

# Profile extrema for visualizing and quantifying uncertainties on excursion regions. Application to coastal flooding

Dario Azzimonti<sup>\*†</sup>

Istituto Dalle Molle di studi sull'Intelligenza Artificiale (IDSIA)  
Galleria 2, Via Cantonale 2c, 6928 Manno, Switzerland,

and

David Ginsbourger<sup>\*</sup>

UQOD, Idiap Research Institute

Centre du Parc, Rue Marconi 19, 1920 Martigny, Switzerland; and  
IMSV, Department of Mathematics and Statistics, University of Bern  
Alpeneggstrasse 22, 3012 Bern, Switzerland,

and

Jérémy Rohmer<sup>\*‡</sup>

BRGM

3 Av. C. Guillemin, BP 36009, 45060 Orléans Cedex 2, France,  
and

Déborah Idier<sup>\*‡</sup>

BRGM

3 Av. C. Guillemin, BP 36009, 45060 Orléans Cedex 2, France.

---

<sup>\*</sup>All authors thank F. Paris and R. Pedreros (BRGM) for providing the Boucholeurs flooding model.

<sup>†</sup>The author acknowledges the financial support of the *Hasler Foundation*, grant number 16065 and the *SNSF*, grant number 167199.

<sup>‡</sup>The authors would like to thank the OQUAIDO chair in applied mathematics and acknowledge financial support from BRGM via the PROSUB project

## Abstract

We consider the problem of describing excursion sets of a real-valued function  $f$ , i.e. the set of inputs where  $f$  is above a fixed threshold. Such regions are hard to visualize if the input space dimension,  $d$ , is higher than 2. For a given projection matrix from the input space to a lower dimensional (usually 1, 2) subspace, we introduce profile sup (inf) functions that associate to each point in the projection's image the sup (inf) of the function constrained over the pre-image of this point by the considered projection. Plots of profile extrema functions convey a simple, although intrinsically partial, visualization of the set. We consider expensive to evaluate functions where only a very limited number of evaluations,  $n$ , is available, e.g.  $n < 100d$ , and we surrogate  $f$  with a posterior quantity of a Gaussian process (GP) model. We first compute profile extrema functions for the posterior mean given  $n$  evaluations of  $f$ . We quantify the uncertainty on such estimates by studying the distribution of GP profile extrema with posterior quasi-realizations obtained from an approximating process. We control such approximation with a bound inherited from the Borell-TIS inequality. The technique is applied to analytical functions ( $d = 2, 3$ ) and to a 5-dimensional coastal flooding test case for a site located on the Atlantic French coast. Here  $f$  is a numerical model returning the area of flooded surface in the coastal region given some offshore conditions. Profile extrema functions allowed us to better understand which offshore conditions impact large flooding events.

*Keywords:* excursion set; set estimation; Gaussian process; profile extrema function; Graphical Methods.

# 1 Introduction

## 1.1 Problem statement

In many domains of application (structural reliability, [Dubourg et al. \(2013\)](#), nuclear safety, [Chevalier et al. \(2014\)](#), environmental problems, [Bolin and Lindgren \(2015\)](#), natural risk, [Rohmer and Idier \(2012\)](#); [Bayarri et al. \(2009\)](#)) the region in the parameter space leading to response values above or below a given threshold is of crucial interest. In particular here, we assume that a phenomenon can be modelled by a function  $f : D \subset \mathbb{R}^d \rightarrow \mathbb{R}$ , where  $D$  is a compact set and we focus on excursion sets of the form

$$\Gamma = \{\mathbf{x} \in D : f(\mathbf{x}) \geq \tau\},$$

where  $\tau \in \mathbb{R}$  is a prescribed threshold depending on the application at hand. This work presents a method to gain visual insights on  $\Gamma$  when the input space dimension  $d \geq 3$ .

