A stepwise uncertainty reduction strategy for the estimation of small quantile sets

Romain Ait Abdelmalek-Lomenech,
Julien Bect & Emmanuel Vazquez

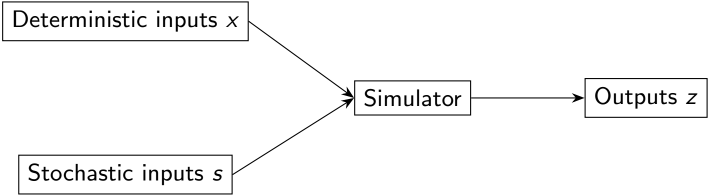
Université Paris-Saclay, CNRS, CentraleSupélec, L2S

MASCOT-NUM 2024, Hyères,
April 3rd, 2024

Table of Contents

Consider an **expensive-to-evaluate** numerical simulator, with inputs:

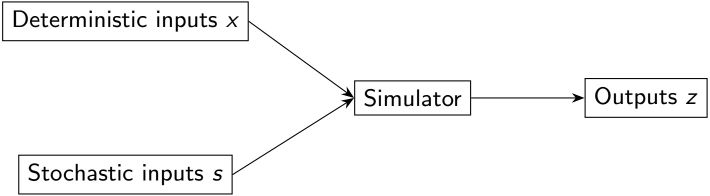
- $x \in \mathbb{X}$ (deterministic design choices).
- $s \in \mathbb{S}$ (stochastic factors).



For simplicity we assume a deterministic simulator $f : \mathbb{X} \times \mathbb{S} \mapsto \mathbb{R}^q$.

Consider an **expensive-to-evaluate** numerical simulator, with inputs:

- $x \in \mathbb{X}$ (deterministic design choices).
- $s \in \mathbb{S}$ (stochastic factors).



For simplicity we assume a deterministic simulator $f : \mathbb{X} \times \mathbb{S} \mapsto \mathbb{R}^q$.

Given:

- $C \subset \mathbb{R}^q$ is a critical/failure region.
- $\alpha \in (0, 1)$ a threshold.
- $\mathbb{P}_S$ a known distribution on $\mathbb{S}$.

We focus on the **quantile set inversion (QSI)** problem:

> **Estimate the set** of all $x \in \mathbb{X}$ such that the system is **robust to uncertainties**, i.e
>
> $$\mathbb{P}\left(f(x, S) \in C\right) \leq \alpha, \qquad S \sim \mathbb{P}_S,$$

by only using a small number N of evaluation points

$$\{(X_1, S_1), \ldots, (X_N, S_N)\}.$$

Given:

- $C \subset \mathbb{R}^q$ is a critical/failure region.
- $\alpha \in (0, 1)$ a threshold.
- $\mathbb{P}_S$ a known distribution on $\mathbb{S}$.

We focus on the **quantile set inversion (QSI)** problem:

Estimate the set of all $x \in \mathbb{X}$ such that the system is **robust to uncertainties**, i.e

$$\mathbb{P}\left(f(x, S) \in C\right) \leq \alpha, \qquad S \sim \mathbb{P}_S,$$

by only using a small number N of evaluation points

$$\{(X_1, S_1), \ldots, (X_N, S_N)\}.$$

Given:

- $C \subset \mathbb{R}^q$ is a critical/failure region.
- $\alpha \in (0, 1)$ a threshold.
- $\mathbb{P}_S$ a known distribution on $\mathbb{S}$.

We focus on the **quantile set inversion (QSI)** problem:

> **Estimate the set** of all $x \in \mathbb{X}$ such that the system is **robust to uncertainties**, i.e
>
> $$\mathbb{P}\left(f(x, S) \in C\right) \leq \alpha, \qquad S \sim \mathbb{P}_S,$$

by only using a small number N of evaluation points

$$\{(X_1, S_1), \ldots, (X_N, S_N)\}.$$

Estimate the quantile set:

$$\Gamma(f) = \{x \in \mathbb{X} : \mathbb{P}\left(f(x, S) \in C\right) \leq \alpha\},$$

Example of function and associated quantile set, with $C = (-\infty, 7.5]$ and $\alpha = 5\%$.

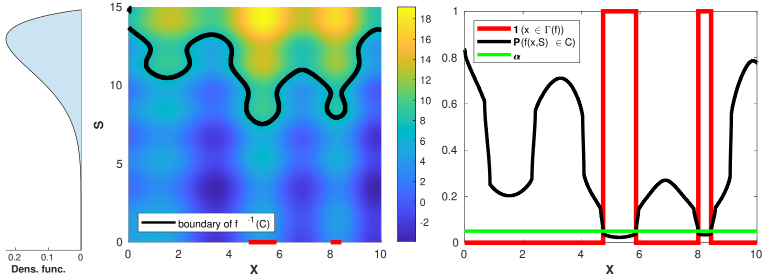


Figure: Representation of the function (middle), the density of $\mathbb{P}_S$ (left) and associated quantile set (right).

Table of Contents

The QSI problem is related to the **estimation of the excursion set**

$$\gamma(f) = \{(x, s) \in \mathbb{X} \times \mathbb{S} \, : \, f(x, s) \notin C\}$$

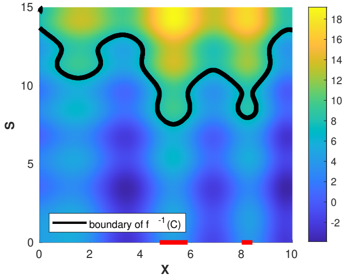


Figure: Example function. The black line delimits the set $\gamma(f)$.

