# Quantifying and reducing uncertainties on target parameter regions with Gaussian Process models

David Ginsbourger [1]

Acknowledgements: a number of co-authors, notably appearing via citations!

[1] IMSV, Department of Mathematics and Statistics, University of Bern

INPUT Research Meeting

February 3rd, 2023

# Preamble

While statistics is increasingly employed to address societal needs, there is a number in situations where data (un)availibility is a potential bottleneck:

- Environmental hazards, especially in a climate change context
- Safety of industrial installations such as power plants
- Personalized/precision medecine
- Exploration of natural resources, surveys, etc.

# Preamble

While statistics is increasingly employed to address societal needs, there is a number in situations where data (un)availibility is a potential bottleneck:

- Environmental hazards, especially in a climate change context
- Safety of industrial installations such as power plants
- Personalized/precision medecine
- Exploration of natural resources, surveys, etc.

It is not always that the absolute amount of data is moderate, but rather that useful data for the question at hand is rare, expensive, confidential. . .

# Preamble

Numerical simulations (also known as “Computer Experiments”) are more and more used to cope with the lack of observational data. Yet they also come with issues as accurate simulations can be quite resource-intensive!

# Preamble

Numerical simulations (also known as “Computer Experiments”) are more and more used to cope with the lack of observational data. Yet they also come with issues as accurate simulations can be quite resource-intensive!

Models inherited from spatial statistics turn out to be of great use in uncertainty quantification to deal with both observational and simulation data.

# Preamble

Numerical simulations (also known as “Computer Experiments”) are more and more used to cope with the lack of observational data. Yet they also come with issues as accurate simulations can be quite resource-intensive!

Models inherited from spatial statistics turn out to be of great use in uncertainty quantification to deal with both observational and simulation data.

In particular, the former offer a neat framework for the definition of sequential design strategies for “uncertainty reduction”, or, in other words, data acquisition policies dedicated to address specific research questions.

# Outline

# Set up of main contributions presented next

We investigate i) a complex system represented by a deterministic function $f : \mathbf{x} \in E \mapsto f(\mathbf{x}) \in F$, ii) and/or quantities relying on $f$, based on a limited number of evaluations of $f$.

# Set up of main contributions presented next

We investigate i) a complex system represented by a deterministic function $f : \mathbf{x} \in E \mapsto f(\mathbf{x}) \in F$, ii) and/or quantities relying on $f$, based on a limited number of evaluations of $f$.

## Two typical examples

- Safety engineering: $\mathbf{x}$ is a vector parametrizing some system and $f$ returns an indicator of dangerousness. It is then crucial to understand which $\mathbf{x}$'s lead to "high" values of $f(\mathbf{x})$.

# Set up of main contributions presented next

We investigate i) a complex system represented by a deterministic function $f : \mathbf{x} \in E \mapsto f(\mathbf{x}) \in F$, ii) and/or quantities relying on $f$, based on a limited number of evaluations of $f$.

## Two typical examples

- Safety engineering: $\mathbf{x}$ is a vector parametrizing some system and $f$ returns an indicator of dangerousness. It is then crucial to understand which $\mathbf{x}$'s lead to "high" values of $f(\mathbf{x})$.
- Flow dynamics: $\mathbf{x}$ stands e.g. for the medium, boundary conditions, etc. and $f$ returns the evolution of a fluid and/or a measure of discrepancy between simulation results and given observation results.

# Function approximation by means of GP models

Typical situation : $f$ was evaluated at a set of "points" $\mathbf{x}_1, \ldots, \mathbf{x}_n \in D \subset E$ and one wishes to estimate a quantity relying on $f$ and/or perform new evaluations in order to improve this estimation.

# Function approximation by means of GP models

Typical situation : $f$ was evaluated at a set of “points” $\mathbf{x}_1, \ldots, \mathbf{x}_n \in D \subset E$ and one wishes to estimate a quantity relying on $f$ and/or perform new evaluations in order to improve this estimation.

$\Rightarrow$ legitimate to rely on some approximation(s) of $f$ knowing $f(\mathbf{x}_i) + \epsilon_i$ $(1 \leq i \leq n)$. A number of approaches do exist. . .

# Function approximation by means of GP models

Typical situation : $f$ was evaluated at a set of "points" $\mathbf{x}_1, \ldots, \mathbf{x}_n \in D \subset E$ and one wishes to estimate a quantity relying on $f$ and/or perform new evaluations in order to improve this estimation.

$\Rightarrow$ legitimate to rely on some approximation(s) of $f$ knowing $f(\mathbf{x}_i) + \epsilon_i$ $(1 \leq i \leq n)$. A number of approaches do exist...

Principles of the Gaussian Process approach (GP): suppose that, *a priori*, $f$ is a realization of a GP $(Z(\mathbf{x}))_{\mathbf{x} \in D}$ and approximate $f$ and/or the quantities of interest via the **conditional distribution** of $Z$ knowing $Z(\mathbf{x}_i) + \varepsilon_i = f(\mathbf{x}_i) + \epsilon_i$.

# Function approximation by means of GP models

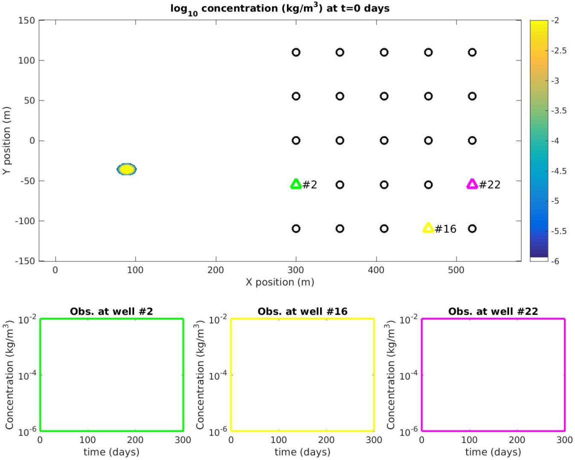
Typical situation : $f$ was evaluated at a set of “points” $\mathbf{x}_1, \ldots, \mathbf{x}_n \in D \subset E$ and one wishes to estimate a quantity relying on $f$ and/or perform new evaluations in order to improve this estimation.

$\Rightarrow$ legitimate to rely on some approximation(s) of $f$ knowing $f(\mathbf{x}_i) + \epsilon_i$ $(1 \leq i \leq n)$. A number of approaches do exist...

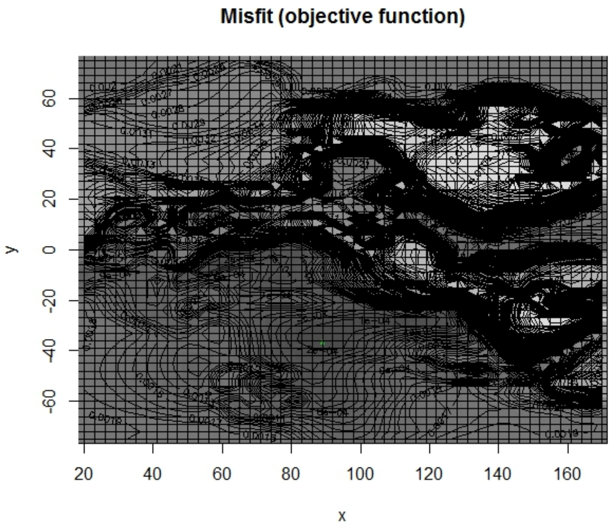
Principles of the Gaussian Process approach (GP): suppose that, *a priori*, $f$ is a realization of a GP $(Z(\mathbf{x}))_{\mathbf{x} \in D}$ and approximate $f$ and/or the quantities of interest via the **conditional distribution** of $Z$ knowing $Z(\mathbf{x}_i) + \varepsilon_i = f(\mathbf{x}_i) + \epsilon_i$.

$\Rightarrow$ very practical for sequential design of experiments.