Visualization of point clouds and manifolds in high dimensions is a very active field of research. Principal components analysis and its kernel version ([Schölkopf et al., 1997](#)) play a fundamental role in dimensionality reduction for high dimensional data. Parallel coordinates ([Inselberg, 1985, 2009](#)) are widely used for exploratory data analysis. In order to visualize more complex structure, techniques relying on the topological concept of Morse-Smale complex ([Edelsbrunner and Harer, 2008](#); [Gerber et al., 2010](#)) provide powerful, but hard to interpret, tools. Projection based techniques, in particular tours (see, e.g., [Asimov, 1985](#); [Cook et al., 1995, 2008](#)), are also a prominent method to visualize data in high-dimensional Euclidean spaces. All these techniques, however, are based on a finite set of data points while, in our case, we provide a representation that exploits the structure of  $\Gamma$ .

The contribution of this paper is two-fold: first we show how to use profile extrema functions for visualizing excursion sets of known deterministic functions; second, we implement these tools for expensive computer experiments with Gaussian process (GP) emulation.

For a given projection matrix  $\Psi \in \mathbb{R}^{d \times p}$ , a profile sup (inf respectively) is a function  $P_\Psi^{\sup} f$  ( $P_\Psi^{\inf} f$ ) that associates to each point  $\eta$  in the image of  $\Psi^T$ , the sup (inf respectively) of  $f$  over all points with projection equal to  $\eta$ , i.e.  $P_\Psi^{\sup} f(\eta) = \sup_{\Psi^T \mathbf{x} = \eta} f(\mathbf{x})$ . For example, if  $\Psi^T = [1, 0, \dots, 0] \in \mathbb{R}^{1 \times d}$  is the canonical projection over the first coordinate, then  $P_\Psi^{\sup} f(\eta)$  is the sup of  $f$  over all points in  $D$  with first coordinate equal to  $\eta$ . In this case we can plot the values of  $P_\Psi^{\sup} f$  and  $P_\Psi^{\inf} f$  over their 1-dimensional ( $p$  dimensional, in general) domain and study such functions. In particular, all  $\tilde{\eta}$  such that  $P_\Psi^{\sup} f(\tilde{\eta}) < \tau$ , i.e. all inputs  $\mathbf{x} \in D$  such that  $\Psi^T \mathbf{x} = \tilde{\eta}$ , identify regions of non-excursion. Analogously  $P_\Psi^{\inf} f(\eta) > \tau$  selects regions of excursion. We compute profile extrema for known, fast-to-evaluate functions  $f$  with numerical optimization, as described in [Section 2](#).

In this paper, however, we focus on the case where the function  $f$  is very expensive to evaluate and thus the number of evaluations,  $n$ , is limited, e.g.  $n < 100d$ . Following the classical computer experiments literature (see, e.g., [Sacks et al., 1989](#); [Santner et al., 2003](#)), we define a design of experiments  $X_n = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in D^n$  and denote with  $\mathbf{y}_n = (f(\mathbf{x}_1), \dots, f(\mathbf{x}_n))^T$  the corresponding vector of evaluations. We assume that  $f$  is a realization of a GP  $(Z_{\mathbf{x}})_{\mathbf{x} \in D} \sim GP(\mu, \mathfrak{K})$  with prior mean function  $\mu : D \subset \mathbb{R}^d \rightarrow \mathbb{R}$  and covariance kernel  $\mathfrak{K} : D \times D \rightarrow \mathbb{R}$ . The set  $\Gamma$  can be estimated from the posterior GP

distribution (see, e.g. [Bingham et al., 2014](#); [Bolin and Lindgren, 2015](#); [Azzimonti, 2016](#)), in particular, here we consider the plug-in estimate

$$\hat{\Gamma}_n = \{\mathbf{x} \in D : \gamma_n(\mathbf{x}) \geq \tau\},$$

where  $\gamma_n$  is a posterior quantity related to  $Z$ . Examples of interesting quantities  $\gamma_n$  are the posterior mean, the  $1 - \alpha$  posterior quantiles or a posterior realization of  $Z$ . If  $d = 1, 2$ , a visualization of  $\hat{\Gamma}_n$  only requires evaluations of  $\gamma_n$  on a grid; however, when  $d \geq 3$ , this approach is problematic. We propose a technique to compute and plot profile extrema functions for  $\gamma_n$  and we detail how to use profile extrema to identify regions of interest.

