Asymptotic normality of a Sobol index estimator for stochastic simulators

Loic Le Gratiet[1 2]
PhD supervisor: Josselin Garnier[1]

[1]Université Denis-Diderot, Paris, France

[2]CEA, DAM, DIF, F-91297 Arpajon, France

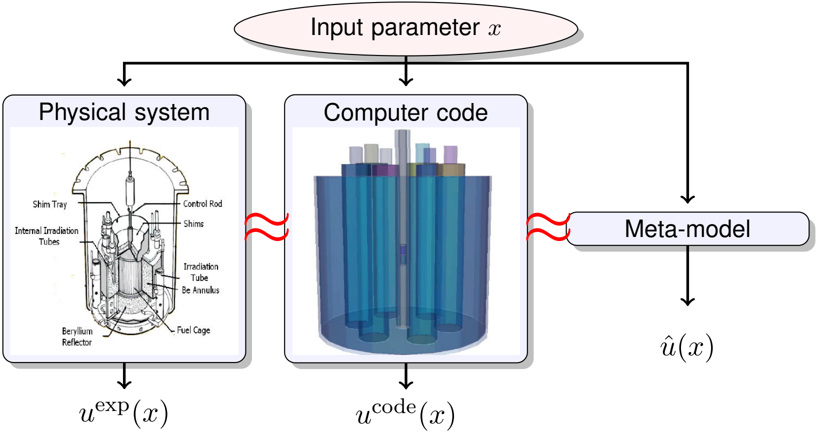
July 1 2013
GdR MASCOT-NUM & SAMO

Outline

Outline

1. Introduction
2. Sensitivity analysis and Gaussian process regression
3. Asymptotic normality of a Sobol index estimator

Introduction to stochastic simulators



- Examples of stochastic simulators: MCNP, TRIPOLI, MORET
- Examples of outputs: maximal energy, neutron multiplication coefficient.

Academic example: the heat equation

Let us consider the heat equation:

$$\frac{\partial u}{\partial t}(x,t) - \frac{1}{2}\Delta u(x,t) = 0,$$

where $x \in \mathbb{R}^d$ and $u(x,0) = g(x) = \exp\left(-\sum_{i=1}^{d} x_i^2/(2\theta_{g,i}^2)\right)$.

Probabilistic representation of the solution: $u(x,t) = \mathbb{E}\left[g(x+W_t)\right]$.

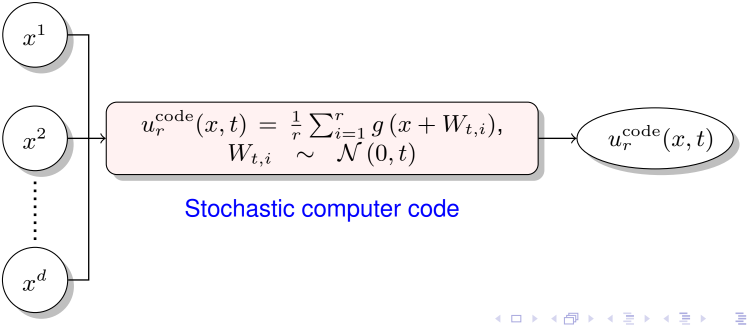
Academic example: the heat equation

Let us consider the heat equation:

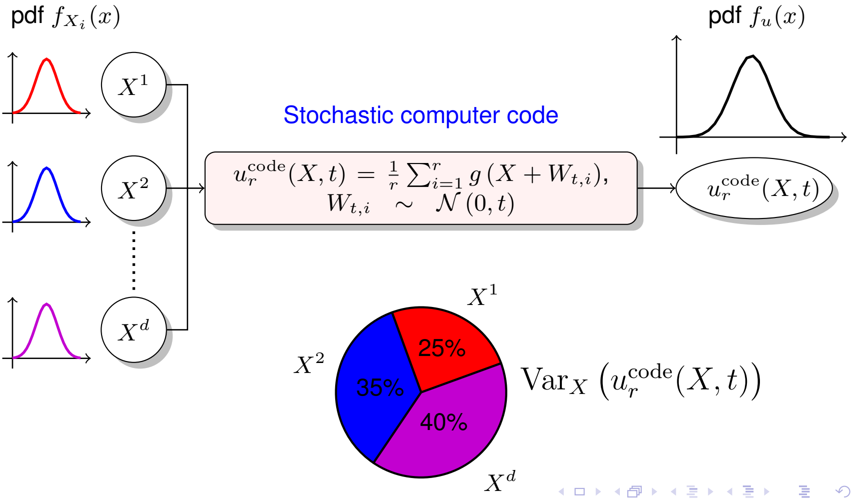
$$\frac{\partial u}{\partial t}(x,t) - \frac{1}{2}\Delta u(x,t) = 0,$$

where $x \in \mathbb{R}^d$ and $u(x,0) = g(x) = \exp\left(-\sum_{i=1}^d x_i^2/(2\theta_{g,i}^2)\right)$.

Probabilistic representation of the solution: $u(x,t) = \mathbb{E}\left[g(x+W_t)\right]$.



The sensitivity analysis



Outline

Sobol index

- Assumptions: $X = (X^i)_{i=1,\dots,d}$ are independent and defined on $(\Omega_X, \mathcal{F}_X, \mathbb{P}_X)$ with probability measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and t is fixed.
- Sensitivity index: We consider the Sobol index of X^1:

$$S^1 = \frac{\mathrm{var}_X\left(\mathbb{E}_X\left[u^{\mathrm{code}}_{r=\infty}(X)|X^1\right]\right)}{\mathrm{var}_X\left(u^{\mathrm{code}}_{r=\infty}(X)\right)}.$$

- Sobol index estimator: Monte-Carlo integration [Sobol, (1993), Sobol et al. (2007), Janon et al. (2012)].

$$S^1_m = \frac{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_i) - m^{-2}\sum_{i,j=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_j)}{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)^2 - m^{-2}(\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i))^2}.$$

- We substitute $u^{\mathrm{code}}_{r=\infty}(X)$ with a meta-model built from noisy observations at points $(D_i)_{i=1,\dots,n}$.
- $(D_i)_{i=1,\dots,n}$ are independent and defined on $(\Omega_D, \mathcal{F}_D, \mathbb{P}_D)$ with probability measure μ.