Knowing $\gamma(f) \implies$ knowing $\Gamma(f)$.

Indeed, $\Gamma(f) = \{x \in \mathbb{X} \, : \, \mathbb{P}((x, S) \in \gamma(f)) > 1 - \alpha\}$.

The QSI problem is related to the **estimation of the excursion set**

$$\gamma(f) = \{(x, s) \in \mathbb{X} \times \mathbb{S} \, : \, f(x, s) \notin C\}$$

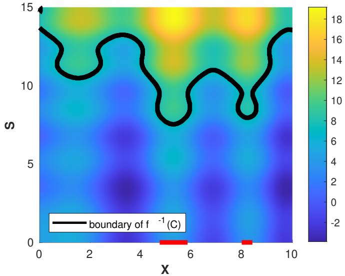


Figure: Example function. The black line delimits the set $\gamma(f)$.

Knowing $\gamma(f) \implies$ knowing $\Gamma(f)$.

Indeed, $\Gamma(f) = \{x \in \mathbb{X} \, : \, \mathbb{P}((x, S) \in \gamma(f)) > 1 - \alpha\}$.

Bayesian approach: Consider $\xi \sim \text{GP}(\mu, k)$ a prior on f. We denote:

- $\mathbb{P}_n$ the distribution of ξ given $\{(X_i, S_i, f(X_i, S_i)),\ i \leq n\}$.
- $\mathbb{E}_n$ the expectation w.r.t. $\mathbb{P}_n$.
- $p_n(x, s) = \mathbb{P}_n(\xi(x, s) \notin C)$ the cond. probability of $(x, s) \in \gamma(\xi)$, with $\gamma(\xi)$ the random excursion set associated to ξ.

Several Bayesian methods focus on **estimating** $\gamma(f)$. For example:

- **Maximal uncertainty sampling methods**:
  - Maximum misclassification probability [Bryan et al. (2005)]:

$$(X_{n+1}, S_{n+1}) \in \underset{(x,s) \in \mathbb{X} \times \mathbb{S}}{\text{argmax}} \min(p_n(x, s), 1 - p_n(x, s))$$

  - [Ranjan et al. (2008); Echard et al. (2011), ...]

Bayesian approach: Consider $\xi \sim \mathrm{GP}(\mu, k)$ a prior on f. We denote:

- $\mathbb{P}_n$ the distribution of ξ given $\{(X_i, S_i, f(X_i, S_i)),\ i \leq n\}$.
- $\mathbb{E}_n$ the expectation w.r.t. $\mathbb{P}_n$.
- $p_n(x, s) = \mathbb{P}_n(\xi(x, s) \notin C)$ the cond. probability of $(x, s) \in \gamma(\xi)$, with $\gamma(\xi)$ the random excursion set associated to ξ.

Several Bayesian methods focus on **estimating $\gamma(f)$**. For example:

- **Maximal uncertainty sampling methods**:
  - Maximum misclassification probability [Bryan et al. (2005)]:

$$(X_{n+1}, S_{n+1}) \in \underset{(x,s)\in\mathbb{X}\times\mathbb{S}}{\operatorname{argmax}} \min(p_n(x, s), 1 - p_n(x, s))$$

  - [Ranjan et al. (2008); Echard et al. (2011), ...]

- **Stepwise uncertainty reduction (SUR)** methods:
  - For instance [Bect et al. (2012); Chevalier et al. (2014)]:

$$(X_{n+1}, S_{n+1}) \in \underset{(x,s)\in\mathbb{X}\times\mathbb{S}}{\operatorname{argmin}} \mathbb{E}_n(\mathcal{H}_{n+1} \mid (X_{n+1}, S_{n+1}) = (x, s))$$

    with $\mathcal{H}_n = \int_{\mathbb{X}\times\mathbb{S}} \min(p_n(x,s), 1 - p_n(x,s))\, \mathrm{d}x\mathrm{d}s$.

  - [Picheny et al. (2010); Marques et al. (2018), ...]

- **Stepwise uncertainty reduction (SUR)** methods:
  - For instance [Bect et al. (2012); Chevalier et al. (2014)]:

$$(X_{n+1}, S_{n+1}) \in \underset{(x,s)\in\mathbb{X}\times\mathbb{S}}{\operatorname{argmin}} \ \mathbb{E}_n(\mathcal{H}_{n+1} \mid (X_{n+1}, S_{n+1}) = (x, s))$$

with $\mathcal{H}_n = \int_{\mathbb{X}\times\mathbb{S}} \min(p_n(x,s), 1 - p_n(x,s))\,\mathrm{d}x\mathrm{d}s$.

  - [Picheny et al. (2010); Marques et al. (2018), ...]

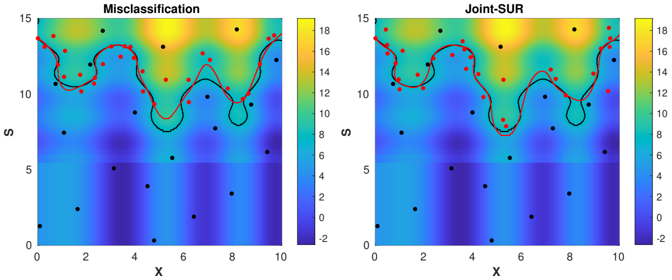


Figure: Examples of designs (red dots) obtained after $n = 30$ steps with the maximum misclassification and the 'joint-SUR' criteria.

