Sequential Bayesian inversion of black-box functions in presence of uncertainties

Laboratoire Signaux & Systèmes

Romain Ait Abdelmalek-Lomenech, Julien Bect, Emmanuel Vazquez

Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des Signaux et Systèmes

Context & motivations

In engineering, many system design problems come down to finding design parameters for the system to operate under desired conditions. When time-consuming numerical simulations are used to assess the conditions of operation of such systems, it is essential to treat this problem –broadly known as **"inversion"**– using search methods that are very sparing in the number of simulations

With this constraint in mind, we propose here a **sequential strategy** for a particular robust inversion problem that we call **"reliability-based inversion"**, in which the system has both **deterministic and uncertain inputs**. In this formulation, the objective is to retrieve the set of deterministic inputs such that the probability of the outputs exceeding a given level is greater than a threshold.

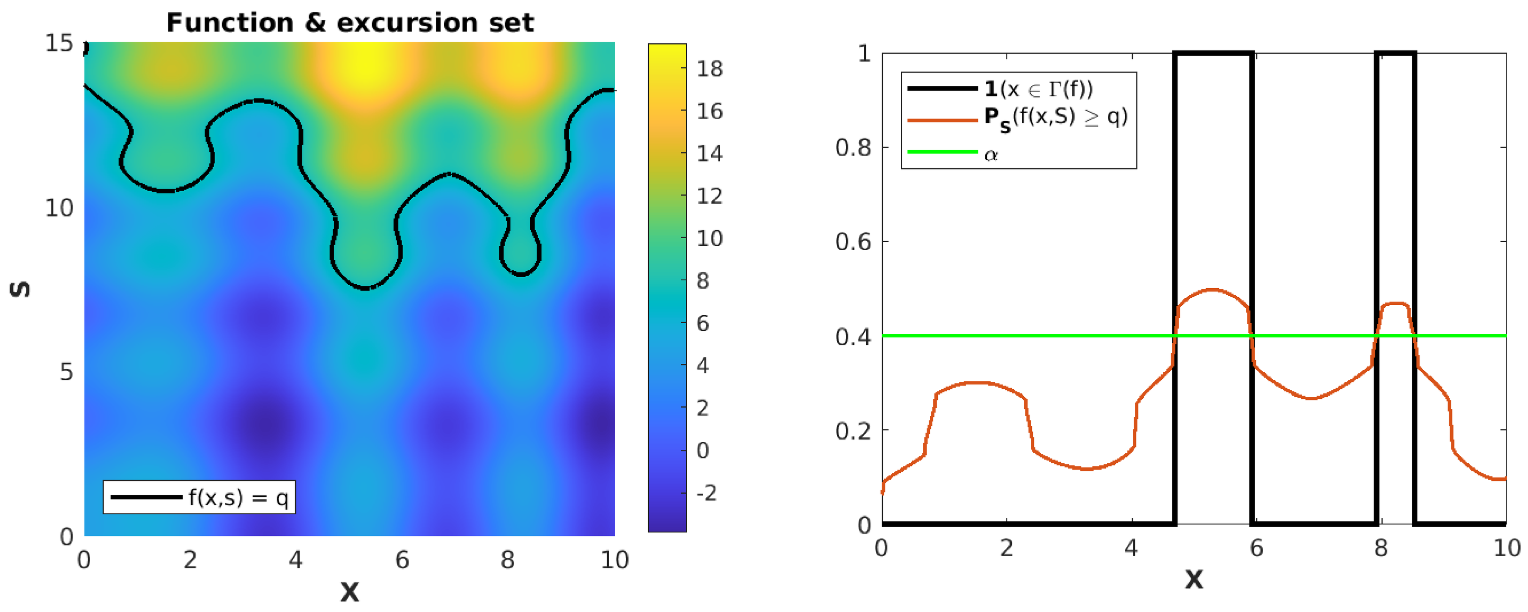
The reliability-based inversion (RBI) problem

We view our system as a **black-box function** $f\ :\ \mathbb{X}\times\mathbb{S}\longmapsto\mathbb{R}$. In this setting, $\mathbb{X}$ is the space of deterministic inputs and $\mathbb{S}$ the space of **uncertain** inputs.

