Efficient estimation of multiple expectations with the same sample by adaptive importance sampling and control variates

Julien Demange-Chryst, PhD student
julien.demange-chryst@onera.fr

Ph.D. supervisors: Jérôme Morio[1], François Bachoc[2]

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

Context

Uncertainty quantification

Numerical code

$$\phi : \mathbb{X} \subseteq \mathbb{R}^d \longrightarrow \mathbb{R}$$

Characteristics of the numerical code ϕ:

- black-box model
- deterministic
- expensive to evaluate
  $\Longrightarrow$ cost of an algorithm : number of calls to ϕ

Context

Uncertainty quantification

Inputs
$\mathbf{X} = (X_1, \ldots, X_d)^\top \sim f$

$\longrightarrow$

Numerical code
$\phi : \mathbb{X} \subseteq \mathbb{R}^d \longrightarrow \mathbb{R}$

Characteristics of the numerical code ϕ:

- black-box model
- deterministic
- expensive to evaluate
  $\Longrightarrow$ cost of an algorithm : number of calls to ϕ

Context

Uncertainty quantification

Uncertainty propagation

$$\Longrightarrow$$

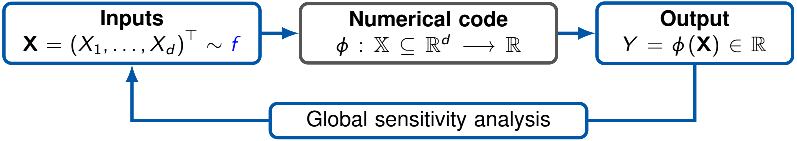
Inputs	Numerical code	Output
$\mathbf{X} = (X_1, \ldots, X_d)^\top \sim f$	$\phi : \mathbb{X} \subseteq \mathbb{R}^d \longrightarrow \mathbb{R}$	$Y = \phi(\mathbf{X}) \in \mathbb{R}$

Characteristics of the numerical code ϕ:

- black-box model
- deterministic
- expensive to evaluate
  $\Longrightarrow$ cost of an algorithm : number of calls to ϕ

Variance-based global sensitivity analysis

Introduction



Goal of global sensitivity analysis

Identify the most influential input variables of $\mathbf{X}$ on the variability of the output Y

Some motivations of sensitivity analysis:

- input prioritisation
- factor fixing
- variance cutting

Independent inputs: Sobol' indices [1]

Definition, properties and estimation

With **independent** input variables, the functional variance decomposition (or ANOVA) is unique and leads to the Sobol' indices [1]:

First order Sobol' indices:

$$S_i = \frac{\mathbb{V}\left[\mathbb{E}\left(\phi(\mathbf{X})|X_i\right)\right]}{\mathbb{V}\left(\phi(\mathbf{X})\right)}$$

Important properties:

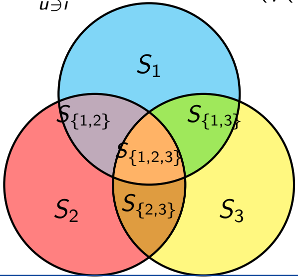
$$0 \leq S_i \leq S_{T_i} \leq 1 \text{ and } \sum_{u \in [\![1,d]\!]} S_u = 1$$

Some estimation methods:

- Pick-Freeze [1]
- estimators based on rank statistics [3]
- surrogate-based methods (PCE) [4]

Total order Sobol' indices [2]:

$$S_{T_i} = \sum_{u \ni i} S_u = 1 - \frac{\mathbb{V}\left[\mathbb{E}\left(\phi(\mathbf{X})|\mathbf{X}_{-i}\right)\right]}{\mathbb{V}\left(\phi(\mathbf{X})\right)}$$



Estimation of multiple expectations

Problem presentation

For $J \geq 2$, the main goal is to **efficiently** and **commonly** estimate with the same N-sample the expectations:

$$\forall j \in [\![1, J]\!],\ I_j = \mathbb{E}_{f_j}\left(\phi_j\left(\mathbf{X}\right)\right)$$

The distributions $(f_j)_{j \in [\![1,J]\!]}$ are OpenTURNS distributions and the computer models $(\phi_j)_{j \in [\![1,J]\!]}$ are OpenTURNS PythonFunction.

Application examples :

- estimation of sensitivity indices
- sensitivity analysis under epistemic uncertainty
- estimation of covariance matrices

Estimation of multiple expectations

Problem presentation

For $J \geq 2$, the main goal is to **efficiently** and **commonly** estimate with the same N-sample the expectations:

$$\forall j \in [\![1, J]\!],\ I_j = \mathbb{E}_{f_j}\left(\phi_j\left(\mathbf{X}\right)\right)$$

Criterion to minimise

The quality of the common estimation of a family of expectations is quantified by:

$$\sum_{j=1}^{J} w_j \mathbb{V}\left(\widehat{I}_j\right)$$

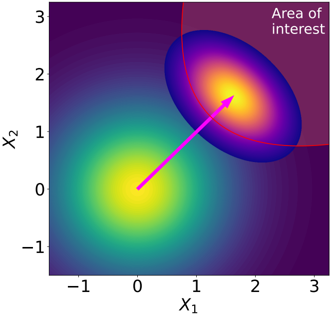
where $\widehat{I}_1, \ldots, \widehat{I}_J$ are **unbiased** estimators based on the same N-sample.