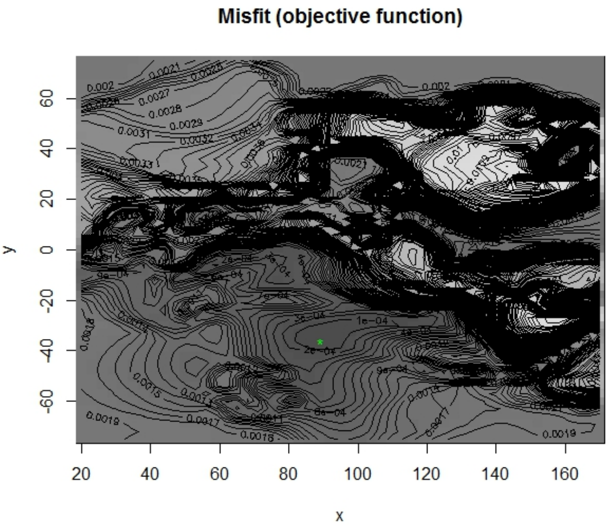
# Example inverse problem in hydrogeology



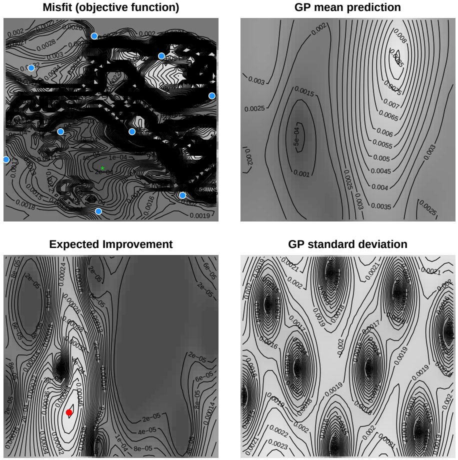
# A costly full factorial experimental design!



# A costly full factorial experimental design!



# Application of GP-based optimization



The previous example was produced in the framework of a collaboration with G. Pirot (University of Lausanne), T. Krityakierne (now at Mahidol University, Bangkok) and P. Renard (University of Neuchâtel).

$\Rightarrow$ See recently published Hydrol. Earth Syst. Sci. paper (2019).

The previous example was produced in the framework of a collaboration with G. Pirot (University of Lausanne), T. Krityakierne (now at Mahidol University, Bangkok) and P. Renard (University of Neuchâtel).

$\Rightarrow$ See recently published Hydrol. Earth Syst. Sci. paper (2019).

**Main focus in most of the following**

In a related set-up, how to estimate parameter regions where $f$ takes target values using such models and dedicated sequential design strategies?

The previous example was produced in the framework of a collaboration with G. Pirot (University of Lausanne), T. Krityakierne (now at Mahidol University, Bangkok) and P. Renard (University of Neuchâtel).

$\Rightarrow$ See recently published Hydrol. Earth Syst. Sci. paper (2019).

**Main focus in most of the following**

In a related set-up, how to estimate parameter regions where $f$ takes target values using such models and dedicated sequential design strategies?

As a transition, let us review a few selected seminal references about GP modelling and GP-based "Bayesian Optimization".

# A few references on GP modelling ...

A. O'Hagan (1978).
Curve fitting and optimal design for prediction.
Journal of the Royal Statistical Society, Series B, 40(1):1-42.

J. Sacks, W.J. Welch, T.J. Mitchell, and H. P. Wynn (1989).
Design and Analysis of Computer Experiments
Statist. Sci. 4(4), 409-423.

H. Omre and K. Halvorsen (1989).
The bayesian bridge between simple and universal kriging.
Mathematical Geology, 22 (7):767-786.

M. S. Handcock and M. L. Stein (1993).
A bayesian analysis of kriging.
Technometrics, 35(4):403-410.

A.W. Van der Vaart and J. H. Van Zanten (2008).
Rates of contraction of posterior distributions based on Gaussian process priors.
Annals of Statistics, 36:1435-1463.

# . . . and on GP-based Optimization

H.J. Kushner (1964).
A new method of locating the maximum of an arbitrary multi-peak curve in the presence of noise.
Journal of Basic Engineering, 86:97-106.

J. Mockus (1972).
On Bayesian methods for seeking the extremum.
Automatics and Computers (Avtomatika i Vychislitel'naya Tekhnika), 4(1):53-62.

J. Mockus, V. Tiesis, and A. Zilinskas (1978).
The application of Bayesian methods for seeking the extremum.
In Dixon, L. C. W. and Szegö, G. P., editors, Towards Global Optimisation, volume 2, pages 117-129. Elsevier Science Ltd., North Holland, Amsterdam.

J.M. Calvin (1997).
Average performance of a class of adaptive algorithms for global optimization.
The Annals of Applied Probability, 7(3):711-730.

M. Schonlau, W.J. Welch and D.R. Jones (1998).
Efficient Global Optimization of Expensive Black-box Functions.
*Journal of Global Optimization*.

# Outline

# Our main topic today: background and motivations

A number of practical problems boil down to determining sets of the form

$$\Gamma^{\star} = \{\mathbf{x} \in D : f(\mathbf{x}) \in T\} = f^{-1}(T)$$

where $f : D \longrightarrow \mathbb{R}^k$ ($k \geq 1$), $D \subset \mathbb{R}^d$ ($d \geq 1$), and $T \subset \mathbb{R}^k$.

# Our main topic today: background and motivations

A number of practical problems boil down to determining sets of the form

$$\Gamma^{\star} = \{\mathbf{x} \in D : f(\mathbf{x}) \in T\} = f^{-1}(T)$$

where $f : D \longrightarrow \mathbb{R}^k$ ($k \geq 1$), $D \subset \mathbb{R}^d$ ($d \geq 1$), and $T \subset \mathbb{R}^k$.

## Examples

- Contour lines
- Excursion/sojourn sets above/below thresholds
- Admissible regions in constrained optimization
- High gradient/high curvature regions, etc.
- (Pareto sets in multi-objective optimization. . . but then $T$ depends on $f$!)

# Background and motivations

We essentially focus today on the case where $k = 1$, $D$ is compact, $f$ is continuous, and $T = [t, +\infty)$ or $(-\infty, t]$ for some prescribed $t \in \mathbb{R}$.

$\Gamma^\star = \{\mathbf{x} \in D : f(\mathbf{x}) \geq t\}$ is then referred to as the excursion set of f above $t$.

# Background and motivations

We essentially focus today on the case where $k = 1$, $D$ is compact, $f$ is continuous, and $T = [t, +\infty)$ or $(-\infty, t]$ for some prescribed $t \in \mathbb{R}$.

$\Gamma^{\star} = \{\mathbf{x} \in D : f(\mathbf{x}) \geq t\}$ is then referred to as the excursion set of f above $t$.

Our aim is to estimate $\Gamma^{\star}$ and quantify uncertainty on it when $f$ can solely be evaluated at a few points, both in static and sequential cases.

# Test case from safety engineering

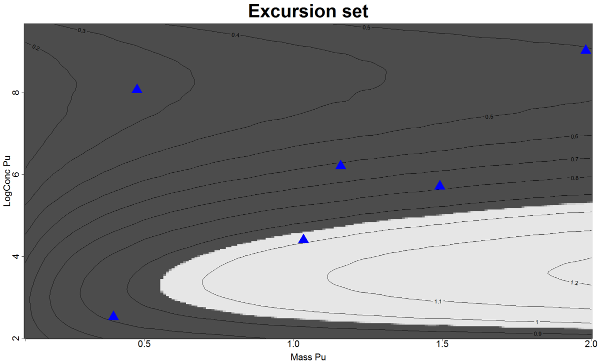



Figure: Excursion set (light gray) of a nuclear criticality safety coefficient depending on two design parameters. Blue triangles: initial experiments.

C. Chevalier (2013).
Fast uncertainty reduction strategies relying on Gaussian process models.
Ph.D. thesis, University of Bern.

Making a sensible estimation of $\Gamma^{\star}$ based on a drastically limited number of evaluations $f(\mathbf{X}_n) = (f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n))'$ calls for additional assumptions on $f$.

Making a sensible estimation of $\Gamma^{\star}$ based on a drastically limited number of evaluations $f(\mathbf{X}_n) = (f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n))'$ calls for additional assumptions on $f$.

As before, we consider the Bayesian framework where a Gaussian Process (GP) prior is put on $f$, i.e. $f$ is seen as one realization of a GP $(Z(\mathbf{x}))_{\mathbf{x} \in D}$ (characterized in distribution by a mean $m$ and a covariance kernel $k$).

Making a sensible estimation of $\Gamma^\star$ based on a drastically limited number of evaluations $f(\mathbf{X}_n) = (f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n))'$ calls for additional assumptions on $f$.

As before, we consider the Bayesian framework where a Gaussian Process (GP) prior is put on $f$, i.e. $f$ is seen as one realization of a GP $(Z(\mathbf{x}))_{\mathbf{x} \in D}$ (characterized in distribution by a mean $m$ and a covariance kernel $k$).

In the GP set-up, the main object of interest is represented by

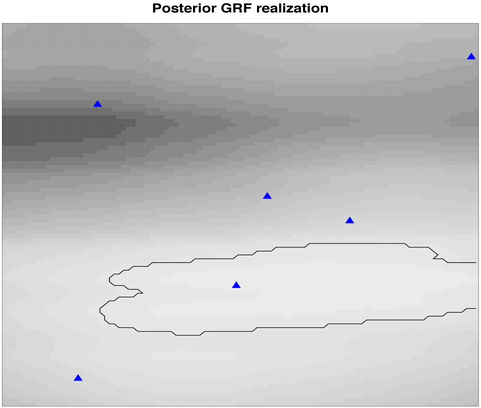
$$\Gamma = \{\mathbf{x} \in D : Z(\mathbf{x}) \in T\} = Z^{-1}(T)$$

Under our previous assumptions on $T$ and assuming that is chosen $Z$ with continuous paths, $\Gamma$ is a Random Closed Set (See thesis below for detail).

