Point process-based approaches for the reliability analysis of systems modeled by costly simulators

G. Perrin

UQSay seminar #10 | June 11th 2020

CEA/DAM/DIF, F-91297, Arpajon, France

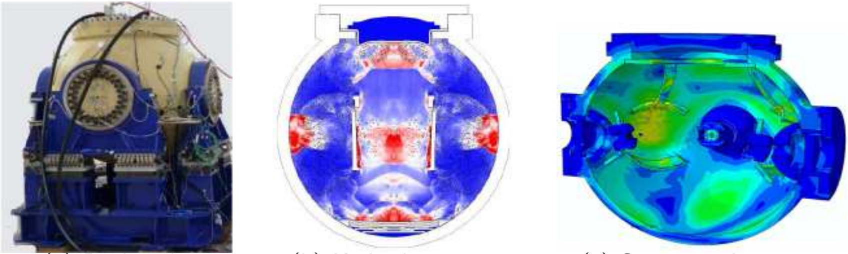
Introduction

- Simulation plays a key role in the reliability analysis of complex systems.
- Most of the time, these analyses can be reduced to estimating the probability of occurrence of an undesirable event, using a stochastic model of the system.
- If the considered event is rare, sophisticated sample-based procedures are generally introduced to get a relevant estimate of the failure probability.

Problematic

Based on a reduced number of model evaluations, how to bound this failure probability with a prescribed confidence ?

Example (1/2)



(a) Real tank (b) Hydrodynamics (c) Structure dynamics

FIGURE: Pressure tank under dynamic pressure

Example (1/2)

Inp.	Distribution	Parameters
Internal radius (m)	Normal	$(0.720, 0.005)$
Thickness (m)	Log-normal	$(0.073, 0.0015)$
In. press. uncert.	Weibull	$(24.95, 1.022)$
Time uncert.	Weibull	$(24.95, 1.022)$
YM of the tank (Pa)	Log-normal	$(2.1 \times 10^{11}, 2.1 \times 10^{10})$
σ_{elast} of the tank (Pa)	Normal	$(7 \times 10^{8}, 3 \times 10^{7})$
YM of the tap (Pa)	Log-normal	$(2.1 \times 10^{11}, 2.1 \times 10^{10})$
σ_{elast} of the tap (Pa)	Normal	$(8.6 \times 10^{8}, 3 \times 10^{7})$

TABLE: Stochastic model of the pressure tank.

Problematic

How to certify that the maximum value in time and space of the cumulative equivalent plastic strain is less than q?

General framework

Notations

- $\mathcal{S} \leftrightarrow$ system of interest,
- $\boldsymbol{x} \in \mathbb{X} \subset \mathbb{R}^D \leftrightarrow$ system characteristics (dimensions, boundary conditions, material properties...),
- $\boldsymbol{x} \mapsto y(\boldsymbol{x}) \in \mathbb{R} \leftrightarrow$ quantity of interest for the monitoring of $\mathcal{S}$,
- $\mathcal{F} = \{\boldsymbol{x} \in \mathbb{X} \mid y(\boldsymbol{x}) < 0\} \leftrightarrow$ system's failure domain.

Asumptions

- $\boldsymbol{x}$ is not perfectly known $\Rightarrow$ it is modeled by a random vector with PDF $f_{\boldsymbol{x}}$.
- Each evaluation of y is **time-demanding** (several hours for the considered application).

$\Rightarrow$ $p(y) := \mathbb{P}(y(\boldsymbol{x}) < 0) = \int_{\mathcal{F}} f_{\boldsymbol{x}}(\boldsymbol{x}) d\boldsymbol{x} \leftrightarrow$ system failure probability of interest.

Outline

Failure probability estimation

$$p(y) := \mathbb{P}(y(\boldsymbol{x}) < 0) = \int_{\mathcal{F}} f_{\boldsymbol{x}}(\boldsymbol{x}) d\boldsymbol{x}.$$

- $y \leftrightarrow$ output of a numerically **expensive deterministic "black box"** : for each $\boldsymbol{x}$, $y(\boldsymbol{x})$ is unique, and can be calculated by using a simulator that can take a long time to evaluate.

⇒ In this type of configuration, the calculation of $p(y)$ is usually based on **sampling techniques**.

Failure probability estimation

$$p(y) := \mathbb{P}(y(\boldsymbol{x}) < 0) = \int_{\mathcal{F}} f_{\boldsymbol{x}}(\boldsymbol{x}) d\boldsymbol{x}.$$

- $y \leftrightarrow$ output of a numerically **expensive deterministic "black box"** : for each $\boldsymbol{x}$, $y(\boldsymbol{x})$ is unique, and can be calculated by using a simulator that can take a long time to evaluate.

⇒ In this type of configuration, the calculation of $p(y)$ is usually based on **sampling techniques**.