Principle of importance sampling

Consider an auxiliary sampling distribution g to draw more samples in interesting zones of the input domain $\mathbb{X}$.

Principle of importance sampling

Consider an auxiliary sampling distribution g to draw more samples in interesting zones of the input domain $\mathbb{X}$.



Rewriting I according to g:

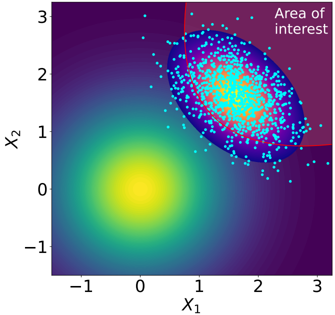
$$I = \mathbb{E}_f \left[\phi\left(\mathbf{X}\right) \right] = \mathbb{E}_g \left[\phi\left(\mathbf{X}\right) \frac{f\left(\mathbf{X}\right)}{g\left(\mathbf{X}\right)} \right]$$

Variance reduction methods

Method 1 : importance sampling

Principle of importance sampling

Consider an auxiliary sampling distribution g to draw more samples in interesting zones of the input domain $\mathbb{X}$.



Rewriting I according to g:

$$I = \mathbb{E}_f \left[\phi\left(\mathbf{X}\right)\right] = \mathbb{E}_g \left[\phi\left(\mathbf{X}\right) \frac{f\left(\mathbf{X}\right)}{g\left(\mathbf{X}\right)}\right]$$

Importance sampling estimator of I:

$$\widehat{I}_N^{\mathrm{IS}} = \frac{1}{N} \sum_{n=1}^{N} \phi\left(\mathbf{X}^{(n)}\right) \frac{f\left(\mathbf{X}^{(n)}\right)}{g\left(\mathbf{X}^{(n)}\right)} \text{ with } \left(\mathbf{X}^{(n)}\right)_{n \in [\![1,N]\!]} \sim g$$

Variance reduction methods

Choice of the auxiliary distribution g

Variance of $\widehat{I}_N^{\mathrm{IS}}$: $$\mathbb{V}_g\left(\widehat{I}_N^{\mathrm{IS}}\right) \propto \mathbb{V}_g\left(\phi(\mathbf{X})\frac{f_{\mathbf{X}}(\mathbf{X})}{g(\mathbf{X})}\right)$$

$\Longrightarrow \mathbb{V}_g\left(\widehat{I}_N^{\mathrm{IS}}\right)$ strongly depends on g

Variance reduction methods

Choice of the auxiliary distribution g

Variance of $\widehat{I}_N^{\text{IS}}$: $$\mathbb{V}_g\left(\widehat{I}_N^{\text{IS}}\right) \propto \mathbb{V}_g\left(\phi(\mathbf{X})\frac{f_{\mathbf{X}}(\mathbf{X})}{g(\mathbf{X})}\right)$$

$\Longrightarrow \mathbb{V}_g\left(\widehat{I}_N^{\text{IS}}\right)$ strongly depends on g

Optimal IS auxiliary distribution [5]: $$g_{\text{opt}}(\mathbf{x}) = \frac{\phi(\mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})}{I} \text{ when } \phi \geq 0$$

$\Longrightarrow$ leads to zero-variance estimator

$\Longrightarrow$ impossible to use because it depends on I

Variance reduction methods

Choice of the auxiliary distribution g

Variance of $\widehat{I}_N^{\text{IS}}$: $$\mathbb{V}_g\left(\widehat{I}_N^{\text{IS}}\right) \propto \mathbb{V}_g\left(\phi(\mathbf{X})\frac{f_{\mathbf{X}}(\mathbf{X})}{g(\mathbf{X})}\right)$$

$\Longrightarrow \mathbb{V}_g\left(\widehat{I}_N^{\text{IS}}\right)$ strongly depends on g

Optimal IS auxiliary distribution [5]: $$g_{\text{opt}}(\mathbf{x}) = \frac{\phi(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})}{I} \text{ when } \phi \geq 0$$

$\Longrightarrow$ leads to zero-variance estimator

$\Longrightarrow$ impossible to use because it depends on I

Existing approximation techniques:

- parametric importance sampling [6]
- non-parametric importance sampling [7]

Variance reduction methods

Method 2 : Control variates [8]

Principle of control variates

Exploiting the known value of the integral of a **control function** h and a **control parameter** $\beta \in \mathbb{R}$ to control the variance of the corresponding estimator.