- Given a level $q\in\mathbb{R}$ and a threshold $\alpha\in]0,1[$, **our objective is to approximate the set**

$$\Gamma(f)=\{x\in\mathbb{X}\ :\ \mathbb{P}(f(x,S)\geq q)\geq\alpha\}$$

where S is a random variable with **known distribution** $\mathbb{P}_S$ on $\mathbb{S}$.



- Equivalently, we can formulate this problem as a **classification problem**, in which we wish to **approximate the classifier**

$$c(x)\quad=\quad\begin{cases}1 & \text{if}\quad x\in\Gamma(f)\\ 0 & \text{otherwise}\end{cases}$$

The Stepwise Uncertainty Reduction (SUR) principle

SUR strategies are a subset of the Bayesian strategies in which the points of evaluation are sequentially chosen by minimizing the expected future uncertainty on the object of interest, using a probabilistic surrogate of the unknown function.

- Consider $\xi\sim GP(\mu,k)$ a **Gaussian Process prior** on f, with μ and k respectively the mean and covariance functions of the process.

- We define the **current information** $\mathcal{I}_n$:

$$\mathcal{I}_n=\{(X_1,S_1,\xi\left(X_1,S_1\right)),\ \ldots,(X_n,S_n,\xi\left(X_n,S_n\right))\}$$

- A SUR strategy choose the point **minimizing the expected future uncertainty**:

$$(X_{n+1},S_{n+1})\in\underset{(x,s)\in\mathbb{X}\times\mathbb{S}}{\arg\min}\,\mathbb{E}_n\left(U_{n+1}\mid X_{n+1}=x,\ S_{n+1}=s\right)$$

where:

- U_{n+1} is an uncertainty metric on the object of interest, depending on $\mathcal{I}_{n+1}$
- $\mathbb{E}_n$ expectation with respect to $\mathbb{P}_n$, the distribution of ξ given $\mathcal{I}_n$

SUR approaches have been successfully used for the resolution of several problems regarding expensive-to-evaluate black box, including optimization [VVW09; MTZ78], function fitting, and more standard inversion problems. [Bec+12]

Construction of a SUR criterion for the RBI problem

- Let us define the **integrated process** τ and $\Gamma(\xi)$ the associated **random excursion set** above the threshold α:

$$\tau(\xi,x)=\int_{\mathbb{S}}\mathbb{1}(\xi(x,s)\geq q)\mathbb{P}_S(ds)\quad\text{ and }\quad\Gamma(\xi)=\{x\in\mathbb{X},\ \tau(\xi,x)\geq\alpha\}$$

- We consider the uncertainty metric U_n defined by

$$U_n=\int_{\mathbb{X}}\tilde{p}_n(x)(1-\tilde{p}_n(x))dx$$

with $\tilde{p}_n(x)=\mathbb{P}_n(\tau(\xi,x)\geq\alpha)$

The link between the metric and the **uncertainty of the random set** $\Gamma(\xi)$ can be seen more clearly by remarking that:

$$U_n=\int_{\mathbb{X}}\mathbb{V}_n\left(\mathbb{1}(x\in\Gamma(\xi)\right)dx$$

with $\mathbb{V}_n$ the conditional variance given $\mathcal{I}_n$

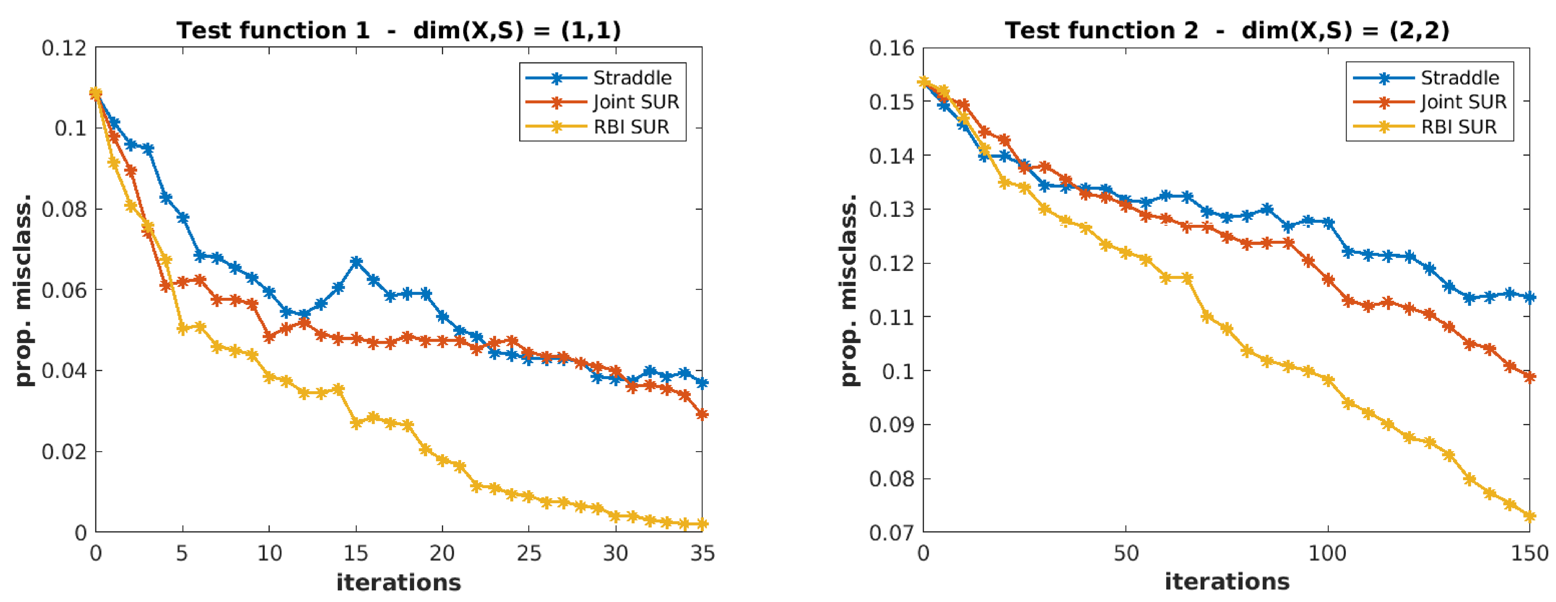
- The conditional distributions of $\tau(\xi,x)$ are **intractable**: the SUR criterion does not admit an analytic expression.
- **Monte Carlo simulations**: ξ being Gaussian, we can generate $\xi_1,\ldots,\xi_M$ sample paths of ξ under $\mathbb{P}_n$

$$\frac{1}{M}\sum_{m=1}^{M}\mathbb{1}(\tau(\xi_m,x)\geq\alpha)\underset{M\to+\infty}{\xrightarrow{a.s}}\mathbb{P}_n(\tau(\xi,x)\geq\alpha)=\tilde{p}_n(x)$$

For each candidate point (x,s), the sample paths are **reconditioned** according to values sampled from the distribution of $\xi(x,s)$ given $\mathcal{I}_n$

Numerical results

We compare, iteration by iteration, the proportion of misclassified points obtained with our method to the ones of methods focusing on the excursion set in the joint space $\mathbb{X}\times\mathbb{S}$ of design and uncertainties: the **straddle criterion** [Bry+05] and a **SUR approach** to standard inversion problems. [Bec+12]



Empirical evidences show that our method **generally outperforms** state of the art joint space inversion methods. Moreover, in all the test cases, the new SUR strategy never lags behind the others approaches considered.

References

[Bec+12] J. Bect et al. "Sequential design of computer experiments for the estimation of a probability of failure". In: *Statistics and Computing* 22.3 (2012), pp. 773–793.

[Bry+05] B. Bryan et al. "Active Learning for Identifying Function Threshold Boundaries". In: *Proceedings of the 18th International Conference on Neural Information Processing Systems*. NIPS'05. 2005.

[MTZ78] J. Mockus, V. Tiesis, and A. Zilinskas. "The Application of Bayesian Methods for Seeking the Extremum". In: *Towards Global Optimization* 2.117-129 (1978), p. 2.

[VVW09] J. Villemonteix, E. Vazquez, and E. Walter. "An Informational Approach to the Global Optimization of Expensive-to-Evaluate Functions". In: *J. of Global Optimization* 44.4 (Aug. 2009), pp. 509–534.

This work has been funded by the National French Research Agency (ANR) in the context of the project SAMOURAI (ANR-20-CE46-0013).

CentraleSupélec | université PARIS-SACLAY | ANR | SAMOURAI