Monte Carlo approach

- $p(y) \approx \widehat{p}(y) := \frac{1}{N} \sum_{n=1}^{N} 1_{y(\boldsymbol{x}^{(n)})<0}$, $\left\{\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(N)}\right\} \leftrightarrow N$ iid copies of $\boldsymbol{x}$,
- $\delta^2 = \frac{\mathrm{Var}(\widehat{p}(y))}{\mathbb{E}[\widehat{p}(y)]^2} = \frac{1-p(y)}{Np(y)}$,

⇒ to get a "correct" estimation of $p(y)$, very large values of N are needed when $p(y)$ is low (say $p(y) < 10^{-3}$ for a rare event).

Failure probability estimation

Estimation of $p(y)$ using approximations of $\mathcal{F}$ (or of its boundary)

- the well-known first- or second-order reliability methods (FORM/SORM)
- approximations based on support vector machine (SVM) techniques or on the generalized least squares linear regression.

$\Rightarrow$ Practical problem : given an approximation of $\mathcal{F}$, it is generally difficult to evaluate the difference between $p(y)$ and its approximation.

Variance reduction techniques

- importance sampling approach,
- splitting techniques.

$\Rightarrow$ Practical problems : the choice of the random vector for importance sampling, the conditional sampling for splitting techniques.

Surrogate modeling

- To bypass some of the former problems, most techniques for estimating rare events are currently based on a **surrogate model**.
- We focus here on the Gaussian process regression (GPR), which models y as a particular realization of a Gaussian process $Y \sim \mathrm{GP}(\mu, C)$.
- Under that formalism,

$$p(y) = \mathbb{P}\left(Y(\boldsymbol{x}) < 0 \mid Y = y\right),$$

⇒ $p(y)$ is a particular realization of the random variable

$$p(Y) := \mathbb{P}\left(Y(\boldsymbol{x}) < S \mid Y\right).$$

Surrogate modeling

- To bypass some of the former problems, most techniques for estimating rare events are currently based on a **surrogate model**.
- We focus here on the Gaussian process regression (GPR), which models y as a particular realization of a Gaussian process $Y \sim \mathrm{GP}(\mu, C)$.
- Under that formalism,

$$p(y) = \mathbb{P}\left(Y(\boldsymbol{x}) < 0 \mid Y = y\right),$$

⇒ $p(y)$ is a particular realization of the random variable

$$p(Y) := \mathbb{P}\left(Y(\boldsymbol{x}) < S \mid Y\right).$$

Remark

$$\mathbb{E}\left[p(Y)\right] = \mathbb{E}\left[\Phi\left(-\frac{\mu(\boldsymbol{x})}{\sigma(\boldsymbol{x})}\right)\right].$$

Surrogate modeling and reliability analysis

$$p(y) = \mathbb{P}\left(Y(\boldsymbol{x}) < 0 \mid Y = y\right), \quad p(Y) := \mathbb{P}\left(Y(\boldsymbol{x}) < S \mid Y\right).$$

⇒ To correctly anticipate the risks, rather than focusing on the estimation of $\mathbb{E}\left[p(Y)\right]$, which may strongly overestimate or underestimate $p(y)$ depending on the quality of the surrogate, we propose to construct an estimator $\widehat{q}_{\alpha,\beta}$ of the **(1-α) quantile** of $p(Y)$, noted q_α, so that :

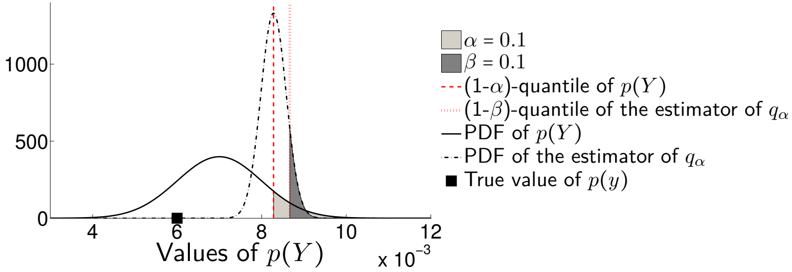
$$\mathbb{P}(p(Y) < q_\alpha) = 1 - \alpha,$$

$$\mathbb{P}\left(\mathbb{P}\left(p(Y) \le \widehat{q}_{\alpha,\beta} \mid \widehat{q}_{\alpha,\beta}\right) \ge 1 - \alpha\right) \ge 1 - \beta.$$

- α characterizes the risk associated to the replacement of y by Y,
- β controls the fact that only finite-dimensional samples of $Y(\boldsymbol{x})$ are available for its construction.

Surrogate modeling and reliability analysis

$$\mathbb{P}(p(Y) < q_\alpha) = 1 - \alpha, \quad \mathbb{P}\left(\mathbb{P}\left(p(Y) \leq \widehat{q}_{\alpha,\beta} \mid \widehat{q}_{\alpha,\beta}\right) \geq 1 - \alpha\right) \geq 1 - \beta.$$



- α characterizes the risk associated to the replacement of y by Y,
- β controls the fact that only finite-dimensional samples of $Y(\boldsymbol{x})$ are available for its construction.