Variance reduction methods

Method 2 : Control variates [8]

Principle of control variates

Exploiting the known value of the integral of a **control function** h and a **control parameter** $\beta \in \mathbb{R}$ to control the variance of the corresponding estimator.

Estimator of I with control variates and importance sampling [8] :

$$\widehat{I}^{\text{IS-CV}}_{N,g,h,\beta} = \frac{1}{N} \sum_{n=1}^{N} \frac{\phi\left(\mathbf{X}^{(n)}\right) f\left(\mathbf{X}^{(n)}\right) - \beta h\left(\mathbf{X}^{(n)}\right)}{g\left(\mathbf{X}^{(n)}\right)} + \beta \int_{\mathbb{X}} h(\mathbf{x})\, d\mathbf{x}$$

where $\left(\mathbf{X}^{(n)}\right)_{n \in [\![1,N]\!]} \sim g$ is an unbiased estimator of $I = \mathbb{E}_f\left(\phi\left(\mathbf{X}\right)\right)$.

Principle of control variates

Exploiting the known value of the integral of a **control function** h and a **control parameter** $\beta \in \mathbb{R}$ to control the variance of the corresponding estimator.

Minimisation of the variance of the estimator according to β:

$$\beta^* = \underset{\beta \in \mathbb{R}}{\arg\min} \mathbb{V}_g \left(\widehat{I}^{\text{IS-CV}}_{N,g,h,\beta} \right) = \mathbb{V}_g \left(\frac{h(\mathbf{X})}{g(\mathbf{X})} \right)^{-1} \mathrm{Cov}_g \left(\frac{\phi(\mathbf{X}) f(\mathbf{X})}{g(\mathbf{X})}, \frac{h(\mathbf{X})}{g(\mathbf{X})} \right)$$

Then, $\mathbb{V}\left(\widehat{I}^{\text{IS-CV}}_{N,g,h,\beta^*} \right) \leq \mathbb{V}\left(\widehat{I}^{\text{IS}}_{N,g} \right)$.

Estimation of multiple expectations

Optimal estimator family

Theorem [9]

The family of estimator

$$\widehat{I}^{\text{IS-CV}}_{N,g_{\alpha},g_j^*,\beta^*} = \frac{1}{N}\sum_{n=1}^{N} \frac{\phi_j\left(\mathbf{X}^{(n)}\right) f_j\left(\mathbf{X}^{(n)}\right) - \beta_j^* g_j^*\left(\mathbf{X}^{(n)}\right)}{g_{\alpha}\left(\mathbf{X}^{(n)}\right)} + \beta_j^*$$

with $\left(\mathbf{X}^{(n)}\right)_{n\in[\![1,N]\!]} \sim g_{\alpha} = \sum_{j=1}^{J} \alpha_j g_j^*$ and $\beta_j^* = I_j$, satisfies

$$\sum_{j=1}^{J} w_j \mathbb{V}\left(\widehat{I}^{\text{IS-CV}}_{N,g_{\alpha},g_j^*,\beta_j^*}\right) = 0.$$

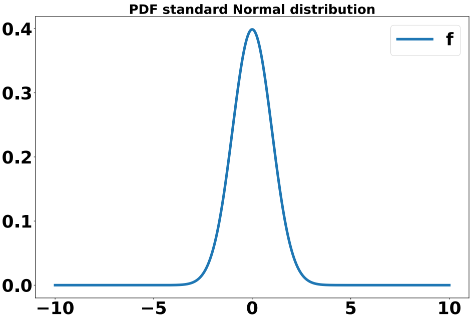
These optimal estimators cannot be used in practice
$\Longrightarrow$ require a procedure to approach these optimal estimators: **ME-aISCV** algorithm [9]

ME-aISCV algorithm [9]

Example

Estimation of the first $J = 10$ even moments of the 1-dimensional standard Normal distribution $\mathcal{N}_1(0, 1)$:

$$\forall j \in [\![1, J]\!], \begin{cases} f_j & = & f \equiv \mathcal{N}_1(0,1) \\ I_j & = & \mathbb{E}_f\left(X^{2j}\right) \end{cases}$$



e

ME-aISCV algorithm [9]

Description of the algorithm

Initialisation $k = 0$
Initial MC sample

Initialisation k = 0:

- Draw an initial MC sample $\left(\mathbf{X}^{(0,n)}\right)_{n \in [\![1, N_0]\!]} \sim f$ and estimate :

$$\widehat{I}_j^{(0)} = \frac{1}{N} \sum_{n=1}^{N} \phi_j \left(\mathbf{X}^{(0,n)} \right)$$