D. Azzimonti (2016).
Contributions to Bayesian set estimation relying on random field priors.
Ph.D. thesis, University of Bern.

# Simulating excursion sets under a GRF model

Posterior simulations on a $50 \times 50$ grid of $Z$ and $\Gamma$ knowing $Z(\mathbf{X}_n) = f(\mathbf{X}_n)$.



# How to quantify the uncertainty on $\Gamma$?

There are many ways to quantify uncertainties on sets!

# How to quantify the uncertainty on $\Gamma$?

There are many ways to quantify uncertainties on sets!

This will be one of the recurring questions throughout the talk, but we will not be exhaustive by far. For more detail see, e.g.,

I. Molchanov (2005)
Theory of Random Sets.
*Springer*.

D. Azzimonti, J. Bect, C. Chevalier and D. Ginsbourger (2016).
Quantifying uncertainties on excursion sets under a Gaussian random field prior.
SIAM/ASA Journal on Uncertainty Quantification.

# How to quantify the uncertainty on $\Gamma$?

There are many ways to quantify uncertainties on sets!

This will be one of the recurring questions throughout the talk, but we will not be exhaustive by far. For more detail see, e.g.,

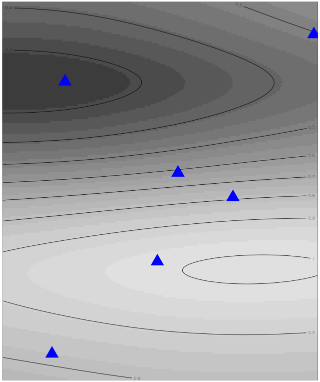
I. Molchanov (2005)
Theory of Random Sets.
*Springer*.

D. Azzimonti, J. Bect, C. Chevalier and D. Ginsbourger (2016).
Quantifying uncertainties on excursion sets under a Gaussian random field prior.
SIAM/ASA Journal on Uncertainty Quantification.

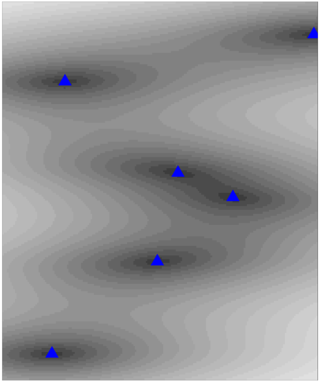
Before moving to random set-related concepts, a first spontaneous idea is to "scalarize" the problem, for instance by looking at $\Gamma$'s volume. Let us make a detour through some GP basics in order to do so.

# Kriging (Gaussian Process Interpolation)

Posterior mean

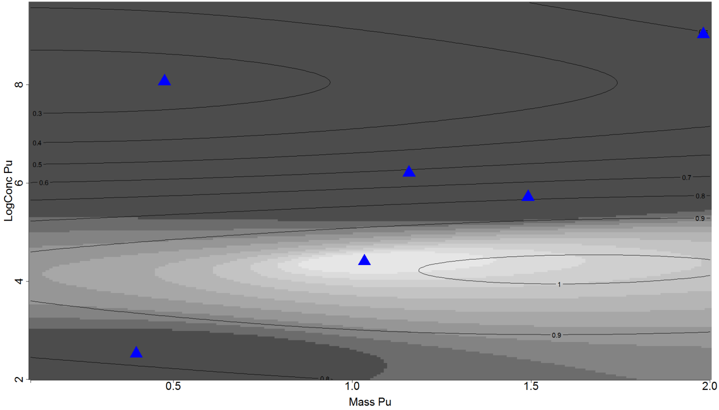


Posterior variance



$$\begin{cases} m_n(\mathbf{x}) = m(\mathbf{x}) + k(\mathbf{X}_n, \mathbf{x})^T k(\mathbf{X}_n, \mathbf{X}_n)^{-1}(f(\mathbf{X}_n) - m(\mathbf{X}_n)) \\ s_n^2(\mathbf{x}) = k(\mathbf{x}, \mathbf{x}) - k(\mathbf{X}_n, \mathbf{x})^T k(\mathbf{X}_n, \mathbf{X}_n)^{-1} k(\mathbf{X}_n, \mathbf{x}) \end{cases}$$

## Conditional probability of excursion




From $\mathcal{L}_n(Z(\mathbf{x})) = \mathcal{N}(m_n(\mathbf{x}), s_n^2(\mathbf{x}))$, the "coverage probability" of $\Gamma$ (or conditional/posterior probability of excursion, here) can be expanded as

$$p_n(\mathbf{x}) = \mathbb{P}_n(\mathbf{x} \in \Gamma) = \mathbb{P}_n(Z(\mathbf{x}) \geq t) = \Phi\left(\frac{m_n(\mathbf{x}) - t}{s_n(\mathbf{x})}\right)$$

# Outline

1. Introduction
   - Settings
   - From GP-based optimization to set estimation

2. Focus on (reducing) uncertainties on sets
   - Introduction to Sequential Uncertainty Reduction strategies
   - Towards conservative excursion set estimation

# From $p_n$ to moments of $\Gamma$'s volume

Denote by $\mu$ a finite measure on $(D, \mathcal{B}(D))$ [one can think here of $\mu$ as $\mathrm{Vol}$] and set $\alpha^* = \mu(\Gamma^*)$, i.e. the "volume of excursion" in the considered case.

The GP model leads to a random analogue $\alpha = \mu(\Gamma)$, and by Robbins' theorem, the posterior expectation of $\alpha$ can be written in terms of $p_n$:

$$\mathbb{E}_n[\mu(\Gamma)] = \mathbb{E}_n\left[\int_D \mathbf{1}_\Gamma(\mathbf{u})d\mu(\mathbf{u})\right] = \int_D p_n(\mathbf{u})d\mu(\mathbf{u})$$

# From $p_n$ to moments of Γ's volume

Denote by $\mu$ a finite measure on $(D, \mathcal{B}(D))$ [one can think here of $\mu$ as $\mathrm{Vol}$] and set $\alpha^* = \mu(\Gamma^*)$, i.e. the "volume of excursion" in the considered case.

The GP model leads to a random analogue $\alpha = \mu(\Gamma)$, and by Robbins' theorem, the posterior expectation of $\alpha$ can be written in terms of $p_n$:

$$\mathbb{E}_n[\mu(\Gamma)] = \mathbb{E}_n\left[\int_D \mathbf{1}_\Gamma(\mathbf{u}) d\mu(\mathbf{u})\right] = \int_D p_n(\mathbf{u}) d\mu(\mathbf{u})$$

However, the (posterior) distribution of $\alpha$ has been considered analytically intractable.

R.J. Adler (2000)
On excursion sets, tube formulas and maxima of random fields.
Annals of Applied Probability, 10(1):1-74.

E. Vazquez and M. Piera Martinez (2006).
Estimation of the volume of an excursion set of a Gaussian process using intrinsic Kriging.
arXiv:math/0611273 [math.ST].

# About conditional moments of $\alpha$

Fortunately, as already pointed out in Molchanov 2005 in more general settings, $\mathbb{E}_n[\alpha^r]$ can also be worked out for $r \geq 2$), at the price of calculating integrals. In our framework, we have indeed:

$$
\begin{aligned}
\mathbb{E}_n[\alpha^r] &= \mathbb{E}_n\left[\left(\int_D \mathbf{1}_\Gamma(\mathbf{u})d\mu(\mathbf{u})\right)^r\right] \\
&= \mathbb{E}_n\left[\left(\int_D \mathbf{1}_\Gamma(\mathbf{u}_1)d\mu(\mathbf{u}_1)\right)\dots\left(\int_D \mathbf{1}_\Gamma(\mathbf{u}_r)d\mu(\mathbf{u}_r)\right)\right] \\
&= \int_D \cdots \int_D \mathbb{E}_n\left[\mathbf{1}_\Gamma(\mathbf{u}_1)\dots\mathbf{1}_\Gamma(\mathbf{u}_r)\right] d\mu(\mathbf{u}_1)\dots d\mu(\mathbf{u}_r) \\
&= \int_D \cdots \int_D \mathbb{P}_n(Z(\mathbf{u}_1) \geq t, \dots, Z(\mathbf{u}_r) \geq t) d\mu(\mathbf{u}_1)\dots d\mu(\mathbf{u}_r)
\end{aligned}
$$

Hence, recalling the GP assumption, $\mathbb{E}_n[\alpha^r]$ writes as an $r$-dimensional integral which integrand involves a $r$-dimensional Gaussian CDF.