Specific objectives

For $\alpha, \beta \in (0,1)$ and a fixed number of evaluations of y,

- **First objective** : propose an algorithm allowing us to construct estimator $\widehat{q}_{\alpha,\beta}$. Key elements :
  1. order statistics,
  2. the Gaussian process regression formalism,
  3. a particular Marked Poisson Process.

Specific objectives

For $\alpha, \beta \in (0,1)$ and a fixed number of evaluations of y,

- **First objective** : propose an algorithm allowing us to construct estimator $\widehat{q}_{\alpha,\beta}$. Key elements :
  1. order statistics,
  2. the Gaussian process regression formalism,
  3. a particular Marked Poisson Process.
- **Second objective** : propose a strategy adapted to the former algorithm to sequentially minimize the dependence of $\widehat{q}_{\alpha,\beta}$ on the replacement of y by Y, while managing the cases where :
  1. no point of the initial experimental design for the construction of Y belongs to the failure domain ,
  2. the failure domain is multimodal.

Outline

Initial exploration of the input space

Context reminder

- Input random vector : $\boldsymbol{x} \in \mathbb{X} \subset \mathbb{R}^d$ with PDF $f_{\boldsymbol{x}}$,
- Quantity of interest : $\boldsymbol{x} \mapsto y(\boldsymbol{x}) \in \mathbb{R}$,
- Failure probability : $p(y) = \mathbb{P}(y(\boldsymbol{x}) < 0)$.

Gaussian process regression

- Model y has been evaluated in L (the value of L is assumed relatively small) points of $\mathbb{X}$, $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(L)}$ (space filling LHS).
- y is seen as a sample path of a Gaussian process defined on $(\Omega, \mathcal{A}, \mathbb{P})$.
- Let $Y \sim \text{GP}(\mu, C)$ be this Gaussian process conditioned by the L available code evaluations.

Order statistics (1/3)

Let $\mathcal{X}^N$ be a random set of $N > 1$ points chosen (independently or not) in $\mathbb{X}$. For $\omega, \theta \in \Omega$, let :

- $Y(\omega)$ be a realization of Y and $Y(\boldsymbol{x};\omega)$ be the projection of $Y(\omega)$ in random vector $\boldsymbol{x}$,
- $\widehat{p}(Y(\omega), \mathcal{X}^N)$ be an estimator of the realization $p(Y(\omega))$ of $p(Y)$,
- $\widehat{p}(Y, \mathcal{X}^N)$ be the random variable so that $\widehat{p}(Y(\omega), \mathcal{X}^N(\theta))$ is a realization of $\widehat{p}(Y, \mathcal{X}^N)$ with $\mathcal{X}^N(\theta)$ a realization of $\mathcal{X}^N$,
- $Y_1, \dots, Y_M$ and $\mathcal{X}_1^N, \dots, \mathcal{X}_M^N$ be $M \geq 1$ independent copies of Y and $\mathcal{X}^N$ respectively.

For $1 \leq m \leq M$, we then write $\widehat{p}_m := \widehat{p}(Y_m, \mathcal{X}_m^N)$. These random variables are supposed to be sorted in **ascending order** :

$$\widehat{p}_1 \leq \widehat{p}_2 \leq \dots \leq \widehat{p}_M.$$

Order statistics (2/3)

For $1 \leq k \leq M$ and $\alpha \in (0,1)$, as $\mathbb{P}(p(y) < q_\alpha) = 1 - \alpha$ we have :

$$\mathbb{P}(\widehat{p}_k > q_\alpha) = \sum_{u=0}^{k-1} \binom{M}{u} (1-\gamma)^{M-u} \gamma^u,$$

$$\begin{aligned}
\gamma &:= \mathbb{P}(\widehat{p}(Y, \mathcal{X}^N) \leq q_\alpha) \\
&= \mathbb{P}(\widehat{p}(Y, \mathcal{X}^N) \leq q_\alpha \mid p(Y) \leq q_\alpha)\mathbb{P}(p(Y) \leq q_\alpha) + \mathbb{P}(\widehat{p}(Y, \mathcal{X}^N) \leq q_\alpha \leq p(Y)) \\
&\leq 1 \times (1-\alpha) + \mathbb{P}(\widehat{p}(Y, \mathcal{X}^N) \leq p(Y) \ \cap \ p(Y) \geq q_\alpha) \\
&\leq 1 - \alpha(1 - \mathbb{P}(\widehat{p}(Y, \mathcal{X}^N) \leq p(Y) \mid p(Y) \geq q_\alpha)) =: \gamma_\star,
\end{aligned}$$

so that :

$$\mathbb{P}(\widehat{p}_k > q_\alpha) \geq \sum_{u=0}^{k-1} \binom{M}{u} (1-\gamma_\star)^{M-u} \gamma_\star^u.$$

Order statistics (3/3)

Choosing $k^\star(\alpha,\beta)$ as the minimal index such that

$$\sum_{u=0}^{k^\star(\alpha,\beta)-1} \binom{M}{u} (1-\gamma_\star)^{M-u}\gamma_\star^u \geq 1-\beta,$$

it can be seen that $\mathbb{P}(\widehat{p}_{k^\star(\alpha,\beta)} > q_\alpha) \geq 1-\beta$, and we obtain :

$$\mathbb{P}(\mathbb{P}\left(p(Y) \leq \widehat{p}_{k^\star(\alpha,\beta)} \mid \widehat{p}_{k^\star(\alpha,\beta)}\right) \geq 1-\alpha) \geq 1-\beta$$

$\Rightarrow$ Choosing $\widehat{q}_{\alpha,\beta} = \widehat{p}_{k^\star(\alpha,\beta)} = \widehat{p}(Y_{k^\star(\alpha,\beta)}, \mathcal{X}^N_{k^\star(\alpha,\beta)})$ leads to the searched result !

