Sensitivity analysis for optimization under constraints and with uncertainties

Kernel-based sensitivity analysis on (excursion) sets

Noé Fellmann
Céline Helbert & Christophette Blanchet (ECL)
Adrien Spagnol & Delphine Sinoquet (IFPEN)

École Centrale de Lyon & IFP Énergie nouvelles

Rencontres MEXICO, 24 et 25 novembre 2022

System & black-box model

Input parameters :
-component size
-material
-swell height
-...

inputs → **Complex system : Floating wind turbine** → output

-Cost
-energy production
-environmental impact
-...

System & black-box model

$x \in \mathcal{X}$
$u \in \mathcal{U}$ $\longrightarrow$ $\boxed{\text{Black-box functions } f \text{ and } g}$ $\longrightarrow$ $f(x,u) \in \mathbb{R}$
$g(x,u) \in \mathbb{R}^m$

- The x are the deterministic inputs
- The u are uncertain inputs : $u = U(\omega)$ with U a random vector of density ρ_U
- f is the objective function to minimize
- g is the constraint function defining the constraint to respect : $g \leq 0$

Optimization problem

Robust optimization problem

$$x^* = \arg\min_x \mathbb{E}[f(x, U)] \quad \text{s.t. } \mathbb{P}[g(x, U) \leq 0] \geq P_{target} \tag{1}$$

Deterministic strategy

$$x^* = \arg\min_x F(x) \quad \text{s.t. } G(x) \leq 0 \tag{2}$$

- Goal Oriented Sensitivity Analysis to reduce the input dimension : Spagnol 2020
- Huge evaluation cost of F and G
- What about the U ?

How to quantify the impact of the uncertain inputs U on the optimization ?

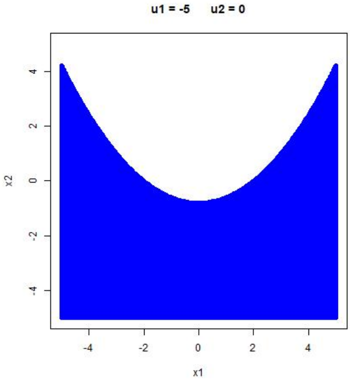
Toy function

Toy function from [El-Amri et al. 2021]

$$\forall (x, u) \in [-5, 5]^4 \;\; g(\boldsymbol{x}, \boldsymbol{u}) = -x_1^2 + 5x_2 - u_1 + u_2^2 - 1$$

$$\Gamma_u = \{x \in [-5, 5]^2, g(x, u) \leq 0\}$$

u_2 fixé



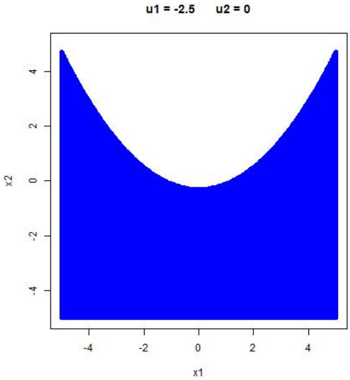
Toy function

Toy function from [El-Amri et al. 2021]

$$\forall (x, u) \in [-5, 5]^4 \;\; g(\boldsymbol{x}, \boldsymbol{u}) = -x_1^2 + 5x_2 - u_1 + u_2^2 - 1$$

$$\Gamma_u = \{x \in [-5, 5]^2, g(x, u) \leq 0\}$$

u_2 fixé



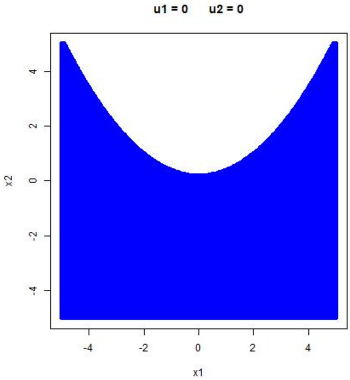
Toy function

Toy function from [El-Amri et al. 2021]

$$\forall (x,u) \in [-5,5]^4 \;\; g(\boldsymbol{x},\boldsymbol{u}) = -x_1^2 + 5x_2 - u_1 + u_2^2 - 1$$

