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

UNCECOMP 2023, 12-14 June 2023

Table of Contents

Table of Contents

Sensitivity analysis

Sensitivity analysis (SA)

$$(U_1, ..., U_d) \overset{f}{\mapsto} Y = f(U_1, ..., U_d)$$

How can the uncertainty of Y be divided and allocated to the uncertainty of the inputs U_i ?

- Sobol indices : $S_i = \frac{\text{Var}\,\mathbb{E}(Y|U_i)}{\text{Var}\,Y}$
- Dependence measures : $S_i = ||\mathbb{P}_{(U_i,Y)} - \mathbb{P}_{U_i} \otimes \mathbb{P}_Y||$
  - Density-based indices (Borgonovo 2007)
  - Cramer von mises indices (Gamboa, Klein et Lagnoux 2018)
  - Hilbert Schmidt Independence Criterion : HSIC (Gretton, Bousquet et al. 2005)

Screening : $U_1, ..., U_k$ are influential and $U_{k+1}, ...U_d$ are not influential
Ranking : $U_1 \prec ... \prec U_d$

Sensitivity analysis

Sensitivity analysis (SA)

$$(U_1, ..., U_d) \overset{f}{\mapsto} Y = f(U_1, ..., U_d)$$

How can the uncertainty of Y be divided and allocated to the uncertainty of the inputs U_i ?

- Sobol indices : $S_i = \frac{\text{Var}\,\mathbb{E}(Y|U_i)}{\text{Var}\,Y}$
- Dependence measures : $S_i = ||\mathbb{P}_{(U_i,Y)} - \mathbb{P}_{U_i} \otimes \mathbb{P}_Y||$
    - Density-based indices (Borgonovo 2007)
    - Cramer von mises indices (Gamboa, Klein et Lagnoux 2018)
    - Hilbert Schmidt Independence Criterion : HSIC (Gretton, Bousquet et al. 2005)

Screening : $U_1, ..., U_k$ are influential and $U_{k+1}, ...U_d$ are not influential
Ranking : $U_1 \prec ... \prec U_d$

What if Y is an excursion set ?

A toy excursion set

Excursion sets

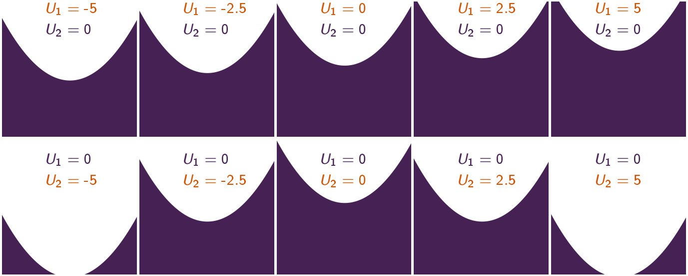
New output :

$$\Gamma_U = \{x \in \mathcal{X}, g(x, U) \leq 0\}, \quad (1)$$

which is called a random excursion set.

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

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



How can we do sensitivity analysis on (excursion) sets ?

Table of Contents

HSIC-based indices

Let $k_{\mathcal{U}_i}$ be input kernels (i.e. positive definite function from $\mathcal{U}_i \times \mathcal{U}_i \to \mathbb{R}$) and $k_{\mathcal{Y}}$ an output kernel.

Hilbert Schmidt Independence Criterion (HSIC), Gretton, Borgwardt et al. 2006

With $K = k_{\mathcal{U}_i} \otimes k_{\mathcal{Y}}$, the HSIC is given by :

$$\begin{aligned} \mathrm{HSIC}_K(U_i, Y) = &\ \mathbb{E}[k_{\mathcal{U}_i}(U_i, U_i')k_{\mathcal{Y}}(Y, Y')] \\ &+ \mathbb{E}[k_{\mathcal{U}_i}(U_i, U_i')]\mathbb{E}[k_{\mathcal{Y}}(Y, Y')] \\ &- 2\mathbb{E}[\mathbb{E}[k_{\mathcal{U}_i}(U_i, U_i')|U_i]\mathbb{E}[k_{\mathcal{Y}}(Y, Y')|Y]]. \end{aligned}$$