Order statistics (3/3)

Choosing $k^\star(\alpha,\beta)$ as the minimal index such that

$$\sum_{u=0}^{k^\star(\alpha,\beta)-1} \binom{M}{u} (1-\gamma_\star)^{M-u} \gamma_\star^u \geq 1-\beta,$$

it can be seen that $\mathbb{P}(\widehat{p}_{k^\star(\alpha,\beta)} > q_\alpha) \geq 1-\beta$, and we obtain :

$$\mathbb{P}(\mathbb{P}\left(p(Y) \leq \widehat{p}_{k^\star(\alpha,\beta)} \mid \widehat{p}_{k^\star(\alpha,\beta)}\right) \geq 1-\alpha) \geq 1-\beta$$

⇒ Choosing $\widehat{q}_{\alpha,\beta} = \widehat{p}_{k^\star(\alpha,\beta)} = \widehat{p}(Y_{k^\star(\alpha,\beta)}, \mathcal{X}^N_{k^\star(\alpha,\beta)})$ leads to the searched result !

⇒ But to construct $\widehat{q}_{\alpha,\beta} = \widehat{p}(Y_{k^\star(\alpha,\beta)}, \mathcal{X}^N_{k^\star(\alpha,\beta)})$, we need an estimator of $p(Y(\omega)) = \mathbb{P}(Y(\omega;\boldsymbol{x}) < 0)$ for several ω in Ω.

Choice of the estimator (1/3)

Monte Carlo estimator

If $\boldsymbol{x}^{(n,m)}$, $1 \leq n \leq N$, $1 \leq m \leq M$ denote $N \times M$ independent copies of $\boldsymbol{x}$, a possible choice is :

$$\widehat{p}_m = \widehat{p}(Y_m, \mathcal{X}_m^N) = \frac{1}{N} \sum_{n=1}^{N} 1_{Y_m(\boldsymbol{x}^{(n,m)})<0}, \quad \mathcal{X}_m^N = \left\{ \boldsymbol{x}^{(n,m)}, \ 1 \leq n \leq N \right\}.$$

Using the central limit theorem :

$$\sqrt{N}(\widehat{p}(Y_m(\omega), \mathcal{X}^N) - p(Y_m(\omega))) \xrightarrow{\mathcal{L}} \mathcal{N}(0, p(Y_m(\omega))(1 - p(Y_m(\omega)))), \ \omega \in \Omega.$$

- → $\mathbb{P}(\widehat{p}(Y, \mathcal{X}^N) \leq p(Y) \mid p(Y) \geq q_\alpha) \approx 1/2$ for sufficiently high values of N, so that $\gamma_\star \approx 1 - \alpha/2$.
- → However, to numerically calculate $\widehat{p}(Y_m, \mathcal{X}^N)$, we need to project Y_m in a very high number of points, which is often not possible due to computational reasons (memory and conditioning problems).

Choice of the estimator (2/3)

Marked Poisson process [Walter, 2015]

Let Z be real-valued random variable with continuous CDF F_Z, and $(Z_i)_{i\geq 0}$ be the decreasing random walk so that :

$$Z_0 = +\infty, \ \mathbb{P}(Z_{i+1} \leq z \mid Z_i) = \mathbb{P}(Z \leq z \mid Z \leq Z_i).$$

Property

For all $t > 0$, $P(t) := \sup\{i; \ -\log\left(\mathbb{P}(Z \leq Z_i|Z_i)\right) \leq t\}$ is a Poisson random process with parameter t.

Corollary

For $t > 0$ and $i \geq 1$, if $P_1, \dots, P_Q$ denote Q independent copies of $P(t)$, the random variable $p_Q := \left(1 - \frac{1}{Q}\right)^{\sum_{q=1}^{Q} P_q}$ verifies :

$$\mathbb{E}\left[p_Q\right] = \mathrm{e}^{-t}, \quad \mathrm{Var}(p_Q) = \mathrm{e}^{-2t}(\mathrm{e}^{t/Q} - 1).$$

Choice of the estimator (3/3)

Marked Poisson process

- Applying the former results with $Z = Y_m(\boldsymbol{x};\omega)$ and $t = -\log\left(\mathbb{P}(Y_m(\boldsymbol{x};\omega) < 0)\right)$, we obtain :

$$P(-\log\left(\mathbb{P}(Y_m(\boldsymbol{x};\omega) < 0)\right)) = \sup\left\{i;\ Z_i \geq 0\right\}.$$

