An introduction to optimization under uncertainties

CIMPA, Ulaanbaatar, Mongolia

Didier Lucor

didier.lucor@lisn.upsaclay.fr

July 10th 2021

LABORATOIRE INTERDISCIPLINAIRE DES SCIENCES DU NUMÉRIQUE

Course III

Organization of the class

1. Course I : Concepts, formalism and some classes of resolution methods for simple optimization under uncertainty (1.5h, Thursday)

2. Course II : quantification of uncertainty for more efficient optimization under uncertainties : metamodeling with Gaussian Processes (1.5h, Friday)

3. Course III : Metamodeling-based Reliability-Based Design Optimization (1.5h, Saturday)
   - Bayesian optimization
   - Bayesian optimization under uncertainties
   - Surrogate modeling deployment in the context of RBDO

Motivation for metamodeling in OUU

Recalling the example of Reliability-Based Design type of optimization (**RBDO**) :

$$\boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}\in\boldsymbol{\mathcal{X}}} \left[f(\boldsymbol{x}, \overline{\boldsymbol{U}}) \right],$$

$$\text{such that } \mathbb{P}\left[g_i(\boldsymbol{x}, \boldsymbol{U}) > 0\right] \leq P_{f_i}, \quad \text{for } i \in \{1, \dots, m\}$$

$$h_j(\boldsymbol{x}) \leq 0, \quad \text{for } j \in \{m+1, \dots, M\}$$

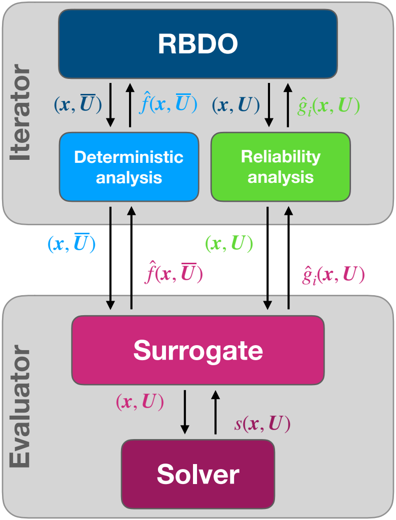
computational cost : evaluation of **expensive functionals** (e.g. constraints g_i) may be very time consuming (call to expensive black-box solvers, e.g. FE, FV solvers) + computation of **statistical estimators** (e.g. probability of failure (PoF) : needs many evaluations 10^{5-8})

e.g. MPP search, sampling methods, outer optimization loop led by GA optimizators

functional complexity : **high dimensionality, nonlinearity**, non-differentiability, non-convex and non-connex domains of failure...

choice of spaces : both design/other parameters may contain deterministic/random quantities... what is the best space over which to build our surrogate ?

Surrogate-based RBDO flowchart



Different (types) of surrogate models (SM) can help

various SM (one or many !) may be used for various terms, e.g. approximating : – the objective function, – **the constraints** with the idea of approximating the full solver with less information/data than needed/provided by the full solver.

- **data fit SM** (local/intermediate/**global** approximation **scale**) ; e.g. polynomial response surfaces, polynomial chaos expansions, SVM, neural networks all relate in part on the **Design of Experiments** (**DoE**)
- multifidelity SM
- Reduced-Order Model (ROM) SM

Once we hold a "good" SM we can use it at "no-cost" to generate **large approximation samples** (e.g. to estimate probabilities of failure), determine some **bounds**, get some insights about **parametric sensitivity**, etc...

When SM nested within an optimization loop, question of whether – building a **new SM every iteration** or a – **single global SM** in general a single **global** surrogate model is constructed for the optimization instead of a bunch of independent models

Two strategies to build SM in the context of OUU

either 1. the SM is built **prior to the optimization process** and is used but **not updated** during the optimization

or 2. the SM is **refined at each iteration** of the optimization process

in any case many variants exist as there are many different optimization processes, with multisteps, etc.

moreover, **the refinement criteria must account for the accuracy** of the SM – **so it is VERY useful if the latter also bears a measure of confidence** – **in relation to the optimization process !**

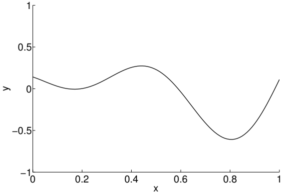
Iterative adaptation of the surrogate : an active learning method

optimization : **iteratively** searches for optimum under a **limited budget of simulations**

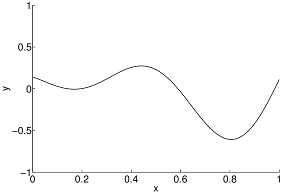
$\rightarrow$ so it makes sense to **adapt** the **(DoE, SM)** as we progress, e.g. increasing its accuracy in important regions

$\Rightarrow$ importance of initial DoE :
$\mathcal{S}_s = \left\{ (\boldsymbol{x}^{(i)}, \boldsymbol{U}^{(i)}, y^{(i)}) \in \mathcal{X} \times \mathcal{A} \times \mathbb{R} \mid y^{(i)} = s(\boldsymbol{x}^{(i)}, \boldsymbol{U}^{(i)}), i = 1, \dots, n \right\}$

$\Rightarrow$ each new simulation must address a compromise btw **exploration** (enriching in $\boldsymbol{x}$ in *new* regions), or **exploitation** (enriching in $\boldsymbol{U}$ in *known* good regions)



(a) Exploitation



(b) Exploration

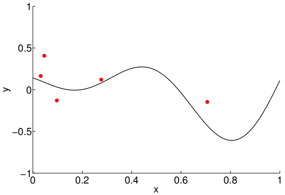
FIGURE: e.g. of objective function to minimize $\boldsymbol{x} \mapsto \boldsymbol{y}(\boldsymbol{x}; z = 0)$

Iterative adaptation of the surrogate : an active learning method

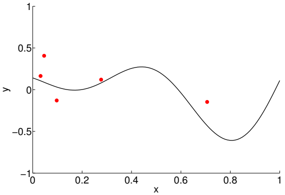
optimization : **iteratively** searches for optimum under a **limited budget of simulations**

$\rightarrow$ so it makes sense to **adapt** the **(DoE, SM)** as we progress, e.g. increasing its accuracy in important regions

- $\Rightarrow$ importance of initial DoE :
  $\mathcal{S}_s = \left\{ (\boldsymbol{x}^{(i)}, \boldsymbol{U}^{(i)}, y^{(i)}) \in \mathcal{X} \times \mathcal{A} \times \mathbb{R} \mid y^{(i)} = s(\boldsymbol{x}^{(i)}, \boldsymbol{U}^{(i)}), i = 1, \ldots, n \right\}$
- $\Rightarrow$ each new simulation must address a compromise btw **exploration** (enriching in x in *new* regions), or **exploitation** (enriching in $\boldsymbol{U}$ in *known* good regions)



(a) Exploitation



(b) Exploration

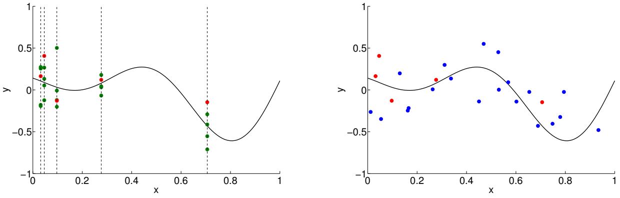
FIGURE: Information : 5 noisy observations of the function.

