

# Estimation of Small Quantile Sets Using a Sequential Bayesian Strategy

Romain Ait Abdelmalek-Lomenech<sup>1</sup>, Julien Bect<sup>1</sup>,  
Vincent Chabridon<sup>2</sup> & Emmanuel Vazquez<sup>1</sup>

<sup>1</sup>Université Paris-Saclay, CNRS, CentraleSupélec, L2S,  
<sup>2</sup>EDF R&D

SIAM UQ 24, Trieste, Italy,  
February 29, 2024



# Table of Contents

Background on Quantile Set Inversion

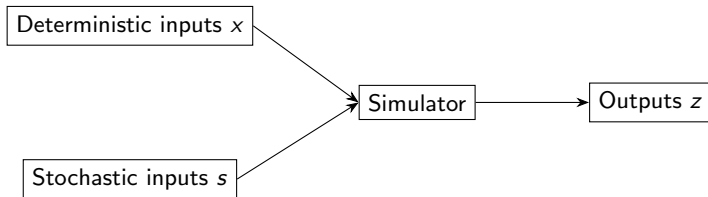
Improving the QSI-SUR strategy

Numerical experiments

Conclusion

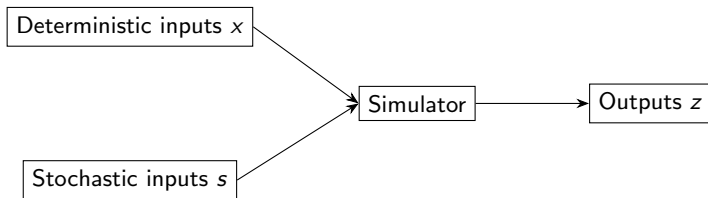
Consider an **expensive-to-evaluate** numerical simulator, with inputs:

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



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.
- ▶  $S \sim \mathbb{P}_S$  a known distribution on  $\mathcal{S}$ ,

Given:

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

We focus on the **QSI** problem, where the goal is to **estimate the set** of all  $x \in \mathbb{X}$  such that

$$\mathbb{P}(f(x, S) \in C) \leq \alpha,$$

Given:

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

We focus on the **QSI** problem, where the goal is to **estimate the set** of all  $x \in \mathbb{X}$  such that

$$\mathbb{P}(f(x, S) \in C) \leq \alpha,$$

by only using a given, small, number of evaluation points

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

## Estimate the quantile set:

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



## Estimate the quantile set:

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

### Example:

- ▶  $f$  simulator representing the production of a wind farm,
- ▶  $x$  geometrical design of the wind turbines,
- ▶  $s$  meteorological conditions,
- ▶  $C = (-\infty, T]$ .

## Estimate the quantile set:

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

### Example:

- ▶  $f$  simulator representing the production of a wind farm,
- ▶  $x$  geometrical design of the wind turbines,
- ▶  $s$  meteorological conditions,
- ▶  $C = (-\infty, T]$ .

Then  $\Gamma(f)$  is the set of all designs  $x \in \mathbb{X}$  such that the production of the wind farm is greater than  $T$  kWh with high probability.

## Estimate the quantile set:

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

Example of function and associated quantile set, with  $C = (-\infty, 7.75]$  and  $\alpha = 5\%$ .

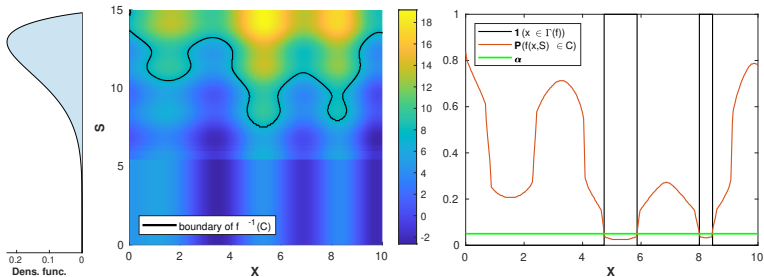


Figure: Representation of the function (middle), the density of  $\mathbb{P}_S$  (left) and associated quantile set (right).

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

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

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

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

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

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

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

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

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

- ▶  $\mathbb{P}_n$  the distribution of  $\xi$  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  $(x, s) \in \gamma(\xi)$ .