Sobol index

- Assumptions: $X = (X^i)_{i=1,\dots,d}$ are independent and defined on $(\Omega_X, \mathcal{F}_X, \mathbb{P}_X)$ with probability measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and t is fixed.
- Sensitivity index: We consider the Sobol index of X^1:

$$S^1 = \frac{\mathrm{var}_X\left(\mathbb{E}_X\left[u^{\mathrm{code}}_{r=\infty}(X)|X^1\right]\right)}{\mathrm{var}_X\left(u^{\mathrm{code}}_{r=\infty}(X)\right)}.$$

- Sobol index estimator: Monte-Carlo integration [Sobol, (1993), Sobol et al. (2007), Janon et al. (2012)].

$$S^1_m = \frac{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_i) - m^{-2}\sum_{i,j=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_j)}{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)^2 - m^{-2}(\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i))^2}.$$

- We substitute $u^{\mathrm{code}}_{r=\infty}(X)$ with a meta-model built from noisy observations at points $(D_i)_{i=1,\dots,n}$.
- $(D_i)_{i=1,\dots,n}$ are independent and defined on $(\Omega_D, \mathcal{F}_D, \mathbb{P}_D)$ with probability measure μ.

Sobol index

- Assumptions: $X = (X^i)_{i=1,\dots,d}$ are independent and defined on $(\Omega_X, \mathcal{F}_X, \mathbb{P}_X)$ with probability measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and t is fixed.
- Sensitivity index: We consider the Sobol index of X^1:

$$S^1 = \frac{\mathrm{var}_X\left(\mathbb{E}_X\left[u^{\mathrm{code}}_{r=\infty}(X)|X^1\right]\right)}{\mathrm{var}_X\left(u^{\mathrm{code}}_{r=\infty}(X)\right)}.$$

- Sobol index estimator: Monte-Carlo integration *[Sobol, (1993), Sobol et al. (2007), Janon et al. (2012)]*.

$$S^1_m = \frac{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_i) - m^{-2}\sum_{i,j=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_j)}{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)^2 - m^{-2}(\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i))^2}.$$

- We substitute $u^{\mathrm{code}}_{r=\infty}(X)$ with a meta-model built from noisy observations at points $(D_i)_{i=1,\dots,n}$.
- $(D_i)_{i=1,\dots,n}$ are independent and defined on $(\Omega_D, \mathcal{F}_D, \mathbb{P}_D)$ with probability measure μ.

Sobol index

- Assumptions: $X = (X^i)_{i=1,\dots,d}$ are independent and defined on $(\Omega_X, \mathcal{F}_X, \mathbb{P}_X)$ with probability measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and t is fixed.
- Sensitivity index: We consider the Sobol index of X^1:

$$S^1 = \frac{\mathrm{var}_X\left(\mathbb{E}_X\left[u^{\mathrm{code}}_{r=\infty}(X)|X^1\right]\right)}{\mathrm{var}_X\left(u^{\mathrm{code}}_{r=\infty}(X)\right)}.$$

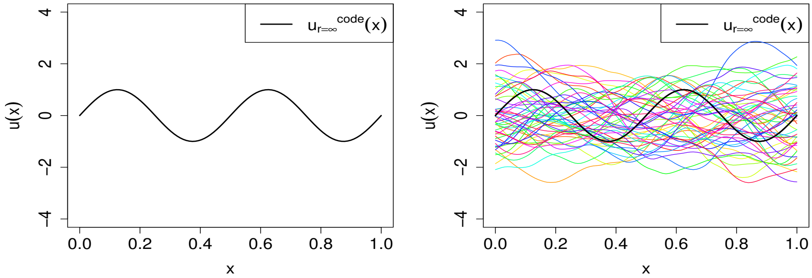
- Sobol index estimator: Monte-Carlo integration *[Sobol, (1993), Sobol et al. (2007), Janon et al. (2012)]*.

$$S^1_m = \frac{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_i) - m^{-2}\sum_{i,j=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_j)}{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)^2 - m^{-2}(\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i))^2}.$$

- We substitute $u^{\mathrm{code}}_{r=\infty}(X)$ with a meta-model built from noisy observations at points $(D_i)_{i=1,\dots,n}$.
- $(D_i)_{i=1,\dots,n}$ are independent and defined on $(\Omega_D, \mathcal{F}_D, \mathbb{P}_D)$ with probability measure μ.

Gaussian process regression

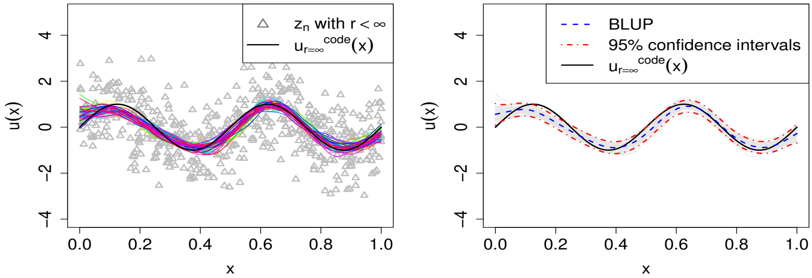
- We model $u^{\text{code}}_{r=\infty}(x)$ with a Gaussian process $Z(x)$ defined on $(\Omega_Z, \mathcal{F}_Z, \mathbb{P}_Z)$ with covariance kernel $k(x, \tilde{x})$.



Gaussian process regression

- Let us consider noisy observations of $u_{r=\infty}^{\text{code}}$ at points $(D_i)_{i=1,\dots,n}$:

$$\mathbf{z}^n = \left(u_{r=\infty}^{\text{code}}(D_i) + \varepsilon_i \right)_{i=1,\dots,n}, \quad \varepsilon_i \sim \mathcal{N}\left(0, \frac{\sigma_g^2}{r}\right).$$



- Remark: $(D_i)_{i=1,\dots,n}$ are defined on $(\Omega_D, \mathcal{F}_D, \mathbb{P}_D)$ and $(\varepsilon_i)_{i=1,\dots,n}$ are defined on $(\Omega_Z, \mathcal{F}_Z, \mathbb{P}_Z)$.

