Estimating a Probability of Failure with an Adaptive Algorithm

Lucie Bernard & Arnaud Guyader (LPSM - Sorbonne Université) &
Philippe Leduc (STMicroelectronis - Tours) & Florent Malrieu (LMPT - Univ. Tours)

Framework & Objectives

Let us consider:

- A cube $A = [a, b]^d$,
- An **expensive-to-evaluate black-box** function $g : A \to \mathbb{R}$,
- A random variable $\mathbf{X} \backsim \mathbf{P_X}$ on A,
- A threshold $T \in \mathbb{R}$.

We want to provide an **estimation of the small probability of failure** p defined by

$$p = \mathbb{P}\left(g(\mathbf{X}) > T\right) = \mathbf{P_X}(\mathcal{S}),$$

where $\mathcal{S} = \{\mathbf{x} \in A,\ g(\mathbf{x}) > T\}$, **performing at most** n **evaluations of** g.

Main assumptions

1. Level set condition

We assume that

$$\lambda\left(\{\mathbf{x} \in A : |g(\mathbf{x}) - T| \le \delta\}\right) \le L\delta, \quad \delta > 0,$$

where $\lambda(E)$ is the Lebesgue measure of a set E. See [1].

2. Lipschitz condition

We assume that g is **Lipschitz with a known constant** $M > 0$, meaning that for all $(\mathbf{x}, \mathbf{x}') \in A \times A$, we have

$$|g(\mathbf{x}) - g(\mathbf{x}')| \le M \left\|\mathbf{x} - \mathbf{x}'\right\|_\infty .$$

Implementation of the algorithm

The algorithm defines a **decreasing sequence** $(\mathcal{A}_k)_{k\ge0}$, such that $\mathcal{A}_k$ is the set of all cubes Q of sidelength $2^{-k}|b-a|$ where g is **still likely to exceed** T. It also defines **increasing sequences** $(\mathcal{S}_k)_{k\ge0}$, and $(\mathcal{O}_k)_{k\ge0}$, such that $\mathcal{S}_k$ (resp. $\mathcal{O}_k$) is the set of all cubes $Q \in A$ where g is **always above** (resp **below**) T. It **stops** when n **evaluations of** g **have been performed**.

- **Step** $k = 0$

  We defined $\mathcal{A}_0 = \{A\}$ and $\mathcal{S}_0 = \mathcal{O}_0 = \{\emptyset\}$.

- **Step** $k = 1, 2, \ldots$

  For all cube $Q \in \mathcal{A}_{k-1}$, we evaluate g at the center c_Q of Q.

  ▷ We define $\tilde{\mathcal{S}}_{k-1}$ (resp. $\tilde{\mathcal{O}}_{k-1}$) as the set of all cubes $Q \in \mathcal{A}_{k-1}$ for which

  $$g(c_Q) - 2^{-k}M|b-a| > T$$

  (resp. $g(c_Q)+2^{-k}M|b-a| \le T$), that is the set of all cubes $Q \in \mathcal{A}_{k-1}$ where g is **always above** (resp. **below**) T.

  ▷ We define $\tilde{\mathcal{A}}_{k-1} = \mathcal{A}_{k-1} \setminus (\tilde{\mathcal{S}}_{k-1} \cup \tilde{\mathcal{O}}_{k-1})$, that is the set of all cubes $Q \in \mathcal{A}_{k-1}$ where g is **still likely to exceed** T.

  ▷ Then, we define:

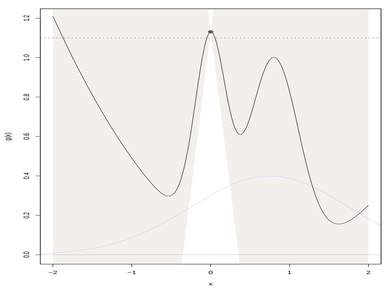
  ⋆ $\mathcal{S}_k = \{\mathcal{S}_{k-1} \cup \tilde{\mathcal{S}}_{k-1}\}$ (resp. $\mathcal{O}_k = \{\mathcal{O}_{k-1} \cup \tilde{\mathcal{O}}_{k-1}\}$) as the set of all cubes $Q \in A$ where g is **always above** (resp. **below**) T.

  ⋆ $\mathcal{A}_k$ as the set of all cubes of sidelength $2^{-k}|b-a|$, which are the children of $\tilde{\mathcal{A}}_{k-1}$, so that $\#(\mathcal{A}_k) = 2^d \#(\tilde{\mathcal{A}}_{k-1})$.