- If $P_1^m, \dots, P_Q^m$ are Q independent copies of $P(-\log\left(\mathbb{P}(Y_m(\boldsymbol{x};\omega) < 0)\right))$, $\widehat{p}(Y_m(\omega), \mathcal{X}^m) := \left(1 - \frac{1}{Q}\right)^{\sum_{q=1}^{Q} P_q^m}$ defines a new estimator of $p(Y_m(\omega)) = \mathbb{P}(Y_m(\boldsymbol{x};\omega) < 0)$
- As for the former Monte Carlo estimator, it can be shown that $\gamma_\star$ tends to $1 - \alpha/2$ when Q is chosen sufficiently high.
- However, while $Y_m(\omega)$ was to be projected in approximately $100/p(Y_m(\omega))$ points for the MC approach, it only needs to be projected in $\mathbb{E}\left[\sum_{q=1}^{Q} P_q^m\right] = -Q\log(p(Y_m(\omega)))$ points in average.

Pseudo-code

Compute $M = k^\star(\alpha, \beta)$, $Y \sim \text{GP}(\mu, C)$ as the GPR of y based on L evaluations of y,

For $1 \leq m \leq M$:

Sample Q indep. real. of $\boldsymbol{x}$: $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(Q)}$,

Initialize : $Y_m := Y \mid Y(\boldsymbol{x}^{(q)}) = y_q, \; 1 \leq q \leq Q,$, $P_q^m = 0$, $n_{\text{iter}} = 0$,

For $1 \leq q \leq Q$: $z = y_q$,

While z>0 and $n_{\text{iter}} \leq N_{\text{max}}$:

increment $n_{\text{iter}} = n_{\text{iter}} + 1$,

draw at random a realization of $\boldsymbol{x}$: $\boldsymbol{x}^\star$,

draw at random a realization of $Y_m(\boldsymbol{x}^\star)$: $y^\star$,

If $y^\star < z$: $z = y^\star$, $P_q^m = P_q^m + 1$,

$\widetilde{p}_m := \left(1 - \frac{1}{Q}\right)^{\sum_{q=1}^Q P_q^m}$,

Return $\max_{1 \leq m \leq M} \widetilde{p}_m$.

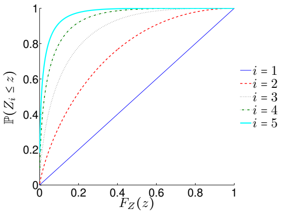
Remarks on the proposed algorithm

- For $\alpha = \beta = 0.1$, $M = 47$.
- The $\widetilde{p}_m$ being independent, their construction can be made completely parallel.
- As we are only interested in the max. of $\widetilde{p}_1, \dots, \widetilde{p}_M$, once the loop m associated with Y_m has converged, it is possible to break each other loop m' as soon as $\sum_{q=1}^{Q} P_q^{m'}$ is greater than $\sum_{q=1}^{Q} P_q^{m}$.
- By stopping a loop of the alg. before its convergence, we obtain a value of P_q^m which will be smaller than what it should have been, leading to an overestimation of $\widehat{q}_{\alpha,\beta}$, i.e. to a more protective bound for the true probability of system failure $\Rightarrow N_{\text{max}}$ certifies the convergence in a reasonable time.
- When considering very small $p(y)$, the convergence can be accelerated by using sequential MC simulation techniques.

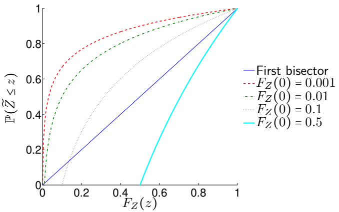
Distribution of the generated points

Property

$$\mathbb{P}(\widetilde{Z} \leq z) = \frac{\sum_{i=1}^{+\infty} \left(\mathbb{P}(Z_i \leq z) - \mathbb{P}(Z_i \leq 0)\right)}{\sum_{i=1}^{+\infty} \left(1 - \mathbb{P}(Z_i \leq 0)\right)}, \quad \mathbb{P}(Z_i \leq z) = F_Z(z) \sum_{k=0}^{i-1} \frac{|\log(F_Z(z))|^k}{k!}.$$



(a) CDF of Z_i



(b) CDF of generated outputs

- ⇒ the generated points will naturally show the regions of $\mathbb{X}$ that need to be particularly well covered by evaluations of y to minimize the impact of replacing y by Y (while being concentrated around y=0).
- ⇒ we denote by $\mathcal{L}(\mathbb{X})$ the (non-trivial) distribution of the generated points.

Outline

Sources of uncertainty

Estimator $\widehat{q}_{\alpha,\beta}$ aggregates at least two sources of uncertainty :

- the variability in the different copies of the decreasing random walk associated with each sample path, which can be reduced by increasing Q,
- the variability in the sample paths of Y due to the substitution of the computer code by its GP-based surrogate model, which can be reduced by adding new code evaluations.