- When K is characteristic (injectivity of the mean embedding),

$$\mathrm{HSIC}_K(U_i, Y) = 0 \text{ iif } U_i \perp Y \to \text{screening}.$$

- Easy to estimate :

$$\widehat{\mathrm{HSIC}}(U_i, Y) = \frac{1}{n^2} \mathrm{Tr}\left(\boldsymbol{L}_i \boldsymbol{H} \boldsymbol{L} \boldsymbol{H}\right)$$

where $L_{i_{j,k}} = k_{\mathcal{U}_i}\left(u_i^j, u_i^k\right)$, $L_{j,k} = k_{\mathcal{Y}}\left(y^j, y^k\right)$ and $H_{j,k} = \left(\delta_{j,k} - \frac{1}{n}\right)_{1\leq j,k\leq n}$.

HSIC-ANOVA indices [daVeiga 2021]

Assuming that the inputs are independent and that the input kernels are ANOVA,

$$\mathrm{HSIC}(\boldsymbol{U}, Y) = \sum_{A \subseteq \{1,\dots,d\}} \sum_{B \subseteq A} (-1)^{|A|-|B|} \, \mathrm{HSIC}\left(\boldsymbol{U}_B, Y\right).$$

HSIC-ANOVA indices are then defined as :

$$S_i^{\mathrm{HSIC}} := \frac{\mathrm{HSIC}(U_i, Y)}{\mathrm{HSIC}(\boldsymbol{U}, Y)},$$

$$S_{T_i}^{\mathrm{HSIC}} := 1 - \frac{\mathrm{HSIC}(\boldsymbol{U}_{-i}, Y)}{\mathrm{HSIC}(\boldsymbol{U}, Y)}$$

and are suited for ranking (and screening).

Strength of the HSIC-ANOVA indices

- suited for ranking and screening
- easy to estimate
- only require to have kernels on the inputs and on the output (whatever the type of output you have)

HSIC on sets : a kernel between sets

With $A \Delta B = A \cup B - B \cap A$ and λ the Lebesgue measure, we define a kernel on the Lebesgue σ-algebra $\mathcal{B}(\mathcal{X})$ by :

$$\forall \Gamma_1, \Gamma_2 \in \mathcal{B}(\mathcal{X}), \; k_{set}(\Gamma_1, \Gamma_2) = \exp\left(-\frac{\lambda(\Gamma_1 \Delta \Gamma_2)}{2\sigma^2}\right).$$

Proposition (A kernel between sets)

k_{set} *is a kernel* [Balança et Herbin 2012] *and is characteristic.*

For a given random excursion set $\Gamma_U = \{x \in \mathcal{X}, g(x, U) \leq 0\}$, we can define HSIC-based indices on sets :

$$S_i^{\mathrm{H_{set}}} := \frac{\mathrm{HSIC}_{k_{set}}(U_i, \Gamma_U)}{\mathrm{HSIC}_{k_{set}}(\boldsymbol{U}, \Gamma_U)},$$

which quantifies how much U_i impacts the excursion set Γ.

Table of Contents

Toy function 1

From El-Amri et al. 2021,

$$\forall x, u \in [-5,5]^2 \times [-5,5]^3 \quad g(x_1, x_2, u_1, u_2, u_3) = -x_1^2 + 5x_2 - u_1 + u_2^2 - 1$$

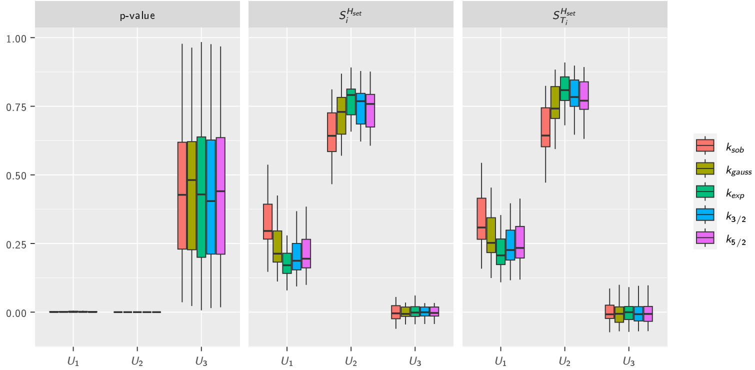


Figure – Estimation of the p-values, $\hat{\hat{S}}_i^{H_{set}}$ and $\hat{\hat{S}}_{T_i}^{H_{set}}$ for the excursion set defined by the constraint $g \leq 0$ computed for 5 input kernels with $n = 100$, $m = 100$ and repeated 20 times

