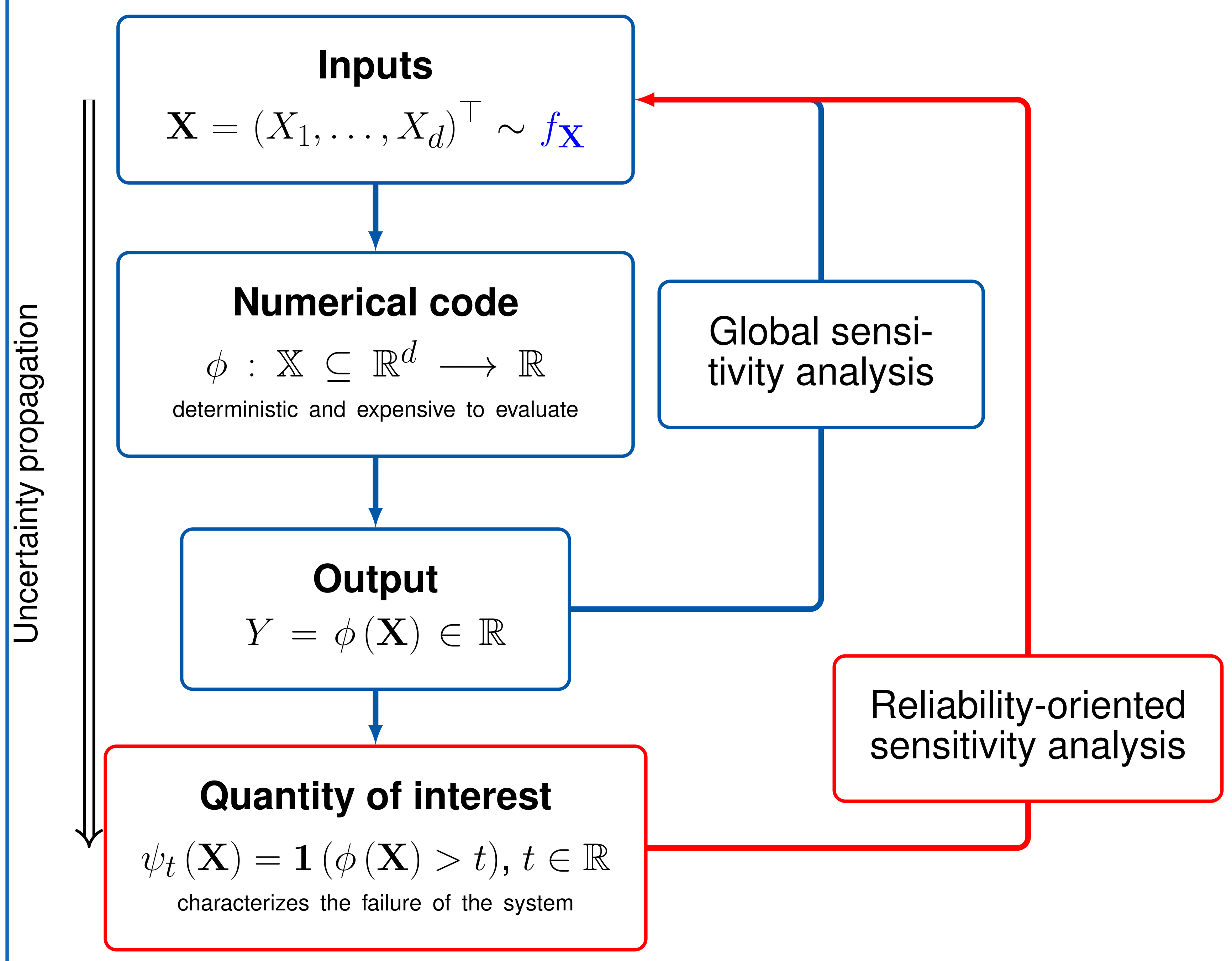
ONERA THE FRENCH AEROSPACE LAB — return on innovation

Shapley effect estimation in reliability-oriented sensitivity analysis with correlated inputs by importance sampling

Julien Demange-Chryst[1,2] – Jérôme Morio[1] – François Bachoc[2]

Context

In the general **uncertainty quantification** framework, the study of a physical system is represented by the following schema:



A major goal in reliability analysis is to efficiently answer the question: **with correlated inputs, which components of $\mathbf{X}$ are most likely to lead to the failure of the system ?**

Reliability-oriented sensitivity analysis

<u>Question</u>: What is the influence of each input component of $\mathbf{X}$ on the random quantity of interest $\psi_t(\mathbf{X}) = \mathbf{1}(\phi(\mathbf{X}) > t)$?