Such a new point $\boldsymbol{x}^\star$ can be chosen as (**Pointwise criterion**) :

$$\boldsymbol{x}^\star := \arg \min_{\boldsymbol{x} \in \mathcal{S}(N)} \frac{|\mu(\boldsymbol{x})|}{\sqrt{C(\boldsymbol{x},\boldsymbol{x})}},$$

where $\mathcal{S}(N)$ gathers N iid samples of $\boldsymbol{x}$ [Echard et al., 2011].

⇒ Problem : this pointwise strategy does not take into account the fact that the new evaluation point will bring additional information in its neighbourhood.

SUR strategy

- In contrast, Stepwise Uncertainty Reduction (SUR) techniques propose to select the new evaluation point that minimizes the expected value of a well chosen measure of the uncertainty.
- For instance, as $1_{Y(\boldsymbol{x})<0}$ has conditional variance $\alpha(\boldsymbol{x})(1-\alpha(\boldsymbol{x}))$ with

$$\alpha(\boldsymbol{x}) := \mathbb{E}\left[1_{Y(\boldsymbol{x})<0} \mid \boldsymbol{x}\right] = \Phi\left(\frac{-\mu(\boldsymbol{x})}{\sqrt{C(\boldsymbol{x},\boldsymbol{x})}}\right),$$

  $\Omega := \mathbb{E}\left[\alpha(\boldsymbol{x})(1-\alpha(\boldsymbol{x}))\right]$ characterizes the uncertainty due to the replacement of y by Y.
- As proposed in [Bect et al., 2012] and [Chevalier et al., 2014], the new point $\boldsymbol{x}^\star$ can be searched as (**SUR criterion**) :

$$\boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}\in\mathcal{S}(N)} \mathbb{E}\left[\mathbb{E}\left[\Omega \mid Z, Y(\boldsymbol{x}) = Z\right]\right], \quad Z \sim \mathcal{N}\left(\mu(\boldsymbol{x}), C(\boldsymbol{x},\boldsymbol{x})\right).$$

Proposed criterion

- Although this two selection criteria have shown their effectiveness on several reliability problems, they are not particularly designed for the estimation of $\widehat{q}_{\alpha,\beta}$ using Marked Poisson processes.
- As the proposed estimation is based on the **classification** (higher/smaller) of different projections of Y, we propose the following adaptation of the SUR criterion (**Proposed criterion**) :

$$\boldsymbol{x}^{\star} = \arg\min_{\widehat{\boldsymbol{x}} \in \mathbb{X}} \mathbb{P}_{\boldsymbol{x},\boldsymbol{x}' \sim \mathcal{L}(\mathbb{X})} \left(Y(\boldsymbol{x}) > Y(\boldsymbol{x}'), \mu(\boldsymbol{x}) < \mu(\boldsymbol{x}') \mid Y(\widehat{\boldsymbol{x}}) = \mu(\widehat{\boldsymbol{x}}) \right),$$

$\mathcal{L}(\mathbb{X}) \leftrightarrow$ distribution of the gen. points $\boldsymbol{x}$ during the est. procedure.

Proposed criterion

- Although this two selection criteria have shown their effectiveness on several reliability problems, they are not particularly designed for the estimation of $\widehat{q}_{\alpha,\beta}$ using Marked Poisson processes.
- As the proposed estimation is based on the **classification** (higher/smaller) of different projections of Y, we propose the following adaptation of the SUR criterion (**Proposed criterion**) :

$$\boldsymbol{x}^\star = \arg\min_{\widehat{\boldsymbol{x}}\in\mathbb{X}} \mathbb{P}_{\boldsymbol{x},\boldsymbol{x}'\sim\mathcal{L}(\mathbb{X})}\left(Y(\boldsymbol{x}) > Y(\boldsymbol{x}'), \mu(\boldsymbol{x}) < \mu(\boldsymbol{x}') \mid Y(\widehat{\boldsymbol{x}}) = \mu(\widehat{\boldsymbol{x}})\right),$$

$\mathcal{L}(\mathbb{X}) \leftrightarrow$ distribution of the gen. points $\boldsymbol{x}$ during the est. procedure.

→ priority is given to the points that are difficult to classify,
→ taking $\mathcal{L}(\mathbb{X})$ as the distribution for $\boldsymbol{x}$, we focus on the areas most used by the draws of the Poisson process,
→ we focus not only on the zone close to 0, but on all positive values to avoid over-sensitivity to a bad initial metamodel.