# A useful bound for the case $r = 2$

In what follows, the case $r = 2$ will be of special importance as we will consider sequential design strategies aiming at reducing $\mathrm{Var}_n[\alpha]$.

The following underlined quantity, that is easier to compute and also comes with a nice interpretation, has been used as well:

$$\begin{aligned}\mathrm{Var}_n[\alpha] &= \mathbb{E}_n\left[\left(\int_D (\mathbf{1}_\Gamma(\mathbf{u}) - p_n(\boldsymbol{u}))d\mu(\boldsymbol{u})\right)^2\right] \\ &\leq \mu(D)^2\mathbb{E}_n\left[\int_D (\mathbf{1}_\Gamma(\mathbf{u}) - p_n(\boldsymbol{u}))^2 d\mu(\boldsymbol{u})\right] \\ &= \mu(D)^2 \underbrace{\int_D p_n(\boldsymbol{u})(1 - p_n(\boldsymbol{u}))d\mu(\boldsymbol{u})}_{\text{Integrated indicator variance}}\end{aligned}$$

# A useful bound for the case $r = 2$

In what follows, the case $r = 2$ will be of special importance as we will consider sequential design strategies aiming at reducing $\mathrm{Var}_n[\alpha]$.

The following underlined quantity, that is easier to compute and also comes with a nice interpretation, has been used as well:

$$\begin{aligned}\mathrm{Var}_n[\alpha] &= \mathbb{E}_n\left[\left(\int_D (\mathbf{1}_\Gamma(\mathbf{u}) - p_n(\boldsymbol{u}))d\mu(\boldsymbol{u})\right)^2\right] \\ &\leq \mu(D)^2 \mathbb{E}_n\left[\int_D (\mathbf{1}_\Gamma(\mathbf{u}) - p_n(\boldsymbol{u}))^2 d\mu(\boldsymbol{u})\right] \\ &= \mu(D)^2 \underbrace{\int_D p_n(\boldsymbol{u})(1 - p_n(\boldsymbol{u}))d\mu(\boldsymbol{u})}_{\text{Integrated indicator variance}}\end{aligned}$$

The excursion volume's variance and the integrated indicator variance are used as two particular "measures of uncertainty" in what follows.

# Towards Stepwise Uncertainty Reduction strategies

Let us informally consider the following 1-step-lookahead scheme:

- For some chosen (say, non-negative) functional defined on GP distributions, define the uncertainty at time $n \geq 0$, $H_n$, as this functional applied to the current posterior GP (E.g., $H_n = var_n(\alpha)$).

- Starting from some intial design $\{\mathbf{x}_1, \ldots, \mathbf{x}_{n_0}\}$, at each iteration $n \geq n_0$, evaluate $f$ at a point $\mathbf{x}^\star_{n+1}$ minimizing the so-called SUR criterion associated with the chosen notion of uncertainty:

$$J_n(\mathbf{x}_{n+1}) := \mathbb{E}_n(H_{n+1}(\mathbf{x}_{n+1}))$$

# Towards Stepwise Uncertainty Reduction strategies

Let us informally consider the following 1-step-lookahead scheme:

- For some chosen (say, non-negative) functional defined on GP distributions, define the uncertainty at time $n \geq 0$, $H_n$, as this functional applied to the current posterior GP (E.g., $H_n = var_n(\alpha)$).

- Starting from some intial design $\{\mathbf{x}_1, \ldots, \mathbf{x}_{n_0}\}$, at each iteration $n \geq n_0$, evaluate $f$ at a point $\mathbf{x}^{\star}_{n+1}$ minimizing the so-called SUR criterion associated with the chosen notion of uncertainty:

$$J_n(\mathbf{x}_{n+1}) := \mathbb{E}_n(H_{n+1}(\mathbf{x}_{n+1}))$$

See notably the following paper and seminal references therein:

J. Bect, D. Ginsbourger, L. Li, V. Picheny and E. Vazquez.
Sequential design of computer experiments for the estimation of a probability of failure.
*Statistics and Computing, 22(3):773-793, 2012.*

# SUR strategies: Two candidate uncertainties

Two possible definitions for the uncertainty $H_n$ are considered below:

$$H_n := Var_n(\alpha)$$
$$\widetilde{H}_n := \int_D p_n(1 - p_n)\mathrm{d}\mu$$

# SUR strategies: Two candidate uncertainties

Two possible definitions for the uncertainty $H_n$ are considered below:

$$H_n := Var_n(\alpha)$$
$$\widetilde{H}_n := \int_D p_n(1 - p_n)\mathrm{d}\mu$$

**Uncertainties:**

$$H_n := Var_n(\alpha)$$
$$\widetilde{H}_n := \int_{\mathbb{X}} p_n(1 - p_n)\mathrm{d}\mu$$

**SUR criteria:**

$$J_n(\mathbf{x}) := \mathbb{E}_n(Var_{n+1}(\alpha))$$
$$\widetilde{J}_n(\mathbf{x}) := \mathbb{E}_n\left(\int_D p_{n+1}(1 - p_{n+1})\mathrm{d}\mu\right)$$

Main challenge to calculate $\widetilde{J}_n(\mathbf{x})$ (similar for $J_n(\mathbf{x})$): Obtain a closed form expression for $\mathbb{E}_n\left(p_{n+1}(1 - p_{n+1})\right)$ and integrate it.

# Deriving SUR criteria

**Proposition**

$$\mathbb{E}_n(p_{n+1}(\mathbf{x})(1-p_{n+1}(\mathbf{x}))) = \Phi_2\left(\begin{pmatrix} a(\mathbf{x}) \\ -a(\mathbf{x}) \end{pmatrix}, \begin{pmatrix} c(\mathbf{x}) & 1-c(\mathbf{x}) \\ 1-c(\mathbf{x}) & c(\mathbf{x}) \end{pmatrix}\right)$$

- $\Phi_2(\cdot, M)$*: c.d.f. of centred bivariate Gaussian with covariance matrix* $M$
- $a(\mathbf{x}) := (m_n(\mathbf{x}) - t)/s_{n+1}(\mathbf{x})$,
- $c(\mathbf{x}) := s_n^2(\mathbf{x})/s_{n+1}^2(\mathbf{x})$

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

# Deriving SUR criteria

**Proposition**

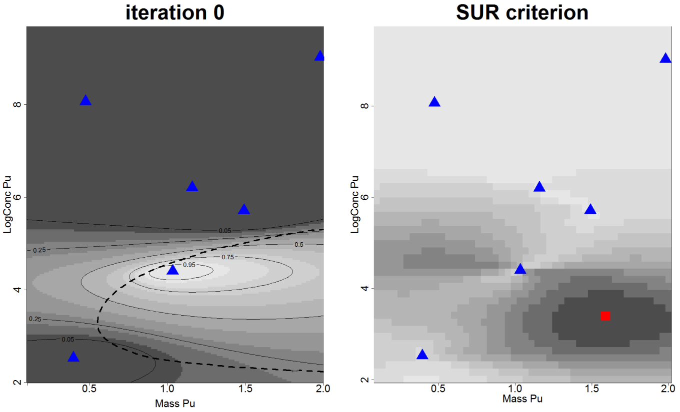
$$\mathbb{E}_n(p_{n+1}(\mathbf{x})(1-p_{n+1}(\mathbf{x}))) = \Phi_2\left(\begin{pmatrix} a(\mathbf{x}) \\ -a(\mathbf{x}) \end{pmatrix}, \begin{pmatrix} c(\mathbf{x}) & 1-c(\mathbf{x}) \\ 1-c(\mathbf{x}) & c(\mathbf{x}) \end{pmatrix}\right)$$

- $\Phi_2(\cdot, M)$*: c.d.f. of centred bivariate Gaussian with covariance matrix* $M$
- $a(\mathbf{x}) := (m_n(\mathbf{x}) - t)/s_{n+1}(\mathbf{x})$,
- $c(\mathbf{x}) := s_n^2(\mathbf{x})/s_{n+1}^2(\mathbf{x})$

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

C. Chevalier, V. Picheny and D. Ginsbourger.
The KrigInv package: An efficient and user-friendly R implementation of Kriging-based inversion algorithms.
*Computational Statistics & Data Analysis, 71:1021-1034, 2014*

# Back to the test case with SUR



# Batch-sequential SUR strategies

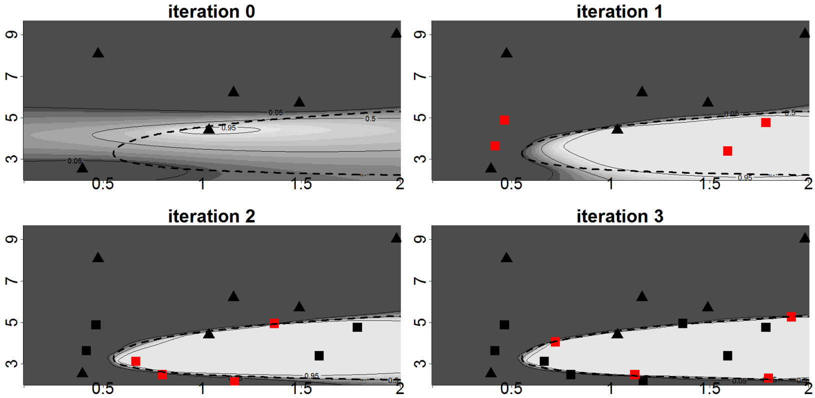



Figure: 3 SUR iterations ($\widetilde{J}_n$ criterion with $q = 4$)

# About the consistency

J. Bect, F. Bachoc and D. Ginsbourger (2019).
A supermartingale approach to Gaussian process based sequential design of experiments.
Bernoulli 25 (4A), 2883-2919.