To estimate $\Gamma(f)$, one only needs to focus on **'interesting parts'** of $\gamma(f)$.

We denote:

- $\Gamma(\xi)$ the random quantile set associated to ξ.
- $\pi_n(x) = \mathbb{P}_n(x \in \Gamma(\xi))$,
- $\mathcal{Q}_n = \int_{\mathbb{X}} \min(\pi_n(x), 1 - \pi_n(x))\, \mathrm{d}x$.

QSI-SUR sampling criterion [Ait Abdelmalek-Lomenech et al. (2023)]:

$$(X_{n+1}, S_{n+1}) \in \underset{(x,s)\in\mathbb{X}\times\mathbb{S}}{\operatorname{argmin}} \mathbb{E}_n(\mathcal{Q}_{n+1} \mid (X_{n+1}, S_{n+1}) = (x, s)),$$

To estimate $\Gamma(f)$, one only needs to focus on **'interesting parts'** of $\gamma(f)$.

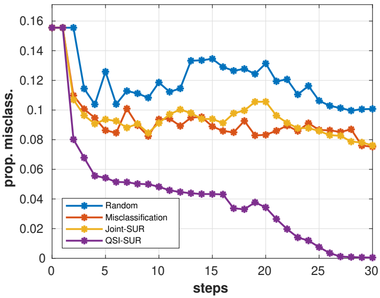
We denote:

- $\Gamma(\xi)$ the random quantile set associated to ξ.
- $\pi_n(x) = \mathbb{P}_n(x \in \Gamma(\xi))$,
- $\mathcal{Q}_n = \int_{\mathbb{X}} \min(\pi_n(x), 1 - \pi_n(x))\,\mathrm{d}x$.

QSI-SUR sampling criterion [Ait Abdelmalek-Lomenech et al. (2023)]:

$$(X_{n+1}, S_{n+1}) \in \underset{(x,s)\in\mathbb{X}\times\mathbb{S}}{\operatorname{argmin}}\ \mathbb{E}_n(\mathcal{Q}_{n+1} \mid (X_{n+1}, S_{n+1}) = (x, s)),$$

The implementation proposed in [Ait Abdelmalek-Lomenech et al. (2023)] produces good results on moderately difficult examples.



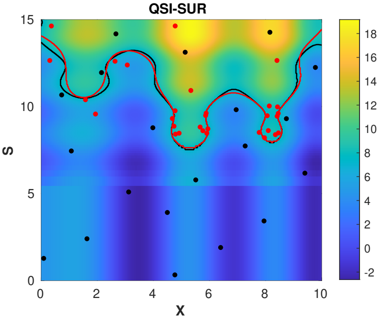


Figure: Median of the proportion of misclassified points vs. number of iterations (left). Example of design obtained (right).

The QSI-SUR criterion is based on

$$\int_{\mathbb{X}} \min(\pi_n(x), 1 - \pi_n(x)) \, \mathrm{d}x.$$

- In practice, both the integral involved **and** the optimization of the criterion are **discretized**.
- Necessity of a collection of points $x \in \mathbb{X}$ such that their **probability of misclassification** is non-null.

Main issue: If $\Gamma(f)$ is 'small' relatively to $\mathbb{X}$, difficulty to sample relevant points in the set $\mathbb{X}$.

The QSI-SUR criterion is based on

$$\int_{\mathbb{X}} \min(\pi_n(x), 1 - \pi_n(x)) \, \mathrm{d}x.$$

- In practice, both the integral involved **and** the optimization of the criterion are **discretized**.
- Necessity of a collection of points $x \in \mathbb{X}$ such that their **probability of misclassification** is non-null.

Main issue: If $\Gamma(f)$ is 'small' relatively to $\mathbb{X}$, difficulty to sample relevant points in the set $\mathbb{X}$.

Table of Contents

Idea: Multilevel splitting/**subset simulation** [Kahn and Harris (1951); Au and Beck (2001)] to efficiently sample points in $\mathbb{X}$.

- Sequentially estimate a sequence of **decreasing quantile sets**

$$\Gamma_0(f) \supset \Gamma_1(f) \supset \ \dots \ \supset \Gamma_K(f) = \Gamma(f),$$

using a QSI-SUR criterion.

- Such sets can be defined by setting

$$\Gamma_k(f) = \{x \in \mathbb{X} \, : \, \mathbb{P}(f(x, S) \in C_k) \leq \alpha_k\},$$

with $\alpha_k \geq \alpha_{k+1}$ and $C_k \subset C_{k+1}$.

Idea: Multilevel splitting/**subset simulation** [Kahn and Harris (1951); Au and Beck (2001)] to efficiently sample points in $\mathbb{X}$.

- Sequentially estimate a sequence of **decreasing quantile sets**

$$\Gamma_0(f) \supset \Gamma_1(f) \supset \ldots \supset \Gamma_K(f) = \Gamma(f),$$

using a QSI-SUR criterion.

- Such sets can be defined by setting

$$\Gamma_k(f) = \{x \in \mathbb{X} : \mathbb{P}(f(x, S) \in C_k) \leq \alpha_k\},$$

with $\alpha_k \geq \alpha_{k+1}$ and $C_k \subset C_{k+1}$.