With correlated inputs, Sobol indices [6] do not allow anymore to identify the origin of the variability of the output.
$\Longrightarrow$ Alternative : consider the **Shapley effects** [5] for global sensitivity analysis [4]:

$$\text{T-Sh}_i = \frac{1}{d}\frac{1}{\mathbb{V}(\psi_t(\mathbf{X}))}\sum_{u \subseteq \{-i\}} \binom{d-1}{|u|}^{-1} \left(\text{T-S}^c_{u\cup\{i\}} - \text{T-S}^c_u\right) \quad (1)$$

with $\text{T-S}^c_u = \mathbb{V}\left[\mathbb{E}\left(\psi_t(\mathbf{X})\,|\mathbf{X}_u\right)\right]$ or $\text{T-S}^c_u = \mathbb{E}\left[\mathbb{V}\left(\psi_t(\mathbf{X})\,|\mathbf{X}_{-u}\right)\right]$ where for $u \subseteq [\![1,d]\!]$, $-u = [\![1,d]\!]\backslash u$ and $\mathbf{X}_u = (X_i)_{i\in u}$.

The estimation of the **target Shapley effects** [3] in (1) can be divided into two steps:

1. estimation of the closed Sobol indices T-S^c_u for $u \subseteq [\![1,d]\!]$
2. aggregation procedure using (1)

A less expensive estimation method requiring only a **unique input/output N-sample** has been introduced by [1].

Problem and scientific goal

<u>Problem</u>: the existing estimators of the closed Sobol indices are based on a Monte Carlo sampling according to the input distribution $f_{\mathbf{X}}$
$\Longrightarrow$ they are not efficient when $p_t = \mathbb{P}(\phi(\mathbf{X}) > t) \ll 1$ because they require too many calls to ϕ to be accurate.
<u>Goal</u>: estimate more efficiently the T-Sh_i in (1) when $p_t \ll 1$.
<u>Idea</u>: estimate the T-S^c_u with **importance sampling**.

[1] ONERA/DTIS, University of Toulouse
[2] Institut de Mathématiques de Toulouse / University Toulouse III - Paul Sabatier

Estimators with importance sampling

- <u>Principle for the estimation of a failure probability</u>: rewriting p_t according to an auxiliary sampling distribution g, the corresponding natural estimator becomes:

$$\widehat{p}^{\text{IS}}_{t,N} = \frac{1}{N}\sum_{n=1}^{N} w^g_t\left(\mathbf{X}^{(n)}\right) \text{ with } \begin{cases} \left(\mathbf{X}^{(n)}\right)_{n\in[1,N]} \sim g \\ w^g_t(\mathbf{x}) = \psi_t(\mathbf{x})\dfrac{f_{\mathbf{X}}(\mathbf{x})}{g(\mathbf{x})} \end{cases} \quad (2)$$

- <u>Adaptation to the estimation of T-S^c_u [2]</u>: rewriting T-S^c_u according to an auxiliary sampling distribution g, the corresponding natural estimators become:
  1. double Monte-Carlo method

$$\widehat{\text{T-S}}^{\text{IS}}_{u,\text{MC}} = \widehat{p}^{\text{IS}}_{t,N} - \frac{1}{N_u}\sum_{n=1}^{N_u}\left(\overline{\psi^{\text{IS}}_{t,N_I}}(\mathbf{X}^{(n)}_{-u})\right)^2 \frac{g_{\mathbf{X}_{-u}}\left(\mathbf{X}^{(n)}_{-u}\right)}{f_{\mathbf{X}_{-u}}\left(\mathbf{X}^{(n)}_{-u}\right)} \quad (3)$$

with $\overline{\psi^{\text{IS}}_{t,N_I}}\left(\mathbf{X}^{(n)}_{-u}\right) = \frac{1}{N_I}\sum_{k=1}^{N_I} w^g_t\left(\mathbf{X}^{(n,k)}_u, \mathbf{X}^{(n)}_{-u}\right)$

  2. Pick-Freeze method

$$\widehat{\text{T-S}}^{\text{IS}}_{u,\text{PF}} = \frac{1}{N_u}\sum_{n=1}^{N_u} w^g_t\left(\mathbf{X}^{(n)}_u, \mathbf{X}^{(n,1)}_{-u}\right) w^g_t\left(\mathbf{X}^{(n)}_u, \mathbf{X}^{(n,2)}_{-u}\right)\frac{g_{\mathbf{X}_u}(\mathbf{X}^{(n)}_u)}{f_{\mathbf{X}_u}(\mathbf{X}^{(n)}_u)} - \left(\widehat{p}^{\text{IS}}_{t,N}\right)^2 \quad (4)$$

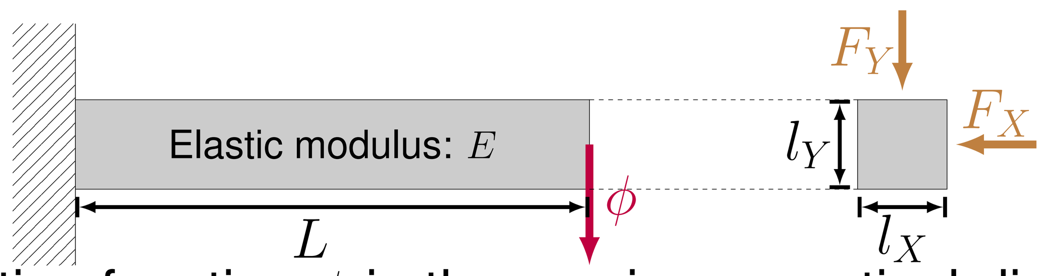
Theorem 1. *[2] Using $g_{opt}(\mathbf{x}) \propto \psi_t(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})$ as the importance sampling auxiliary density, we have:*

$$\mathbb{V}_{g_{opt}}\left(\widehat{T\text{-}S}^{IS}_{u,PF}\right) \le \frac{p_t^2}{N_u} \underset{p_t\to 0}{\le} \mathbb{V}_{f_{\mathbf{X}}}\left(\widehat{T\text{-}S}_{u,PF}\right) \quad (5)$$