Several Bayesian methods exist for this estimation problem, taking into account the **expensive nature** of  $f$ . For example:

- ▶ Maximum misclassification probability [Bryan et al. (2005)]:

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

Several Bayesian methods exist for this estimation problem, taking into account the **expensive nature** of  $f$ . For example:

- ▶ Maximum misclassification probability [Bryan et al. (2005)]:

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

- ▶ 'Joint-SUR' [Bect et al. (2012); Chevalier et al. (2014)]:

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

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

Several Bayesian methods exist for this estimation problem, taking into account the **expensive nature** of  $f$ . For example:

- ▶ Maximum misclassification probability [Bryan et al. (2005)]:

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

- ▶ 'Joint-SUR' [Bect et al. (2012); Chevalier et al. (2014)]:

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

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

- ▶ [Ranjan et al. (2008); Picheny et al. (2010); Echard et al. (2011); Marques et al. (2018), ... ]



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

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

We denote:

- ▶  $\pi_n(x) = \mathbb{P}_n(x \in \Gamma(\xi)),$
- ▶  $Q_n = \int_{\mathbb{X}} \min(\pi_n(x), 1 - \pi_n(x)) dx.$

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

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

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

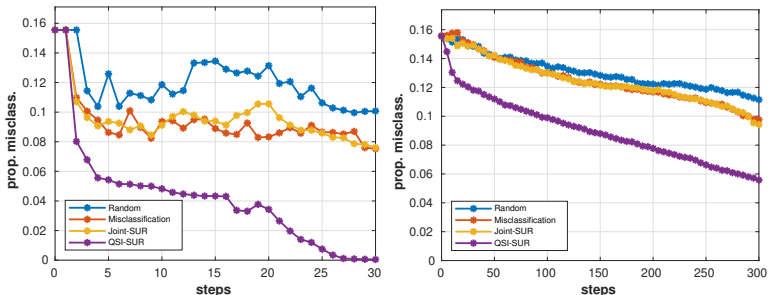


Figure: Median of the proportion of misclassified points vs. number of iterations, for 100 repetitions of the algorithms on two test functions.

# Table of Contents

Background on Quantile Set Inversion

Improving the QSI-SUR strategy

Numerical experiments

Conclusion

QSI-SUR criterion is based on  $\min(\pi_n(x), 1 - \pi_n(x))$ .

$\implies$  Necessity of a collection of points  $x \in \mathbb{X}$  such that this **misclassification probability** is non-null.

QSI-SUR criterion is based on  $\min(\pi_n(x), 1 - \pi_n(x))$ .

⇒ Necessity of a collection of points  $x \in \mathbb{X}$  such that this **misclassification probability** is non-null.

### Previously:

A simple approach (importance sampling) allowed good estimation of the quantile set in 'simple' cases.

### Main issue:

If  $\Gamma(f)$  is 'small' ( $\alpha$  small or/and  $C$  big), difficulty to sample relevant points in the set  $\mathbb{X}$ .

⇒  $\{x \in \Gamma(f)\}$  is a rare event w.r.t. the uniform distribution.

**Idea:** Multilevel splitting/subset simulation [Kahn and Harris (1951); Au and Beck (2001)].

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

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

**Idea:** Multilevel splitting/subset simulation [Kahn and Harris (1951); Au and Beck (2001)].

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

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

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 the **BSS** algorithm [Bect et al. (2017)]

It alternates two distinct phases:

▶ **Estimation phase**

- ▶ Define a new quantile set to estimate.
- ▶ Select the points to evaluate using approximated QSI-SUR criterion.

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

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

It alternates two distinct phases:

▶ **Estimation phase**

- ▶ Define a new quantile set to estimate.
- ▶ Select the points to evaluate using approximated QSI-SUR criterion.

▶ **Resample/move phase**

- ▶ Concentrate the particles towards the previously estimated set (resampling + MHRW).

Given  $\alpha_k$ ,  $C_k$  and a collection  $\tilde{X}_k$  of particles targeting  $\Gamma_k(f)$ .  
We denote  $q_{n,k}$  a density targeting  $\Gamma_k(f)$  at step  $n$ .

### Estimation phase:

- ▶ Set  $C_{k+1}$  and  $\alpha_{k+1}$  such that  $\text{ESS} \left( \frac{q_{n,k+1}(x)}{q_{n,k}} \right) \cong 30\%$ .
- ▶ Until a stopping condition is satisfied:
  - ▶ Select point  $(X_n, S_n)$  according to  $\text{argmin } J_n(x, s)$
  - ▶ Evaluate  $f$  at  $(X_n, S_n)$ .