Estimation of the proposed selection criterion

If $(\boldsymbol{x}_r, \boldsymbol{x}'_r)_{r=1}^R$ is a set of $R \geq 1$ independent realizations of $(\boldsymbol{x}, \boldsymbol{x}')$ under $\mathcal{L}(\mathbb{X})$,

$$\mathbb{P}_{\boldsymbol{x},\boldsymbol{x}' \sim \mathcal{L}(\mathbb{X})} \left(Y(\boldsymbol{x}) > Y(\boldsymbol{x}'), \mu(\boldsymbol{x}) < \mu(\boldsymbol{x}') \mid Y(\widehat{\boldsymbol{x}}) = \mu(\widehat{\boldsymbol{x}}) \right) \approx \frac{1}{2R} \sum_{r=1}^{R} \Phi \left(- \frac{|\mu(\boldsymbol{x}_r) - \mu(\boldsymbol{x}'_r)|}{\sigma_\star(\boldsymbol{x}_r, \boldsymbol{x}'_r, \widehat{\boldsymbol{x}})} \right),$$

$$\sigma_\star^2(\boldsymbol{x}, \boldsymbol{x}', \widehat{\boldsymbol{x}}) := C(\boldsymbol{x}, \boldsymbol{x}) + C(\boldsymbol{x}', \boldsymbol{x}') - 2C(\boldsymbol{x}, \boldsymbol{x}') - \frac{\left(C(\boldsymbol{x}, \boldsymbol{x}^\star) - C(\boldsymbol{x}', \widehat{\boldsymbol{x}})\right)^2}{C(\boldsymbol{x}^\star, \widehat{\boldsymbol{x}})}.$$

Estimation of the proposed selection criterion

If $(\boldsymbol{x}_r, \boldsymbol{x}'_r)_{r=1}^R$ is a set of $R \geq 1$ independent realizations of $(\boldsymbol{x}, \boldsymbol{x}')$ under $\mathcal{L}(\mathbb{X})$,

$$\mathbb{P}_{\boldsymbol{x},\boldsymbol{x}' \sim \mathcal{L}(\mathbb{X})} \left(Y(\boldsymbol{x}) > Y(\boldsymbol{x}'), \mu(\boldsymbol{x}) < \mu(\boldsymbol{x}') \mid Y(\widehat{\boldsymbol{x}}) = \mu(\widehat{\boldsymbol{x}})\right)$$
$$\approx \frac{1}{2R} \sum_{r=1}^{R} \Phi\left(-\frac{|\mu(\boldsymbol{x}_r) - \mu(\boldsymbol{x}'_r)|}{\sigma_\star(\boldsymbol{x}_r, \boldsymbol{x}'_r, \widehat{\boldsymbol{x}})}\right),$$

$$\sigma_\star^2(\boldsymbol{x}, \boldsymbol{x}', \widehat{\boldsymbol{x}}) := C(\boldsymbol{x}, \boldsymbol{x}) + C(\boldsymbol{x}', \boldsymbol{x}') - 2C(\boldsymbol{x}, \boldsymbol{x}') - \frac{\left(C(\boldsymbol{x}, \boldsymbol{x}^\star) - C(\boldsymbol{x}', \widehat{\boldsymbol{x}})\right)^2}{C(\boldsymbol{x}^\star, \widehat{\boldsymbol{x}})}.$$

- → the selection criterion is just a post-processing of the generated points using an explicit expression,
- → the evaluation of the criterion is also completely parallelizable.

Outline

Objectives of the application section

1. Check that bound $\widehat{q}_{\alpha,\beta}$ is higher than $p(y)$ with high probability, when $\mathbb{E}\left[p(Y)\right]$ may strongly underestimate it. To this end, we also introduce the following estimator of $\mathbb{E}\left[p(Y)\right] = \mathbb{E}\left[\Phi\left(-\frac{\mu(\boldsymbol{x})}{\sqrt{C(\boldsymbol{x},\boldsymbol{x})}}\right)\right]$:
$$\widehat{m}_{N_m} = \frac{1}{N_m}\sum_{i=1}^{N_m}\Phi\left(-\frac{\mu(\boldsymbol{x}^{(i)})}{\sqrt{C(\boldsymbol{x}^{(i)},\boldsymbol{x}^{(i)})}}\right), \quad \boldsymbol{x}^{(i)} \leftrightarrow \text{ iid copies of } \boldsymbol{x}.$$
2. Analyze the trade-off between exploration and exploitation of the proposed procedure.
3. Quantify the efficiency of the proposed selection criterion to make $\widehat{q}_{\alpha,\beta}$ converge to $p(y)$ from above.

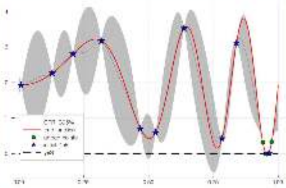
Presentation of the test functions

Example	Function y	d	$f_{\boldsymbol{x}}$	$p(y)$	L	N_{add}
1	Oscillating function	1	$1_{[0,1]}$	0.0109	10	12
2	Branin	2	$1_{[0,1]^2}$	0.00196	20	100
3	Hartmann	6	$1_{[0,1]^6}$	0.000949	300	100

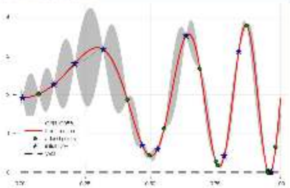
TABLE: Summary of the three test functions

- All the code evaluations of the initial DOE are strictly positive (no points in the failure domain).
- Each test function presents several local minima.
- The value of L is adapted to the dimension of $\mathbb{X}$ and to the complexity of y.
- The following results are based on 10 repetitions of the whole procedure.
- The compared selection criteria are initialized with the same DOE.