We now assume $C = (-\infty, T]$.

We propose a **SMC-based** algorithm inspired by **BSS** [Li (2012); Bect et al. (2017)]

It alternates two distinct phases:

- **Estimation phase**
  - Define a new intermediary quantile set to estimate.
  - Sample points (X_n, S_n) using a QSI-SUR criterion.
- **Move phase**
  - Concentrate the particles towards the previously estimated set.

We now assume $C = (-\infty, T]$.

We propose a **SMC-based** algorithm inspired by **BSS** [Li (2012); Bect et al. (2017)]

It alternates two distinct phases:

- **Estimation phase**
  - Define a new intermediary quantile set to estimate.
  - Sample points (X_n, S_n) using a QSI-SUR criterion.
- **Move phase**
  - Concentrate the particles towards the previously estimated set.

Let $q_{n,k}$ a density targeting $\Gamma_k(f) = \{x \in \mathbb{X} \,:\, \mathbb{P}(f(x,S) \in C_k) \leq \alpha_k\}$ at step n.

Estimation phase:

- Set C_{k+1} and α_{k+1} such that

$$\mathrm{ESS}\left(\frac{q_{n,k+1}}{q_{n,k}}(x)\right) \approxeq 30\%.$$

- Sample points

$$(X_n, S_n) \in \operatorname{argmin} J_n(x,s),$$

with J_n a QSI-SUR criterion targeting $\Gamma_{k+1}(f)$.

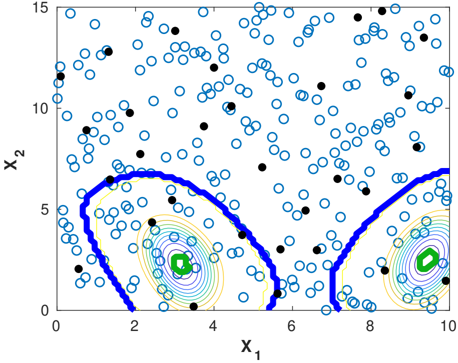


Figure: Temporary quantile set (blue line), final quantile set (green line), particles (blue dots). - $n = 0$.

Let $q_{n,k}$ a density targeting $\Gamma_k(f) = \{x \in \mathbb{X} : \mathbb{P}(f(x, S) \in C_k) \leq \alpha_k\}$ at step n.

Estimation phase:

- Set C_{k+1} and α_{k+1} such that

$$\mathrm{ESS}\left(\frac{q_{n,k+1}}{q_{n,k}}(x)\right) \approxeq 30\%.$$

- Sample points

$$(X_n, S_n) \in \operatorname{argmin} J_n(x, s),$$

with J_n a QSI-SUR criterion targeting $\Gamma_{k+1}(f)$.

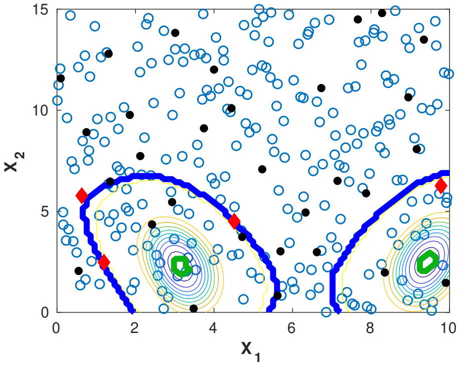


Figure: Temporary quantile set (blue line), final quantile set (green line), particles (blue dots) and projection of the sequential design (red dots). - $n = 4$.

Let $q_{n,k}$ a density targeting $\Gamma_k(f) = \{x \in \mathbb{X} : \mathbb{P}(f(x,S) \in C_k) \leq \alpha_k\}$ at step n.

Move phase:

When stopping condition is met:

- Residual resampling.
- Move particles to $\Gamma_{k+1}(f)$ using MHRW with target density $q_{n,k+1}$.
- Adapt walk's variance to target acceptation rate 25%.

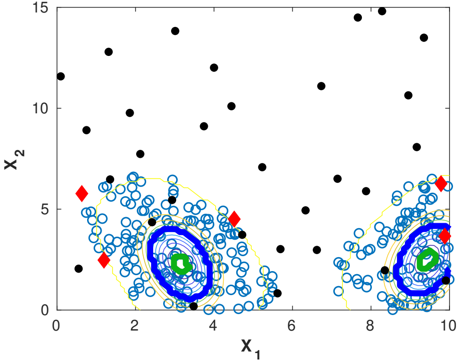


Figure: Temporary quantile set (blue line), final quantile set (green line), particles (blue dots) and projection of the sequential design (red dots). - $n = 5$.

Choice of the target densities:

Natural idea (in the spirit of [Dubourg et al. (2013); Bect et al. (2017)]):

$$q_{n,k}(x) \propto \pi_n^k(x) = \mathbb{P}_n(x \in \Gamma_k(\xi))$$

- Does not admit a closed-form expression.
- Expensive to estimate.

Idea: Replace $\pi_n^k(x)$ by $\mathbb{1}(x \in \Gamma_{n,k}^+)$. How to define $\Gamma_{n,k}^+$?