$$\Gamma_u = \{x \in [-5,5]^2, g(x,u) \leq 0\}$$

u_2 fixé



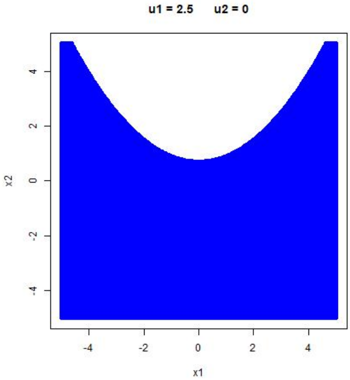
Toy function

Toy function from [El-Amri et al. 2021]

$$\forall (x,u) \in [-5,5]^4 \;\; g(\boldsymbol{x},\boldsymbol{u}) = -x_1^2 + 5x_2 - u_1 + u_2^2 - 1$$

$$\Gamma_u = \{x \in [-5,5]^2, g(x,u) \leq 0\}$$

u_2 fixé



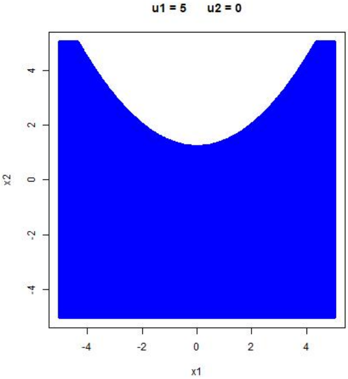
Toy function

Toy function from [El-Amri et al. 2021]

$$\forall (x,u) \in [-5,5]^4 \;\; g(\boldsymbol{x},\boldsymbol{u}) = -x_1^2 + 5x_2 - u_1 + u_2^2 - 1$$

$$\Gamma_u = \{x \in [-5,5]^2, g(x,u) \leq 0\}$$

u_2 fixé



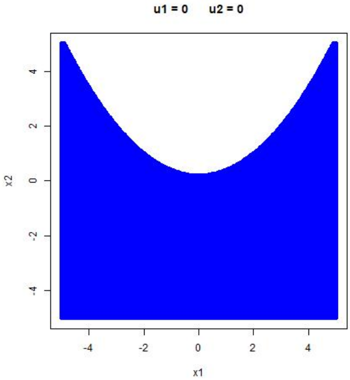
Toy function

Toy function from [El-Amri et al. 2021]

$$\forall (x,u) \in [-5,5]^4 \;\; g(\boldsymbol{x},\boldsymbol{u}) = -x_1^2 + 5x_2 - u_1 + u_2^2 - 1$$

$$\Gamma_u = \{x \in [-5,5]^2, g(x,u) \leq 0\}$$

u_1 fixé



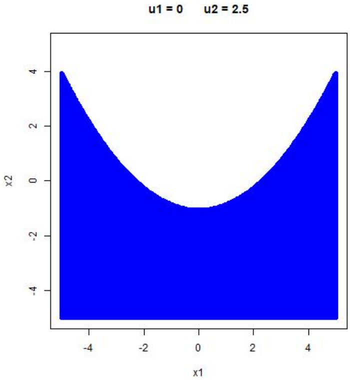
Toy function

Toy function from [El-Amri et al. 2021]

$$\forall (x, u) \in [-5, 5]^4 \;\; g(\boldsymbol{x}, \boldsymbol{u}) = -x_1^2 + 5x_2 - u_1 + u_2^2 - 1$$

$$\Gamma_u = \{x \in [-5, 5]^2, g(x, u) \leq 0\}$$

u_1 fixé



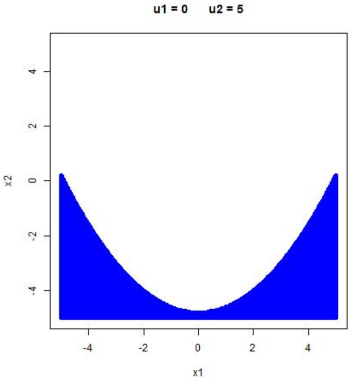
Toy function

Toy function from [El-Amri et al. 2021]

$$\forall (x, u) \in [-5, 5]^4 \;\; g(\boldsymbol{x}, \boldsymbol{u}) = -x_1^2 + 5x_2 - u_1 + u_2^2 - 1$$

$$\Gamma_u = \{x \in [-5, 5]^2, g(x, u) \leq 0\}$$

u_1 fixé



Subproblem : Excursion sets

Excursion sets

New output :

$$\Gamma_U = \{x \in \mathcal{X}, g(x, U) \leq 0\}, \tag{3}$$

which is called a random excursion set.