Pollutant concentration maps : Maps of Sobol indices

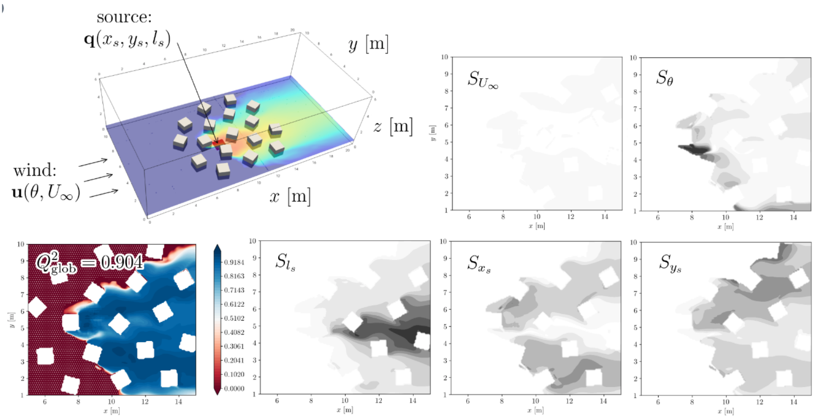


Figure – Maps of the sobol indices of pollutant dispersion (Mathis Pasquier)

Interpretation :

$$S_{l_s} \gg S_{x_s} \approx S_{y_s} \approx S_\theta \gg S_{U_\infty}$$

Kernel-based SA on pollutant concentration maps

Sobol map interpretation : $S_{l_s} \gg S_{x_s} \approx S_{y_s} \approx S_\theta \gg S_{U_\infty}$.
$\forall (x, y) \in [5, 15] \times [1, 10]$, $g(x, y, U)$ is the pollutant concentration at the point (x, y) for a given uncertain parameter U. What is the set-valued output ?

- Test 1 : $\Gamma_U = \{(x, y) \in [5, 15] \times [1, 10], g(x, y, U) \geq C_{seuil}\}$. C_{seuil} to choose (toxicity threshold).

	θ	U_∞	x_s	y_s	l_s
P-value	$6.6.10^{-4}$	0.11	0	0	0
$S_i^{\mathrm{H}_{set}}$	0.069	0.016	0.25	0.15	0.48

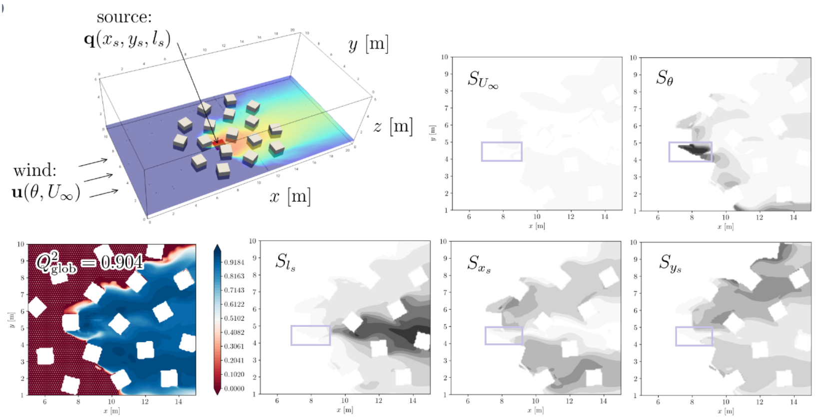
- Test 2 : $\Gamma_U = \{(x, y, z) \in [5, 15] \times [1, 10] \times [C_{min}, C_{max}], z \leq g(x, y, U)\}$