Given $x_0 \in \mathbb{X}$, μ_n and σ_n the posterior mean and standard deviation of ξ and $\beta \sim 1$, consider the **quantile function**:

$$\xi_n^+(x_0, \cdot) = \mu_n(x_0, \cdot) + \Phi^{-1}(\beta)\sigma_n(x_0, \cdot),$$

Choice of the target densities:

Natural idea (in the spirit of [Dubourg et al. (2013); Bect et al. (2017)]):

$$q_{n,k}(x) \propto \pi_n^k(x) = \mathbb{P}_n(x \in \Gamma_k(\xi))$$

- Does not admit a closed-form expression.
- Expensive to estimate.

Idea: Replace $\pi_n^k(x)$ by $\mathbb{1}(x \in \Gamma_{n,k}^+)$. How to define $\Gamma_{n,k}^+$?

Given $x_0 \in \mathbb{X}$, μ_n and σ_n the posterior mean and standard deviation of ξ and $\beta \sim 1$, consider the **quantile function**:

$$\xi_n^+(x_0, \cdot) = \mu_n(x_0, \cdot) + \Phi^{-1}(\beta)\sigma_n(x_0, \cdot),$$

$C = (-\infty, T]$ and $\xi(x_0, \cdot)$ is a **high quantile**

- $\mathbb{P}(\xi_n^+(x_0, S) \in C_k)$ is an **optimistic estimation of the probability of failure** at point x_0.

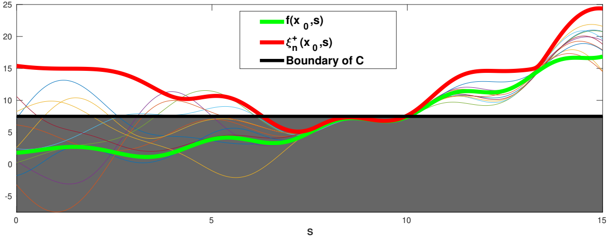


Figure: Example of quantile function $\xi_n^+(x_0, \cdot)$, with a fixed x_0.

Setting $\Gamma_{n,k}^+ = \Gamma_k(\xi_n^+)$ eliminates x_0 if $\{x_0 \in \Gamma_k(\xi)\}$ is **very improbable**.

We define the target densities as

$$q_{n,k}(x) \propto \mathbb{1}(x \in \Gamma_k(\xi_n^+))$$

NB: The MHRW step becomes a constrained random walk.

$C = (-\infty, T]$ and $\xi(x_0, \cdot)$ is a **high quantile**

- $\mathbb{P}(\xi_n^+(x_0, S) \in C_k)$ is an **optimistic estimation of the probability of failure** at point x_0.

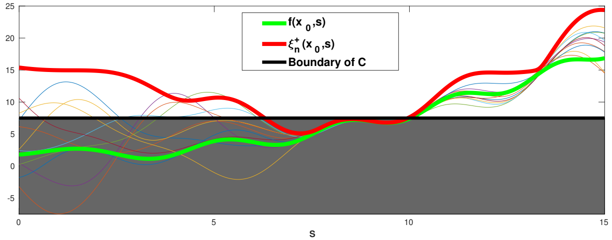


Figure: Example of quantile function $\xi_n^+(x_0, \cdot)$, with a fixed x_0.

Setting $\Gamma_{n,k}^+ = \Gamma_k(\xi_n^+)$ eliminates x_0 if $\{x_0 \in \Gamma_k(\xi)\}$ is **very improbable**.

We define the target densities as

$$q_{n,k}(x) \propto \mathbb{1}(x \in \Gamma_k(\xi_n^+))$$

NB: The MHRW step becomes a constrained random walk.

$C = (-\infty, T]$ and $\xi(x_0, \cdot)$ is a **high quantile**

- $\mathbb{P}(\xi_n^+(x_0, S) \in C_k)$ is an **optimistic estimation of the probability of failure** at point x_0.

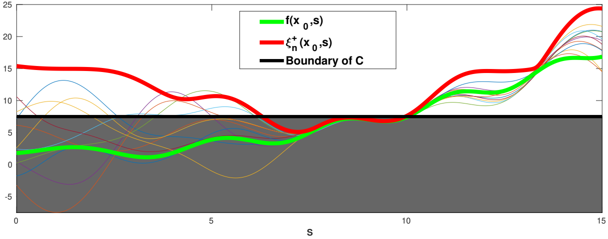


Figure: Example of quantile function $\xi_n^+(x_0, \cdot)$, with a fixed x_0.

Setting $\Gamma_{n,k}^+ = \Gamma_k(\xi_n^+)$ eliminates x_0 if $\{x_0 \in \Gamma_k(\xi)\}$ is **very improbable**.

We define the target densities as

$$q_{n,k}(x) \propto \mathbb{1}(x \in \Gamma_k(\xi_n^+))$$

NB: The MHRW step becomes a constrained random walk.

Table of Contents

For illustration purposes, we take interest in two examples functions of the form

$$f_i(x, s) = g_i(x_1, x_2) + s, \qquad i = 1, 2.$$

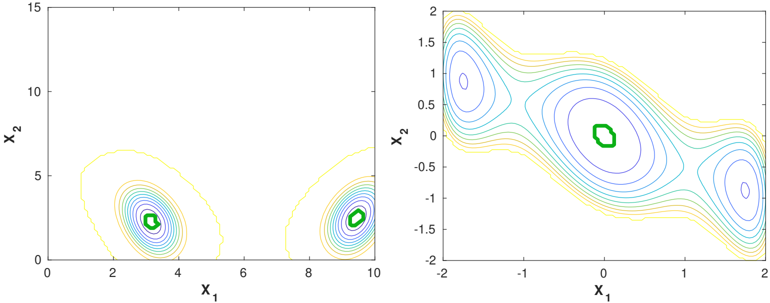


Figure: Representation of $\Gamma(f_1)$ (left - green curve) and $\Gamma(f_2)$ (right - green curve).