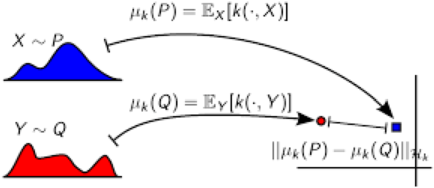
Influence of the uncertain inputs U on Γ_U ? $\Rightarrow$ SA on excursion sets.
Yet most SA techniques have scalar or vectorial outputs.
How then can we do sensitivity analysis on (excursion) sets ?

- SA on sets using random set theory notably Vorob'ev expectation and deviation
- SA on sets using universal indices from Fort, Klein et Lagnoux 2021
- **SA on sets using RKHS theory**

Table of Contents

Table of Contents

Distribution embedding into a RKHS



Distance between distributions as expectation of kernels, Gretton et al. 2006

$$\gamma_k^2(\mathbb{P}, \mathbb{Q})) := ||\mu_{\mathbb{P}} - \mu_{\mathbb{Q}}||^2_{\mathcal{H}_k} = \mathbb{E}_{\substack{X,X' \sim \mathbb{P} \\ Y,Y' \sim \mathbb{Q}}}[k(X, X') + k(Y, Y') - 2k(X, Y)]$$

$\rightarrow$ is a distance between distribution iif k is *characteristic* (i.e. injectivity of the embedding)

MMD-based index

MMD-based index

$$\begin{aligned} S_i := & \mathbb{E}_{X_i}[\gamma_k^2(\mathbb{P}_Y, \mathbb{P}_{Y|X_i})] \\ & = \mathbb{E}_{X_i}\mathbb{E}_{\xi\sim\mathbb{P}_{Y|X_i}}\mathbb{E}_{\xi'\sim\mathbb{P}_{Y|X_i}}[k(\xi,\xi')] - \mathbb{E}_{Y\sim\mathbb{P}_Y}\mathbb{E}_{Y'\sim\mathbb{P}_Y}[k(Y,Y')]. \end{aligned}$$

$\rightarrow$ has a ANOVA decomposition (daVeiga 2021) : can be used to rank the inputs by influence.

Estimation

- Pick & freeze estimation
- Rank-based estimation

HSIC-based index Gretton et al. 2006

HSIC-based index

$$\begin{aligned} HSIC_k(X_i, Y) :&= \gamma_k^2(\mathbb{P}_{X_i,Y}, \mathbb{P}_{X_i} \otimes \mathbb{P}_Y) \\ &= \mathbb{E}_{X_i,X_i',Y,Y'} k_{\mathcal{X}}(X_i, X_i')k(Y, Y') \\ &+ \mathbb{E}_{X_i,X_i'} k_{\mathcal{X}}(X_i, X_i')\mathbb{E}_{Y,Y'} k(Y, Y') \\ &- 2\mathbb{E}_{X_i,Y}[\mathbb{E}_{X_i'} k_{\mathcal{X}}(X_i, X_i')\mathbb{E}_{Y'} k(Y, Y')] \end{aligned}$$

with (X_i', Y') an independent copy of (X_i, Y).

$HSIC_k(X_i, Y) = 0$ iif $X_i \perp Y$ (when k characteristic)$\rightarrow$ suited to identify the negligible inputs through independence testing.

Estimation

- Biased or unbiased classic estimators of $HSIC_k(X_i, Y)$
- p-values computable through asymptotic results or bootstrap methods

.