Iterative adaptation of the surrogate : an active learning method

optimization : **iteratively** searches for optimum under a **limited budget of simulations**

$\rightarrow$ so it makes sense to **adapt** the **(DoE, SM)** as we progress, e.g. increasing its accuracy in important regions

- $\Rightarrow$ importance of initial DoE : $\mathcal{S}_s = \left\{ (\boldsymbol{x}^{(i)}, \boldsymbol{U}^{(i)}, y^{(i)}) \in \mathcal{X} \times \mathcal{A} \times \mathbb{R} \mid y^{(i)} = s(\boldsymbol{x}^{(i)}, \boldsymbol{U}^{(i)}), i = 1, \ldots, n \right\}$
- $\Rightarrow$ each new simulation must address a compromise btw **exploration** (enriching in x in *new* regions), or **exploitation** (enriching in $\boldsymbol{U}$ in *known* good regions)



(a) Exploitation (b) Exploration

FIGURE: Different choices for new sampling

Sampling for better design or to reduce uncertainties of current design

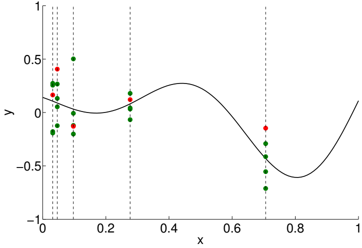
why repeating a simulation for the same design in optimization ?

Theory, bad idea : information gain seems to be no better than the previous simulation at the same point...

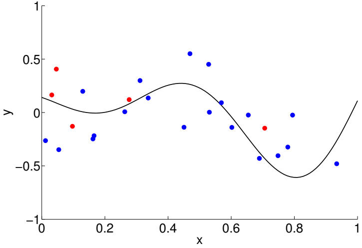
In practice, by capitalizing the knowledge about the noise/uncertainty, chance to better guide future exploration (long term benefit)

$\Rightarrow$ for instance, one compromise that maybe efficient is to slowly increase the number of exploitations during optimization

getting back to previous example...



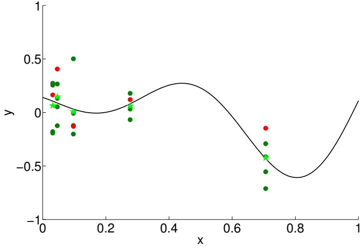
(a) Exploitation



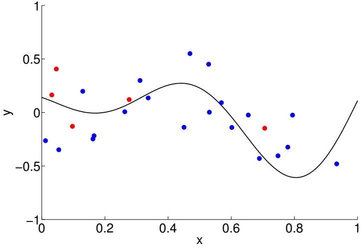
(b) Exploration

FIGURE: Initial information.

getting back to previous example...



(a) Duplication



(b) Exploration

FIGURE: Condensing of exploitation information to guide further search (here we take the mean : green stars).

getting back to previous example...

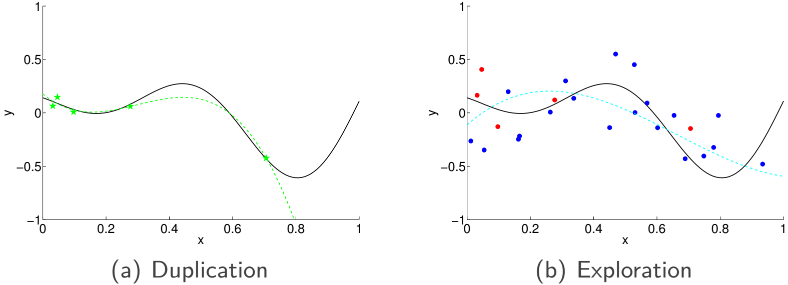


FIGURE: Polynomial predictions for new simulations.

Table of contents

1 Course I : Concepts, formalism and some classes of resolution methods for simple optimization under uncertainty (1.5h, Thursday)

2 Course II : quantification of uncertainty for more efficient optimization under uncertainties : metamodeling with Gaussian Processes (1.5h, Friday)

3 Course III : Metamodeling-based Reliability-Based Design Optimization (1.5h, Saturday)

- Bayesian optimization
- Bayesian optimization under uncertainties
- Surrogate modeling deployment in the context of RBDO

Bayesian optimization

Unconstrained deterministic optimization problem

$$x^\star = \arg \min_{x \in \mathcal{X} \subset \mathbb{R}^d} f(x, \overline{U}),$$
$$x_{\text{min}} \leq x \leq x_{\text{max}}$$

Assumptions

objective function $s \equiv f$ is a **"black-box"** solver : ie **lacks known special structure** like concavity or linearity that would make it easy to optimize using techniques that leverage such structure to improve efficiency

f is **expensive** to evaluate

only $f(x)$ is observed : **"derivative-free"** (no first- or second-order derivatives : ie no methods like gradient descent, Newton's method)

Leverage

take advantage of GP capabilities to **sequentially refine** our metamodel

Recall : GP regression (Kriging)

based on the knowledge of the **current DoE** of size n :
$\mathcal{S}_s = \left\{(\boldsymbol{x}^{(i)}, \overline{\boldsymbol{U}}, y^{(i)}) \in \mathcal{X} \times \mathcal{A} \times \mathbb{R} \mid y^{(i)} = s(\boldsymbol{x}^{(i)}, \overline{\boldsymbol{U}}, i = 1, \dots, n\right\}$, our function is approximated as a realization of the following **conditioned** stochastic process :

$$s_{\mathcal{SCGP}}(\boldsymbol{x}) = s(\boldsymbol{x}) \mid s(\boldsymbol{X}) = \boldsymbol{y},$$

which is a GP $\sim \mathcal{N}(\hat{\mu}(\boldsymbol{x}), \hat{\sigma}^2(\boldsymbol{x}))$, with :

$$\widehat{\mu}(\boldsymbol{x}) = \boldsymbol{f}^T(\boldsymbol{x})\widehat{\boldsymbol{\beta}} + \boldsymbol{r}(\boldsymbol{x})^T R^{-1}(\boldsymbol{y} - \boldsymbol{F}\widehat{\boldsymbol{\beta}}),$$

$$\widehat{\sigma}^2(\boldsymbol{x}) = C(\boldsymbol{x}, \boldsymbol{x}) - \boldsymbol{r}(\boldsymbol{x})^T R^{-1}\boldsymbol{r}(\boldsymbol{x}) + \boldsymbol{u}(\boldsymbol{x})^T \left(\boldsymbol{F}^T R^{-1} \boldsymbol{F}\right)^{-1} \boldsymbol{u}(\boldsymbol{x}),$$