Relative size of the quantile sets:

$$\lambda_{\mathbb{X}}(\Gamma(f_1)) = 0.0035 \text{ and } \lambda_{\mathbb{X}}(\Gamma(f_2)) = 0.0039.$$

We can first observe that the strategy indeed concentrates the particles and sample relevant points.

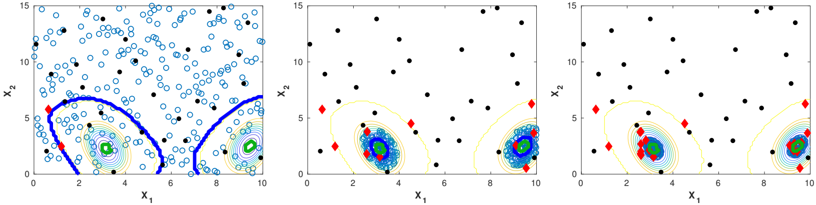


Figure: Temporary quantile set (blue line), final quantile set (green line), particles (blue dots), projections of the initial design (black dots) and sequential design (red dots). - $n = 2, 10, 20$.

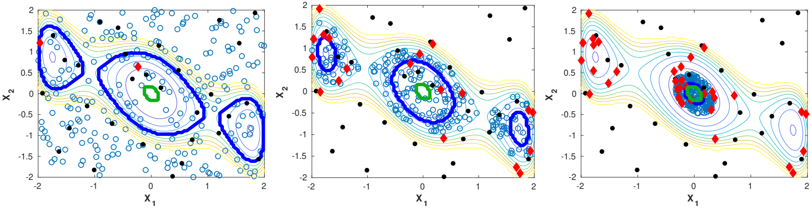


Figure: Temporary quantile set (blue line), final quantile set (green line), particles (blue dots), projections of the initial design (black dots) and sequential design (red dots). - $n = 2, 15, 35$.

We compare the accuracy of the estimation obtained by our method against BSS, which focus on the estimation of the joint excursion set

$$\gamma(f) = \{(x,s) \in \mathbb{X} \times \mathbb{S} \;:\; f(x,s) \notin C\}$$

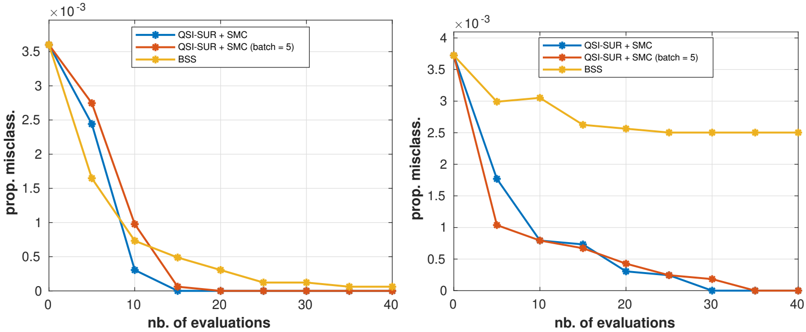


Figure: Median of the proportion of misclassified points vs. number of evaluations (initial design excluded).

The results obtained are **at least** similar to BSS.

In some difficult cases, the necessity of estimating several intermediary quantile sets before focusing on $\Gamma(f)$ leads to slow convergence.

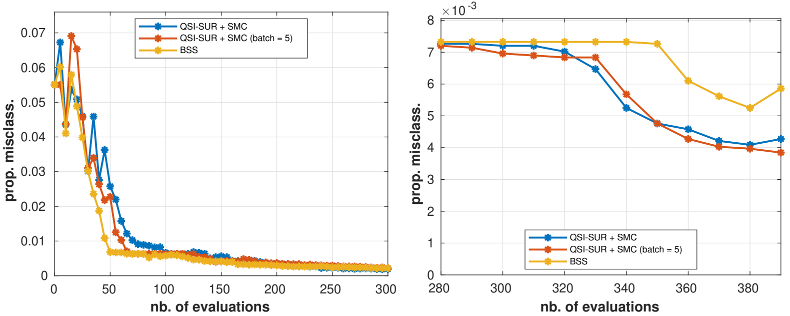


Figure: Median of the proportion of misclassified points vs. number of evaluations on **two other test functions f_3 and f_4**, with $\lambda_{\mathbb{X}}(\Gamma(f_3)) = 0.0058$ and $\lambda_{\mathbb{X}}(\Gamma(f_4)) = 0.007$ (initial design excluded).

Table of Contents

Conclusion:

- The proposed method allows to accurately estimate small quantile sets.
- The target densities chosen, although simple, concentrate efficiently the particles in $\mathbb{X}$ towards regions of interest.
- Batch sequential designs can also be obtained by adapting the criterion.

- However, this strategy remains computationally complex.
- For now, the QSI-SUR criterion is not adapted to threshold $\alpha \sim 0$.