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

## Resampling phase:

When stopping condition is met:

- ▶ Residual resampling of the particles.
- ▶ Move the particles to  $\Gamma_{k+1}(f)$  according to a Metropolis-Hastings with Gaussian random walk and target density  $q_{n,k+1}$ .
- ▶ Adapt variance of the steps to target acceptance rate  $\approx 25\%$ .

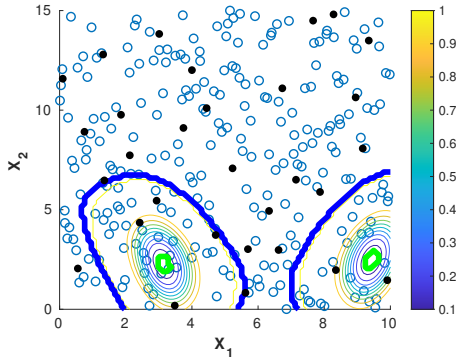


Figure: Temporary quantile set (blue line), final quantile set (green line), particles (blue dots) and projection of the initial design (black dots). -  $n = 0$ .

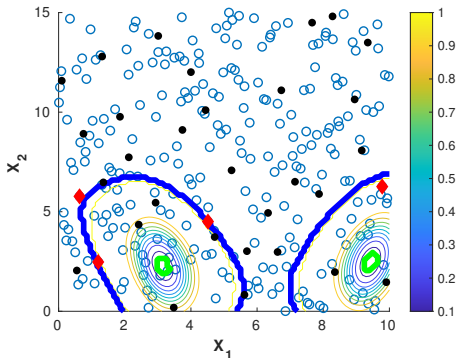


Figure: Temporary quantile set (blue line), final quantile set (green line), particles (blue dots), projections of the initial design (black dots) and sequential design (red dots). -  $n = 4$ .

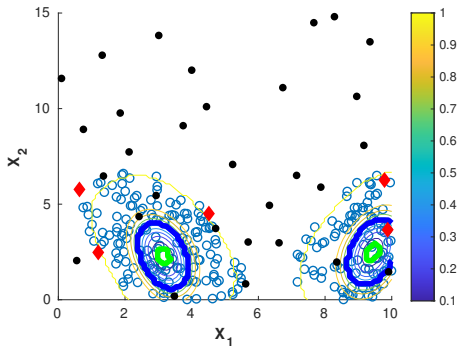


Figure: Temporary quantile set (blue line), final quantile set (green line), particles (blue dots), projections of the initial design (black dots) and sequential design (red dots). -  $n = 5$ .

## Choice of the target densities:

Natural idea (in the spirit of BSS): set

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

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



## Choice of the target densities:

Natural idea (in the spirit of BSS): set

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

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

**Idea:** Let us define the **'quantile' process**, with  $\beta \sim 1$ :

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

where  $\mu_n$  and  $\sigma_n$  are the posterior mean and standard deviation of  $\xi$ .

## 'Quantile' process:

$$\xi_n^+(x, s) = \mu_n(x, s) + \Phi^{-1}(\beta)\sigma_n(x, s) \quad \beta \sim 1,$$

Because  $C = (-\infty, T]$ , we can interpret  $\mathbb{P}(\xi_n^+(x, S) \in C_k)$  as a **very optimistic (low) estimation** of  $\mathbb{P}_n(\xi(x, S) \in C_k)$ .

## 'Quantile' process:

$$\xi_n^+(x, s) = \mu_n(x, s) + \Phi^{-1}(\beta)\sigma_n(x, s) \quad \beta \sim 1,$$

Because  $C = (-\infty, T]$ , we can interpret  $\mathbb{P}(\xi_n^+(x, S) \in C_k)$  as a **very optimistic (low) estimation** of  $\mathbb{P}_n(\xi(x, S) \in C_k)$ .

We define the target densities as

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

$\implies$  The MHRW step becomes a simple constrained Gaussian walk.