A GP model not only provides an estimate for  $\Gamma$ , but it also allows for a quantification of the associated uncertainty. In particular, here we study the distribution of the stochastic object  $P_\Psi^{\text{sup}}Z$  and we provide point-wise confidence statements on profile extrema functions. The quantities  $P_\Psi^{\text{sup}}Z$ ,  $P_\Psi^{\text{inf}}Z$  are strongly related to extreme values of Gaussian processes indexed by compact sets which have been widely studied in probability theory (see, e.g., [Adler and Taylor, 2007](#); [Azaïs and Wschebor, 2009](#)) and have been the subject of recent interest in the computer experiments literature, see, e.g. [Chevalier \(2013, Chapter 6\)](#) and [Ginsbourger et al. \(2014\)](#). Here we obtain point-wise confidence intervals for profile extrema functions by using posterior quasi-realizations of an approximating process obtained as in [Azzimonti et al. \(2016\)](#). Moreover we introduce a probabilistic bound for the quantiles of the posterior profile function based on the sample quantiles of the approximating process. This procedure enabled the identification of regions of interests for the coastal flooding test case presented in [Section 4](#).

## 1.2 Outline of the paper

In [Section 2](#) we introduce profile extrema functions for a generic function  $\gamma$  and we apply them to a synthetic test case. In this section, we also discuss implementation details. [Section 3](#) describes the GP model and the quantification of uncertainty on the profiles with approximate posterior GP simulations. Moreover we provide a probabilistic bound ([Theorem 1](#) and [Corollary 1](#), proofs in [Appendix A](#)) for this approximation. In [Section 4](#) we apply the procedure on a coastal flooding problem where  $f$  is evaluated with an expensive-to-evaluate computer experiment. Further details are in [Appendix C](#). Finally, in [Section 5](#), we discuss advantages and drawbacks of the current method and propose possible extensions.

# 2 Profile extrema functions

## 2.1 Definitions

In this section, we introduce the concept of profile extrema for a continuously differentiable function  $\gamma : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^d$  compact. While  $\gamma$  can be any function, in this paper we mainly consider  $\gamma$  as the posterior mean, a quantile or a posterior realization of a GP. The first example of profile extrema are coordinate profiles defined below, i.e.  $2d$  univariate functions describing the extremal behavior of  $\gamma$  for each coordinate.

**Definition 1** (Coordinate profile extrema). *For each  $i \in \{1, \dots, d\}$ , let us denote with  $\mathbf{e}_i$  the  $i$ -th canonical coordinate vector, then the  $i$ -th coordinate profile extrema functions are*

$$P_{\mathbf{e}_i}^{\sup} \gamma = P_i^{\sup} \gamma : \eta \in E_{\mathbf{e}_i} \longrightarrow P_i^{\sup} \gamma(\eta) := \sup_{x_i = \eta} \gamma(\mathbf{x}) \quad (1)$$

$$P_{\mathbf{e}_i}^{\inf} \gamma = P_i^{\inf} \gamma : \eta \in E_{\mathbf{e}_i} \longrightarrow P_i^{\inf} \gamma(\eta) := \inf_{x_i = \eta} \gamma(\mathbf{x}). \quad (2)$$

where  $E_{\mathbf{e}_i} = \{\eta \in \mathbb{R} : \mathbf{e}_i^T \mathbf{x} = \eta, \mathbf{x} \in D\}$  and  $\mathbf{x} = (x_1, \dots, x_i, \dots, x_d)$ .