Gaussian process regression

- We consider a fixed number of experiments $T = nr$
- Best Linear Unbiased Predictor (BLUP):

$$\hat{z}_{T,n}(x) = \mathbf{k}'(x)\left(\mathbf{K} + \frac{n\sigma_g^2}{T}\mathbf{I}\right)^{-1}\mathbf{z}^n,$$

where $\mathbf{k}'(x) = [k(x, D_i)]_{1 \le i \le n}$ and $\mathbf{K} = [k(D_i, D_j)]_{1 \le i,j \le n}$.
Note: $\hat{z}_{T,n}(x)$ is defined on $(\Omega_D \times \Omega_Z, \sigma(\mathcal{F}_D \times \mathcal{F}_Z), \mathbb{P}_D \times \mathbb{P}_Z)$

- Mean Squared Error (MSE) of the BLUP (the mean is w.r.t $\mathbb{E}_Z$):

$$\sigma^2_{T,n}(x) = k(x,x) - \mathbf{k}'(x)\left(\mathbf{K} + \frac{n\sigma_g^2}{T}\mathbf{I}\right)^{-1}\mathbf{k}(x).$$

- Let us consider the case $n \gg 1$:

$$\sigma^2_{T,n}(x) \xrightarrow[n\to\infty]{\mathbb{P}_D} \sum_{p\ge 0} \frac{\sigma_g^2 \lambda_p / T}{\sigma_g^2/T + \lambda_p}\phi_p(x)^2,$$

where $k(x,\tilde{x}) = \sum_{p\ge 0} \lambda_p \phi_p(x)\phi_p(\tilde{x})$ is the Mercer decomposition of $k(x,\tilde{x})$ (see [Le Gratiet Loic et al. (2012)]).

Gaussian process regression

- We consider a fixed number of experiments $T = nr$
- Best Linear Unbiased Predictor (BLUP):

$$\hat{z}_{T,n}(x) = \mathbf{k}'(x)\left(\mathbf{K} + \frac{n\sigma_g^2}{T}\mathbf{I}\right)^{-1}\mathbf{z}^n,$$

where $\mathbf{k}'(x) = [k(x, D_i)]_{1 \leq i \leq n}$ and $\mathbf{K} = [k(D_i, D_j)]_{1 \leq i,j \leq n}$.
Note: $\hat{z}_{T,n}(x)$ is defined on $(\Omega_D \times \Omega_Z, \sigma(\mathcal{F}_D \times \mathcal{F}_Z), \mathbb{P}_D \times \mathbb{P}_Z)$

- Mean Squared Error (MSE) of the BLUP (the mean is w.r.t $\mathbb{E}_Z$):

$$\sigma^2_{T,n}(x) = k(x,x) - \mathbf{k}'(x)\left(\mathbf{K} + \frac{n\sigma_g^2}{T}\mathbf{I}\right)^{-1}\mathbf{k}(x).$$

- Let us consider the case $n \gg 1$:

$$\sigma^2_{T,n}(x) \xrightarrow[n\to\infty]{\mathbb{P}_D} \sum_{p\geq 0} \frac{\sigma_g^2 \lambda_p / T}{\sigma_g^2/T + \lambda_p}\phi_p(x)^2,$$

where $k(x,\tilde{x}) = \sum_{p\geq 0} \lambda_p \phi_p(x)\phi_p(\tilde{x})$ is the Mercer decomposition of $k(x,\tilde{x})$ (see [Le Gratiet Loic et al. (2012)]).

Gaussian process regression

- We consider a fixed number of experiments $T = nr$
- Best Linear Unbiased Predictor (BLUP):

$$\hat{z}_{T,n}(x) = \mathbf{k}'(x)\left(\mathbf{K} + \frac{n\sigma_g^2}{T}\mathbf{I}\right)^{-1}\mathbf{z}^n,$$

where $\mathbf{k}'(x) = [k(x, D_i)]_{1\leq i\leq n}$ and $\mathbf{K} = [k(D_i, D_j)]_{1\leq i,j\leq n}$.
Note: $\hat{z}_{T,n}(x)$ is defined on $\left(\Omega_D \times \Omega_Z, \sigma\left(\mathcal{F}_D \times \mathcal{F}_Z\right), \mathbb{P}_D \times \mathbb{P}_Z\right)$

- Mean Squared Error (MSE) of the BLUP (the mean is w.r.t $\mathbb{E}_Z$):

$$\sigma^2_{T,n}(x) = k(x,x) - \mathbf{k}'(x)\left(\mathbf{K} + \frac{n\sigma_g^2}{T}\mathbf{I}\right)^{-1}\mathbf{k}(x).$$

- Let us consider the case $n \gg 1$:

$$\sigma^2_{T,n}(x) \xrightarrow[n\to\infty]{\mathbb{P}_D} \sum_{p\geq 0} \frac{\sigma_g^2\lambda_p/T}{\sigma_g^2/T + \lambda_p}\phi_p(x)^2,$$

where $k(x,\tilde{x}) = \sum_{p\geq 0}\lambda_p\phi_p(x)\phi_p(\tilde{x})$ is the Mercer decomposition of $k(x,\tilde{x})$ (see [Le Gratiet Loic et al. (2012)]).