# Further questions about SUR and UQ on sets

Of course, in operational conditions, asymptotic results are worthwhile. However, concrete finite-sample outputs such as estimates of $\Gamma^{\star}$ and quantifications of the associated uncertainty are required as well.

# Further questions about SUR and UQ on sets

Of course, in operational conditions, asymptotic results are worthwhile. However, concrete finite-sample outputs such as estimates of $\Gamma^{\star}$ and quantifications of the associated uncertainty are required as well.

Now, $n$ being fixed, how to estimate $\Gamma^{\star}$ and to assess/represent the variability of the corresponding estimate(s)?

# Outline

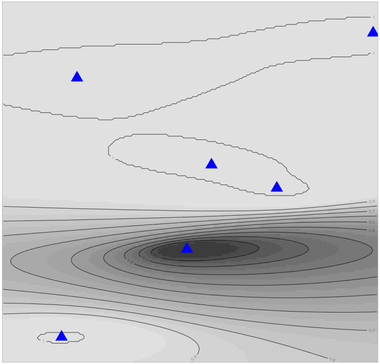
1. Introduction
   - Settings
   - From GP-based optimization to set estimation
2. Focus on (reducing) uncertainties on sets
   - Introduction to Sequential Uncertainty Reduction strategies
   - Towards conservative excursion set estimation

# How to summarize the posterior distribution of sets?

For application purposes, let us reverse the perspective and focus on the sojourn/excursion case **below** $t$, where $\Gamma = \{\mathbf{x} \in D : Z(\mathbf{x}) \leq t\}$ [say a "safe set"] and $p_n : \mathbf{x} \in D \to p_n(x) = P_n(Z(\mathbf{x}) \leq t)$.

# How to summarize the posterior distribution of sets?

For application purposes, let us reverse the perspective and focus on the sojourn/excursion case **below** $t$, where $\Gamma = \{\mathbf{x} \in D : Z(\mathbf{x}) \leq t\}$ [say a "safe set"] and $p_n : \mathbf{x} \in D \to p_n(x) = P_n(Z(\mathbf{x}) \leq t)$.




Define the (conditional) quantiles of $\Gamma$ as $\rho-$level sets of $p_n$:

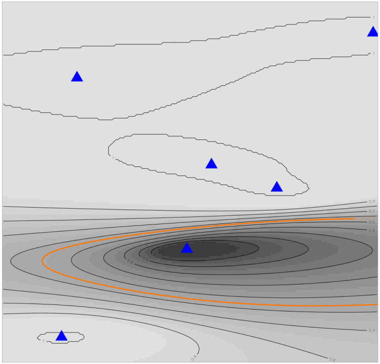
$$Q_\rho := \{\mathbf{x} \in D : p_n(\mathbf{x}) \geq \rho\}$$
$$= \{\mathbf{x} \in D : P_n(Z(\mathbf{x}) \leq t) \geq \rho\}.$$

How well $Q_\rho$ estimates $\Gamma$ can be quantified for instance through the "expected deviation":

$$\mathbb{E}_n\left(\mu(Q_\rho \Delta \Gamma)\right)$$

# Estimates of $\Gamma^{\star}$: the Vorob'ev expectation

Coverage probability function (Vorob'ev expectation)



The **Vorob'ev expectation of** $\Gamma \mid (Z(\mathbf{x}_1) = f(x_1), \ldots, Z(\mathbf{x}_n) = f(x_n))$ is the $\rho^{\star}$ level set of $p_n$ such that

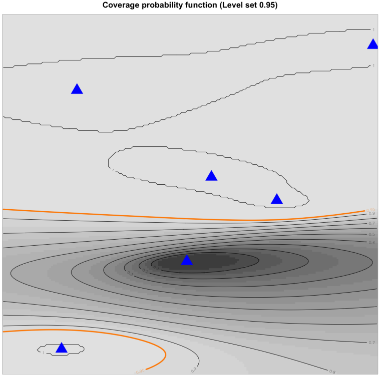
$$\mu(Q_{\rho^{\star}}) = \mathbb{E}_n[\mu(\Gamma)].$$

It is a state of the art result that $Q_{\rho^{\star}}$ minimizes $S \to \mathbb{E}_n\left(\mu(S\Delta\Gamma)\right)$ among all closed sets $S \subset \mathbb{R}^d$ with volume $\mathbb{E}_n[\mu(\Gamma)]$.

C. Chevalier, D. Ginsbourger, J. Bect, and Molchanov, I.
Estimating and quantifying uncertainties on level sets using the Vorob'ev expectation and deviation with Gaussian process models.
*mODa 10 Advances in Model-Oriented Design and Analysis, Physica-Verlag HD, 2013.*

# Estimates of $\Gamma^{\star}$: some limitations of $Q_{\rho}$ quantiles

In practice one often wish to give **confidence statements** on the estimates.

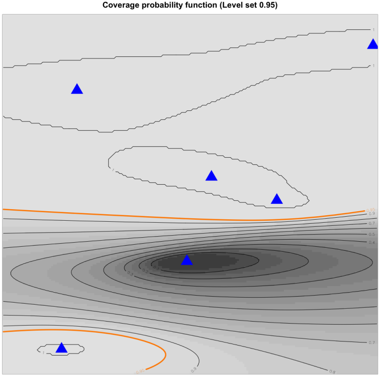



$Q_{\rho}$ contains points which have marginal probability at least $\rho$ of being in $\Gamma$.

$\Rightarrow$ no confidence statement on the probability of the actual excursion set containing this specific estimate.

# Estimates of $\Gamma^{\star}$: some limitations of $Q_\rho$ quantiles

In practice one often wish to give **confidence statements** on the estimates.




$Q_\rho$ contains points which have marginal probability at least $\rho$ of being in $\Gamma$.

$\Rightarrow$ no confidence statement on the probability of the actual excursion set containing this specific estimate.

E.g., the probabilities of $Q_\rho$ containing the excursion set (computed on a grid) are

- 0.67 for $\rho = 0.95$
- 0.009 for $\rho = 0.5$
- 0.019 for $\rho = 0.56$ (Vorob'ev)

# Conservative Estimates of $\Gamma^\star$

We denote by **conservative estimate** for $\Gamma \mid (Z(\mathbf{x}_1) = f(x_1), \ldots, Z(\mathbf{x}_n) = f(x_n))$ at level $\beta$ the largest $Q_\rho$ such that $\mathbb{P}_n(Q_\rho \subset \Gamma) \geq \beta$:

$$E_{t,\alpha} = \arg\max_{Q_\rho} \{\mu(Q_\rho) : \mathbb{P}_n(Q_\rho \subset \Gamma) \geq \beta\}$$

D. Bolin, F. Lindgren.
Excursion and contour uncertainty regions for latent Gaussian models.
Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2014.

# Conservative Estimates of $\varGamma^{\star}$

We denote by **conservative estimate** for $\varGamma \mid (Z(\mathbf{x}_1) = f(x_1), \ldots, Z(\mathbf{x}_n) = f(x_n))$ at level $\beta$ the largest $Q_\rho$ such that $\mathbb{P}_n(Q_\rho \subset \Gamma) \geq \beta$:

$$E_{t,\alpha} = \arg\max_{Q_\rho} \{\mu(Q_\rho) : \mathbb{P}_n(Q_\rho \subset \Gamma) \geq \beta\}$$

D. Bolin, F. Lindgren.
Excursion and contour uncertainty regions for latent Gaussian models.
Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2014.