	θ	U_∞	x_s	y_s	l_s
P-value	$6.6.10^{-3}$	0.77	0.026	0.010	0
$S_i^{\mathrm{H}_{set}}$	0.11	0.0081	0.20	0.17	0.50

In both cases we obtain :

$$S_{l_s} > S_{x_s} > S_{y_s} > S_\theta > S_{U_\infty}$$

Kernel-based SA on pollutant concentration maps : subspace



$$\Gamma_U = \{(x, y, C) \in [7, 9] \times [4, 5] \times [C_{min}, C_{max}], C \leq g(x, y, U)\}$$

	θ	U_∞	x_s	y_s	l_s
P-value	0	0.002	0	0	0.03
$S_i^{\text{H}set}$	0.59	0.09	0.14	0.13	0.04

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

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.

Borgonovo, E. (2007). “A new uncertainty importance measure”. In : *Reliability Engineering and System Safety* 92.6, p. 771-784. issn : 0951-8320. doi : https://doi.org/10.1016/j.ress.2006.04.015. url : https://www.sciencedirect.com/science/article/pii/S0951832006000883.

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.

Gamboa, Fabrice, Thierry Klein et Agnès Lagnoux (2018). “Sensitivity Analysis Based on Cramér–von Mises Distance”. In : *SIAM/ASA Journal on Uncertainty Quantification* 6.2, p. 522-548. doi : 10.1137/15M1025621. eprint : https://doi.org/10.1137/15M1025621. url : https://doi.org/10.1137/15M1025621.

Gretton, Arthur, Karsten Borgwardt 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.

Gretton, Arthur, Olivier Bousquet et al. (2005). "Measuring Statistical Dependence with Hilbert-Schmidt Norms". In : *Algorithmic Learning Theory*. Sous la dir. de Sanjay Jain, Hans Ulrich Simon et Etsuji Tomita. Berlin, Heidelberg : Springer Berlin Heidelberg, p. 63-77. isbn : 978-3-540-31696-1.

Ziegel, Johanna, David Ginsbourger et Lutz Dümbgen (2022). *Characteristic kernels on Hilbert spaces, Banach spaces, and on sets of measures*. arXiv : 2206.07588 [stat.ML].

k_{set} is characteristic, sketch of proof

- $\mathcal{B}(\mathcal{X}) \to \mathcal{B} = \mathcal{B}(\mathcal{X})/\sim_\delta$ where δ is the volume of the symmetric difference and $\sim_\delta$ the equivalent relation $A \sim_\delta B$ iif $\delta(A, B) = 0$ i.e. A and B are equal except on a λ-negligible set.
- We show that $(\mathcal{B}, \delta)$ is a Polish space (separable completely metrizable topological space). $(\mathcal{B}, \delta)$ is a metric space so we just need separability and completeness. Separability holds as "it's a subspace of $L_2(\mathcal{X})$" and we show that it's closed.
- We use a Proposition from Ziegel, Ginsbourger et Dümbgen 2022,

Proposition

Let $\mathcal{B}$ be a Polish space, H a separable Hilbert space, T a measurable and injective mapping from $\mathcal{B}$ to H, and $\varphi \in \Phi_\infty^+$. Then, the kernel k on $\mathcal{B}$ defined by

$$k(\gamma, \gamma') := \varphi\left(\|T(\gamma) - T(\gamma')\|_H^2\right), \quad \gamma, \gamma' \in \mathcal{B}$$

is integrally strictly positive definite with respect to $\mathcal{M}(\mathcal{B})$ (which implies that it is characteristic).

with $H = L_2(\mathcal{X})$, $\varphi = \exp(-\frac{\cdot}{2\sigma^2})$ and T defined by $T(\gamma) := x \mapsto \mathbb{1}_\gamma(x)$ for any $\gamma \in \mathcal{B}$ so that $\|T(\gamma) - T(\gamma')\|_H^2 = \lambda(\gamma \Delta \gamma')$.