Sketch of proof *[Le Gratiet Loic et al. (2012)]*

- Lower bound. Let us consider $\bar{\sigma}^2_{T,n}$ the MSE associated to $\bar{Z}(x) = \sum_{p \leq \bar{p}} z_p \phi_p(x)$, $z_p \sim \mathcal{N}(0, \lambda_p)$, we have almost surely:

$$\liminf_n \sigma^2_{T,n} \geq \lim_n \bar{\sigma}^2_{T,n} = \sum_{p \geq 0} \frac{\sigma_g^2 \lambda_p / T}{\sigma_g^2 / T + \lambda_p} \phi_p(x)^2.$$

- Upper bound. Let us consider $\tilde{\sigma}^2_{T,n}$ the MSE associated to the LUP

$$\tilde{z}_{T,n}(x) = \mathbf{k}'(x)\mathbf{A}\mathbf{z}^n,$$

$\mathbf{A} = \mathbf{L}^{-1} + \sum_{k \leq q} (-1)^k (\mathbf{L}^{-1}\mathbf{M})^k \mathbf{L}^{-1}$, p^* such that $\lambda_{p^*} < \sigma_g^2/T$, $\mathbf{M} = \sum_{p > p^*} \lambda_p [\phi_p(D_i)\phi_p(D_j)]_{1 \leq i,j \leq n}$ and $\mathbf{L} = n\sigma_g^2/T + \sum_{p \leq p^*} \lambda_p [\phi_p(D_i)\phi_p(D_j)]_{1 \leq i,j \leq n}$.
We have in probability:

$$\limsup_n \sigma^2_{T,n} \leq \lim_n \tilde{\sigma}^2_{T,n} = \sum_{p \geq 0} \frac{\sigma_g^2 \lambda_p / T}{\sigma_g^2 / T + \lambda_p} \phi_p(x)^2.$$

Sketch of proof *[Le Gratiet Loic et al. (2012)]*

- Lower bound. Let us consider $\bar{\sigma}^2_{T,n}$ the MSE associated to $\bar{Z}(x) = \sum_{p \leq \bar{p}} z_p \phi_p(x)$, $z_p \sim \mathcal{N}(0, \lambda_p)$, we have almost surely:

$$\liminf_n \sigma^2_{T,n} \geq \lim_n \bar{\sigma}^2_{T,n} = \sum_{p \geq 0} \frac{\sigma_g^2 \lambda_p / T}{\sigma_g^2 / T + \lambda_p} \phi_p(x)^2.$$

- Upper bound. Let us consider $\tilde{\sigma}^2_{T,n}$ the MSE associated to the LUP

$$\tilde{z}_{T,n}(x) = \mathbf{k}'(x)\mathbf{A}\mathbf{z}^n,$$

$\mathbf{A} = \mathbf{L}^{-1} + \sum_{k \leq q} (-1)^k (\mathbf{L}^{-1}\mathbf{M})^k \mathbf{L}^{-1}$, p^* such that $\lambda_{p^*} < \sigma_g^2 / T$,
$\mathbf{M} = \sum_{p > p^*} \lambda_p [\phi_p(D_i)\phi_p(D_j)]_{1 \leq i,j \leq n}$ and
$\mathbf{L} = n\sigma_g^2 / T + \sum_{p \leq p^*} \lambda_p [\phi_p(D_i)\phi_p(D_j)]_{1 \leq i,j \leq n}$.
We have in probability:

$$\limsup_n \sigma^2_{T,n} \leq \lim_n \tilde{\sigma}^2_{T,n} = \sum_{p \geq 0} \frac{\sigma_g^2 \lambda_p / T}{\sigma_g^2 / T + \lambda_p} \phi_p(x)^2.$$

Idealized Gaussian process regression

- Let us consider the predictor $\hat{z}_{T,n}(x)$ defined on $(\Omega_D \times \Omega_Z, \sigma(\mathcal{F}_D \times \mathcal{F}_Z), \mathbb{P}_D \times \mathbb{P}_Z)$ and built from noisy observations at points $(D_i)_{i=1,\ldots,n}$ defined on $(\Omega_D, \mathcal{F}_D, \mathbb{P}_D)$ with measure μ.
- The convergence of the MSE implies the following one in distribution:

$$\hat{z}_{T,n}(x) \overset{n}{\rightsquigarrow} \hat{z}_T(x) = \sum_{p \geq 0} \frac{\lambda_p}{\lambda_p + \sigma_g^2/T} z_p \phi_p(x), \quad z_p \sim \mathcal{N}(0, \lambda_p).$$

- $(z_p)_{p\geq 0}$ comes from the Karhunen-Loeve decomposition of $Z(x)$:

$$Z(x) = \sum_{p \geq 0} \phi_p(x) z_p$$

- Therefore, $\hat{z}_T(x)$ is defined on the probability space $(\Omega_Z, \mathcal{F}_Z, \mathbb{P}_Z)$.

Idealized Gaussian process regression

- Let us consider the predictor $\hat{z}_{T,n}(x)$ defined on $(\Omega_D \times \Omega_Z, \sigma(\mathcal{F}_D \times \mathcal{F}_Z), \mathbb{P}_D \times \mathbb{P}_Z)$ and built from noisy observations at points $(D_i)_{i=1,\dots,n}$ defined on $(\Omega_D, \mathcal{F}_D, \mathbb{P}_D)$ with measure μ.
- The convergence of the MSE implies the following one in distribution:

$$\hat{z}_{T,n}(x) \overset{n}{\rightsquigarrow} \hat{z}_T(x) = \sum_{p\geq 0} \frac{\lambda_p}{\lambda_p + \sigma_g^2/T} z_p \phi_p(x), \quad z_p \sim \mathcal{N}(0, \lambda_p).$$

- $(z_p)_{p\geq 0}$ comes from the Karhunen-Loeve decomposition of $Z(x)$:

$$Z(x) = \sum_{p\geq 0} \phi_p(x) z_p$$

- Therefore, $\hat{z}_T(x)$ is defined on the probability space $(\Omega_Z, \mathcal{F}_Z, \mathbb{P}_Z)$.

