})\right)^2.$$

MC estimation of 1st order and total sensitivity indices:

- requires $M \times (2d+1)$ **simulations**
- convergence of estimators in $\mathcal{O}(1/\sqrt{M})$
- slow convergence, but d-independent
- convergence not related to smoothness of f
- works also (in practice) with advanced sampling schemes (QMC, LHS, . . .). How to proceed with QMC/LHS?

S-H decomposition from PC expansions.

Consider the model-output $f : \boldsymbol{\xi} \in \Xi \subset \mathbb{R}^d \mapsto \mathbb{R}$, where $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d)$ are independent real-valued r.v. with joint-probability density function

$$p_{\boldsymbol{\xi}}(x_1, \dots, x_d) = \prod_{i=1}^{d} p_i(x_i).$$

Let $\{ \Psi_{\boldsymbol{\alpha}} \}$ be the set of d-variate orthogonal polynomials,

$$\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \prod_{i=1}^{d} \psi_{\alpha_i}^{(i)}(\xi_i),$$

with $\psi_{l \geq 0}^{(i)} \in \pi_l$ the uni-variate polynomials mutually orthogonal with respect to the density p_i.

If $f \in L_2(\Xi, p_{\boldsymbol{\xi}})$, it has a convergent PC expansion

$$f(\boldsymbol{\xi}) = \lim_{\mathrm{No} \to \infty} \sum_{|\boldsymbol{\alpha}| \leq \mathrm{No}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) f_{\boldsymbol{\alpha}}, \quad |\boldsymbol{\alpha}| = \sum_{i=1}^{d} |\alpha_i|.$$

S-H decomposition from PC expansions.

Given the truncated PC expansion of f,

$$\hat{f}(\boldsymbol{\xi}) = \sum_{|\boldsymbol{\alpha}| \leq \mathcal{A}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) f_{\boldsymbol{\alpha}},$$

one can readily obtain the PC approximation of the S-H functionals through

$$\hat{f}_{\mathfrak{i}}(\boldsymbol{\xi}_{\mathfrak{i}}) = \sum_{|\boldsymbol{\alpha}| \leq \mathcal{A}(\mathfrak{i})} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}_{\mathfrak{i}}) f_{\boldsymbol{\alpha}},$$

where the multi-index set $\mathcal{A}(\mathfrak{i})$ is given by

$$\mathcal{A}(\mathfrak{i}) = \{\boldsymbol{\alpha} \in \mathcal{A}; \alpha_i > 0 \text{ for } i \in \mathfrak{i}, \alpha_i = 0 \text{ for } i \notin \mathfrak{i}\} \subsetneq \mathcal{A}.$$

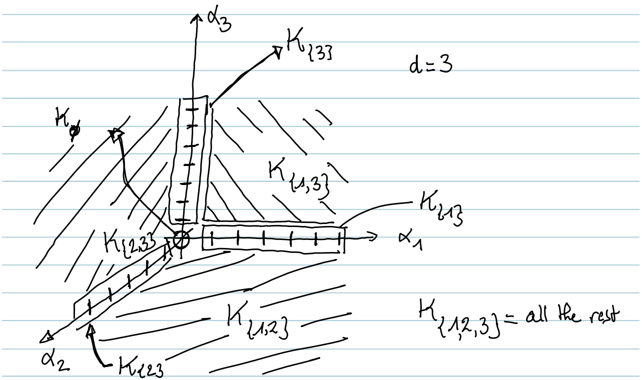
For the sensitivity indices it comes

$$S_{\mathfrak{i}}(\hat{f}) = \frac{\sum_{\boldsymbol{\alpha} \in \mathcal{A}(\mathfrak{i})} f_{\boldsymbol{\alpha}}^2 \langle \Psi_{\boldsymbol{\alpha}}, \Psi_{\boldsymbol{\alpha}} \rangle}{\sum_{\boldsymbol{\alpha} \in \mathcal{A}} f_{\boldsymbol{\alpha}}^2 \langle \Psi_{\boldsymbol{\alpha}}, \Psi_{\boldsymbol{\alpha}} \rangle}, \quad T_{\{i\}}(\hat{f}) = \frac{\sum_{\boldsymbol{\alpha} \in \mathcal{T}(\mathfrak{i})} f_{\boldsymbol{\alpha}}^2 \langle \Psi_{\boldsymbol{\alpha}}, \Psi_{\boldsymbol{\alpha}} \rangle}{\sum_{\boldsymbol{\alpha} \in \mathcal{A}} f_{\boldsymbol{\alpha}}^2 \langle \Psi_{\boldsymbol{\alpha}}, \Psi_{\boldsymbol{\alpha}} \rangle},$$