By studying these functions we can split the input space  $D$  in simple regions providing a marginal visualization of  $\Gamma$ . Consider the first coordinate profile sup function  $P_1^{\sup} \gamma$ . If there exists  $\eta^*$  such that  $P_1^{\sup} \gamma(\eta^*) < \tau$ , the function  $\gamma$  is below  $\tau$  for all  $\mathbf{x} \in D_{i, \eta^*} = \{\mathbf{x} \in D : x_i = \eta^*\}$ . In a symmetric way, if  $\eta^*$  is such that  $P_1^{\inf} \gamma(\eta^*) \geq \tau$ , the function  $\gamma$  is above  $\tau$  over the whole  $D_{i, \eta^*}$ . Since  $P_1^{\sup}$  and  $P_1^{\inf}$  are univariate functions we can evaluate them over a discretization of  $E_{\mathbf{e}_1}$  and with a simple plot we can locate those regions. Profile functions could also lead to more undetermined situations: for example, for  $\tilde{\eta}$  such that  $P_1^{\sup} \gamma(\tilde{\eta}) \geq \tau$ , there exists at least one point  $\bar{\mathbf{x}}^{\tilde{\eta}} = (\tilde{\eta}, \bar{x}_2, \dots, \bar{x}_d) \in D_{1, \tilde{\eta}}$  such that  $\gamma(\bar{\mathbf{x}}^{\tilde{\eta}}) \geq \tau$  and we cannot flag the region  $D_{1, \tilde{\eta}}$  neither as excursion nor as non-excursion. Coordinate profile extrema are the results of constrained optimization problems over particular projections on the coordinate axes. Profile extrema functions further generalize this concept.

**Definition 2** (Profile extrema). *Consider a full column rank matrix  $\Psi \in \mathbb{R}^{d \times p}$  with  $p < d$ . Let  $E_{\Psi} := \{\Psi^T \mathbf{x} : \mathbf{x} \in D\}$  and  $D_{\Psi, \eta} := \{\mathbf{x} \in D : \Psi^T \mathbf{x} = \eta\}$ , then the functions*

$$P_{\Psi}^{\sup} \gamma : \eta \in E_{\Psi} \longrightarrow P_{\Psi}^{\sup} \gamma(\eta) := \sup_{\mathbf{x} \in D_{\Psi, \eta}} \gamma(\mathbf{x}) \quad (3)$$

$$P_{\Psi}^{\inf} \gamma : \eta \in E_{\Psi} \longrightarrow P_{\Psi}^{\inf} \gamma(\eta) := \inf_{\mathbf{x} \in D_{\Psi, \eta}} \gamma(\mathbf{x}), \quad (4)$$

are called Profile sup and Profile inf functions of  $\gamma$  respectively. The set  $E_{\Psi}$  is the image of  $D$  under  $\Psi^T$  and  $D_{\Psi, \eta}$  is the pre-image of  $\eta$  under  $\Psi^T$  restricted to  $D$ .

Since we are interested in visualization here we only consider  $p \in \{1, 2\}$  and  $\Psi$  as orthogonal projections because they have a more direct interpretation.

## 2.2 Analytical example

Let us review the concepts just introduced on an analytical test case. Consider

$$\gamma(\mathbf{x}) = \sin(a \mathbf{v}_1^T \mathbf{x} + b) + \cos(c \mathbf{v}_2^T \mathbf{x} + d) \quad \mathbf{x} \in [0, 1]^2, \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2 \quad (5)$$

where  $a, b, c, d \in \mathbb{R}$ ,  $\mathbf{v}_1 = [\cos(\theta), \sin(\theta)]^T$ ,  $\mathbf{v}_2 = [\cos(\theta + \pi/2), \sin(\theta + \pi/2)]^T$  and  $\theta = \pi/6$ .

Figure 1 shows contour lines of  $\gamma$  with parameters chosen as  $a = 1, b = 0, c = 10, d = 0$ . We fix a threshold  $\tau = 0$  and we compute the coordinate profile extrema, plotted in Figure 2. In this case,  $P_1^{\sup} \gamma(\eta) < \tau$  for  $\eta \in [0, 0.13]$  and  $P_2^{\sup} \gamma(\eta) < \tau$  for  $\eta \in [0, 0.25] \cup [0.70, 0.80]$ . The second coordinate seems more informative on the excursion:  $P_1^{\sup} \gamma$  flags as non-excursion a region with volume 0.13 while  $P_2^{\sup} \gamma$  flags a region with volume 0.35. The

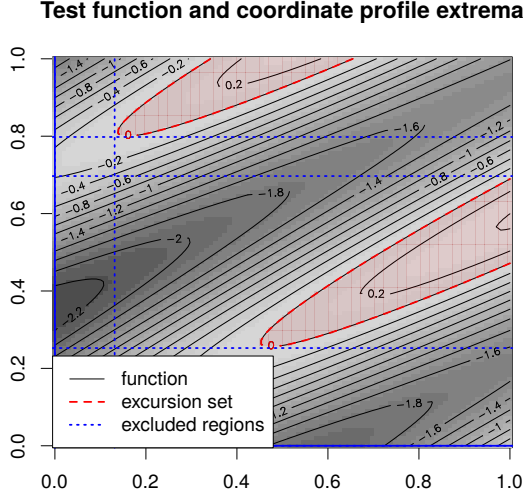


Figure 1: Example of 2-dimensional test case in equation (5). Threshold  $\tau$  in red dashed. The regions excluded with coordinate profiles are delimited by dotted blue lines.