HSIC-ANOVA on sets, estimation

- $\mathrm{H}_{set}(U_I, \Gamma) := \mathrm{HSIC}_{k_I, k_{set}}(U_I, \Gamma) = \mathbb{E}\left[(k_I(U_I, U_I{}') - 1) k_{set}(\Gamma, \Gamma')\right]$

HSIC-ANOVA on sets, estimation

- $\mathrm{H}_{set}(U_I, \Gamma) := \mathrm{HSIC}_{k_I, k_{set}}(U_I, \Gamma) = \mathbb{E}\left[(k_I(U_I, U_I{}') - 1)k_{set}(\Gamma, \Gamma')\right]$
- $\widehat{\mathrm{H}_{set}}\,(U_I, \Gamma) = \frac{2}{n(n-1)} \sum_{i<j}^{n} \left(k_I\left(U_I^{(i)}, U_I^{(j)}\right) - 1\right) k_{set}\left(\Gamma^{(i)}, \Gamma^{(j)}\right)$

HSIC-ANOVA on sets, estimation

- $H_{set}(U_l, \Gamma) := \mathrm{HSIC}_{k_l, k_{set}}(U_l, \Gamma) = \mathbb{E}\left[(k_l(U_l, U_l') - 1) k_{set}(\Gamma, \Gamma')\right]$
- $\widehat{H_{set}}\,(U_l, \Gamma) = \frac{2}{n(n-1)} \sum_{i<j}^{n} \left(k_l\left(U_l^{(i)}, U_l^{(j)}\right) - 1\right) k_{set}\left(\Gamma^{(i)}, \Gamma^{(j)}\right)$
- $\widehat{k_{set}}(\Gamma^{(i)}, \Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X_{i,j}^{(k)})) \rightarrow$ $n(n-1)m$ tests of $X \in \Gamma$

HSIC-ANOVA on sets, estimation

- $\mathrm{H}_{set}(U_l,\Gamma) := \mathrm{HSIC}_{k_l,k_{set}}(U_l,\Gamma) = \mathbb{E}\left[(k_l(U_l,U_l{}') - 1)k_{set}(\Gamma,\Gamma')\right]$
- $\widehat{\mathrm{H}_{set}}\left(U_l,\Gamma\right) = \frac{2}{n(n-1)}\sum_{i<j}^{n}\left(k_l\left(U_l^{(i)},U_l^{(j)}\right)-1\right)k_{set}\left(\Gamma^{(i)},\Gamma^{(j)}\right)$
- $\widehat{k_{set}}(\Gamma^{(i)},\Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2}\frac{1}{m}\sum_{k=1}^{m}\mathbb{1}_{\Gamma^{(i)}\Delta\Gamma^{(j)}}(X_{i,j}^{(k)}))\rightarrow n(n-1)m$ tests of $X\in\Gamma$
- $\widehat{k_{set}}(\Gamma^{(i)},\Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2}\frac{1}{m}\sum_{k=1}^{m}\mathbb{1}_{\Gamma^{(i)}\Delta\Gamma^{(j)}}(X^{(k)}))$

HSIC-ANOVA on sets, estimation

- $\mathrm{H}_{set}(U_I, \Gamma) := \mathrm{HSIC}_{k_I, k_{set}}(U_I, \Gamma) = \mathbb{E}\left[(k_I(U_I, U_I') - 1)k_{set}(\Gamma, \Gamma')\right]$
- $\widehat{\mathrm{H}_{set}}(U_I, \Gamma) = \frac{2}{n(n-1)} \sum_{i<j}^{n} \left(k_I\left(U_I^{(i)}, U_I^{(j)}\right) - 1\right) k_{set}\left(\Gamma^{(i)}, \Gamma^{(j)}\right)$
- $\widehat{k_{set}}(\Gamma^{(i)}, \Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X_{i,j}^{(k)})) \rightarrow n(n-1)m$ tests of $X \in \Gamma$
- $\widehat{k_{set}}(\Gamma^{(i)}, \Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X^{(k)}))$
- $\widehat{\widehat{\mathrm{H}_{set}}}(U_I, \Gamma) = \frac{2}{n(n-1)} \sum_{i<j}^{n} \left(k_I\left(U_I^{(i)}, U_I^{(j)}\right) - 1\right) \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X^{(k)}))$