Notations recall

$$\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{F}^T R^{-1} \boldsymbol{F}\right)^{-1} \boldsymbol{F}^T R^{-1} \boldsymbol{y}, \quad \boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{F}^T R^{-1} \boldsymbol{r}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}),$$

with $\widehat{\beta}$ is the **maximum likelihood** approximation of β

Infill criteria (UNCONSTRAINED optim)

Probability Improvement (PI) : new sample at $\boldsymbol{x}^{n+1} = \arg\max_{\boldsymbol{x} \in \mathcal{X}} \mathrm{PI}(\boldsymbol{x})$

$$\mathrm{PI}(\boldsymbol{x}) = \Phi\left(\boldsymbol{w}\right),$$

Expected Improvement (EI) : new sample at $\boldsymbol{x}^{n+1} = \arg\max_{\boldsymbol{x} \in \mathcal{X}} \mathrm{EI}(\boldsymbol{x})$

$I(\boldsymbol{x}) = \max\left(0, \min(\boldsymbol{y}) - \hat{s}(\boldsymbol{x})\right)$ with $\mathrm{PI}(\boldsymbol{x}) = \int_{\mathbb{R}} I(\boldsymbol{x}) d\hat{s}(\boldsymbol{x})$

$$\mathrm{PI}(\boldsymbol{x}) = \left(\min\left(\boldsymbol{y}\right) - \hat{s}(\boldsymbol{x})\right)\Phi\left(\boldsymbol{w}\right) + \hat{\sigma}(\boldsymbol{x})\phi(\boldsymbol{w}),$$

with $\boldsymbol{w} = \frac{\min\left(\boldsymbol{y}\right) - \hat{s}(\boldsymbol{x})}{\hat{\sigma}(\boldsymbol{x})}$, $\Phi \equiv \mathrm{CDF}_{\mathcal{N}}$ and $\phi \equiv \mathrm{PDF}_{\mathcal{N}}$

Efficient Global Optimization (EGO) [Jones et al. 1998]

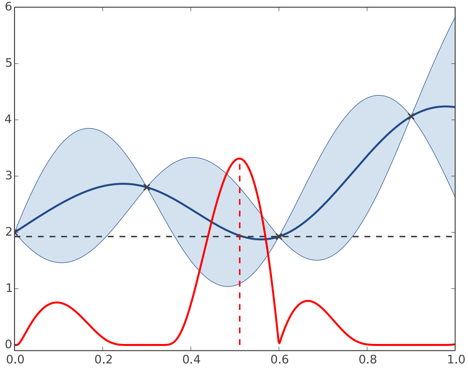
1. built a first GP for the QoI (e.g. objective function) from a **scarcely sampled** DoE of size n (e.g. max. log-likelihood on σ and correlation lengths of the SP)
2. $\boldsymbol{x}^{n+1} = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mathsf{EI}(\boldsymbol{x})$ (in high dimension with another optimizer, e.g. CMA-ES) [Hansen and Ostermeier, 2001]
3. run the expensive solver $y^{n+1} = s(\boldsymbol{x}^{n+1})$ and update the DoE : $n \leftarrow n+1$
4. stop when $(n > N)$ or goto 2

EGO : a **good trade-off btw exploitation/exploration** without arbitrary ad-hoc procedure ; very efficient when combined with **EI criteria**

efficiency : it requires **few function calls** to get close to optima, but the efficiency comes from the order in which points are sampled

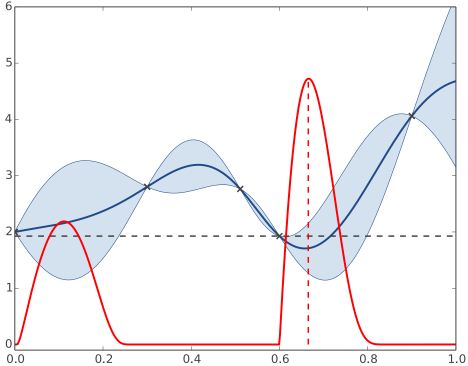
does not converge in the traditional sense : it creates dense samples in the parametric space

EGO example



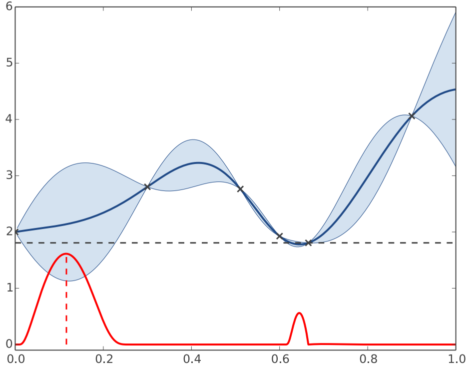
(a)

EGO example



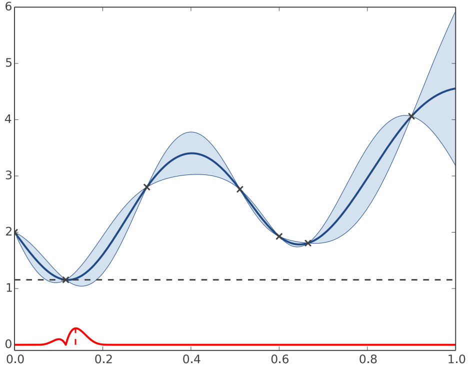
(b)

EGO example



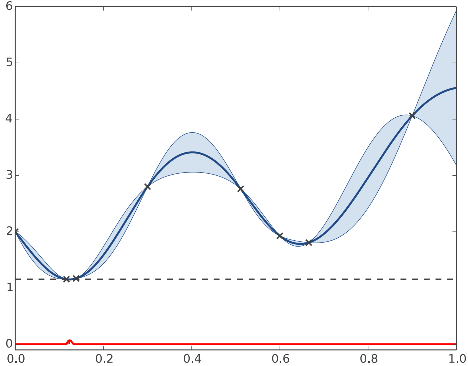
(c)