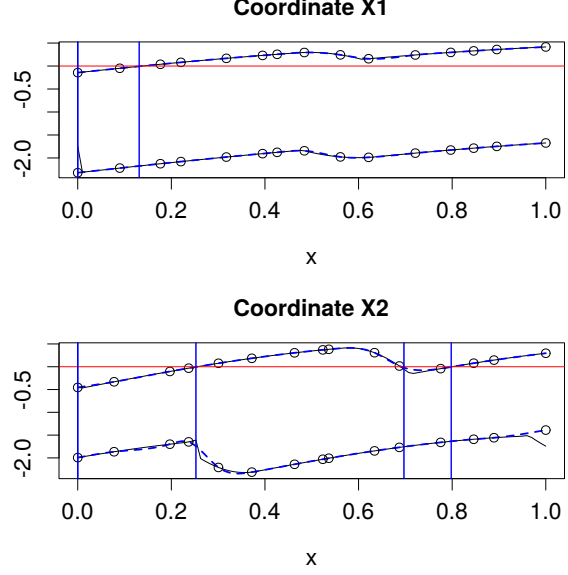


Figure 2: Profile extrema functions for  $\gamma$  (black solid) and their approximation (blue dashed) from  $k = 15$  points (circles). Threshold  $\tau = 0$ , horizontal solid line.

true set has volume 0.127, i.e. the region of non excursion has volume 0.873. This example shows that canonical directions might not be the most appropriate ones to visualize this function. Figure 4 shows profile extrema computed along the (oblique) generating directions  $\mathbf{v}_1, \mathbf{v}_2$ , i.e. profile extrema in Definition 2 with  $\Psi = \mathbf{v}_1$  or  $\mathbf{v}_2$ . By using oblique directions we can say that for  $\mathbf{x}$  such that  $\mathbf{v}_1^T \mathbf{x} \in [0, 0.52] \cup [1.22, 1.37]$  (area excluded 0.33) and for  $\mathbf{x}$  such that  $\mathbf{v}_2^T \mathbf{x} \in [-0.5, -0.1] \cup [0.11, 0.54] \cup [0.71, 0.87]$  (area excluded 0.68) there is no excursion. As shown in Figure 3 profile extrema along this direction allow us to exclude a much larger portion (total area excluded 0.79 versus 0.44 with coordinate profiles) of the input space. In Appendix E we combine coordinate, oblique and bivariate profiles extrema for a finer visualization of the excursion set on a 3-dimensional analytical function inspired by (5).

## 2.3 Implementation

The value of the profile  $\sup P_\Psi^{\sup} \gamma(\eta)$ , for a fixed matrix  $\Psi \in \mathbb{R}^{d \times p}$ , is the solution of a constrained optimization problem. In general,  $\gamma$  is not convex, therefore we are not guaranteed to achieve a global optimum. We implement a local method to obtain profile extrema functions for  $\Psi = \mathbf{e}_i \in \mathbb{R}^d, i \in \{1, \dots, d\}$ , and for a matrix  $\Psi \in \mathbb{R}^{d \times p}$  of rank  $p$ .