- Initialisation of the coefficients $\boldsymbol{\alpha}^{(0)}$ and $\beta_j^{(0)}$

ME-aISCV algorithm [9]

Description of the algorithm

Initialisation $k = 0$
Initial MC sample

$\downarrow k = 1$

Cross-entropy: Approximation of g_j^* by parametric densities $g_{\boldsymbol{\lambda}_j^{(k)}}$

Cross-entropy:

- Update $k = k + 1$
- Approaching the optimal distributions g_j^* within a parametric family of distributions $g_{\boldsymbol{\lambda}_j^{(k)}}$ by solving the optimisation problem:

$$\boldsymbol{\lambda}_j^{(k)} = \arg\min_{\lambda_j} D_{\mathrm{KL}}\left(g_j^*, g_{\boldsymbol{\lambda}_j}\right)$$

In practice, $g_{\boldsymbol{\lambda}_j}$ is picked up in the Gaussian family or in the Gaussian mixture family.

ME-aISCV algorithm [9]

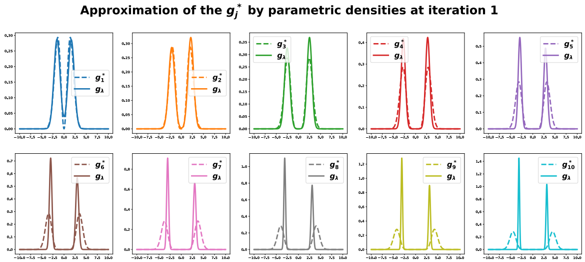
Description of the algorithm

Initialisation $k = 0$
Initial MC sample

$k = 1$

Cross-entropy: Approximation of g_j^* by parametric densities $g_{\boldsymbol{\lambda}_j^{(k)}}$

Cross-entropy:



ME-aISCV algorithm [9]

Description of the algorithm

Initialisation $k = 0$
Initial MC sample

$\downarrow$ $k=1$

Cross-entropy: Approximation of g_j^* by parametric densities $g_{\boldsymbol{\lambda}_j^{(k)}}$

$\downarrow$

Mixture: Update of $\boldsymbol{\alpha}^{(k)}$
Draw a new sample according to $g_{\boldsymbol{\alpha}^{(k)}}$

Mixture:

- Update the weights $\boldsymbol{\alpha}$ and create the new mixture $g_{\boldsymbol{\alpha}^{(k)}} = \sum_{j=1}^{J} \alpha_j^{(k)} g_{\boldsymbol{\lambda}_j^{(k)}}$ by solving the optimisation problem:

$$\boldsymbol{\alpha}^{(k)} = \underset{\sum_{j=1}^{J} \alpha_j = 1}{\arg\min} \sum_{j=1}^{J} w_j \mathbb{V}_{g_{\boldsymbol{\alpha}}} \left(\frac{\phi_j(\mathbf{X}) f_j(\mathbf{X}) - \beta_j^{(k-1)} g_{\boldsymbol{\lambda}_j^{(k)}}(\mathbf{X})}{g_{\boldsymbol{\alpha}^{(k)}}(\mathbf{X})} \right)$$

- Draw a new sample $\left(\mathbf{X}^{(k,n)}\right)_{n \in [\![1, N_k]\!]} \sim g_{\boldsymbol{\alpha}^{(k)}}$

ME-aISCV algorithm [9]

Description of the algorithm

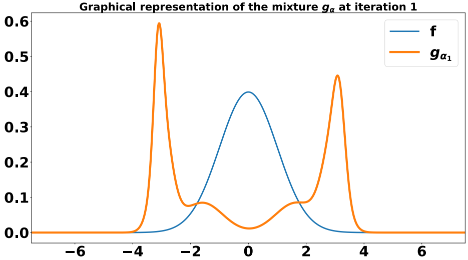
Initialisation $k = 0$
Initial MC sample

$k = 1$

Cross-entropy: Approximation of g_j^* by parametric densities $g_{\boldsymbol{\lambda}_j^{(k)}}$

Mixture: Update of $\boldsymbol{\alpha}^{(k)}$
Draw a new sample according to $g_{\boldsymbol{\alpha}^{(k)}}$

Mixture:



ME-aISCV algorithm [9]

Description of the algorithm

Initialisation $k = 0$
Initial MC sample

↓ $k = 1$

Cross-entropy: Approximation of g_j^* by parametric densities $g_{\boldsymbol{\lambda}_j^{(k)}}$

↓

Mixture: Update of $\boldsymbol{\alpha}^{(k)}$
Draw a new sample according to $g_{\boldsymbol{\alpha}^{(k)}}$