EGO example



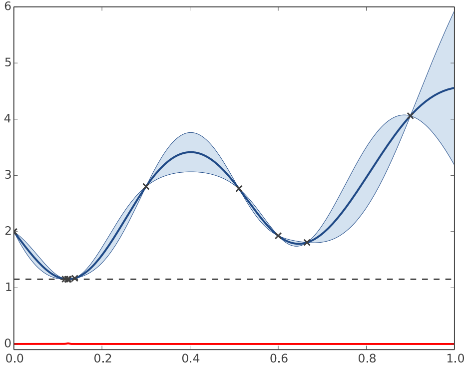
(d)

EGO example



(e)

EGO example



(f)

Constrained deterministic optimization problem

$$\boldsymbol{x}^{\star} = \arg \min_{\boldsymbol{x} \in \boldsymbol{\mathcal{X}} \subset \mathbb{R}^{\boldsymbol{d}}} f(\boldsymbol{x}, \overline{\boldsymbol{U}}),$$

$$\text{such that } g_i(\boldsymbol{x}, \overline{\boldsymbol{U}}) \leq 0, \quad \text{for } i \in \{1, \dots, m\}$$

$$h_j(\boldsymbol{x}) \leq 0, \quad \text{for } j \in \{m+1, \dots, M\}$$

Infill criteria (CONSTRAINED optim) each constraint is approximated by SCGP surrogate.

Probability of Feasibility (PF) [Schonlau et al. 1998, Parr et al. 2012] :

$$\mathsf{EI}(\boldsymbol{x}) \times \mathsf{PF}(\boldsymbol{x}) = \mathsf{EI}(\boldsymbol{x}) \prod_{i=1}^{m} \mathsf{PF}_i(\boldsymbol{x}) \text{ with } \mathsf{PF}_i = \Phi\left(-\boldsymbol{w}\right) \text{ and } \boldsymbol{w} = \frac{\hat{g}_i(\boldsymbol{x})}{\hat{\sigma}(\boldsymbol{x})}$$

Constrained EI [Sasena et al., 2001], solving the EI optimization as a *constrained auxiliary* optimization problem :

$$\begin{aligned} &\arg\max_{\boldsymbol{x}\in\boldsymbol{\mathcal{X}}} \mathsf{EI}(\boldsymbol{x}), \\ &\text{such that } \hat{g}_i(\boldsymbol{x}) \leq 0, \end{aligned}$$

Expected Violation (EV) [Audet et al., 2000], different *constrained auxiliary* optimization :

$$\begin{aligned} &\arg\max_{\boldsymbol{x}\in\boldsymbol{\mathcal{X}}} \mathsf{EI}(\boldsymbol{x}), \\ &\text{such that } \mathsf{EV}_i(\boldsymbol{x}) \geq t_{EV}, \text{ for } i \in \{1,\dots,m\} \text{ with} \\ &\mathsf{EV}_i(\boldsymbol{x}) = -\hat{g}_i(\boldsymbol{x})\mathsf{PF}_i(\boldsymbol{x}) + \hat{\sigma}_{g_i}(\boldsymbol{x})\phi(\boldsymbol{w}) \end{aligned} \quad (1)$$

Expected improvement for contour approximation [Ranjan et al., 2008]

AK-MCS algorithm

In **reliability** (ie. rare event estimation), we wish to evaluate the PoF $\rightarrow$ learning emphasis is put on **vicinity of the limit-state surface**

1. Generation of a **large MC population** $\mathcal{S}$ in the design space
2. Definition of the **initial DoE**(random selection)
3. construction of the **Kriging model** used for : $\hat{g}_i$ prediction + **PoF estimation** on $\mathcal{S}$
4. Identification of **the best next point** $x^\star$ based on criteria $\rightarrow$ evaluate $\hat{g}_i(\boldsymbol{x}^\star)$
   - ✓ **deviation number** : $u(\boldsymbol{x}) = \frac{|\hat{g}(\boldsymbol{x})|}{\sigma_{\hat{g}}(\boldsymbol{x})}$
   - ✓ **expected feasibility function** (EFF), **EGRA** method [Bichon et al. 2008]

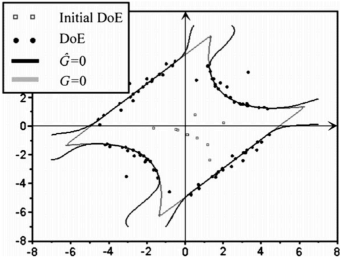


FIGURE: example of AK-MCS with $u-$criteria at 7^{th} iteration.

AK-MCS algorithm for reliability analysis

In **reliability** (ie. rare event estimation), we wish to evaluate the PoF $\rightarrow$ learning emphasis is put on **vicinity of the limit-state surface**

1. Generation of a **large MC population** $\mathcal{S}$ in the design space
2. Definition of the **initial DoE**(random selection)
3. construction of the **Kriging model** used for : $\hat{g}_i$ prediction + **PoF estimation** on $\mathcal{S}$
4. Identification of **the best next point** $\boldsymbol{x}^\star$ based on criteria $\rightarrow$ evaluate $\hat{g}_i(\boldsymbol{x}^\star)$
5. Stopping condition on learning :

   if no : **enrichment of DoE** with new best point and **goto** 2

   else : enlarge MC population $\mathcal{S} \leftarrow \mathcal{S}^+$ and **goto** 1 or **end**

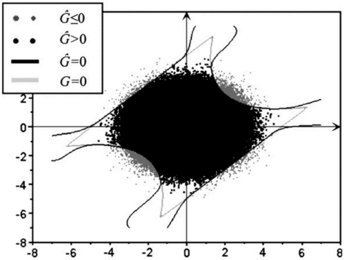


FIGURE: Monte-Carlo population estimated AK-MCS with $u-$criteria at 7^{th} iteration.