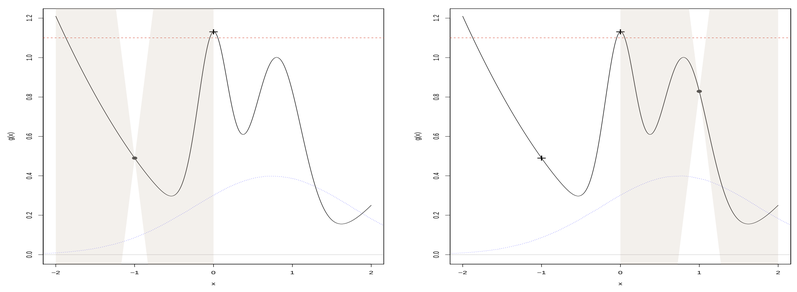
  ▷ At the end of step k, we have the **upper-bound** $p_k \ge p$ defined by:

  $$p_k = \mathbf{P_X}\left(\mathcal{A}_k\right) + \mathbf{P_X}\left(\mathcal{S}_k\right).$$

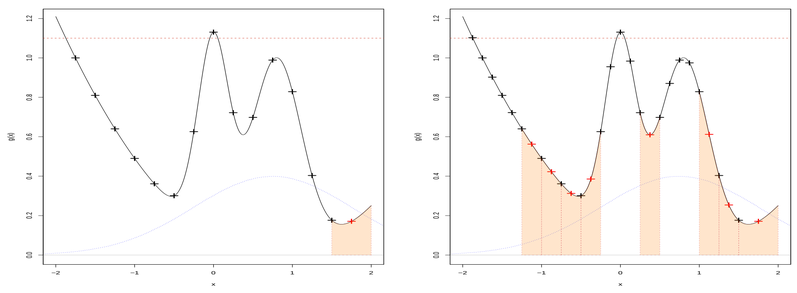
Example in one dimension



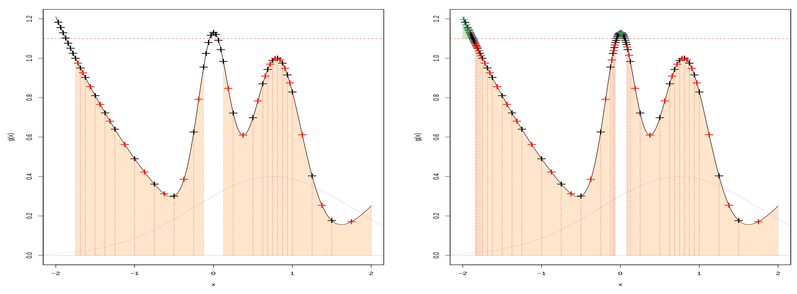
Step $k = 1$.



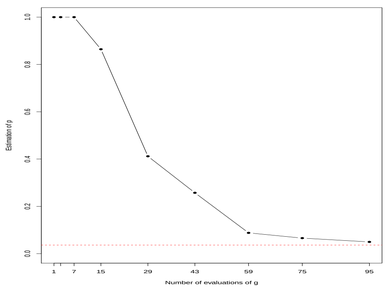
Step $k = 2$.



At the end of steps $k = 4$ and $k = 5$.



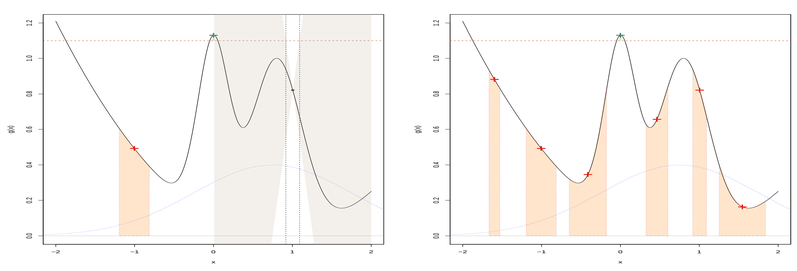
At the end of steps $k = 7$ and $k = 9$.



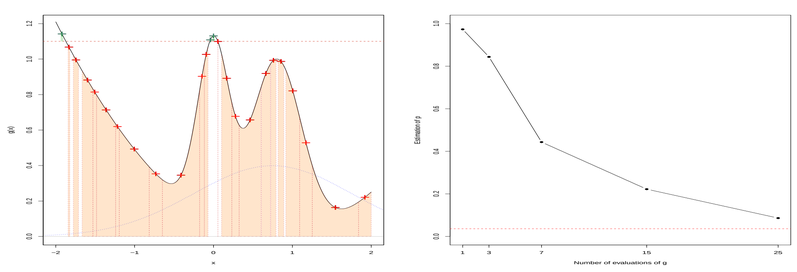
The corresponding estimations of p.

Improvement of the algorithm in practice

- At step $k \ge 1$, one can take into account of **all information** provided by the Lipschitz condition and extract, for any $Q \in \tilde{\mathcal{A}}_{k-1}$, an hyper rectangle H_Q of center c_Q, where g is always above of below T.



At step $k = 2$ and at the end of step $k = 3$.



At the end of step $k = 5$ and the corresponding estimations.