↓

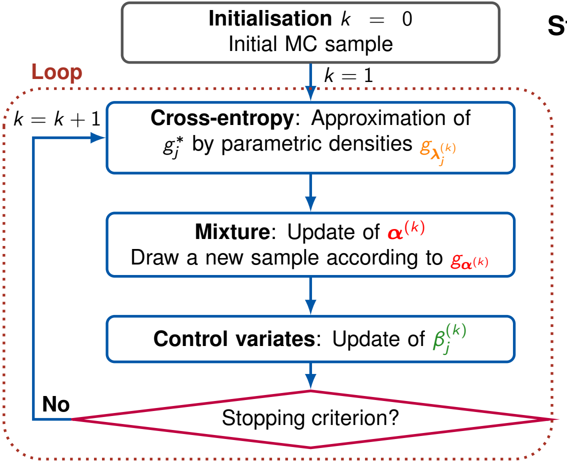
Control variates: Update of $\beta_j^{(k)}$

Control variates:
Update the control parameters β_j by estimating:

$$\beta_j^{(k)} \approx \frac{\mathrm{Cov}_{g_{\boldsymbol{\alpha}^{(k)}}}\left(\frac{\phi_j(\mathbf{X})\, f_j(\mathbf{X})}{g_{\boldsymbol{\alpha}^{(k)}}(\mathbf{X})}, \frac{g_{\boldsymbol{\lambda}_j^{(k)}}(\mathbf{X})}{g_{\boldsymbol{\alpha}^{(k)}}(\mathbf{X})}\right)}{\mathbb{V}_{g_{\boldsymbol{\alpha}^{(k)}}}\left(\frac{g_{\boldsymbol{\lambda}_j^{(k)}}(\mathbf{X})}{g_{\boldsymbol{\alpha}^{(k)}}(\mathbf{X})}\right)}$$

ME-aISCV algorithm [9]

Description of the algorithm



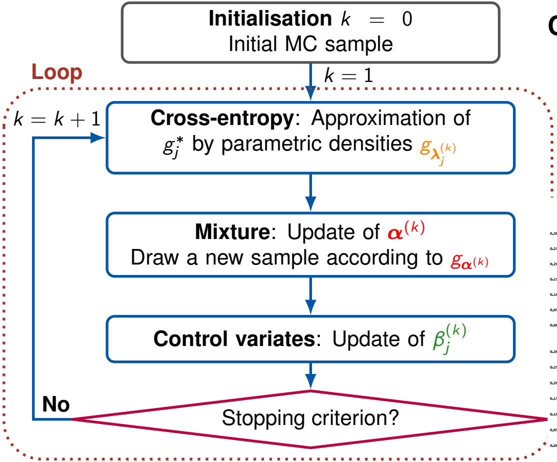
Stopping criterion:

Quality of the approximation of the optimal parameters λ_j, $\boldsymbol{\alpha}$ and β_j

Budget left N_f for the final estimates

ME-aISCV algorithm [9]

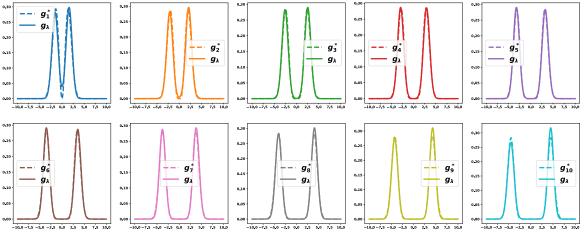
Description of the algorithm



Cross-entropy:

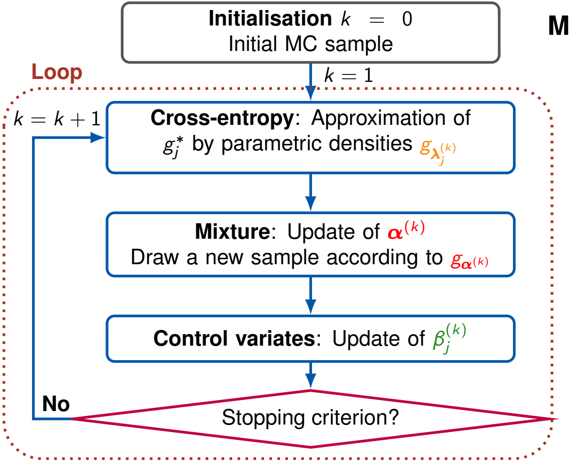
- Update $k = k + 1$
- Approaching the optimal distributions g_j^* by parametric densities $g_{\boldsymbol{\lambda}_j^{(k)}}$

Approximation of the g_j^* by parametric densities at iteration 4



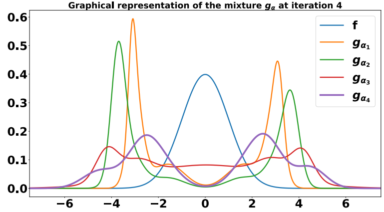
ME-aISCV algorithm [9]