Table of contents

1 Course I : Concepts, formalism and some classes of resolution methods for simple optimization under uncertainty (1.5h, Thursday)

2 Course II : quantification of uncertainty for more efficient optimization under uncertainties : metamodeling with Gaussian Processes (1.5h, Friday)

3 Course III : Metamodeling-based Reliability-Based Design Optimization (1.5h, Saturday)
- Bayesian optimization
- Bayesian optimization under uncertainties
- Surrogate modeling deployment in the context of RBDO

$$
\begin{aligned}
\boldsymbol{x}^{\star} &= \arg\min_{\boldsymbol{x}\in\boldsymbol{\mathcal{X}}} \mathbb{L}\left[f(\boldsymbol{x},\boldsymbol{U})\right], \\
\text{such that } \mathbb{K}_i\left[g_i(\boldsymbol{x},\boldsymbol{U})\right] &\le 0 \text{ for } i \in \{1,\dots,m\} \\
\boldsymbol{x}_{\text{min}} \le \boldsymbol{x} &\le \boldsymbol{x}_{\text{max}}
\end{aligned}
\tag{2}
$$

different potential sources of uncertainties...

even when $\boldsymbol{U} \equiv \overline{\boldsymbol{U}}$ is **nominal**, the measure of risk $(\mathbb{L}, \mathbb{K})$ used to close the OUU problem is often *approximated* by a **statistical estimator** (e.g. Monte-Carlo estimator of the mean), introducing **noise** or **bias**.
In this case it is useful to – consider version of **Kriging with noise** (keyword : nugget) [Le Riche & Durrande, 2019], – adjust EI-EGO criteria not reliable anymore

when $\boldsymbol{U}$ is **random**, assuming we will rely on SMs → reducing statistical estimator errors (e.g. large samples), the chosen formulation will have to account for the complexity induced by the $\boldsymbol{U}$ variability and dimensionality

"common-sense" algorithm

1. Create an initial **sparse** DoE, $(\boldsymbol{x}^{(j)}, \boldsymbol{U}^{(j)}, f(\cdot,\cdot)$ or $g_i(\boldsymbol{x}^{(j)}, \boldsymbol{U}^{(j)}))$ and use it to initialize 1/many Gaussian Process(es) (in $\mathcal{X}$ or in augmented space)
2. Use the GP(s) to choose the **next best** $\boldsymbol{x}^{(n+1)}$
3. Use the GP(s) to choose the next $\boldsymbol{U}^{(n+1)}$ knowing $\boldsymbol{x}^{(n+1)}$ (steps 2 & 3 may be simultaneous !)
4. Evaluate $f(\boldsymbol{x}^{(n+1)}, \boldsymbol{U}^{(n+1)})$ and $g_i(\boldsymbol{x}^{(n+1)}, \boldsymbol{U}^{(n+1)})$, update the GPs, **stop** or **goto** 2

"common-sense" algorithm

1. Create an initial **sparse** DoE, $(\boldsymbol{x}^{(j)}, \boldsymbol{U}^{(j)}, f(\cdot,\cdot)$ or $g_i(\boldsymbol{x}^{(j)}, \boldsymbol{U}^{(j)}))$ and use it to initialize 1/many Gaussian Process(es) (in $\mathcal{X}$ or in augmented space)
2. Use the GP(s) to choose the **next best** $\boldsymbol{x}^{(n+1)}$
3. Use the GP(s) to choose the next $\boldsymbol{U}^{(n+1)}$ knowing $\boldsymbol{x}^{(n+1)}$ (steps 2 & 3 may be simultaneous !)
4. Evaluate $f(\boldsymbol{x}^{(n+1)}, \boldsymbol{U}^{(n+1)})$ and $g_i(\boldsymbol{x}^{(n+1)}, \boldsymbol{U}^{(n+1)})$, update the GPs, **stop** or **goto** 2

Infill criteria in both design and random parameter spaces is **problem-dependent !** (statistical measure and its estimator)

what is the best metamodeling space ?

choice of metamodeling space

1. x is **deterministic** and **independent** from $\boldsymbol{U} \Rightarrow \boldsymbol{W} = \{\boldsymbol{x}, \boldsymbol{U}\}^T$ is the vector gathering all random/deterministic input parameters of the system to optimize
   $\Rightarrow$ the **hybrid space** $\mathcal{W} \equiv \mathcal{X} \otimes \mathcal{A}$ of dimension is $(d + n)$, is the tensor product btw the **design** deterministic space and the **random parameter** space [Kharmanda 2002]

choice of metamodeling space

1. $\boldsymbol{x}$ is **deterministic** and **independent** from $\boldsymbol{U} \Rightarrow \boldsymbol{W} = \{\boldsymbol{x}, \boldsymbol{U}\}^T$ is the vector gathering all random/deterministic input parameters of the system to optimize
   $\Rightarrow$ the **hybrid space** $\mathcal{W} \equiv \mathcal{X} \otimes \mathcal{A}$ of dimension is $(d+n)$, is the tensor product btw the **design** deterministic space and the **random parameter** space [Kharmanda 2002]
2. $\boldsymbol{x}$ taken **uniformly distributed** and **ind.** from $\boldsymbol{U}$, all variables fitting within a proba. framework, $\boldsymbol{W} = \{\boldsymbol{x}, \boldsymbol{U}\}^T$: new vector of variables of an **augmented space**

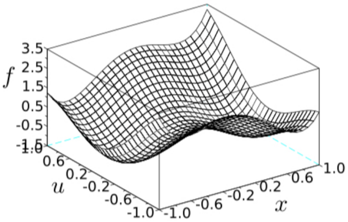


FIGURE: Example of objective : $\arg\min_{x \in \mathcal{X}} \mathbb{E}\left[f(x, U)\right]$; formulation :
$f(x, U) \approx \hat{f}(x, U)$: (GP) so $\mathbb{E}\left[f(x, U)\right] \approx \mathbb{E}_U\left[\hat{f}(x, U)\right] \equiv \tilde{f}(x)$: (GP) ; infill criteria on $\tilde{f}$...
[Janusevskis & Le Riche, 2012]

choice of metamodeling space for other form of Reliability-Based Design type of optimization (**RBDO**) :

$$\boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}\in\boldsymbol{\mathcal{X}}} \left[f(\boldsymbol{x},\overline{\boldsymbol{U}})\right],$$

$$\text{such that } \mathbb{P}\left[g_i(\boldsymbol{D}(\boldsymbol{x}),\boldsymbol{U}) > 0\right] \leq P_{f_i}, \quad \text{for } i \in \{1,\dots,m\}$$

$$h_j(\boldsymbol{x}) \leq 0, \quad \text{for } j \in \{m+1,\dots,M\}$$

3 sometimes the OUU problem involves **design variables**, now noted $\boldsymbol{D}$, which are **also considered random** but with a **known distribution** depending on a (**deterministic !**) unknown hyperparameter $\boldsymbol{x}$ to be optimized, e.g. $f_{\boldsymbol{D}|\boldsymbol{x}} \sim \mathcal{N}(\mu,\sigma)$. In this case :