Such conservative estimate $E_{t,\beta}$ is hence

- the largest quantile such that, with probability $\beta$, the response is below the threshold simultaneously at each of its locations.
- based on a confidence statement on the whole set

# Computing conservative estimates

The computation of a conservative estimate

$$E_{t,\beta} = \arg\max_{Q_\rho} \{\mu(Q_\rho) : \mathbb{P}_n(Q_\rho \subset \Gamma) \geq \beta\}$$

presents two (nested) computational bottlenecks:

1. find the set with the maximum volume;
2. compute $\mathbb{P}_n(Q_\rho \subset \Gamma)$.

# Computing conservative estimates

The computation of a conservative estimate

$$E_{t,\beta} = \arg\max_{Q_\rho} \{\mu(Q_\rho) : \mathbb{P}_n(Q_\rho \subset \Gamma) \geq \beta\}$$

presents two (nested) computational bottlenecks:

1. find the set with the maximum volume;
2. compute $\mathbb{P}_n(Q_\rho \subset \Gamma)$.

For recent work on computing the last term, see for instance

D. Azzimonti and D. Ginsbourger (2018).
Estimating orthant probabilities of high dimensional Gaussian vectors with an application to set estimation.
Journal of Computational and Graphical Statistics, 27:2, 255-267

# Computing $\mathbb{P}_n(Q_\rho \subset \Gamma)$

If $Q_\rho$ is discretized over a grid $W = \{w_1, \ldots, w_m\}$, then

$$\mathbb{P}_n(Q_\rho \subset \Gamma) = \mathbb{P}_n(Z(\mathbf{w}_1) \leq t, \ldots, Z(\mathbf{w}_m) \leq t) = 1 - \mathbb{P}_n\left(\max_{i=1,\ldots,m} Z(\mathbf{w}_i) > t\right)$$

# Computing $\mathbb{P}_n(Q_\rho \subset \Gamma)$

If $Q_\rho$ is discretized over a grid $W = \{w_1, \ldots, w_m\}$, then

$$\mathbb{P}_n(Q_\rho \subset \Gamma) = \mathbb{P}_n(Z(\mathbf{w}_1) \leq t, \ldots, Z(\mathbf{w}_m) \leq t) = 1 - \mathbb{P}_n\left(\max_{i=1,\ldots,m} Z(\mathbf{w}_i) > t\right)$$

There exists a number of algorithms to estimate $\mathbb{P}_n(Z(\mathbf{w}_1) \leq t, \ldots, Z(\mathbf{w}_m) \leq t)$:

1. quasi-MC integration techniques
   - very fast and reliable in small dimensions;
   - hardly usable for dimensions higher than 1000.
2. pure MC techniques:
   - dimension independent;
   - high number of simulations for small variance.

# Computing $\mathbb{P}_n(Q_\rho \subset \Gamma)$

If $Q_\rho$ is discretized over a grid $W = \{w_1, \dots, w_m\}$, then

$$\mathbb{P}_n(Q_\rho \subset \Gamma) = \mathbb{P}_n(Z(\mathbf{w}_1) \leq t, \dots, Z(\mathbf{w}_m) \leq t) = 1 - \mathbb{P}_n\left(\max_{i=1,\dots,m} Z(\mathbf{w}_i) > t\right)$$

There exists a number of algorithms to estimate $\mathbb{P}_n(Z(\mathbf{w}_1) \leq t, \dots, Z(\mathbf{w}_m) \leq t)$:

1. quasi-MC integration techniques
   - very fast and reliable in small dimensions;
   - hardly usable for dimensions higher than 1000.
2. pure MC techniques:
   - dimension independent;
   - high number of simulations for small variance.

**IRSN test case**

- an estimate with a good resolution requires an $100 \times 100$ grid for $D$;
- $W$ consists of $+1000$ grid points for some $Q_\rho$.

# $\mathbb{P}_n(\max_{w\in W} Z_w > T)$: proposed hybrid algorithm

**Algorithm:**

1. select $q$ grid points, denoted $W_q \subset W$;
2. compute $p' = P(\max_{w\in W_q} Z_w > t)$ with qMC quadrature;
3. estimate $\mathbb{P}_n(\max_{w\in W} Z_w > t)$ with

$$\hat{p} = p' + (1 - p')\hat{R}_q$$

where $\hat{R}_q$ is a MC estimator of

$$R_q = \mathbb{P}_n\left(\max_{w\in W\setminus W_q} Z_w > t \,\middle|\, \max_{w\in W_q} Z_w \le t\right)$$

# $\mathbb{P}_n(\max_{w \in W} Z_w > T)$: proposed hybrid algorithm

**Algorithm:**

1. select $q$ grid points, denoted $W_q \subset W$;
2. compute $p' = P(\max_{w \in W_q} Z_w > t)$ with qMC quadrature;
3. estimate $\mathbb{P}_n(\max_{w \in W} Z_w > t)$ with

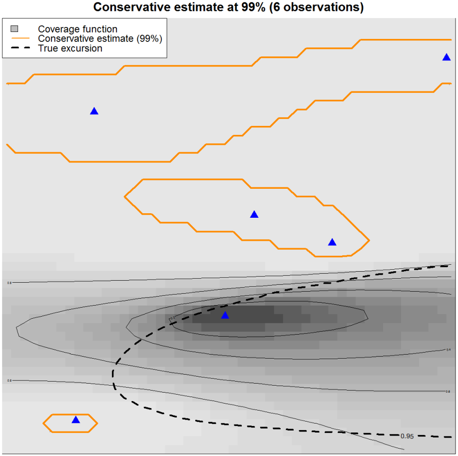
$$\hat{p} = p' + (1 - p')\hat{R}_q$$

where $\hat{R}_q$ is a MC estimator of

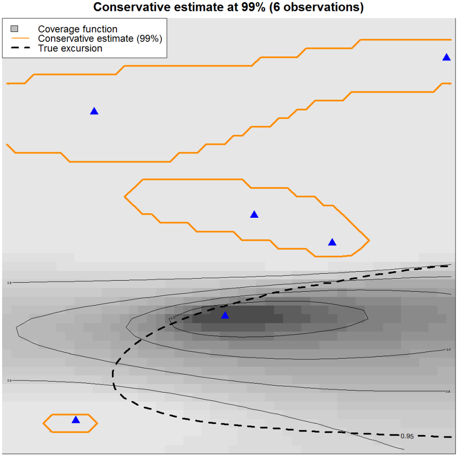
$$R_q = \mathbb{P}_n\left(\max_{w \in W \setminus W_q} Z_w > t \,\middle|\, \max_{w \in W_q} Z_w \leq t\right)$$

An asymmetric nested Monte Carlo scheme was developed for improved efficiency in $R_q$'s estimation. (See ”orthant” paper and **anMC** R package).

# Back to the test case with a conservative estimate. . .



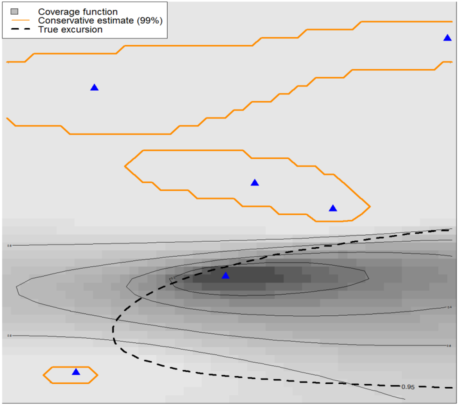
# Back to the test case with a conservative estimate. . .




NB: here, $\rho = 99.88829\%$ for a confidence of $99.12178\%$.

# . . . and associated sequential strategies

Conservative estimate at 99% (6 observations)



# For more on sequential conservative estimation

D. Azzimonti, D. Ginsbourger, C. Chevalier, J. Bect, Y. Richet (2021).
Adaptive Design of Experiments for Conservative Estimation of Excursion Sets.
Technometrics, 63:1, 13-26.

# For more on sequential conservative estimation

D. Azzimonti, D. Ginsbourger, C. Chevalier, J. Bect, Y. Richet (2021).
Adaptive Design of Experiments for Conservative Estimation of Excursion Sets.
Technometrics, 63:1, 13-26.

**Acknowledgements**:

Drs Yann Richet and Grégory Caplin (French Nuclear Safety Institute) for providing the criticality safety test case.

Special thanks to Drs. Dario Azzimonti and Clément Chevalier for numerous invaluable inputs, and more generally, to all co-authors involved.