Description of the algorithm



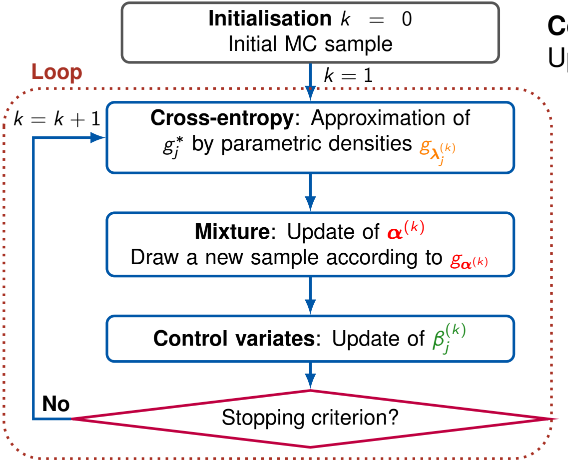
Mixture:

- Update the weights $\boldsymbol{\alpha}$ and create the new mixture $g_{\boldsymbol{\alpha}^{(k)}} = \sum_{j=1}^{J} \alpha_j^{(k)} g_{\boldsymbol{\lambda}_j^{(k)}}$



- Draw a new sample $\left(\mathbf{X}^{(k,n)}\right)_{n \in [\![1, N_k]\!]} \sim g_{\boldsymbol{\alpha}^{(k)}}$

ME-aISCV algorithm [9]



Control variates:

Update the control parameters β_j by estimating:

$$\beta_j^{(k)} \approx \frac{\mathrm{Cov}_{g_{\boldsymbol{\alpha}^{(k)}}}\left(\frac{\phi_j(\mathbf{X})\, f_j(\mathbf{X})}{g_{\boldsymbol{\alpha}^{(k)}}(\mathbf{X})}, \frac{g_{\boldsymbol{\lambda}_j^{(k)}}(\mathbf{X})}{g_{\boldsymbol{\alpha}^{(k)}}(\mathbf{X})}\right)}{\mathbb{V}_{g_{\boldsymbol{\alpha}^{(k)}}}\left(\frac{g_{\boldsymbol{\lambda}_j^{(k)}}(\mathbf{X})}{g_{\boldsymbol{\alpha}^{(k)}}(\mathbf{X})}\right)}$$

ME-aISCV algorithm [9]

Initialisation $k = 0$
Initial MC sample

$k = 1$

Loop

$k = k + 1$

Cross-entropy: Approximation of g_j^* by parametric densities $g_{\boldsymbol{\lambda}_j^{(k)}}$

Mixture: Update of $\boldsymbol{\alpha}^{(k)}$
Draw a new sample according to $g_{\boldsymbol{\alpha}^{(k)}}$

Control variates: Update of $\beta_j^{(k)}$

Stopping criterion?

No

Yes

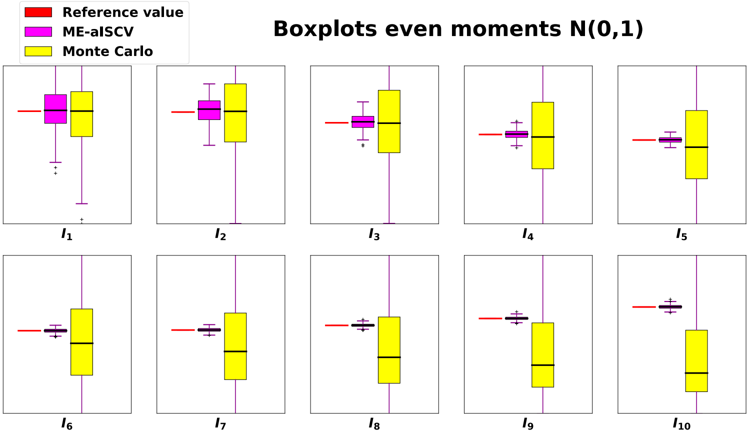
End : Draw a new sample according to $g_{\boldsymbol{\alpha}^{(k)}}$
Final estimates $\widehat{I}^{\text{IS-CV}}_{N, g_{\boldsymbol{\alpha}^{(k)}}, g_{\boldsymbol{\lambda}_j^{(k)}}, \beta_j^{(k)}}$

End :
Draw a new sample independent from the previous ones $\left(\mathbf{X}^{(n)}\right)_{n \in [\![1, N_f]\!]} \sim g_{\boldsymbol{\alpha}^{(k)}}$ and return:

$$\widehat{I}^{\text{IS-CV}}_{N_f, g_{\boldsymbol{\alpha}^{(k)}}, g_{\boldsymbol{\lambda}_j^{(k)}}, \beta_j^{(k)}} = \frac{1}{N_f} \sum_{n=1}^{N_f} \frac{\phi_j\left(\mathbf{X}^{(n)}\right) f_j\left(\mathbf{X}^{(n)}\right) - \beta_j^{(k)} g_{\boldsymbol{\lambda}_j^{(k)}}\left(\mathbf{X}^{(n)}\right)}{g_{\boldsymbol{\alpha}^{(k)}}\left(\mathbf{X}^{(n)}\right)} + \beta_j^{(k)}$$