- $\boldsymbol{x}$ refers to this **"nominal"** dimension of the design parameters
- $\boldsymbol{D}(\boldsymbol{x}) \sim f_{\boldsymbol{D}|\boldsymbol{x}}$ denotes the **uncertain design parameters** (e.g. due to manufacturing tolerances), **conditioned** on the design parameters $\boldsymbol{x}$
- $\boldsymbol{W} = \{\boldsymbol{D}|\boldsymbol{x},\boldsymbol{U}\}^T \sim f_{\boldsymbol{W}}$ is the augmented vector
- the **augmented reliable space** is random of dimension (dim $[\boldsymbol{D}] + n$) and can be composed of **tensorization of hyperrectangular confidence regions** (constructed from quantiles)

example of a 2D augmented reliable space

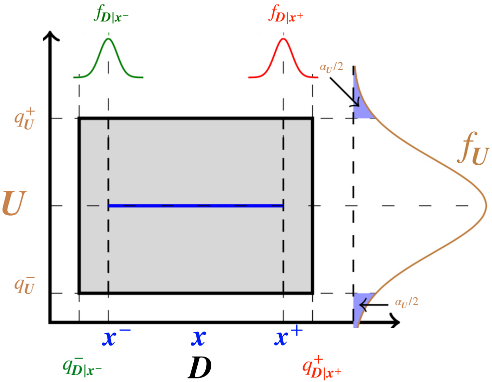


FIGURE: $\mathcal{X} = \prod_{k=1}^{d} \left[x_k^-, x_k^+\right]$; $\mathcal{D} = \prod_{k=1}^{d} \left[q^-_{D_k|x_k^-}, q^+_{D_k|x_k^+}\right]$; $\mathcal{A} = \prod_{l=1}^{n} \left[q^-_{U_l}, q^+_{U_l}\right]$ with $q^{\pm}_{\cdot} \equiv F_{\cdot}^{-1}(0/1 \pm \alpha_{\cdot}/2)$.

Table of contents

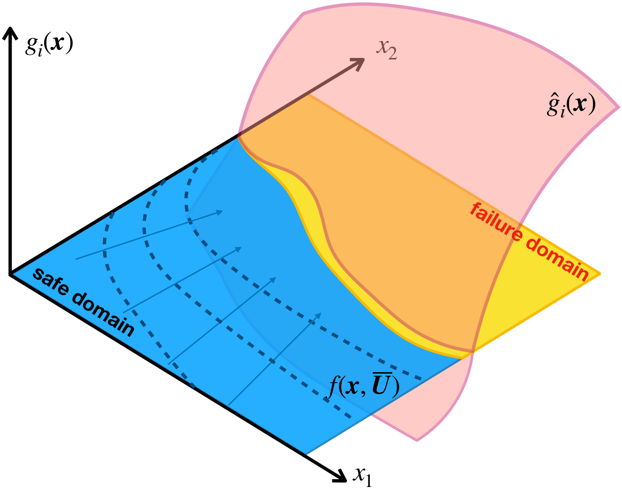
1 Course I : Concepts, formalism and some classes of resolution methods for simple optimization under uncertainty (1.5h, Thursday)

2 Course II : quantification of uncertainty for more efficient optimization under uncertainties : metamodeling with Gaussian Processes (1.5h, Friday)

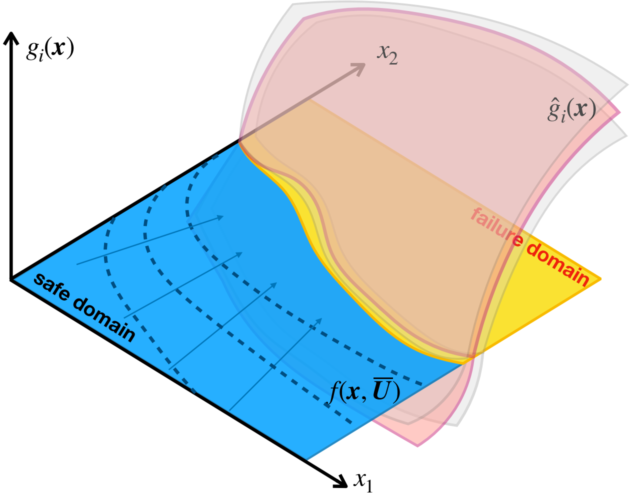
3 Course III : Metamodeling-based Reliability-Based Design Optimization (1.5h, Saturday)

- Bayesian optimization
- Bayesian optimization under uncertainties
- Surrogate modeling deployment in the context of RBDO