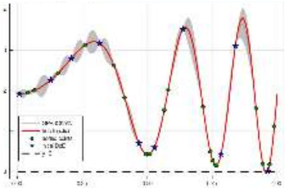
A different trade-off between exploration and exploitation



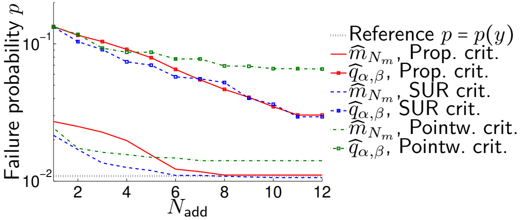
(c) Pointw., $Q^2 = 0.916$



(d) SUR, $Q^2 = 0.982$

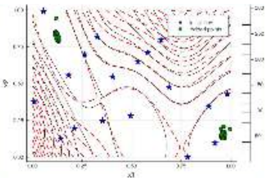


(e) Prop., $Q^2 = 0.992$

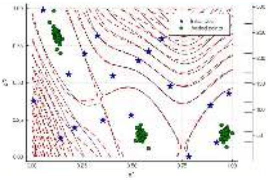


- The results associated with the pointwise crit. do not converge properly.
- SUR criterion seems to allow a quicker convergence of $\widehat{m}_{N_m}$ and $\widehat{q}_{\alpha,\beta}$ when the proposed criterion leads to a better exploration of the input space.

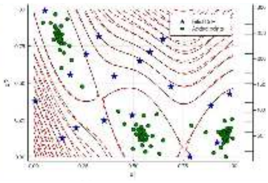
Some expected differences when analyzing the 2D case



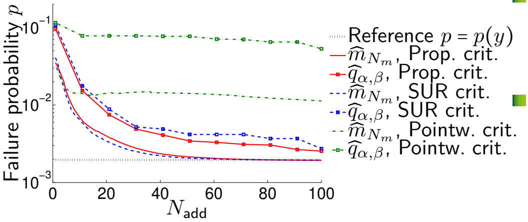
(f) Pointw., $Q^2 = 0.924$



(g) SUR, $Q^2 = 0.966$

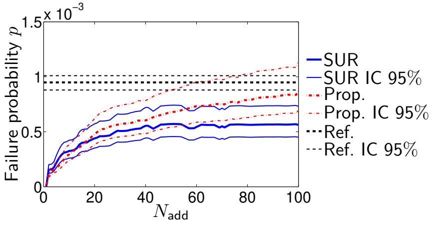


(h) Prop., $Q^2 = 0.986$

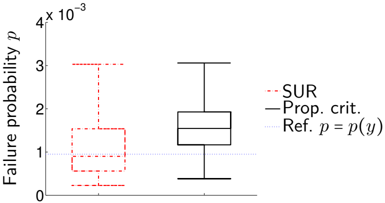


- The pointwise crit. seems to mismanage the presence of several local minima.
- When SUR crit. makes the mean quickly converge, the Prop. crit. focuses instead on the convergence of the bound.

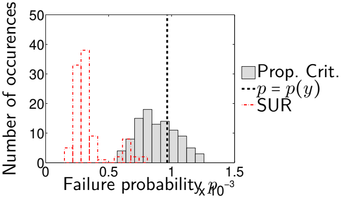
Difficulties in higher dimensions



(i) Conv. of the mean



(j) Distribution of the bound values



(k) 100 estimations of $p(y)$ after 100 iterations

- The SUR results seem to be stuck in local minima, when the Prop. approach seems to provide better estimates.
- The bound values are higher than $p(y)$ with high probability (as requested).
- Estimated values are well distributed around the true value.

Outline

Conclusions and prospects

- This presentation introduces a formalism for estimating probabilities of failure.
- This approach is based on : GPR, order statistics, a marked Poisson process.
- In order to ensure the security of systems of interest, it is proposed to focus on the estimation of quantiles rather than the mean.
- A sequential enrichment criterion particularly dedicated to the estimation method is proposed.
- One of the objectives of the method is to avoid forgetting pathological configurations in the risk analysis.

Bibliography I

Bect, J., Ginsbourger, D., Li, L., Picheny, V., and Vasquez, E. (2012).
Sequential design of computer experiments for the estimation of a probability of failure.
Statistics and Computing, 22 :773–797.

Chevalier, C., Bect, J., Ginsburger, D., Vasquez, E., Picheny, V., and Richet, Y. (2014).
Fast Kriging-based stepwise uncertainty reduction with application to the identification of an excursion set.
Technometrics, 56 (4) :455–465.

Echard, B., Gayton, N., and Lemaire, M. (2011).
AK-MCS : an active learning reliability method combining Kriging and Monte Carlo simulation.
Structural Safety, 22 :145–154.

Bibliography II

Walter, C. (2015).
Moving particles : A parallel optimal multilevel splitting method with application in quantiles estimation and meta-model based algorithms.
Structural Safety, 55 :10–25.

Thank you for your attention.