ME-aISCV algorithm [9]

Results on the example



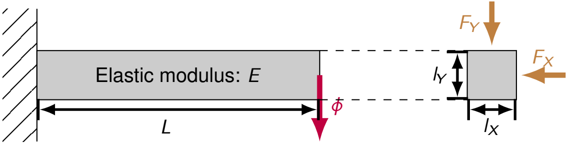
Results

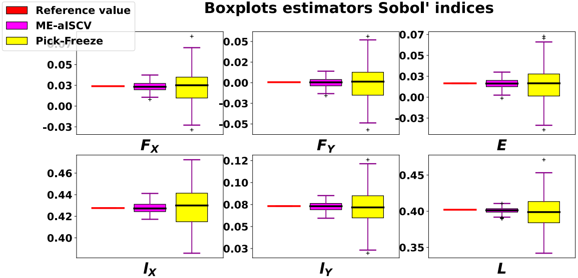
Estimation of the first order Sobol' indices for the cantilever beam problem

	Input	Distribution	Mean	C.o.V.
1	F_X	LogNormal	556.8 N	0.08
2	F_Y	LogNormal	453.6 N	0.08
3	E	LogNormal	200.10^9 Pa	0.06
4	l_X	Normal	0.062 m	0.1
5	l_Y	Normal	0.0987 m	0.1
6	L	Normal	4.29 m	0.1

Maximal displacement of the tip section:

$$\phi\,(F_X, F_Y, E, l_X, l_Y, L) = \frac{4L^3}{E\, l_X\, l_Y}\sqrt{\left(\frac{F_X}{l_X^2}\right)^2 + \left(\frac{F_Y}{l_Y^2}\right)^2}$$





Results

Epistemic uncertainty on the input distribution for the cantilever beam problem

	Input	Distribution	Mean	C.o.V.
1	F_X	LogNormal	m_1	0.08
2	F_Y	LogNormal	m_2	0.08
3	E	LogNormal	m_3	0.06
4	I_X	Normal	m_4	0.1
5	I_Y	Normal	m_5	0.1
6	L	Normal	m_6	0.1

Correlation between some inputs:

$$\rho_{I_X,I_Y} = m_7 \text{ and } \rho_{L,I_X} = m_8 \text{ and } \rho_{L,I_Y} = m_9.$$

Maximal displacement of the tip section:

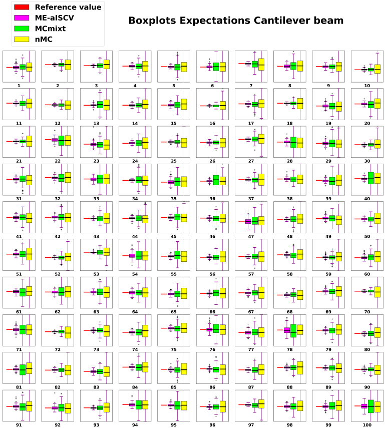
$$\phi(F_X, F_Y, E, I_X, I_Y, L) = \frac{4L^3}{E I_X I_Y}\sqrt{\left(\frac{F_X}{I_X^2}\right)^2 + \left(\frac{F_Y}{I_Y^2}\right)^2}$$

Goal of the test case: commonly estimate the family of expectations $\left(\mathbb{E}_{f_{\boldsymbol{m}^{(j)}}}\left(\phi\left(\mathbf{X}\right)\right)\right)_{j\in[\![1,J]\!]}$ where

- $J = 100$ (!)
- each component of the vector $\boldsymbol{m}^{(j)} = (m_i)_{i\in[\![1,9]\!]}$ is drawn according to a Uniform distribution centered on the "true" value.

Results

Epistemic uncertainty on the input distribution for the cantilever beam problem



Final results:

	ME-aISCV	MC	MC mixture
$\sum_{j=1}^{100} \mathbb{V}\left(\widehat{I_j}\right)$	$\mathbf{4.379 \times 10^{-6}}$	1.309×10^{-4}	6.103×10^{-5}

Conclusion

What is new?

- introduction of the problem of the estimation of multiple expectations with the same N-sample
- definition of a criterion to quantify the quality of the common estimation
- ME-aISCV algorithm to reduce as much as possible the previous criterion

Accepted paper [9] and codes to reproduce the results are available online !