Idealized Gaussian process regression

- Let us consider the predictor $\hat{z}_{T,n}(x)$ defined on $\left(\Omega_D \times \Omega_Z, \sigma\left(\mathcal{F}_D \times \mathcal{F}_Z\right), \mathbb{P}_D \times \mathbb{P}_Z\right)$ and built from noisy observations at points $(D_i)_{i=1,\dots,n}$ defined on $(\Omega_D, \mathcal{F}_D, \mathbb{P}_D)$ with measure μ.
- The convergence of the MSE implies the following one in distribution:

$$\hat{z}_{T,n}(x) \overset{n}{\rightsquigarrow} \hat{z}_T(x) = \sum_{p\geq 0} \frac{\lambda_p}{\lambda_p + \sigma_g^2/T} z_p \phi_p(x), \quad z_p \sim \mathcal{N}(0, \lambda_p).$$

- $(z_p)_{p\geq 0}$ comes from the Karhunen-Loeve decomposition of $Z(x)$:

$$Z(x) = \sum_{p\geq 0} \phi_p(x) z_p$$

- Therefore, $\hat{z}_T(x)$ is defined on the probability space $(\Omega_Z, \mathcal{F}_Z, \mathbb{P}_Z)$.

Idealized Gaussian process regression

- Let us consider the Integrated Mean Squared Error (IMSE) of $\hat{z}_T(x)$:

$$\mathrm{IMSE} = \int \sigma_T^2 \, d\mu(x) = \sum_{p \geq 0} \frac{\sigma_g^2 \lambda_p / T}{\sigma_g^2 / T + \lambda_p}$$

- The IMSE of $\hat{z}_T(x)$ is bounded by:

$$B_T^2/2 \leq \mathrm{IMSE}_T \leq B_T^2,$$

where $B_T^2 = \sum_{p \,\mathrm{s.t.}\, \lambda_p \leq \sigma_g^2/T} \lambda_p + \frac{\sigma_g^2}{T} \#\{p \,\mathrm{s.t.}\, \lambda_p > \sigma_g^2/T\}$.

- Example: for $k(x, \tilde{x}) = \exp\left(-||x - \tilde{x}||^2/\theta^2\right)$, we have:

$$B_T^2 \approx C \frac{\sigma_g^2}{T} \log\left(\frac{T}{\sigma_g^2}\right)^d.$$

- Remark: B_T^2 controls the accuracy of the meta-model $\hat{z}_T(x)$.

Idealized Gaussian process regression

- Let us consider the Integrated Mean Squared Error (IMSE) of $\hat{z}_T(x)$:

$$\text{IMSE} = \int \sigma_T^2 \, d\mu(x) = \sum_{p \geq 0} \frac{\sigma_g^2 \lambda_p / T}{\sigma_g^2/T + \lambda_p}$$

- The IMSE of $\hat{z}_T(x)$ is bounded by:

$$B_T^2/2 \leq \text{IMSE}_T \leq B_T^2,$$

where $B_T^2 = \sum_{p \,\mathrm{s.t.}\, \lambda_p \leq \sigma_g^2/T} \lambda_p + \frac{\sigma_g^2}{T} \#\{p \,\mathrm{s.t.}\, \lambda_p > \sigma_g^2/T\}$.

- Example: for $k(x, \tilde{x}) = \exp\left(-||x - \tilde{x}||^2/\theta^2\right)$, we have:

$$B_T^2 \approx C \frac{\sigma_g^2}{T} \log\left(\frac{T}{\sigma_g^2}\right)^d.$$

- Remark: B_T^2 controls the accuracy of the meta-model $\hat{z}_T(x)$.

Idealized Gaussian process regression

- Let us consider the Integrated Mean Squared Error (IMSE) of $\hat{z}_T(x)$:

$$\mathrm{IMSE} = \int \sigma_T^2 \, d\mu(x) = \sum_{p \geq 0} \frac{\sigma_g^2 \lambda_p / T}{\sigma_g^2 / T + \lambda_p}$$

- The IMSE of $\hat{z}_T(x)$ is bounded by:

$$B_T^2 / 2 \leq \mathrm{IMSE}_T \leq B_T^2,$$

where $B_T^2 = \sum_{p \,\mathrm{s.t.}\, \lambda_p \leq \sigma_g^2/T} \lambda_p + \frac{\sigma_g^2}{T} \#\{p \,\mathrm{s.t.}\, \lambda_p > \sigma_g^2/T\}$.

- Example: for $k(x, \tilde{x}) = \exp\left(-||x - \tilde{x}||^2 / \theta^2\right)$, we have:

$$B_T^2 \approx C \frac{\sigma_g^2}{T} \log\left(\frac{T}{\sigma_g^2}\right)^d.$$

- Remark: B_T^2 controls the accuracy of the meta-model $\hat{z}_T(x)$.

Outline

Sobol index estimator

- Sensitivity index: $S^1 = \mathrm{var}_X\left(\mathbb{E}_X\left[u^{\mathrm{code}}_{r=\infty}(X)|X^1\right]\right)/\mathrm{var}_X\left(u^{\mathrm{code}}_{r=\infty}(X)\right)$.
- Monte-Carlo integration:

$$S^1_m = \frac{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_i) - m^{-2}\sum_{i,j=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_j)}{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)^2 - m^{-2}(\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i))^2}.$$

- Monte-Carlo integration + Meta-model approximation:

$$S^1_{T,m} = \frac{m^{-1}\sum_{i=1}^m \hat{z}_T(X_i)\hat{z}_T(\tilde{X}_i) - m^{-2}\sum_{i,j=1}^m \hat{z}_T(X_i)\hat{z}_T(\tilde{X}_j)}{m^{-1}\sum_{i=1}^m \hat{z}_T^2(X_i) - m^{-2}(\sum_{i=1}^m \hat{z}_T(X_i))^2}$$

- $S^1_{T,m}$ is defined on $(\Omega_Z \times \Omega_X, \sigma(\mathcal{F}_Z \times \mathcal{F}_X), \mathbb{P}_Z \times \mathbb{P}_X)$ and:
  - m controls the Monte-Carlo convergence.
  - T controls the Meta-model convergence through B_T^2.