where

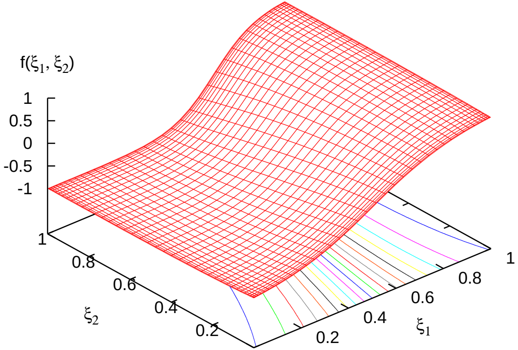
$$\mathcal{T}(\mathfrak{i}) = \{\boldsymbol{\alpha} \in \mathcal{A}; \alpha_i > 0 \text{ for } i \in \mathfrak{i}\}$$

Subsets of PC multi-indices for S-H functionals

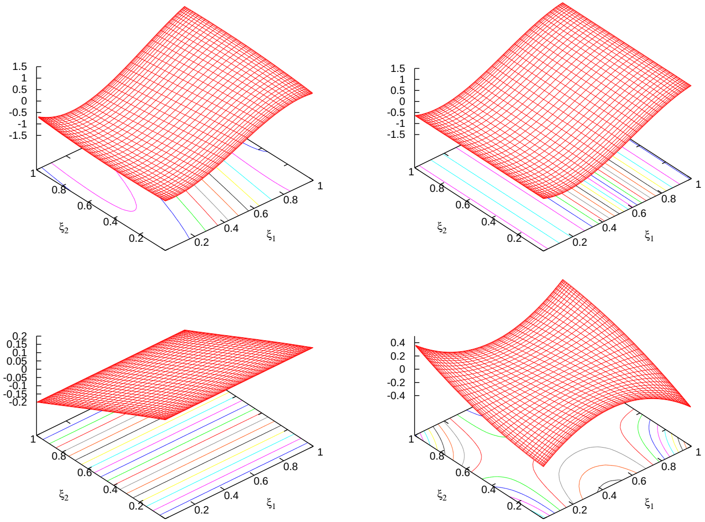


Example: Let (ξ_1, ξ_2) be two independent random variables with uniform distributions on the unit interval. Consider the model-output $f : (\xi_1, \xi_2) \in [0,1]^2 \mapsto \mathbb{R}$ given by

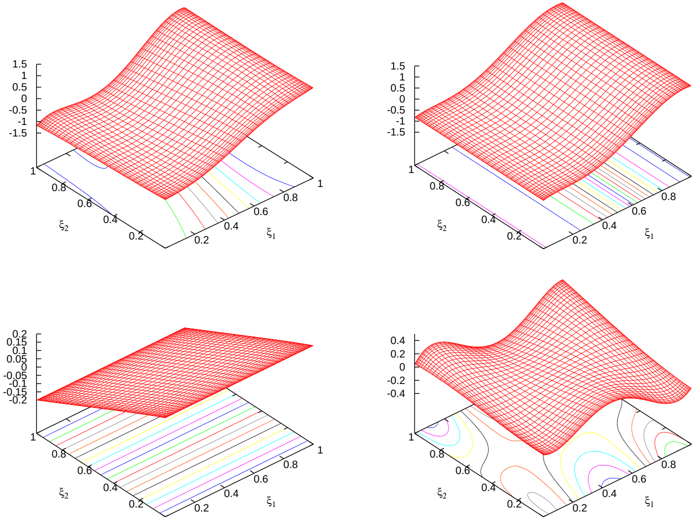
$$f(\xi_1, \xi_2) = \tanh(3(\xi_1 - 0.2\xi_2 - .5)(1 + \xi_2)).$$



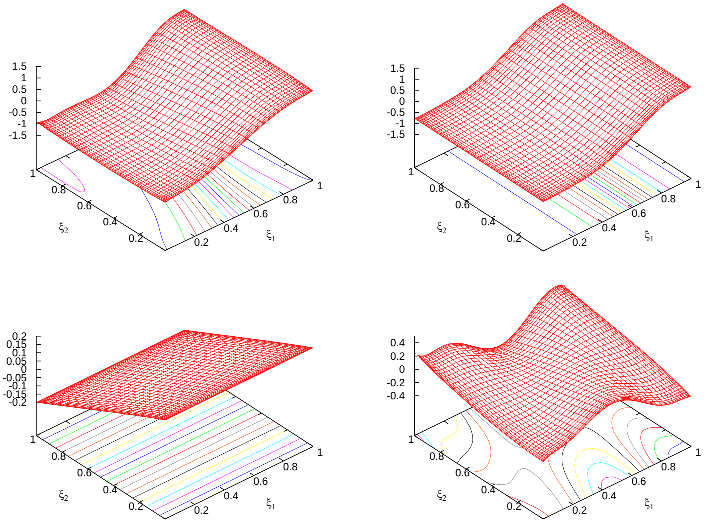
S-H decomposition at $\mathrm{No} = 3$ $(P = 10)$:



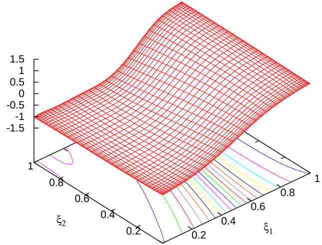
S-H decomposition at $\mathrm{No} = 5$: $(P = 21)$

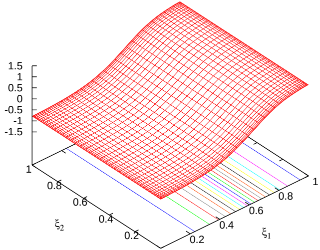


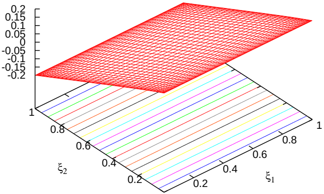
S-H decomposition at $\mathrm{No} = 7$: ($P = 36$)

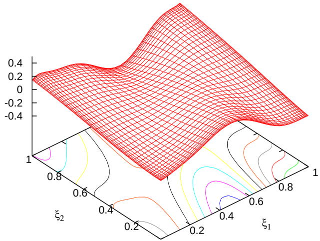


S-H decomposition at $\mathrm{No} = 9$: $(P = 55)$





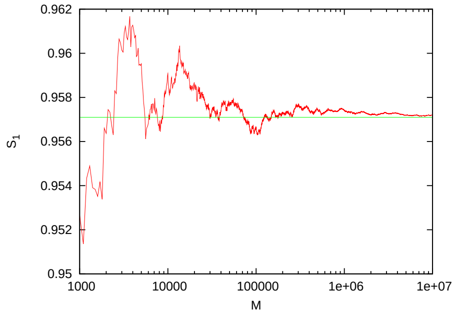




Convergence of sensitivity indices:

No	Card($\mathcal{A}$)	$S_{\{1\}}(\hat{f})$	$S_{\{2\}}(\hat{f})$	$S_{\{1,2\}}(\hat{f})$	$\|\hat{f} - f\|_2$
3	10	9.604 (−1)	2.567 (−2)	1.390 (−2)	7.1 (−2)
5	21	9.578 (−1)	2.545 (−2)	1.676 (−2)	2.6 (−2)
7	36	9.572 (−1)	2.542 (−2)	1.734 (−2)	9.6 (−3)
9	55	9.571 (−1)	2.542 (−2)	1.745 (−2)	3.6 (−3)
15	136	9.571 (−1)	2.542 (−2)	1.748 (−2)	1.9 (−4)

Convergence of sensitivity indices for **naive** MC sampling:



Limsi

Homework: Homma-Saltelli function

$$f(x_1, x_2, x_3) = \sin(x_1) + a\sin^2(x_2) + bx_3^4\sin(x_1),$$

where x_1, x_2 and x_3 are i.i.d. random variables uniformly distributed on $[-\pi, \pi]$. Letting $\langle\cdot\rangle$ be the expectation (over $[-\pi, \pi]^3$), we have

$$\langle\sin\rangle = 0, \quad \left\langle\sin^2\right\rangle = 1/2, \quad \left\langle\sin^4\right\rangle = 3/8, \quad \langle x^n\rangle = \begin{cases} \frac{\pi^n}{n+1}, & n \text{ even}, \\ 0, & n \text{ odd}. \end{cases}$$

1. Compute the SH decomposition of f and the associated sensitivity indices. Solution: $f(x_1, x_2, x_3) = a/2 + \sin(x_1)(1 + b\pi^2/4) + a(\sin^2(x_2) - 1/2) + b\sin(x_1)(x_3^4 - \pi^4/5))$.
2. Compute PC expansion of f by non-intrusive spectral projection for polynomial degrees 1 to 6 and extract 1st order and total sensitivity indices for $a = 7$ and $b = 0.1$.
3. Compute 1-st order and total sensitivity indices by Monte-Carlo simulation.

Further readings:

- Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *The annals of Mathematical Statistics*, **19**, pp. 293-325.
- Sobol, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Wiley*, **1**, pp. 407-414.
- Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulations*, **55**, pp. 271-281.
- Homma T., Saltelli A., (1996). Importance measures in globla sensitivity analysis of nonlinear models. *Reliab. Eng. Syst. Safety*, **52**:1, pp. 1-17.
- Crestaux T., Le Maître O. and Martinez J.M., (2009). Polynomial Chaos expansion for sensitivity analysis. *Reliabg. Eng. Syst. Safety*, **94**:7, pp. 1161-1172.