In the first case we obtain the value for  $P_i^{\sup} \gamma$ , for a coordinate  $i$  with the following procedure. For each  $\eta$ , define a new function  $\tilde{\gamma}_{i,\eta} : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$  such that  $\tilde{\gamma}_i(\mathbf{z}) := \gamma(z_1, \dots, z_{i-1}, \eta, z_i, \dots, z_{d-1})$  for any  $\mathbf{z} = (z_1, \dots, z_{d-1}) \in D_{-i}$ , where  $D_{-i}$  is the  $d-1$  dimensional restriction of  $D$  without the  $i$ th coordinate. We can estimate the value  $P_i^{\sup} \gamma(\eta)$

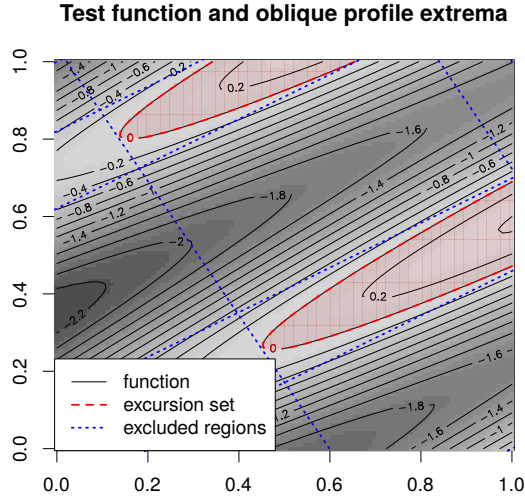


Figure 3: Example of 2-dimensional test case in equation (5). Regions excluded by oblique profile extrema delimited by dotted lines.

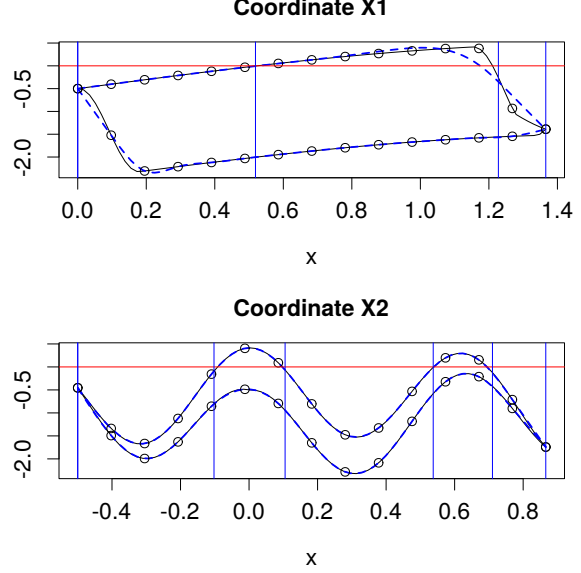


Figure 4: Oblique profile extrema functions for  $\gamma$  (black solid lines) and their approximation (blue dashed lines) from  $k = 15$  points.

with a L-BFGS-B algorithm. If the gradient of  $\gamma$  is available it is used in the optimization, otherwise numerical derivatives are used. This method depends on the starting point chosen for the optimization. In general we are interested in the evaluation of the functions  $P_i^{\sup} \gamma(\eta)$  for all  $\eta$  and for all  $i \in \{1, \dots, d\}$ . By exploiting the smoothness of  $\gamma$  we re-use the previously obtained points of optimum as starting point for the subsequent optimizations.

In the second case, we assume that the matrix  $\Psi$  has full column rank and that  $D$  is a hyper-cube. The value  $P_{\Psi}^{\sup} \gamma(\eta)$  is the solution of the constrained optimization problem

$$\text{maximize } \gamma(\mathbf{x}) \quad \text{subject to } \Psi^T \mathbf{x} = \eta, \mathbf{x} \in D. \quad (6)$$

We transform (6) in an unconstrained optimization problem first by computing an arbitrary  $\xi \in \mathbb{R}^d$  such that  $\Psi^T \xi = \eta$  and then by observing that if  $\mathbf{x}$  satisfies the equality constraint then  $\mathbf{x} = \xi + \text{Null}(\Psi^T) \mathbf{z}$ , where  $\text{Null}(\Psi^T) \in \mathbb{R}^{d \times (d-p)}$  is the matrix representing the null space of  $\Psi^T$  and  $\mathbf{z} \in \mathbb{R}^{d-p}$ . We can then maximize  $\gamma$  over the  $d-p$  dimensional space using a (hard) barrier function for the hyper-cube inequality constraints. Both  $\xi$  and the starting point(s) for the interior point method are computed by solving auxiliary linear programs.