Sobol index estimator

- Sensitivity index: $S^1 = \mathrm{var}_X\left(\mathbb{E}_X\left[u^{\mathrm{code}}_{r=\infty}(X)|X^1\right]\right)/\mathrm{var}_X\left(u^{\mathrm{code}}_{r=\infty}(X)\right)$.
- Monte-Carlo integration:

$$S^1_m = \frac{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_i) - m^{-2}\sum_{i,j=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_j)}{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)^2 - m^{-2}(\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i))^2}.$$

- Monte-Carlo integration + Meta-model approximation:

$$S^1_{T,m} = \frac{m^{-1}\sum_{i=1}^m \hat{z}_T(X_i)\hat{z}_T(\tilde{X}_i) - m^{-2}\sum_{i,j=1}^m \hat{z}_T(X_i)\hat{z}_T(\tilde{X}_j)}{m^{-1}\sum_{i=1}^m \hat{z}_T^2(X_i) - m^{-2}(\sum_{i=1}^m \hat{z}_T(X_i))^2}$$

- $S^1_{T,m}$ is defined on $(\Omega_Z \times \Omega_X, \sigma(\mathcal{F}_Z \times \mathcal{F}_X), \mathbb{P}_Z \times \mathbb{P}_X)$ and:
  - m controls the Monte-Carlo convergence.
  - T controls the Meta-model convergence through B_T^2.

Sobol index estimator

- Sensitivity index: $S^1 = \mathrm{var}_X\left(\mathbb{E}_X\left[u^{\mathrm{code}}_{r=\infty}(X)|X^1\right]\right)/\mathrm{var}_X\left(u^{\mathrm{code}}_{r=\infty}(X)\right)$.
- Monte-Carlo integration:

$$S^1_m = \frac{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_i) - m^{-2}\sum_{i,j=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_j)}{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)^2 - m^{-2}(\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i))^2}.$$

- Monte-Carlo integration + Meta-model approximation:

$$S^1_{T,m} = \frac{m^{-1}\sum_{i=1}^m \hat{z}_T(X_i)\hat{z}_T(\tilde{X}_i) - m^{-2}\sum_{i,j=1}^m \hat{z}_T(X_i)\hat{z}_T(\tilde{X}_j)}{m^{-1}\sum_{i=1}^m \hat{z}^2_T(X_i) - m^{-2}(\sum_{i=1}^m \hat{z}_T(X_i))^2}$$

- $S^1_{T,m}$ is defined on $(\Omega_Z \times \Omega_X, \sigma(\mathcal{F}_Z \times \mathcal{F}_X), \mathbb{P}_Z \times \mathbb{P}_X)$ and:
  - m controls the Monte-Carlo convergence.
  - T controls the Meta-model convergence through B^2_T.

Sobol index estimator

- Sensitivity index: $S^1 = \mathrm{var}_X\left(\mathbb{E}_X\left[u^{\mathrm{code}}_{r=\infty}(X)|X^1\right]\right)/\mathrm{var}_X\left(u^{\mathrm{code}}_{r=\infty}(X)\right)$.
- Monte-Carlo integration:

$$S^1_m = \frac{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_i) - m^{-2}\sum_{i,j=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)u^{\mathrm{code}}_{r=\infty}(\tilde{X}_j)}{m^{-1}\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i)^2 - m^{-2}(\sum_{i=1}^m u^{\mathrm{code}}_{r=\infty}(X_i))^2}.$$

- Monte-Carlo integration + Meta-model approximation:

$$S^1_{T,m} = \frac{m^{-1}\sum_{i=1}^m \hat{z}_T(X_i)\hat{z}_T(\tilde{X}_i) - m^{-2}\sum_{i,j=1}^m \hat{z}_T(X_i)\hat{z}_T(\tilde{X}_j)}{m^{-1}\sum_{i=1}^m \hat{z}_T^2(X_i) - m^{-2}(\sum_{i=1}^m \hat{z}_T(X_i))^2}$$

- $S^1_{T,m}$ is defined on $(\Omega_Z \times \Omega_X, \sigma(\mathcal{F}_Z \times \mathcal{F}_X), \mathbb{P}_Z \times \mathbb{P}_X)$ and:
  - m controls the Monte-Carlo convergence.
  - T controls the Meta-model convergence through B_T^2.

Asymptotic normality of $S^1_{T_m,m}$ [Le Gratiet (2013)]

- Let us suppose that T is an increasing function of $m \in \mathbb{N}^*$.

  If $mB^2_{T_m} \overset{m\to\infty}{\longrightarrow} 0$ — i.e. the Monte-Carlo error dominates — we have in $(\Omega_Z \times \Omega_X)$:

$$\sqrt{m}\left(S^1_{T_m,m} - S^1\right) \overset{m\to\infty}{\rightsquigarrow} \mathcal{N}\left(0, \sigma_S^2\right). \qquad (1)$$

  If $mB^2_{T_m} \overset{m\to\infty}{\longrightarrow} \infty$ — i.e. the meta-model error dominates — then $\exists C > 0$ such that $\forall \delta > 0$:

$$\mathbb{P}_Z\left(\left|\mathbb{P}_X\left(B^{-1}_{T_m}\left(S^1_{T_m,m} - S^1\right) \geq C\right) - 1\right| > \delta\right) \overset{m\to\infty}{\longrightarrow} 0. \qquad (2)$$

Sketch of proof

- Skorokhod's representation theorem: to deal with the convergences in the product probability space $(\Omega_Z \times \Omega_X, \sigma(\mathcal{F}_Z \times \mathcal{F}_X), \mathbb{P}_Z \times \mathbb{P}_X)$ using a common probability space $(\tilde{\Omega}_Z \times \Omega_X, \sigma\left(\tilde{\mathcal{F}}_Z \times \mathcal{F}_X\right), \tilde{\mathbb{P}}_Z \times \mathbb{P}_X)$.
- Lindeberg-Feller Theorem + Delta method + Slutsky's theorem.
- In the probability space $(\Omega_Z \times \Omega_X, \sigma(\mathcal{F}_Z \times \mathcal{F}_X), \mathbb{P}_Z \times \mathbb{P}_X)$:

$$\forall I \in \mathbb{R},\ \mathbb{P}_Z\left(\left|\mathbb{P}_X\left(\sqrt{m}\left(S^1_{T_m,m} - S^1\right) \in I\right) - \int_I f(x)dx\right| > \delta\right) \xrightarrow{m} 0.$$

where $f(x)$ is a Gaussian pdf with zero-mean and an explicit variance.