# Outline

3 Miscellaneous
- Complements on Asymmetric Nested Monte Carlo
- Around profile extrema for excursion set visualization
- About Vorob'ev quantiles

# Computation of the remainder

$$R_q = P_n(\max_{\mathbf{x}\in E\setminus E_q} Z(\mathbf{x}) > t \mid \max_{\mathbf{x}\in E_q} Z(\mathbf{x}) \leq t)$$

**Standard Monte Carlo:**

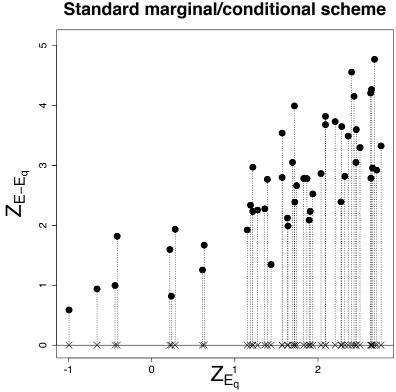
1. draw realizations $z_1^q, \ldots, z_s^q$ from $Z(E_q) \mid \max_{\mathbf{x}\in E_q} Z(\mathbf{x}) \leq t$;
2. for each $z_i^q$, draw a realization from $Z(E \setminus E_q) \mid Z(E_q) = z_i^q$;
3. Estimate $R_q$ with $R_q^{\mathrm{MC}} = \frac{1}{s}\sum_{i=1}^{s} \mathbf{1}_{\max(Z(E\setminus E_q)(\omega_i)\mid Z(E_q)=z_i^q)>t}$

The cost of step 1 is higher than the cost of step 2.

At fixed computational budget we reduce the variance of $R_q^{\mathrm{MC}}$ exploiting this difference with **asymmetric nested Monte Carlo**.

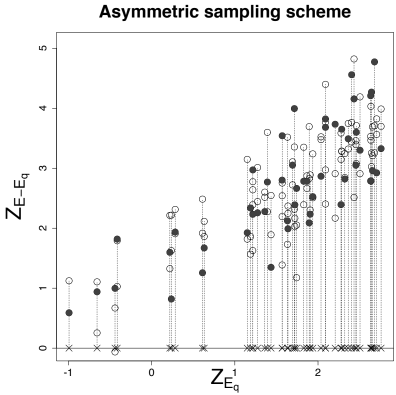
# Computation of the remainder

At fixed computational budget we reduce the variance of $R_q^{\mathrm{MC}}$ drawing many realizations of $Z(E \setminus E_q) \mid Z(E_q) = z_i^q$ for each $z_i$.



# Computation of the remainder

At fixed computational budget we reduce the variance of $R_q^{\mathrm{MC}}$ drawing many realizations of $Z(E \setminus E_q) \mid Z(E_q) = z_i^q$ for each $z_i$.



# Computation of the remainder: asymmetric nested MC

$R_q = P_n(\max_{\mathbf{x}\in E\setminus E_q} Z(\mathbf{x}) > t \mid \max_{\mathbf{x}\in E_q} Z(\mathbf{x}) \leq t)$

1. *draw realizations* $z_1^q, \dots, z_s^q$ *from* $Z(E_q) \mid \max_{\in\in E_q} Z(\mathbf{x}) \leq t$;
2. *for each* $z_i^q$, **draw** $m^* > 1$ **samples** from $Z(E \setminus E_q) \mid Z(E_q) = z_i^q$;
3. $R_q^{\text{anMC}} = \frac{1}{s}\frac{1}{m^*}\sum_{i=1}^{s}\sum_{j=1}^{m^*} \mathbf{1}_{\max(Z(E\setminus E_q;\omega_{i,j})|Z(E_q)=z_i^q)>t}$

$Var(R_q^{\text{anMC}})$ **is optimally reduced if:** $m^* = \sqrt{\frac{(\alpha+c)B}{\beta(A-B)}}$,
where $A = Var(\mathbf{1}_{\max(Z(E\setminus E_q)|Z(E_q))>t})$,
$B = \mathbb{E}\left[Var(\mathbf{1}_{\max(Z(E\setminus E_q)|Z(E_q))>t} \mid \max_{\mathbf{x}\in E_q} Z(\mathbf{x}) \leq t)\right]$ and $\alpha, \beta, c$ system dependent constants.

Azzimonti, D. and Ginsbourger D. (2018). *Estimating orthant probabilities of high dimensional Gaussian vectors with an application to set estimation. Journal of Computational and Graphical Statistics* 27:2, 255-267.

# Properties of the anMC estimator

For each realization $w_i$ we sample $m$ realizations $y_{i,1}, \ldots, y_{i,m}$ from $Y \mid w = W_i$, leading to a realization $\frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} g(w_i, y_{i,j})$ of the estimator

$$\widetilde{G} = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} g(W_i, Y_{i,j}).$$

**Proposition**

Consider $n$ independent copies $W_1, \ldots, W_n$ of $W$ and, for each $W_i$, $m$ copies $Y_{i,j} = Y_j \mid W_i \; j = 1, \ldots, m$, independent conditionally on $W_i$. Then,

$$Var(\widetilde{G}) = \frac{1}{n} Var(g(W_1, Y_{1,1})) - \frac{m-1}{nm}\mathbb{E}\big[Var(g(W_1, Y_{1,1}) \mid W_1)\big].$$

### Optimality for anMC

Under the same assumptions, $\widetilde{G}$ has minimal variance when

$$m = \widetilde{m} = \sqrt{\frac{(\alpha + c)B}{\beta(A - B)}},$$

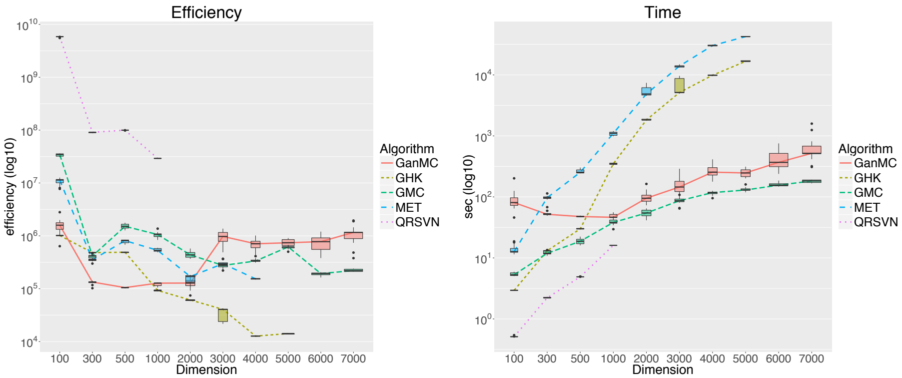
where $A = Var(g(W_1, Y_{1,1}))$ and $B = \mathbb{E}\big[Var(g(W_1, Y_{1,1}) \mid W_1)\big]$. Moreover denote with $\varepsilon = \widetilde{m} - \lfloor \widetilde{m} \rfloor$, then the optimal integer is $m^* = \lfloor \widetilde{m} \rfloor$ if

$$\varepsilon < \frac{(2\widetilde{m} + 1) - \sqrt{4(\widetilde{m})^2 + 1}}{2}$$

or $m^* = \lceil \widetilde{m} \rceil$ otherwise.

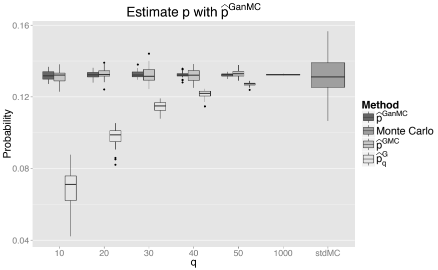
If $m^* > \frac{2(\alpha+c)B}{(c+\alpha)B+\beta(A-B)}$ then $Var(\widetilde{G}) = Var(\widehat{G})\,[1 - \eta]$, with $\eta \in (0, 1)$, where $\widehat{G}$ is the standard MC estimator.

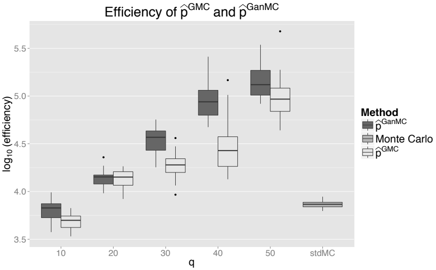
# Comparison GanMC/GMC with state-of-the-art




**Benchmark:** 6d GRF, discretization: 1000 Sobol' points, $k$ Matérn ($\nu = 5/2$) with $\theta = [0.5, 0.5, 1, 1, 0.5, 0.5]^T$ and $\sigma^2 = 8$, $m$ constant, $t = 5$.

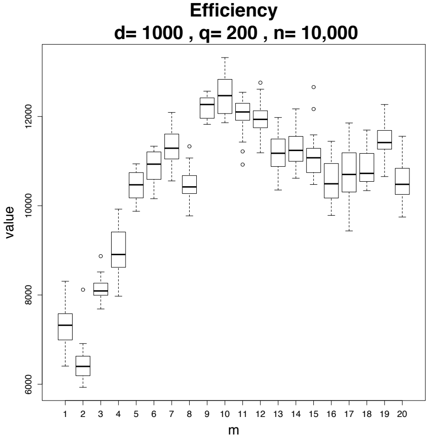
# More comparisons anMC/MC




**Benchmark:** 6d GRF, discretization: 1000 Sobol' points, *k* Matérn ($\nu = 5/2$) with $\theta = [0.5, 0.5, 1, 1, 0.5, 0.5]^T$ and $\sigma^2 = 8$, *m* constant.