Table of Contents

SA on sets : a kernel between sets

Proposition (A kernel between sets, Balança et Herbin 2012)

The function $k_{set} : \mathcal{F}(\mathcal{X}) \times \mathcal{F}(\mathcal{X}) \to \mathbb{R}$ *defined by*

$$\forall \Gamma, \Gamma' \in \mathcal{F}(\mathcal{X}), \quad k_{set}(\Gamma, \Gamma') = \frac{\sigma^2}{2\lambda} e^{-\lambda \mu(\Gamma \Delta \Gamma')},$$

is positive definite for any positive scalars σ *and* λ*.*

Moore-Aronszajn theorem (Aronszajn (1950)) gives then the existence of a unique RKHS $\mathcal{H} \subset \mathcal{F}(\mathcal{X})^{\mathbb{R}}$ of reproducing kernel k_{set}.

Estimation with set-valued outputs

In the MMD- and HSIC-based indices expressions, we need to estimate quantities of the form $\mathbb{E}k_{set}(\Gamma_1, \Gamma_2)$.
With sets, $k_{set}(\Gamma_1, \Gamma_2)$ also require an estimation.

$$k_{set}(\Gamma_1, \Gamma_2) = e^{-\mu(\Gamma_1 \Delta \Gamma_2)}] \tag{4}$$
$$= e^{-\mathbb{E}[\mathbb{1}_{\Gamma\Delta\Gamma'}(X)]} \text{ with } X \sim \mathcal{U}(\mathcal{X}) \tag{5}$$
$$\simeq e^{-\mu(\mathcal{X})\frac{1}{N_x}\sum_{i=1}^{N_x} \mathbb{1}_{\Gamma_1\Delta\Gamma_2}(X^i)} \tag{6}$$
$$= \widehat{k_{set}}(\Gamma_1, \Gamma_2) \tag{7}$$

Then we inject it in the indices estimators. For instance the normalized MMD-based index estimated through pick and freeze method is :

$$\widehat{S^{MMD}_{A,p.f}} = \frac{\sum_{i=1}^{n} \widehat{k_{set}}(\Gamma^{(i)}, \tilde{\Gamma}^{A,(i)}) - \widehat{k_{set}}(\Gamma^{(i)}, \Gamma'^{(i)})}{\sum_{i=1}^{n} \widehat{k_{set}}(\Gamma^{(i)}, \Gamma^{(i)}) - \widehat{k_{set}}(\Gamma^{(i)}, \Gamma'^{(i)})}. \tag{8}$$

Asymptotic behaviour of the indices on sets

Proposition (Quadratic error of a nested Monte Carlo estimator)

With the previous notations, using Rainforth *et al. 2018, we have*

$$\mathbb{E}\left(\frac{1}{n}\sum_{j=1}^{n}\widehat{k_{set}}(\Gamma_1^{(j)},\Gamma_2^{(j)}) - \mathbb{E}k_{set}(\Gamma_1,\Gamma_2)\right)^2 = \mathcal{O}(\frac{1}{n}+\frac{1}{N_x^2}). \quad (9)$$

With this result, we can show that each quadratic error of our indices on sets has the same asymptotic behavior with rate $\mathcal{O}(\frac{1}{n}+\frac{1}{N_x^2})$

Test case : oscillator, Cousin 2021

$$g_1(\boldsymbol{x}, \boldsymbol{u}) = u_{r_1} - \max_{t \in [0,T]} \mathcal{Y}'(x_1 + u_1, x_2 + u_2, u_p; t), \tag{10}$$

$$g_2(\boldsymbol{x}, \boldsymbol{u}) = u_{r_2} - \max_{t \in [0,T]} \mathcal{Y}''(x_1 + u_1, x_2 + u_2, u_p; t), \tag{11}$$

with $\mathcal{Y}$ the solution of the harmonic oscillator defined by :

$$(x_1 + u_1)\mathcal{Y}''(t) + u_p\mathcal{Y}'(t) + (x_2 + u_2)\mathcal{Y}(t) = \eta(t). \tag{12}$$

Uncertainty	Distribution	Uncertainty	Distribution
U_1	$\mathcal{U}\left[-0.3, 0.3\right]$	U_{r_1}	$\mathcal{N}\left(1, 0.1^2\right)$
U_2	$\mathcal{U}\left[-1, 1\right]$	U_{r_2}	$\mathcal{N}\left(2.5, 0.25^2\right)$
U_p	$\mathcal{U}[0.5, 1.5]$	U_{r_3}	$\mathcal{N}\left(15, 3^2\right)$