Sketch of proof

- Skorokhod's representation theorem: to deal with the convergences in the product probability space $(\Omega_Z \times \Omega_X, \sigma(\mathcal{F}_Z \times \mathcal{F}_X), \mathbb{P}_Z \times \mathbb{P}_X)$ using a common probability space $(\tilde{\Omega}_Z \times \Omega_X, \sigma\left(\tilde{\mathcal{F}}_Z \times \mathcal{F}_X\right), \tilde{\mathbb{P}}_Z \times \mathbb{P}_X)$.
- Lindeberg-Feller Theorem + Delta method + Slutsky's theorem.
- In the probability space $(\Omega_Z \times \Omega_X, \sigma(\mathcal{F}_Z \times \mathcal{F}_X), \mathbb{P}_Z \times \mathbb{P}_X)$:

$$\forall I \in \mathbb{R},\ \mathbb{P}_Z\left(\left|\mathbb{P}_X\left(\sqrt{m}\left(S^1_{T_m,m} - S^1\right) \in I\right) - \int_I f(x)dx\right| > \delta\right) \xrightarrow{m} 0.$$

where $f(x)$ is a Gaussian pdf with zero-mean and an explicit variance.

Sketch of proof

- Skorokhod's representation theorem: to deal with the convergences in the product probability space $(\Omega_Z \times \Omega_X, \sigma(\mathcal{F}_Z \times \mathcal{F}_X), \mathbb{P}_Z \times \mathbb{P}_X)$ using a common probability space $(\tilde{\Omega}_Z \times \Omega_X, \sigma\left(\tilde{\mathcal{F}}_Z \times \mathcal{F}_X\right), \tilde{\mathbb{P}}_Z \times \mathbb{P}_X)$.
- Lindeberg-Feller Theorem + Delta method + Slutsky's theorem.
- In the probability space $(\Omega_Z \times \Omega_X, \sigma(\mathcal{F}_Z \times \mathcal{F}_X), \mathbb{P}_Z \times \mathbb{P}_X)$:

$$\forall I \in \mathbb{R}, \; \mathbb{P}_Z\left(\left|\mathbb{P}_X\left(\sqrt{m}\left(S^1_{T_m,m} - S^1\right) \in I\right) - \int_I f(x)dx\right| > \delta\right) \xrightarrow{m} 0.$$

where $f(x)$ is a Gaussian pdf with zero-mean and an explicit variance.

Numerical illustration

- Let us come back to the heat equation:

$$\frac{\partial u}{\partial t}(x,t) - \frac{1}{2}\Delta u(x,t) = 0,$$

with $u(x,0) = \exp(-\sum_{i=1}^d x_i^2/\theta_{g,i}^2)$.

- We consider that $X \sim \mathcal{N}\left(0, \sigma_\mu^2 \mathbf{I}_d\right)$.
- The exact Sobol indices are:

$$S^j = B_j - 1/\left(\prod_{i=1}^d B_i\right) - 1, \quad j = 1, \ldots, d,$$

where $B_j =$

$$\sigma_\mu \left(\frac{2}{t} - \frac{2}{t^2}\left(\frac{1}{t} + \frac{1}{\theta_{g,i}^2}\right)^{-1} + \frac{1}{\sigma_\mu^2} \right)^{-\frac{1}{2}} \left(\frac{1}{t} + \frac{1}{\sigma_\mu^2} - \frac{1}{t^2}\left(\frac{1}{t} + \frac{1}{\theta_{g,i}^2}\right)^{-1} \right).$$

Numerical illustration

- Let us come back to the heat equation:

$$\frac{\partial u}{\partial t}(x,t) - \frac{1}{2}\Delta u(x,t) = 0,$$

with $u(x,0) = \exp(-\sum_{i=1}^d x_i^2/\theta_{g,i}^2)$.

- We consider that $X \sim \mathcal{N}\left(0, \sigma_\mu^2 \mathbf{I}_d\right)$.
- The exact Sobol indices are:

$$S^j = B_j - 1/\left(\prod_{i=1}^d B_i\right) - 1, \quad j = 1, \dots, d,$$

where $B_j =$

$$\sigma_\mu \left(\frac{2}{t} - \frac{2}{t^2}\left(\frac{1}{t} + \frac{1}{\theta_{g,i}^2}\right)^{-1} + \frac{1}{\sigma_\mu^2}\right)^{-\frac{1}{2}} \left(\frac{1}{t} + \frac{1}{\sigma_\mu^2} - \frac{1}{t^2}\left(\frac{1}{t} + \frac{1}{\theta_{g,i}^2}\right)^{-1}\right).$$

- To build the meta-model $\hat{z}_{T,n}(x)$, let us consider the Gaussian kernel:

$$k(x,\tilde{x}) = \sigma^2 \exp\left(-\frac{1}{2}\sum_{i=1}^{d}\frac{(x^i-\tilde{x}^i)^2}{\theta_i^2}\right).$$

- We consider $n = 3000$ and $T_0 = n$ (i.e. $r = 1$). Parameter estimates: $\hat{\sigma}^2 = 1.46$, $\hat{\theta} = (1.01, 1.02, 1.03, 1.00, 1.07)$.
- The critical ratio $mB_T^2 = 1$ leads to the following budget:

$$T = \sigma_g^2 \frac{m}{C} \log\left(\frac{m}{C}\right),$$

with $C = \sigma_g^2 \log\left(T_0/\sigma_g^2\right)/T_0$ and $\hat{\sigma_g}^2 = 6.74.10^{-2}$.
- Let us consider

$$T = \sigma_g^2 \frac{m^\alpha}{C} \log\left(\frac{m}{C}\right),$$

with $\alpha \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$.