Conclusion:

- The proposed method allows to accurately estimate small quantile sets.
- The target densities chosen, although simple, concentrate efficiently the particles in $\mathbb{X}$ towards regions of interest.
- Batch sequential designs can also be obtained by adapting the criterion.

- However, this strategy remains computationally complex.
- For now, the QSI-SUR criterion is not adapted to threshold $\alpha \sim 0$.

Thank you for your attention!

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

References

Ait Abdelmalek-Lomenech, R., Bect, J., Chabridon, V., and Vazquez, E. (2023). Bayesian sequential design of computer experiments for quantile set inversion. arXiv preprint arXiv:2021.01008v3, submitted to Technometrics (in review).

Au, S. and Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16:263–277.

Azzimonti, D., Bect, J., Chevalier, C., and Ginsbourger, D. (2016). Quantifying uncertainties on excursion sets under a gaussian random field prior. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):850–874.

Bect, J., Ginsbourger, D., Li, L., Picheny, V., and Vazquez, E. (2012). Sequential design of computer experiments for the estimation of a probability of failure. *Statistics and Computing*, 22:773–793.

Bect, J., Li, L., and Vazquez, E. (2017). Bayesian Subset Simulations. *SIAM/ASA Journal on Uncertainty Quantification*, 5:762–786.

Bect, J., Vazquez, E., et al. (2022). STK: a Small (Matlab/Octave) Toolbox for Kriging. Release 2.7.0.

Branin, F. H. and Hoo, S. K. (1972). A method for finding multiple extrema of a function of n variables. In Lootsma, F. A., editor, *Numerical methods of Nonlinear Optimization*, pages 231–237. Academic Press.

Bryan, B., Nichol, R. C., Genovese, C. R., Schneider, J., Miller, C. J., and Wasserman, L. (2005). Active learning for identifying function threshold boundaries. In Weiss, Y., Schölkopf, B., and Platt, J., editors, *Advances in Neural Information Processing Systems*, volume 18. MIT Press.

References (cont.)

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

Dixon, L. and Szegö, G. P. (1978). The global optimization problem: an introduction. In Dixon, L. C. W. and Szegö, G. P., editors, *Towards Global Optimization 2*. North Holland.

Dubourg, V., Sudret, B., and Deheeger, F. (2013). Metamodel-based importance sampling for structural reliability analysis. *Probabilistic Engineering Mechanics*, 33:47–57.

Echard, B., Gayton, N., and Lemaire, M. (2011). AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. *Structural Safety*, 33(2):145–154.

Feliot, P. (2017). *Une approche Bayésienne pour l'optimisation multi-objectif sous contraintes*. Theses, Université Paris Saclay (COmUE).

Kahn, H. and Harris, T. E. (1951). Estimation of particle transmission by random sampling. *National Bureau of Standards applied mathematics series*, 12:27–30.

Li, L. (2012). *Sequential Design of Experiments to Estimate a Probability of Failure.* PhD thesis.

Marques, A., Lam, R., and Willcox, K. (2018). Contour location via entropy reduction leveraging multiple information sources. In *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, pages 1–11.

Picheny, V., Ginsbourger, D., Roustant, O., Haftka, R. T., and Kim, N.-H. (2010). Adaptive designs of experiments for accurate approximation of a target region. *Journal of Mechanical Design*, 132(7):071008.

References (cont.)

Picheny, V., Wagner, T., and Ginsbourger, D. (2013). A benchmark of kriging-based infill criteria for noisy optimization. *Structural and Multidisciplinary Optimization*, 48(3):607–626.

Ranjan, P., Bingham, D., and Michailidis, G. (2008). Sequential experiment design for contour estimation from complex computer codes. *Technometrics*, 50(4):527–541.

Approximated QSI-SUR criterion:

To reduce the cost, we define $J_n^k(x,s)$ as the SUR criterion based on

$$\mathcal{Q}_n^k = \int_{\mathbb{X}} \min(\pi_n^k(x), 1 - \pi_n^k(x))\,\mathrm{d}x,$$

where $\pi_n^k(x) = \mathbb{P}_n\left(x \in \Gamma_k(\tilde{\xi})\right)$ and, given a subset of simulation points $\Theta_{sim} \subset \mathbb{X} \times \mathbb{S}$,

$$\tilde{\xi}(x,s) = \mathbb{E}_n[\xi(x,s) \,|\, \xi(\Theta_{sim})].$$

NB: a close idea is exploited in [Azzimonti et al. (2016)].

Extension to batch designs: (inspired by [Feliot (2017)])

Given a batch size parameter r, for $1 \leq j \leq r$:

- Select (X_{n+j}, S_{n+j}) according to QSI-SUR criterion.
- Sample a **random realization** z_j of $\xi(X_{n+j}, S_{n+j})$ according to $\mathbb{P}_{n+j-1}$.
- Consider z_j as value of $f(X_{n+j}, S_{n+j}))$ until $j = r$.

When $j = r$: evaluate f at $\{(X_{n+j}, S_{n+j}), 1 \leq j \leq b\}$.

NB: This procedure produces 'approximated' batchs. The exact batchs

$$\{(X_{n+j}, S_{n+j}), j = 1, .., r\} \in \underset{(x_j, s_j) \in \mathbb{X} \times \mathbb{S}}{\operatorname{argmin}} \mathbb{E}_n(\mathcal{Q}_{n+r} \mid (X_{n+j}, S_{n+j}) = (x_j, s_j), j = 1, .., r)$$

being to computationally expensive (see, e.g [Chevalier et al. (2014)]).