Table – Definition of the uncertain inputs

Results on the oscillator case : MMD-based index

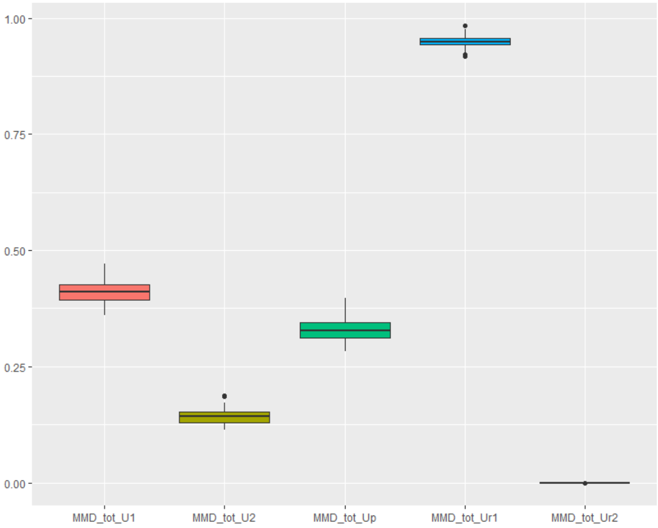


Figure – MMD-based total index for the constraint $g_1 \leq 0$

Results on the oscillator case : MMD-based index

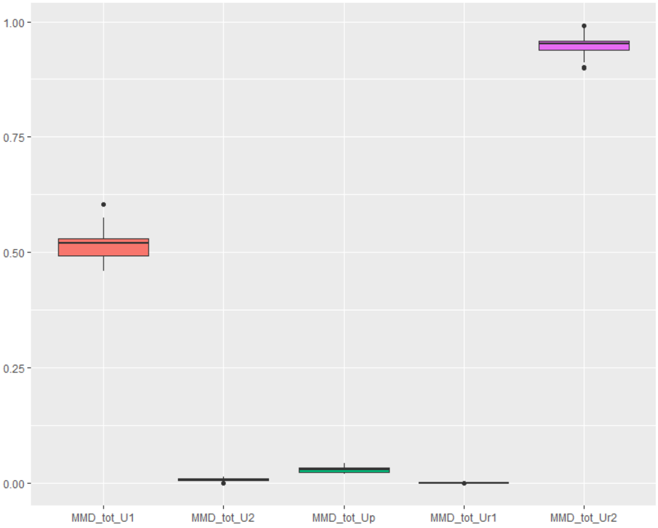


Figure – MMD-based total index for the constraint $g_2 \leq 0$

Results on the oscillator case : MMD-based index

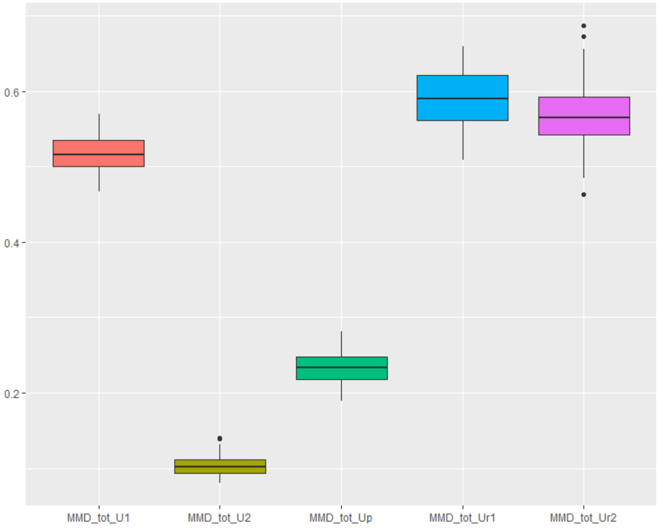


Figure – MMD-based total index for the constraint $g1 \leq 0$ and $g_2 \leq 0$

Results on the oscillator case : MMD-based index

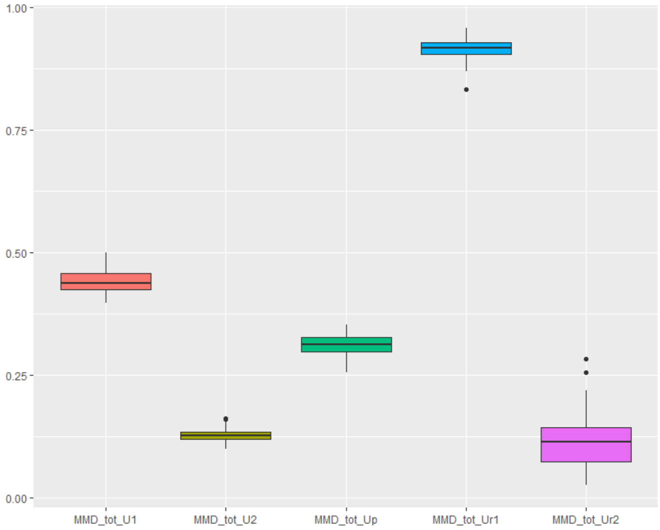


Figure – MMD-based total index for the couple $(g_1 \leq 0, g_2 \leq 0)$

Results on the oscillator case : HSIC-based index

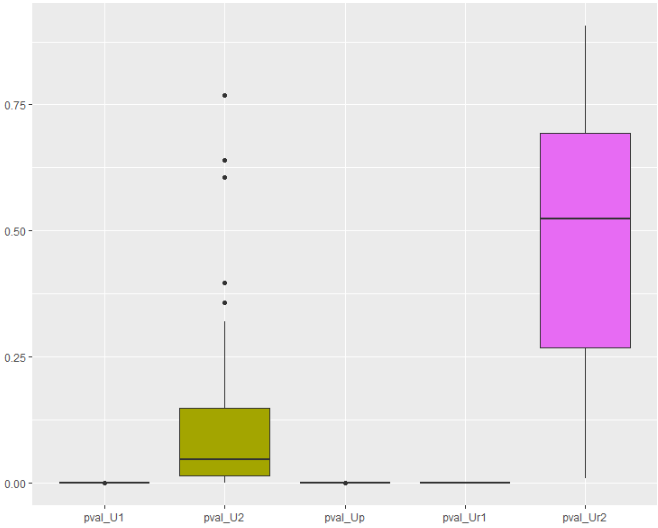


Figure – p-value of the HSIC-based index for the constraint $g_1 \leq 0$

Results on the oscillator case : HSIC-based index

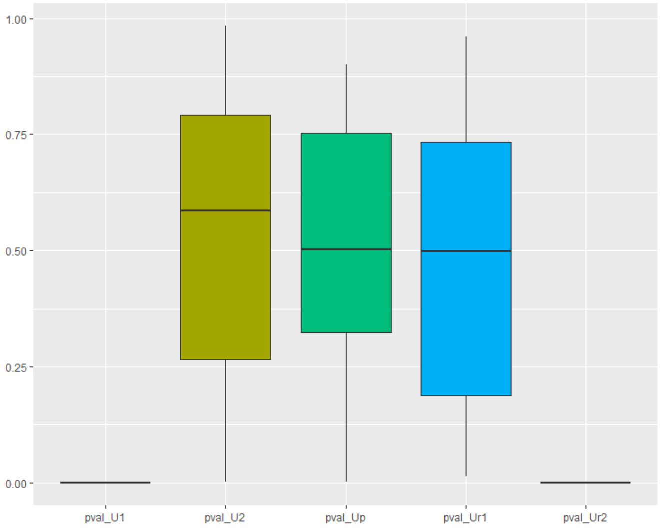


Figure – p-value of the HSIC-based index for the constraint $g_2 \leq 0$

Results on the oscillator case : HSIC-based index