- To build the meta-model $\hat{z}_{T,n}(x)$, let us consider the Gaussian kernel:

$$k(x,\tilde{x}) = \sigma^2 \exp\left(-\frac{1}{2}\sum_{i=1}^{d}\frac{(x^i-\tilde{x}^i)^2}{\theta_i^2}\right).$$

- We consider $n = 3000$ and $T_0 = n$ (i.e. $r = 1$). Parameter estimates: $\hat{\sigma}^2 = 1.46$, $\hat{\theta} = (1.01, 1.02, 1.03, 1.00, 1.07)$.
- The critical ratio $mB_T^2 = 1$ leads to the following budget:

$$T = \sigma_g^2 \frac{m}{C}\log\left(\frac{m}{C}\right),$$

with $C = \sigma_g^2 \log\left(T_0/\sigma_g^2\right)/T_0$ and $\hat{\sigma_g}^2 = 6.74.10^{-2}$.

- Let us consider

$$T = \sigma_g^2 \frac{m^\alpha}{C}\log\left(\frac{m}{C}\right),$$

with $\alpha \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$.

- To build the meta-model $\hat{z}_{T,n}(x)$, let us consider the Gaussian kernel:

$$k(x,\tilde{x}) = \sigma^2 \exp\left(-\frac{1}{2}\sum_{i=1}^{d}\frac{(x^i-\tilde{x}^i)^2}{\theta_i^2}\right).$$

- We consider $n = 3000$ and $T_0 = n$ (i.e. $r = 1$). Parameter estimates: $\hat{\sigma}^2 = 1.46$, $\hat{\theta} = (1.01, 1.02, 1.03, 1.00, 1.07)$.
- The critical ratio $mB_T^2 = 1$ leads to the following budget:

$$T = \sigma_g^2 \frac{m}{C} \log\left(\frac{m}{C}\right),$$

with $C = \sigma_g^2 \log\left(T_0/\sigma_g^2\right)/T_0$ and $\hat{\sigma_g}^2 = 6.74.10^{-2}$.
- Let us consider

$$T = \sigma_g^2 \frac{m^\alpha}{C} \log\left(\frac{m}{C}\right),$$

with $\alpha \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$.

m	α	S^1	S^2	S^3	S^4	S^5
1,000	0.8	0.93	0.90	0.93	0.92	0.93
1,000	0.9	0.91	0.91	0.93	0.90	0.90
1,000	1.0	0.92	0.92	0.96	0.92	0.94
1,000	1.1	0.96	0.92	0.94	0.92	0.92
1,000	1.2	0.92	0.94	0.93	0.93	0.91
3,000	0.8	0.92	0.92	0.89	0.93	0.95
3,000	0.9	0.95	0.94	0.91	0.92	0.94
3,000	1.0	0.95	0.96	0.95	0.94	0.95
3,000	1.1	0.97	0.97	0.97	0.96	0.96
3,000	1.2	0.97	0.97	0.98	0.96	0.98
5,000	0.8	0.94	0.96	0.95	0.93	0.92
5,000	1.0	0.97	0.96	0.95	0.95	0.96
5,000	1.2	0.98	0.98	0.97	0.97	0.97
10,000	0.8	0.94	0.92	0.93	0.93	0.92
10,000	1.0	0.96	0.95	0.94	0.95	0.95
10,000	1.2	0.96	0.97	0.98	0.95	0.96

Sensitivity Analysis

- Sobol, I.M. (1993). *Sensitivity estimates for non linear mathematical models*. Mathematical Modelling and Computational Experiments, 1:407-414.
- Sobol, I.M., Tarantola, D., Gatellu, D., Kucherenko, S. and Mauntz, W. (2007). *Estimating the approximation error when fixing unessential factors in global sensitivity analysis*. Reliability Engineering & System Safety, 92:957-960.
- Janon, A., Klein, T., Lagnoux, A., Nodet, M. and Prieur, C. (2012). *Asymptotic normality and efficiency of two Sobol index estimators*.
- Oakley, J.E. and O'Hagan, A. (2004). *Probabilistic analysis of complex models a Bayesian approach*. Journal of the Royal Statistical Society series B, 66: 751-769.
- Marrel, A., Iooss, B. Laurent, B. and Roustant, O. (2009). *Calculations of sobol indices for the Gaussian process metamodel*. REliability Engineering and System Safety, 94:742-751.

Cited contributions

Le Gratiet, L. and Garnier, J. (2012). *Regularity dependence of the rate of convergence of the learning curve for Gaussian process regression.* arXiv:1210.2879v1, submitted to Journal of Machine Learning.

Le Gratiet, L. (2013). *Asymptotic normality of a Sobol index estimator in Guassian process regression framework*. arXiv:1305.7406, submitted to ESAIM probability and statistics.

Multi-fidelity sensitivity analysis

- Question: How to adapt the presented method in a non-asymptotic case for multi-fidelity computer codes ?
- Response: See Claire CANNAMELA, Thursday on Session 8 “Multi-fidelity sensitivity analysis”.

Le Gratiet, L., Cannamela, C. and Iooss, B. (2013). *A Bayesian approach for global sensitivity analysis of (multi-fidelity) computer codes.* submitted to SIAM/ASA.

Multi-fidelity computer codes

- What is a multi-fidelity computer code ?

- Kennedy, M. C. & O'Hagan, A. (2000). Predicting the output from a complex computer code when fast approximations are available. *Biometrika* **87**, 1-13.
- Qian, Z. & Jeff Wu, C. F. (2008). Bayesian Hierarchical Modeling for Integrating Low-accuracy and High-accuracy Experiments. *Technometrics* **50**, 192-204.
- Le Gratiet, L. (2013). *Bayesian analysis of hierarchical multi-fidelity codes*. Accepted in SIAM/ASA UQ.
- Le Gratiet, L. & Garnier, J. (2012). *Recursive co-kriging model for Design of Computer experiments with multiple levels of fidelity with an application to hydrodynamic*. arxiv:1210.0686 , submitted to International Journal of Uncertainty Quantification.
- Le Gratiet, L. & Cannamela C. (2012). *Kriging-based sequential design strategies using fast cross-validation techniques with extensions to multi-fidelity computer codes*. arxiv:1210.6187, submitted to TECHNOMETRICS.

- We implement a R CRAN package: MuFiCokriging.

Thank you for your attention

Questions ?