# Table of Contents

Background on Quantile Set Inversion

Improving the QSI-SUR strategy

**Numerical experiments**

Conclusion

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

$$f(x, s) = g(x_1, x_2) + s$$

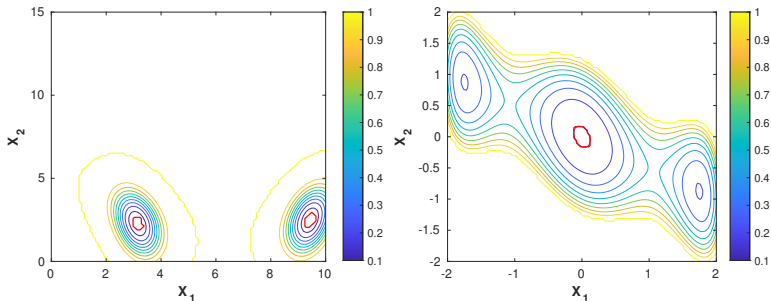


Figure: Representation of  $\Gamma(f)$  (red curve) for the examples functions.

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

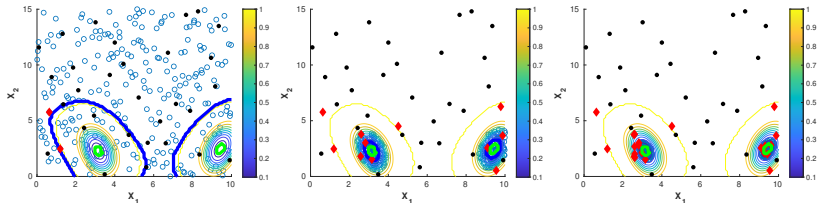


Figure: Temporary quantile set (blue line), final quantile set (green line), particles (blue dots), projections of the initial design (black dots) and sequential design (red dots). -  $n = 2, 10, 20$ .

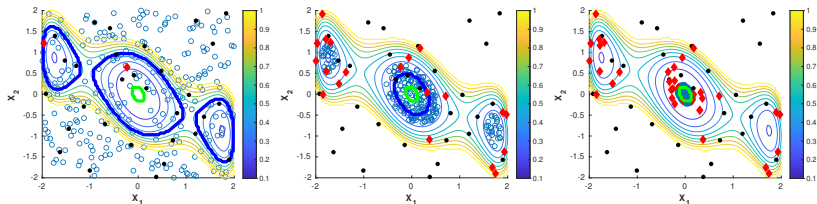


Figure: Temporary quantile set (blue line), final quantile set (green line), particles (blue dots), projections of the initial design (black dots) and sequential design (red dots). -  $n = 2, 15, 35$ .

We compare against BSS (which focus on the estimation of  $\gamma(f)$ ) the accuracy of the prediction obtained step after step.

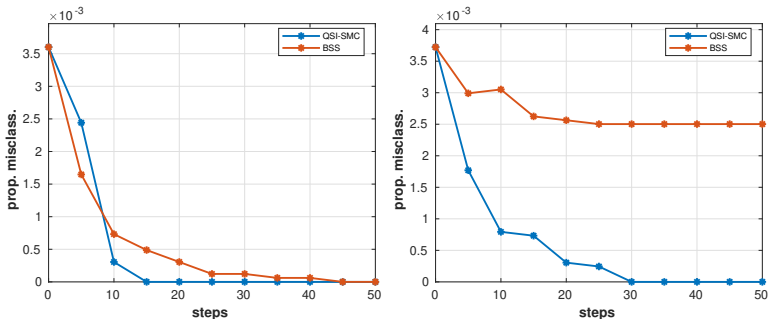


Figure: Median of the proportion of misclassified points vs. number of steps. (100 runs)



# Table of Contents

Background on Quantile Set Inversion

Improving the QSI-SUR strategy

Numerical experiments

**Conclusion**

## Conclusion:

- ▶ The proposed method allows to accurately estimate small quantile sets.
- ▶ The simple target densities chosen efficiently concentrate the particles in  $\mathbb{X}$  towards regions of interest.
- ▶ The criterion permits 'pseudo-batches' sequential designs.
  
- ▶ 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!**

# 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.
- 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.
- 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.
- 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.

## References (cont.)

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

$$Q_n^k = \int_{\mathbb{X}} \min(\pi_n^k(x), 1 - \pi_n^k(x)) dx,$$

where  $\pi_n^k(x) = \mathbb{P}_n(x \in \Gamma_k(\tilde{\xi}))$  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 'pseudo-batches':

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

- ▶ 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  $j = b$ .

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