Complementary details on numerical experiments

GP prior ξ trained on an initial design of size $10 * \dim(\mathbb{X} \times \mathbb{S})$.

Parameters are fitted using reML with:

- Constant mean function μ.
- Matérn covariance function k, with regularity parameter $\nu \in \{1/2, 3/2, 5/2, \infty\}$

All experiments are conducted in Matlab using the STK toolbox [Bect et al. (2022)].

Function f_1:

- $\mathbb{X} = [0, 10] \times [0, 15]$, $\mathbb{S} = [0, 15]$.
- $\mathbb{P}_S$ rescaled Beta$(7.5, 1.9)$
- $C = [15, +\infty)$, $\alpha = 0.05$
- g_1 is the Branin-Hoo function [Branin and Hoo (1972)].

Function f_2:

- $\mathbb{X} = [-2, 2]^2$, $\mathbb{S} = [-1, 1]$.
- $\mathbb{P}_S$ Gaussian $\mathcal{N}(1, 1)$ truncated on $\mathbb{S}$.
- $C = [9.5, +\infty)$, $\alpha = 0.1$
- g_2 is the Camel Back function [Dixon and Szegö (1978)].

Function f_3:

- $\mathbb{X} = [-1, 1]^2$, $\mathbb{S} = [-1, 1]^2$.
- $\mathbb{P}_S$ uniform on $\mathbb{S}$.
- $C = (-\infty, 1.065]$, $\alpha = 0.5$
- f_3 is the Hartmann4 function [Picheny et al. (2013)].

Function f_4:

- $\mathbb{X} = [-2, 2]^2$, $\mathbb{S} = [-1, 1]^2$.
- $\mathbb{P}_S$ uniform on $\mathbb{S}$.
- $C = (-\infty, 1.4]$, $\alpha = 0.1$
- f_4 is a mean of Camel Back functions

$$f_4(x, s) = \frac{1}{2}(g_2(x_1, s_1) + g_2(x_2, s_2))$$

Bayesian Subset Simulation - general idea:

Given a function $f : \mathrm{U} \mapsto \mathbb{R}$ and a critical region $C = (-\infty, T]$, the BSS [Bect et al. (2017)] algorithm aims at estimating the excursion set

$$\gamma(f) = \{u \in \mathrm{U} \,:\, f(x) \notin C\}.$$

The algorithm sequentially estimates a sequence of decreasing sets

$$\gamma_1(f) \supset ... \supset \gamma_K(f) = \gamma(f)$$

using the 'joint-SUR' criterion combined with SMC based on the target densities

$$q_{n,k}(u) = \mathbb{P}_n(u \in \gamma_k(\xi))$$

Heuristic: When does QSI-SUR outperforms methods focusing on $\gamma(f)$?

Empirically, it appears that the QSI problem must respect two conditions:

- f is not 'too linear'.
- Setting

$$\gamma_{restrict}(f) = \{(x, s) \in \mathbb{X} \times \mathbb{S} \, : \, f(x, s) \notin C \text{ and } x \in \Gamma(f)\},$$

the ratio $\frac{\lambda_{\mathbb{X}\times\mathbb{S}}(\gamma_{restrict}(f))}{\lambda_{\mathbb{X}\times\mathbb{S}}(\gamma(f))}$ is small.

Complementary results on QSI-SUR (from [Ait Abdelmalek-Lomenech et al. (2023)]) - 1/2.

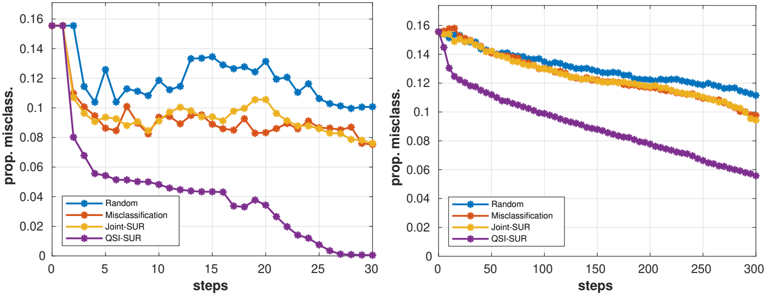


Figure: Median of the proportion of misclassified points vs. number of steps. (100 runs)

Complementary results on QSI-SUR (from [Ait Abdelmalek-Lomenech et al. (2023)]) - 2/2.

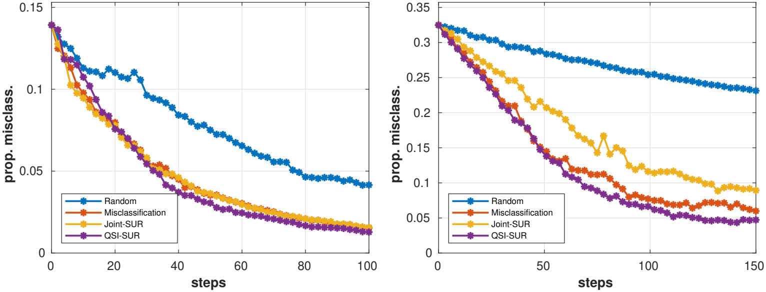


Figure: Median of the proportion of misclassified points vs. number of steps. (100 runs)