Some remarks

1. At the end of step $k \ge 2$, the number n_k of evaluations of g satifies:

$$n_k = 1 + 2^d \sum_{j=1}^{k-1} \#(\tilde{A}_{j-1})$$
$$\le \begin{cases} 1 + LM\,|b-a|\,2(k-1) & \text{if } d = 1, \\ 1 + LM\,|b-a|\,2^d(2^{d-1}-1)^{-1}2^{(k-1)(d-1)} & \text{if } d \ge 2. \end{cases}$$

2. For all $k \ge 0$, we have:

$$\mathbb{P}(\mathbf{X} \in \mathcal{S}_k) \le p \le \mathbb{P}(\mathbf{X} \in \mathcal{A}_k) + \mathbb{P}(\mathbf{X} \in \mathcal{S}_k) = p_k.$$

3. Let f_X be the density of $\mathbf{X}$ and $m = \sup_{\mathbf{x}\in A} f_X(\mathbf{x})$. For all $k \ge 1$, we have:

$$\mathbb{P}(\mathbf{X} \in \mathcal{A}_k) \le m\,|b-a|^{d+1}\,LM \times 2^{(1-k)}.$$

In case of $\mathbf{P_X}$ has a density only known up to a normalizing constant...

... we use an **Adaptive Multilevel Splitting** method, also called **Subset Simulation** (see e.g. [2, 3]) to give an **estimation** of p_k. The principle is as follows:

- For $k \ge 1$, let Q_k be a cube in $\mathcal{A}_k$ and Q_{k-1} be its parent, with $Q_0 = A$ and $\mathbb{P}(\mathbf{X} \in Q_0) = 1$.
- Since we have

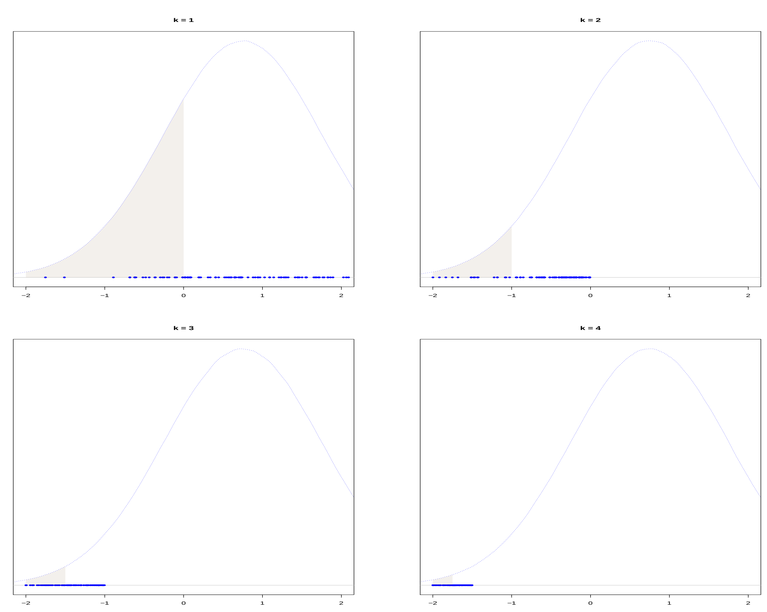
  $$\mathbb{P}(\mathbf{X} \in Q_k) = \mathbb{P}(\mathbf{X} \in Q_k | \mathbf{X} \in Q_{k-1})\mathbb{P}(\mathbf{X} \in Q_{k-1}),$$

  then the probability $\mathbb{P}(\mathbf{X} \in Q_k)$ can be easily estimated:

  ▷ The law $\mathbf{P_X}$ has a density which is known up to a normalizing constant, so that **Metropolis-Hasting algorithm techniques** can be applied to draw a **Monte Carlo sample** $\mathbf{X}_1, \ldots, \mathbf{X}_N$ from the restriction of $\mathbf{P_X}$ to Q_{k-1} (i.e. from $\mathcal{L}(\mathbf{X}|\mathbf{X} \in Q_{k-1})$),

  ▷ At the beginning of the step $k \ge 1$, the probability $\mathbb{P}(\mathbf{X} \in Q_{k-1})$ has already been estimated.

Finally, the **precision** of the estimation of p only depends on the budget n and the size N of the Monte Carlo sample.



- Grey area: The target probability $\mathbb{P}(\mathbf{X} \in Q_k)$.
- Blue points: The Monte Carlo sample used to estimate the probability $\mathbb{P}(\mathbf{X} \in Q_k | \mathbf{X} \in Q_{k-1})$.

References

[1] A. Cohen, R. Devore, G. Petrova, and P. Wojtaszczyk. Finding the minimum of a function. *Methods Appl. Anal.*, 20(4):365–381, 2013.

[2] S.K. Au and J.L. Beck. Estimation of small failure probabilities in high dimensions by subset simulation. *Journal of Probabilistic Engineering Mechanics*, Vol. 16:263–277, 2001.

[3] F. Cérou, P. Del Moral, T. Furon, and A. Guyader. Sequential Monte Carlo for rare event estimation. *Stat. Comput.*, 22(3):795–808, 2012.