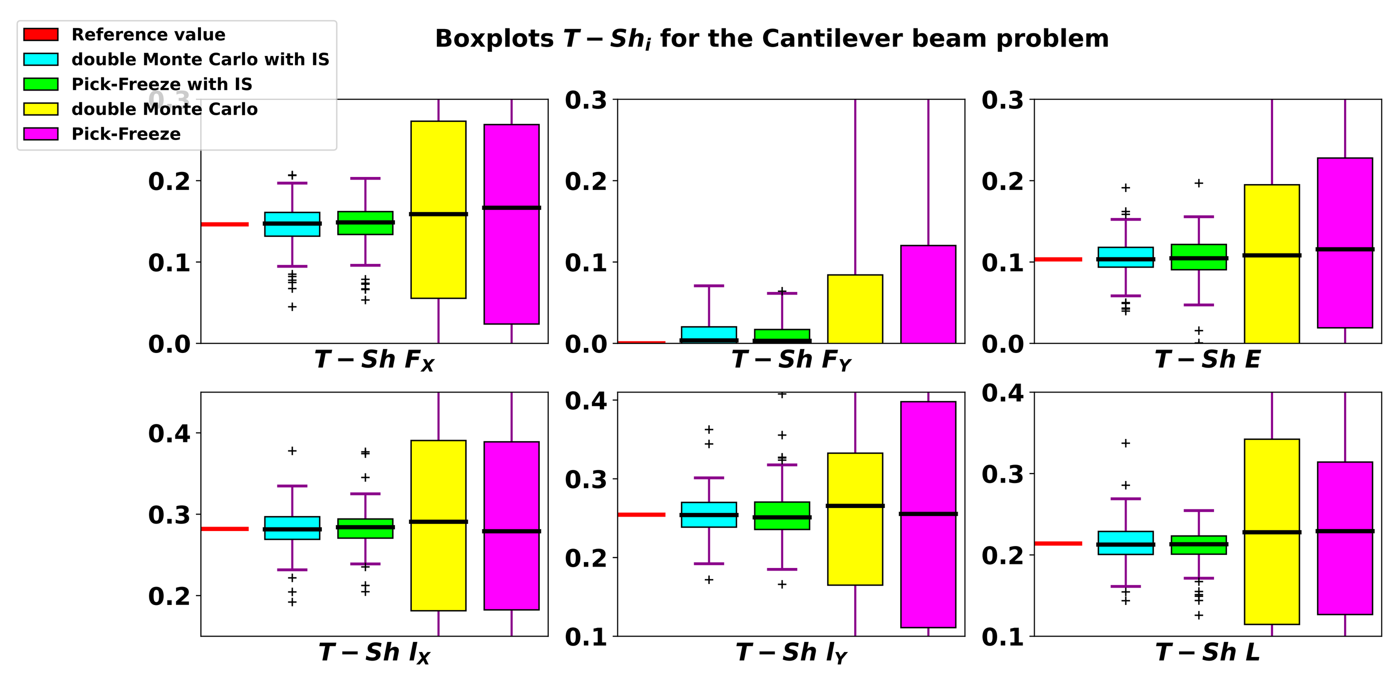
where $\widehat{T\text{-}S}_{u,PF}$ is the existing Pick-Freeze estimator of $T\text{-}S^c_u$.

Results

Consider the following cantilever beam problem:



The objective function ϕ is the maximum vertical displacement of the tip section under both forces F_X and F_Y. The maximal displacement allowed is $t = 0.066\text{cm}$ such that $p_t \approx 0.015$.



Bibliography

[1] Baptiste Broto, François Bachoc, and Marine Depecker. Variance reduction for estimation of shapley effects and adaptation to unknown input distribution. *SIAM/ASA Journal on Uncertainty Quantification*, 8(2):693–716, 2020.

[2] Julien Demange-Chryst, François Bachoc, and Jérôme Morio. Shapley effect estimation in reliability-oriented sensitivity analysis with correlated inputs by importance sampling. *arXiv preprint arXiv:2202.12679*, 2022.

[3] Marouane Il Idrissi, Vincent Chabridon, and Bertrand Iooss. Developments and applications of shapley effects to reliability-oriented sensitivity analysis with correlated inputs. *Environmental Modelling & Software*, 143:105115, 2021.

[4] Art B Owen. Sobol'indices and shapley value. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):245–251, 2014.

[5] Shapley. A value for n-person games. *Contributions to the Theory of Games*, (28):307–317, 1953.

[6] Ilya M Sobol. Sensitivity analysis for non-linear mathematical models. *Mathematical modelling and computational experiment*, 1:407–414, 1993.