surrogate of the constraint : limit-state is poorly captured

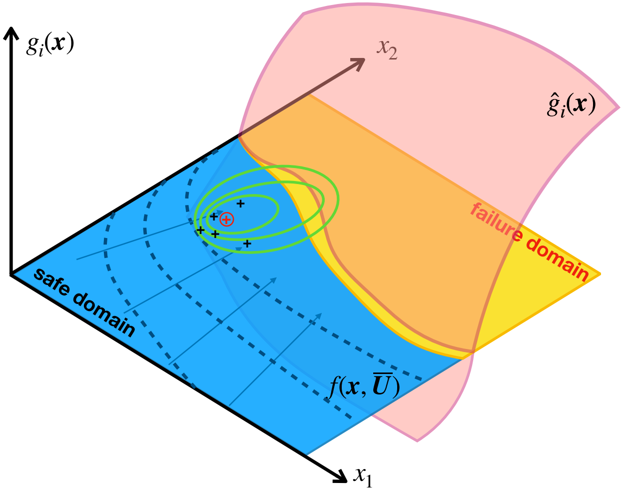


Kriging advantage : holds a local measure of confidence

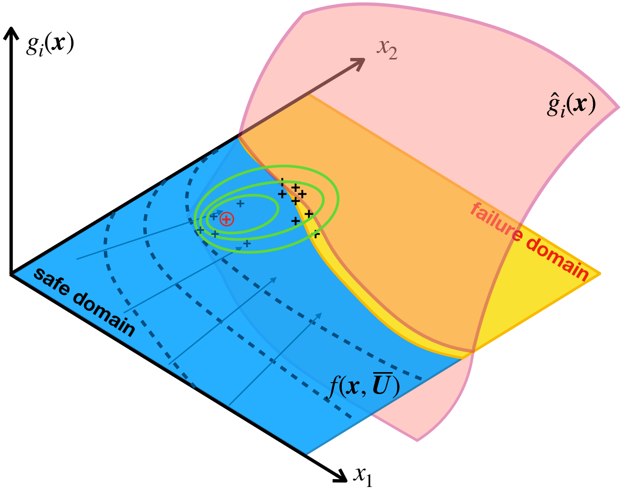


SM deployment in the context of RBDO

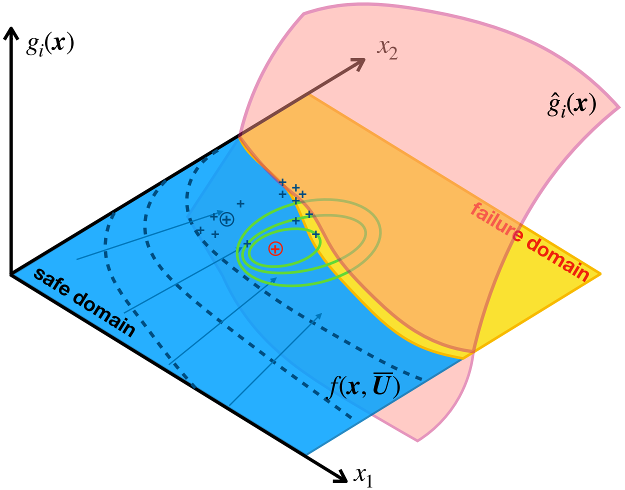
given a design point : enrichment criteria must target limit-state...



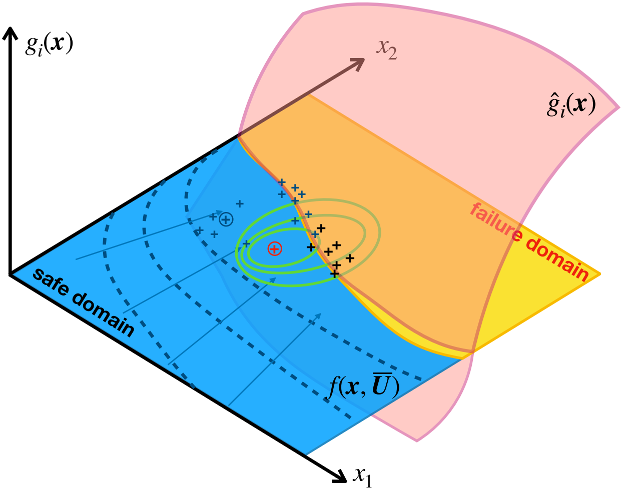
... to refine the SM and better estimate probability of failure



here the metamodel is refined at each iteration of the optimization algo...



... until convergence.



Quantile-based OUU with adaptive Kriging

Reference paper [Moustapha, M., Sudret, B., Bourinet, J. M., & Guillaume, B. (2016). Quantile-based optimization under uncertainties using adaptive Kriging surrogate models. Structural and multidisciplinary optimization, 54(6), 1403-1421.]

They propose a metamodel-based strategy for RBDO problems of this type :

$$\boldsymbol{d}^\star = \arg\min_{\boldsymbol{d}\in\boldsymbol{\mathcal{D}}}(d_1 + d_2)$$

$$\text{such that } \mathbb{P}\left[g_i(\boldsymbol{X}(d_1, d_2)) > 0\right] \leq 1.3\cdot 10^{-3} \text{ for } i \in \{1, \dots, m = 3\}$$

$$(0,0) \leq \boldsymbol{d} = (d_1, d_2) \leq (10, 10)$$

with,

$$\begin{aligned}
g_1(\boldsymbol{X}(\boldsymbol{d})) &= \frac{X_1^2\, X_2}{20} - 1 \\
g_2(\boldsymbol{X}(\boldsymbol{d})) &= \frac{(X_1 + X_2 - 5)^2}{30} + \frac{(X_1 - X_2 - 12)^2}{120} - 1 \\
g_3(\boldsymbol{X}(\boldsymbol{d})) &= \frac{80}{X_1^2 + 8X_2 + 5} - 1
\end{aligned}$$

Reference paper **[Moustapha, M., Sudret, B., Bourinet, J. M., & Guillaume, B. (2016). Quantile-based optimization under uncertainties using adaptive Kriging surrogate models. Structural and multidisciplinary optimization, 54(6), 1403-1421.]**

They propose a **metamodel-based** strategy for RBDO problems :

Thanks to the Kriging variance (measure of local accuracy of the surrogate), they propose a **two-stage DoE enrichment** to construct the surrogate model

1. **global** stage : reduce the Kriging epistemic uncertainty and **adds points in the vicinity of the limit-state surface**
2. **local** stage : checks, and if necessary, improves locally the accuracy of the quantiles estimated along the optimization iterations

Application on a test case with nonlinear constraints in UQLab.

Quantile-based OUU with adaptive Kriging

Reference paper **[Moustapha, M., Sudret, B., Bourinet, J. M., & Guillaume, B. (2016). Quantile-based optimization under uncertainties using adaptive Kriging surrogate models. Structural and multidisciplinary optimization, 54(6), 1403-1421.]**

They propose a **metamodel-based** strategy for RBDO problems :

Thanks to the Kriging variance (measure of local accuracy of the surrogate), they propose a **two-stage DoE enrichment** to construct the surrogate model

1. **global** stage : reduce the Kriging epistemic uncertainty and **adds points in the vicinity of the limit-state surface**

    similar to an AK-MCS refinement strategy on the parameters, BUT in contrast to AK-MCS, the constraint is defined wrt $\boldsymbol{d}$ **in the design space** while the **Kriging model is built in the augmented space** $\boldsymbol{X}(\boldsymbol{d})$.

    IDEA : find new point in augmented space $\rightarrow$ improvement of quantile estimation in design space

2. **local** stage : checks, and if necessary, improves locally the accuracy of the quantiles estimated along the optimization iterations

Relying on computation of probability or quantile ?

Equivalence between RBDO and quantile-based formulation

Standard reliable formulation via PoF defined for particular constraint g_i :

$$\mathbb{K}\left[g_i(\boldsymbol{x},\boldsymbol{U})\right] \equiv \mathbb{P}\left[g_i(\boldsymbol{x},\boldsymbol{U}) > 0\right] \leq \overline{P_f^{g_i}} \leftrightarrow P_f (\ll 1)$$

for a given design : $P_f^{g_i}(\boldsymbol{x}) = \mathbb{P}\left[g_i(\boldsymbol{U}|\boldsymbol{x}) > 0\right] = \int_{g_i(\boldsymbol{U}|\boldsymbol{x})>0} f_{\boldsymbol{U}|\boldsymbol{X}}(\boldsymbol{U}|\boldsymbol{x}) d\boldsymbol{U}$

Relying on computation of probability or quantile ?