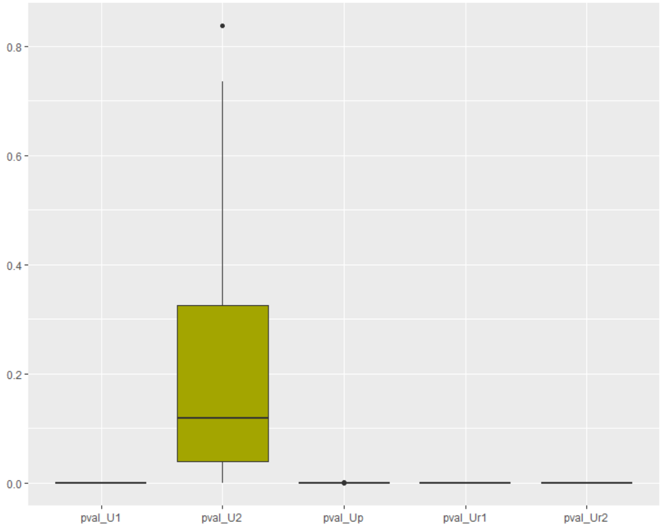


Figure – p-value of the HSIC-based index for the constraint $g1 \leq 0$ and $g_2 \leq 0$

Results on the oscillator case : HSIC-based index

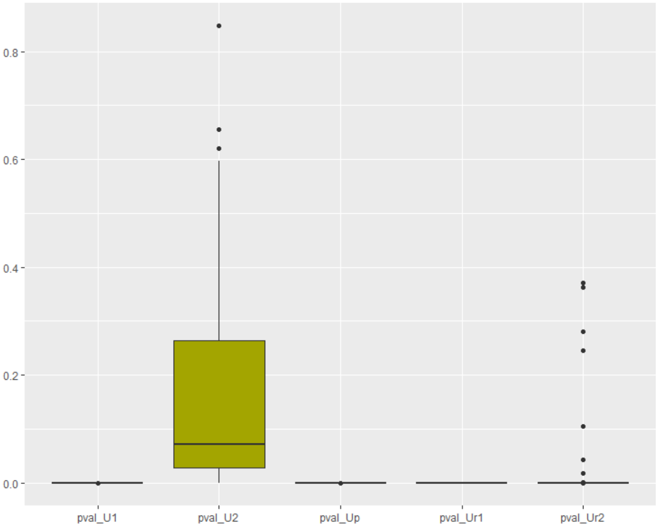


Figure – p-value of the HSIC-based index for the couple $(g_1 \leq 0, g_2 \leq 0)$

Conclusion

Kernel-based SA on set-valued outputs (paper "soon")

- A way to do SA on set-valued outputs
- On excursion sets : An answer to "How to do SA on the uncertain inputs in the context of robust optimization ?"

Future work

- Test the three methods on a real test case (of Adan Reyes Reyes from IFPEN)
- Use it inside an optimization

El-Amri, Reda et al. (2021). *A sampling criterion for constrained Bayesian optimization with uncertainties*. arXiv : 2103.05706 [stat.ML].

Aronszajn, N. (1950). “Theory of reproducing kernels”. In : *Transactions of the American Mathematical Society* 68.3, p. 337-404. issn : 0002-9947, 1088-6850. doi : 10.1090/S0002-9947-1950-0051437-7. url : https://www.ams.org/tran/1950-068-03/S0002-9947-1950-0051437-7/ (visité le 02/11/2022).

Balança, Paul et Erick Herbin (jan. 2012). “A set-indexed Ornstein-Uhlenbeck process”. In : *Electronic Communications in Probability* 17.none. doi : 10.1214/ecp.v17-1903. url : https://doi.org/10.1214%2Fecp.v17-1903.

Cousin, Alexis (oct. 2021). “Chance constraint optimization of a complex system : Application to the design of a floating offshore wind turbine”. Theses. Institut Polytechnique de Paris. url : https://tel.archives-ouvertes.fr/tel-03500604.

daVeiga, Sébastien (jan. 2021). “Kernel-based ANOVA decomposition and Shapley effects - Application to global sensitivity analysis”. working paper or preprint. url : https://hal.archives-ouvertes.fr/hal-03108628.

Fort, Jean-Claude, Thierry Klein et Agnès Lagnoux (jan. 2021). “Global Sensitivity Analysis and Wasserstein Spaces”. In : *SIAM/ASA Journal on Uncertainty Quantification* 9.2. Publisher : Society for Industrial and Applied Mathematics, p. 880-921. doi : 10.1137/20M1354957. url : https://epubs.siam.org/doi/abs/10.1137/20M1354957.