# Choice of $m^*$




**Benchmark:** 6d GRF with $m$ constant, $k$ Matérn ($\nu = 5/2$), $\theta = [0.5, 0.5, 1, 1, 0.5, 0.5]$, $\sigma^2 = 8$. Discretized over first 1000 points of Sobol' sequence.

**Quantity of interest:**
$1 - p = P(X < t)$, $t = 5$.

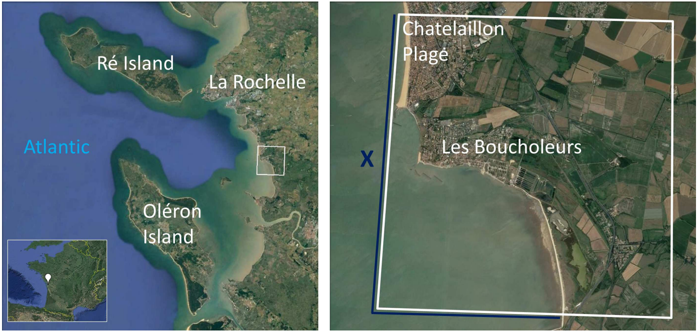
GanMC algorithm with $p_q$ fixed, $R_q$ (anMC) computed for different $m$.

Median estimated value $\widehat{p} = 0.9644$, $\widehat{p}_q = 0.9636$.

# Outline

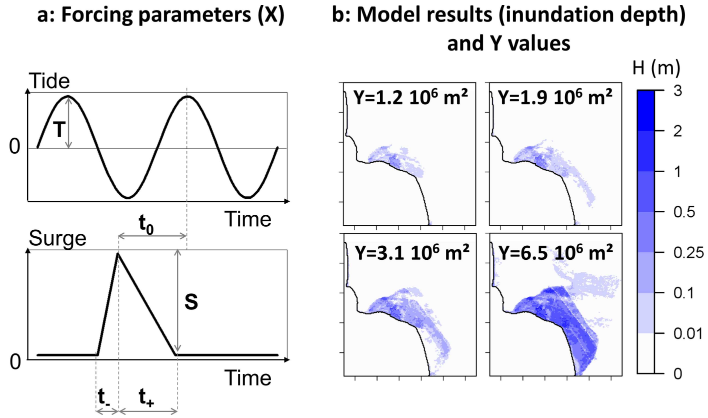
3. Miscellaneous
   - Complements on Asymmetric Nested Monte Carlo
   - Around profile extrema for excursion set visualization
   - About Vorob'ev quantiles

# Simulation of coastal flooding at "Les Boucholeurs"



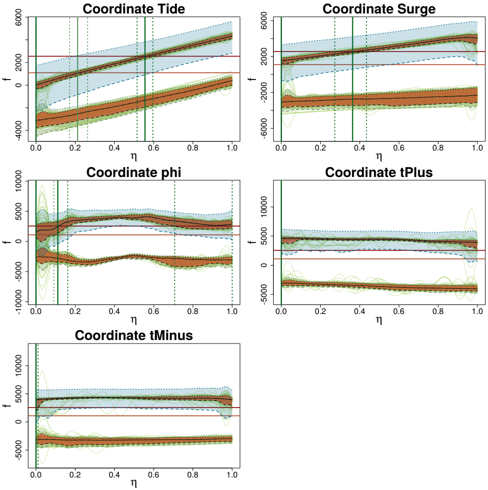

Study site location (left) and computational domain limits (right, in white) with location of the forcing conditions (right, in blue).

# Test case input and output parametrization




(a) Schematic representation of the tide and surge temporal signals and the different parameters describing them. (b) Maps of inland water height for given values of the parameters, and deduced value of flood surface.

# Estimated coordinate profile maxima for the 5 inputs



# Key underlying result

**Theorem**

*Consider* $(Z(\mathbf{x}))_{\mathbf{x}\in D} \sim GP(\mu, \mathfrak{K})$ *and an approximating process of* $Z$, $\widetilde{Z}$, *defined by* $\widetilde{Z}(\mathbf{x}) = a(\mathbf{x}) + \mathbf{b}^T(\mathbf{x})Z_G$ *where the* $\mathbf{a}, \mathbf{b}$ *functions and* $G = \{\mathbf{g}_1, \dots, \mathbf{g}_\ell\} \subset D$ *(*$\ell \geq 1$*) are given. Then, for* $T \subset D$ *and any* $u > \mu_T^{\widetilde{\Delta}}$,

$$\mathbb{P}\left(|\sup_{\mathbf{x}\in T} Z(\mathbf{x}) - \sup_{\mathbf{x}\in T} \widetilde{Z}(\mathbf{x})| > u\right) \leq 2\exp\left(-\frac{(u-\mu_T^{\widetilde{\Delta}})^2}{2(\sigma_T^{\widetilde{\Delta}})^2}\right), \tag{1}$$

*where*

$$\mu_T^{\widetilde{\Delta}} = \sup_{\mathbf{x}\in T}|\mu^{\widetilde{\Delta}}(\mathbf{x})| \text{ and } (\sigma_T^{\widetilde{\Delta}})^2 = \sup_{\mathbf{x}\in T} \mathfrak{K}^{\widetilde{\Delta}}(\mathbf{x},\mathbf{x}) \text{ with} \tag{2}$$

$$\mu^{\widetilde{\Delta}}(\mathbf{x}) = \mathbb{E}[Z(\mathbf{x}) - \widetilde{Z}(\mathbf{x})] = \mu(\mathbf{x}) - a(\mathbf{x}) - \mathbf{b}^T(\mathbf{x})\mu(G)$$

$$\mathfrak{K}^{\widetilde{\Delta}}(\mathbf{x},\mathbf{x}') = \mathfrak{K}(\mathbf{x},\mathbf{x}') - \mathfrak{K}(\mathbf{x}',G)\mathbf{b}(\mathbf{x}) - \mathfrak{K}(\mathbf{x},G)\mathbf{b}(\mathbf{x}') + \mathbf{b}^T(\mathbf{x})\mathfrak{K}(G,G)\mathbf{b}(\mathbf{x}'),$$

*If* $\widetilde{Z} - Z$ *is centred then (1) is valid for any* $u > 0$.

# For more detail

More on the profile maxima approach and its application to the BRGM data can be found in

D. Azzimonti, D. Ginsbourger, J. Rohmer, D. Idier (2019, to appear in Technometrics)
Profile extrema for visualizing and quantifying uncertainties on excursion regions. Application to coastal flooding.
https://arxiv.org/abs/1710.00688

# For more detail

More on the profile maxima approach and its application to the BRGM data can be found in

D. Azzimonti, D. Ginsbourger, J. Rohmer, D. Idier (2019, to appear in Technometrics)
Profile extrema for visualizing and quantifying uncertainties on excursion regions. Application to coastal flooding.
https://arxiv.org/abs/1710.00688

For more on random fields and geometry, see in particular

R. J. Adler and J. E. Taylor (2007)
Random Fields and Geometry.
Springer

and references therein.

# Outline

3 Miscellaneous
- Complements on Asymmetric Nested Monte Carlo
- Around profile extrema for excursion set visualization
- About Vorob'ev quantiles

# Generalized optimality property for Vorob'ev quantiles

**Proposition**

For any $\rho \in [0, 1]$, the Vorob'ev quantile

$$Q_\rho = \{x \in D : p_n(x) \geq \rho\}$$

minimizes the expected distance in measure with $\Gamma$ among measurable sets $M$ such that $\mu(M) = \mu(Q_\rho)$, i.e.,

$$\mathbb{E}_n\left[\mu(Q_\rho \Delta \Gamma)\right] \leq \mathbb{E}_n\left[\mu(M \Delta \Gamma)\right],$$

for any measurable set $M$ such that $\mu(M) = \mu(Q_\rho)$.

# Generalized optimality property for Vorob'ev quantiles

**Proposition**

For any $\rho \in [0, 1]$, the Vorob'ev quantile

$$Q_\rho = \{x \in D : p_n(x) \geq \rho\}$$

minimizes the expected distance in measure with $\Gamma$ among measurable sets $M$ such that $\mu(M) = \mu(Q_\rho)$, i.e.,

$$\mathbb{E}_n \left[\mu(Q_\rho \Delta \Gamma)\right] \leq \mathbb{E}_n \left[\mu(M \Delta \Gamma)\right],$$

for any measurable set $M$ such that $\mu(M) = \mu(Q_\rho)$.

A proof of this property is presented in Dario Azzimonti's PhD thesis (2016).