HSIC-ANOVA on sets, estimation

- $\mathrm{H}_{set}(U_I, \Gamma) := \mathrm{HSIC}_{k_I, k_{set}}(U_I, \Gamma) = \mathbb{E}\left[(k_I(U_I, U_I{}') - 1)k_{set}(\Gamma, \Gamma')\right]$
- $\widehat{\mathrm{H}_{set}}(U_I, \Gamma) = \frac{2}{n(n-1)} \sum_{i<j}^{n} \left(k_I\left(U_I^{(i)}, U_I^{(j)}\right) - 1\right) k_{set}\left(\Gamma^{(i)}, \Gamma^{(j)}\right)$
- $\widehat{k_{set}}(\Gamma^{(i)}, \Gamma^{(j)}) = \exp\left(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X_{i,j}^{(k)})\right) \rightarrow$ $n(n-1)m$ tests of $X \in \Gamma$
- $\widehat{k_{set}}(\Gamma^{(i)}, \Gamma^{(j)}) = \exp\left(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X^{(k)})\right)$
- $\widehat{\widehat{\mathrm{H}_{set}}}(U_I, \Gamma) = \frac{2}{n(n-1)} \sum_{i<j}^{n} \left(k_I\left(U_I^{(i)}, U_I^{(j)}\right) - 1\right) \exp\left(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X^{(k)})\right)$

Proposition

The quadratic risk of the nested estimator $\widehat{\widehat{\mathrm{H}_{set}}}$ *verifies :*

$$\mathbb{E}\left(\widehat{\widehat{\mathrm{H}_{set}}}(U_I, \Gamma) - \mathrm{H}_{set}(U_I, \Gamma)\right)^2 \leq 2\left(\frac{2\sigma_1^2}{n(n-1)} + \frac{4(n-2)\sigma_2^2}{n(n-1)} + \frac{L^2\sigma_3^2}{m}\right).$$

HSIC-ANOVA on sets, estimation

- $\mathrm{H}_{set}(U_l, \Gamma) := \mathrm{HSIC}_{k_l, k_{set}}(U_l, \Gamma) = \mathbb{E}\left[(k_l(U_l, U_l') - 1)k_{set}(\Gamma, \Gamma')\right]$
- $\widehat{\mathrm{H}_{set}}(U_l, \Gamma) = \frac{2}{n(n-1)} \sum_{i<j}^{n} \left(k_l\left(U_l^{(i)}, U_l^{(j)}\right) - 1\right) k_{set}\left(\Gamma^{(i)}, \Gamma^{(j)}\right)$
- $\widehat{k_{set}}(\Gamma^{(i)}, \Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X_{i,j}^{(k)})) \rightarrow n(n-1)m$ tests of $X \in \Gamma$
- $\widehat{k_{set}}(\Gamma^{(i)}, \Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X^{(k)}))$
- $\widehat{\widehat{\mathrm{H}_{set}}}(U_l, \Gamma) = \frac{2}{n(n-1)} \sum_{i<j}^{n} \left(k_l\left(U_l^{(i)}, U_l^{(j)}\right) - 1\right) \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X^{(k)}))$

Proposition

The quadratic risk of the nested estimator $\widehat{\widehat{\mathrm{H}_{set}}}$ *verifies :*

$$\mathbb{E}\left(\widehat{\widehat{\mathrm{H}_{set}}}(U_l, \Gamma) - \mathrm{H}_{set}(U_l, \Gamma)\right)^2 \leq 2\left(\frac{2\sigma_1^2}{n(n-1)} + \frac{4(n-2)\sigma_2^2}{n(n-1)} + \frac{L^2\sigma_3^2}{m}\right).$$

We can now compute $S_i^{\widehat{\widehat{\mathrm{H}_{set}}}}$ or $S_{T_i}^{\widehat{\widehat{\mathrm{H}_{set}}}}$ to perform SA on set-valued outputs.