- Gretton, Arthur et al. (2006). "A Kernel Method for the Two-Sample-Problem". In : *Advances in Neural Information Processing Systems*. T. 19. MIT Press. url : https://proceedings.neurips.cc/paper/2006/hash/e9fb2eda3d9c55a0d89c98d6c54b5b3e-Abstract.html.
- Rainforth, Tom et al. (23 mai 2018). *On Nesting Monte Carlo Estimators*. arXiv : 1709.06181[stat]. url : http://arxiv.org/abs/1709.06181.
- Spagnol, Adrien (juill. 2020). "Kernel-based sensitivity indices for high-dimensional optimization problems". Theses. Université de Lyon. url : https://tel.archives-ouvertes.fr/tel-03173192.

Excursion sets are random sets

For K compact, $h(U) = X = \{x \in \mathcal{X}, g(x, U) \leq 0\}$

$$
\begin{aligned}
\{K \cap X \neq \emptyset\} &=^c \{\omega, K \cap X(\omega) = \emptyset\} \\
&=^c \{\omega, \forall x \in K \; g(x, U(\omega)) > 0\} \\
&=^c \{\omega, \inf_{x \in K} g(x, U(\omega)) > 0\} \text{ as } K \text{ compact and g continuous in } x \\
&=^c U^{-1}(\inf_{x \in K} g(x, \cdot)^{-1}(]0, +\infty[)) \in \mathcal{F}
\end{aligned}
$$

Indices estimation

$$S^{MMD}_{A,p.f} = \frac{\sum_{i=1}^{n} k_{set}(\Gamma^{(i)}, \tilde{\Gamma}^{A,(i)}) - k_{set}(\Gamma^{(i)}, \Gamma'^{(i)})}{\sum_{i=1}^{n} k_{set}(\Gamma^{(i)}, \Gamma^{(i)}) - k_{set}(\Gamma^{(i)}, \Gamma'^{(i)})}. \quad (13)$$

$$S^{MMD}_{l,rank} = \frac{\frac{1}{n}\sum_{i=1}^{n} k_{set}(\Gamma^{(i)}, \Gamma^{\sigma^l_n(i)}) - \frac{1}{n^2}\sum_{i,j=1}^{n} k_{set}(\Gamma^{(i)}, \Gamma^{(j)})}{\frac{1}{n}\sum_{i=1}^{n} k_{set}(\Gamma^{(i)}, \Gamma^{(i)}) - \frac{1}{n^2}\sum_{i,j=1}^{n} k_{set}(\Gamma^{(i)}, \Gamma^{(j)})}. \quad (14)$$

$$\mathrm{HSIC}_u\left(U_A, \Gamma\right) = \frac{1}{n(n-1)} \sum_{i,j=1, i\neq j}^{n} \left(k_{\mathcal{U}_A}\left(U_A^{(i)}, U_A^{(j)}\right) - 1\right) k_{set}\left(\Gamma^{(i)}, \Gamma^{(j)}\right), \quad (15)$$

$$\mathrm{HSIC}_b\left(U_A, \Gamma\right) = \frac{1}{n^2} \sum_{i,j=1}^{n} \left(k_{\mathcal{U}_A}\left(U_A^{(i)}, U_A^{(j)}\right) - 1\right) k_{set}\left(\Gamma^{(i)}, \Gamma^{(j)}\right). \quad (16)$$

Random set distribution embedding

Definition (Capacity functional)

The capacity functional of a random closed set Γ *denoted* T_Γ *is defined by :*

$$T_\Gamma : \begin{array}{ccc} \mathcal{K}(\mathcal{X}) & \rightarrow & [0,1] \\ K & \mapsto & \mathbb{P}(\Gamma \cap K \neq \emptyset). \end{array} \tag{17}$$

Definition (Mean embedding of a capacity functional)

The mean embedding of T_Γ *is defined as*

$$\mu_\Gamma = \mathbb{E}[k_{set}(\Gamma, \cdot)]. \tag{18}$$