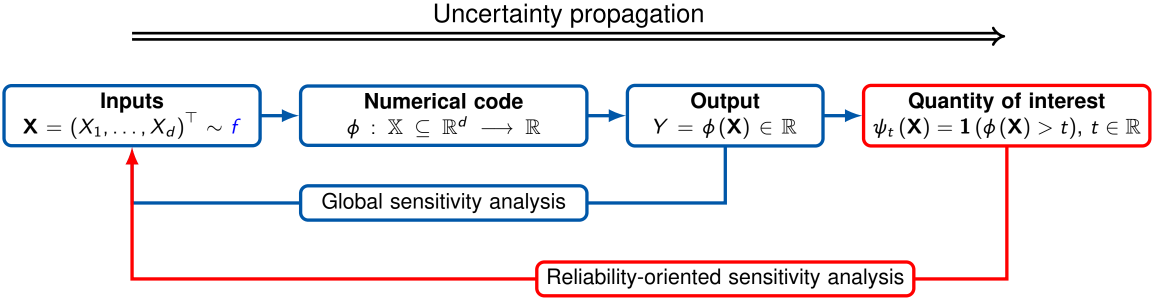
Efficient estimation of multiple expectations with the same sample by adaptive importance sampling and control variates

Julien DEMANGE-CHRYST[a,b,*], François BACHOC[b], Jérôme MORIO[a]

[a] *ONERA/DTIS, Université de Toulouse, F-31055 Toulouse, France*
[b] *Institut de Mathématiques de Toulouse, UMR5219 CNRS, 31062 Toulouse, France*

Estimation of the target Shapley effects by IS

Uncertainty quantification and reliability analysis



Question

With correlated inputs, which components of $\mathbf{X}$ are most likely to lead to the failure of the system ?

Estimation of the target Shapley effects by IS

Importance sampling and the use of OpenTURNS

Reliability-oriented Shapley effects

$$\text{T-Sh}_i = \frac{1}{d}\frac{1}{\mathbb{V}\left(\psi_t\left(\mathbf{X}\right)\right)} \sum_{u\subseteq\{-i\}} \binom{d-1}{|u|}^{-1} \left(\text{T-S}^c_{u\cup\{i\}} - \text{T-S}^c_u\right) \text{ with } \text{T-S}^c_u = \begin{cases} \mathbb{V}\left[\mathbb{E}\left(\psi_t\left(\mathbf{X}\right)|\mathbf{X}_u\right)\right] \text{ or} \\ \mathbb{E}\left[\mathbb{V}\left(\psi_t\left(\mathbf{X}\right)|\mathbf{X}_{-u}\right)\right] \end{cases}$$

Estimation process of T-Sh$_i$:

1. estimation of T-S^c_u for $u \subseteq [\![1,d]\!]$ $\Longrightarrow$ IS here
2. aggregation procedure

Estimation of the T-S^c_u with importance sampling:

- requires the likelihood ratios f/g, but also the marginal likelihood ratios $f_{\mathbf{X}_u}/g_{\mathbf{X}_u}$ for every subset $u \subseteq [\![1,d]\!]$
- OpenTURNS allows me to quickly compute all these ratios, even for complex distributions f or g !

Published article [10] and codes to reproduce the results are available online !

References

[1] I. M. Sobol,
Sensitivity analysis for non-linear mathematical models,
Mathematical modelling and computational experiment, 1, pp. 407–414 (1993).

[2] T. Homma and A. Saltelli,
Importance measures in global sensitivity analysis of nonlinear models,
Reliability Engineering & System Safety, 52 (1), pp. 1–17 (1996).

[3] F. Gamboa, P. Gremaud, T. Klein and A. Lagnoux,
Global sensitivity analysis: A novel generation of mighty estimators based on rank statistics,
Bernoulli, 28 (4), pp. 2345–2374 (2022).

[4] B. Sudret,
Global sensitivity analysis using polynomial chaos expansions,
Reliability engineering & system safety, 93 (7), pp. 964–979 (2008).

[5] J. Bucklew,
Introduction to rare event simulation,
Springer Science & Business Media (2004).

[6] R. Y. Rubinstein and D. P. Kroese,
The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation, and machine learning,
Vol. 133. Springer (2004).

[7] P. Zhang,
Nonparametric importance sampling,
Journal of the American Statistical Association, 91 (435), pp. 1245–1253 (1996).

[8] H. Y. He and A. B. Owen,
Optimal mixture weights in multiple importance sampling,
arXiv preprint arXiv:1411.3954 (2014).

[9] J. Demange-Chryst, F. Bachoc and J. Morio,
Efficient estimation of multiple expectations with the same sample by adaptive importance sampling and control variates,
arXiv preprint arXiv:2212.00568 (2022).

[10] J. Demange-Chryst, F. Bachoc and J. Morio,
Shapley effect estimation in reliability-oriented sensitivity analysis with correlated inputs by importance sampling,
International Journal for Uncertainty Quantification, 13 (3) (2023).