Equivalence between RBDO and quantile-based formulation

Standard reliable formulation via PoF defined for particular constraint g_i :

$$\mathbb{K}\left[g_i(\boldsymbol{x},\boldsymbol{U})\right] \equiv \mathbb{P}\left[g_i(\boldsymbol{x},\boldsymbol{U})>0\right] \leq \overline{P_f^{g_i}} \leftrightarrow P_f(\ll 1)$$

for a given design : $P_f^{g_i}(\boldsymbol{x}) = \mathbb{P}\left[g_i(\boldsymbol{U}|\boldsymbol{x})>0\right] = \displaystyle\int_{g_i(\boldsymbol{U}|\boldsymbol{x})>0} f_{\boldsymbol{U}|\boldsymbol{X}}(\boldsymbol{U}|\boldsymbol{x})d\boldsymbol{U}$

In general we can express the constraint as : $g_i(\boldsymbol{x},\boldsymbol{U}) = \mathcal{M}_i(\boldsymbol{x},\boldsymbol{U}) - \bar{g}_{i(=\text{threshold})}$

$$\begin{aligned} \mathbb{P}\left[g_i(\boldsymbol{x},\boldsymbol{U})>0\right] \leq P_f \quad &\Longleftrightarrow \quad \mathbb{P}\left[\mathcal{M}_i(\boldsymbol{x},\boldsymbol{U}) \geq \bar{g}_i\right] \leq P_f \\ &\Longleftrightarrow \quad \mathbb{P}\left[\mathcal{M}_i(\boldsymbol{x},\boldsymbol{U}) \leq \bar{g}_i\right] \geq 1-P_f \equiv \alpha(\lesssim 1) \end{aligned}$$

we can introduce the quantile : $\mathbb{Q}_\alpha\left[\boldsymbol{x};\cdot\right] = \inf\{q \in \mathbb{R} : \mathbb{P}\left[\cdot \leq q\right] \leq \alpha\} \Rightarrow$

Equivalence between RBDO and quantile-based formulation

$$\begin{aligned} \mathbb{P}\left[g_i(\boldsymbol{x},\boldsymbol{U})>0\right] \leq P_f \quad &\Longleftrightarrow \quad \mathbb{Q}_\alpha\left[\boldsymbol{x};\mathcal{M}_i(\boldsymbol{x},\boldsymbol{U})\right] \leq \bar{g}_i \\ &\Longleftrightarrow \quad \mathbb{Q}_\alpha\left[\boldsymbol{x};g_i(\boldsymbol{x},\boldsymbol{U})\right] \leq 0 \ \text{(put otherwise)} \end{aligned}$$

quantile approach : easier coupling with already existing deterministic design process ; outer loop explores the design space while inner loop simply computes constraints quantiles

Integrating **uncertainties** into the design process **greatly** modifies the optimization problem to be solved.

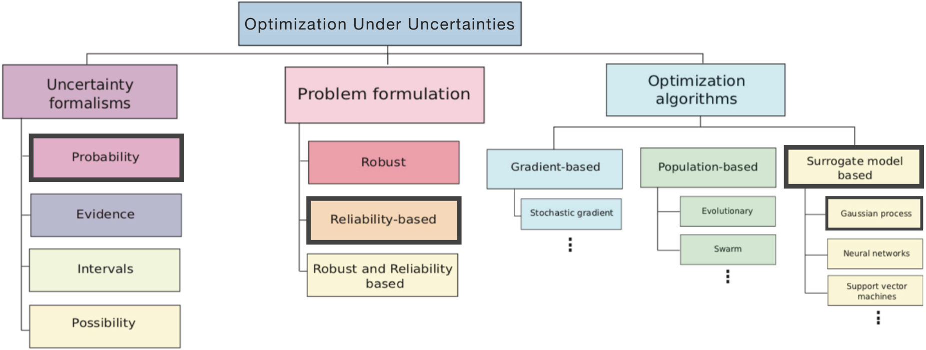
The **statistical measures of risk** must be carefully **adapted to the problem**.

Accounting for the uncertainty induces a **LARGE computational overhead**

Kriging and **Gaussian Processes** are nicely tunable metamodels that reduce the cost but remain limited in terms of dimensionality

machine learning (deep neural networks) may be the future of more efficient metamodeling in this framework

OUU : very broad topic, many things to improve/couple/discover !



Thanks for attending this class !

Бие махбодийн
хувьд байхгүй
байгаадаа
уучлаарай

Big thanks to the organizers !

didier.lucor@lisn.upsaclay.fr

https://perso.limsi.fr/lucor/

Some references including active research teams on the topic

- **Bachoc, F.** (2013). Cross validation and maximum likelihood estimations of hyper-parameters of Gaussian processes with model misspecification. Computational Statistics & Data Analysis, 66, 55-69.
- **Balesdent, M.**, Morio, J., & Brevault, L. (2016). Rare event probability estimation in the presence of epistemic uncertainty on input probability distribution parameters. Methodology and Computing in Applied Probability, 18(1), 197-216.
- **Beyer, H. G.**, & **Sendhoff, B.** (2007). Robust optimization – a comprehensive survey. Computer methods in applied mechanics and engineering, 196(33-34), 3190-3218.
- Dubourg, V., **Sudret, B.**, & Deheeger, F. (2013). Metamodel-based importance sampling for structural reliability analysis. Probabilistic Engineering Mechanics, 33, 47-57.
- Echard, B., **Gayton, N.**, & **Lemaire, M.** (2011). AK-MCS : an active learning reliability method combining Kriging and Monte Carlo simulation. Structural Safety, 33(2), 145-154.
- Janusevskis, J., & **Le Riche, R.** (2013). Simultaneous kriging-based estimation and optimization of mean response. Journal of Global Optimization, 55(2), 313-336.
- Lelièvre, N., Beaurepaire, P., Mattrand, C., **Gayton, N.**, & Otsmane, A. (2016). On the consideration of uncertainty in design : optimization-reliability-robustness. Structural and Multidisciplinary Optimization, 54(6), 1423-1437.
- Lopez, R. H., & **Beck, A. T.** (2012). Reliability-based design optimization strategies based on FORM : a review. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 34, 506-514.
- Moustapha, **M., Sudret, B., Bourinet**, J. M., & Guillaume, B. (2016). Quantile-based optimization under uncertainties using adaptive Kriging surrogate models. Structural and multidisciplinary optimization, 54(6), 1403-1421.
- UQLab user manual on Reliability-Based Design Optimization, at : https ://www.uqlab.com/rbdo-user-manual