In both cases, a multi-start optimization procedure provides a more reliable estimate of the global optimum at an increased computational cost. The two methods are implemented in the R programming language (R Core Team, 2018) in the package `profExtrema`, available on CRAN. The first method uses the L-BFGS-B (Byrd et al., 1995) implementation in the function `optim`, base package; the second method solves the auxiliary linear programs with `lpccdd` from the package `rcdd` (a R interface for `cddlib`, Fukuda (2007)) and optimizes the barrier function with the BFGS implementation in `optim`.



## 2.4 Approximation of profile extrema functions

The cost of evaluating the profile extrema functions at each  $\eta$  increases as  $d-p$  grows larger. Plots of profile extrema functions with  $\Psi \in \mathbb{R}^{d \times p}$  with  $p = 1$  or  $2$  require evaluations over  $E_\Psi \subset \mathbb{R}^p$ . If the input dimension  $d$  is high, each evaluation is a high dimensional optimization and visualizing the function at sufficient resolution becomes expensive. We mitigate this issue by approximating  $P_\Psi^{\text{sup}}\gamma(\eta)$  at any  $\eta$  with an interpolation of profile function evaluations at few space filling points  $\eta_1, \dots, \eta_k \in E_\Psi$ .

Coordinate profile extrema functions can be seen as a special type of optimal value functions. First order characterizations of such objects have been studied in optimization and perturbation theory, see, e.g., [Danskin \(1967\)](#); [Bonnans and Shapiro \(1998\)](#). In particular, consider the  $i$ -th coordinate profile extrema function  $P_i^{\text{sup}}\gamma$ . We can write  $\gamma : D_{|-i} \times D_{|i} \rightarrow \mathbb{R}$  and  $P_i^{\text{sup}}\gamma(\eta) = \sup_{y \in D_{|-i}} \gamma(y, \eta)$ . If  $\gamma(y, \cdot)$  is differentiable with  $\nabla_\eta \gamma(y, \cdot)$  continuous for  $y \in D_{|-i}$  and there exists a unique  $y^* = \arg \max \gamma(y, \eta)$ , then  $\nabla_\eta P_i^{\text{sup}}\gamma(\eta) = \nabla_\eta \gamma(\cdot, \cdot)|_{(y^*, \eta)}$ . We can then evaluate  $P_i^{\text{sup}}\gamma(\eta)$  and  $\nabla_\eta P_i^{\text{sup}}\gamma(\eta)$  at few space filling points  $\eta_1, \dots, \eta_k$ , for a small  $k > 0$ , and interpolate the value  $P_i^{\text{sup}}\gamma(\eta)$  at any  $\eta \in E_\Psi$  using a first order approximation. For arbitrary  $\Psi$ , we approximate  $P_\Psi^{\text{sup}}\gamma(\eta)$  at any  $\eta$  with cubic splines or kriging interpolators as they do not require  $\nabla_\eta P_\Psi^{\text{sup}}$ .

Below we present an example for the case  $p = 1$  with a cubic spline approximation. The blue dashed lines in Figures [2](#), [4](#) show the approximation of coordinate and oblique profile extrema functions, the exact functions are also plotted in black solid lines. We select  $k = 15$  points  $\eta_1, \dots, \eta_k \in E_{\mathbf{e}_i}$  with Latin hyper-cube sampling, the approximate profiles at any point are computed with a cubic spline regression from  $k$  exact calculations. Exact coordinate profiles  $P_i^{\text{sup}}, P_i^{\text{inf}}$ , for  $i = 1, 2$  over a grid of equispaced 100 points on  $[0, 1]$  require 0.25 seconds and oblique profiles  $P_{\mathbf{v}_i}^{\text{sup}}, P_{\mathbf{v}_i}^{\text{inf}}$ ,  $i = 1, 2$  require 0.64 seconds on a laptop with 2.6 GHz Intel Core i5-7300U CPU. On the same laptop, cubic spline approximations plotted on the same grid require 0.04 seconds for coordinate profiles (0.13 oblique). On this example, the maximum absolute median errors achieved with the approximation of  $P^{\text{inf}}\gamma$ ,  $P^{\text{sup}}\gamma$  are 0.07% and 0.54% for coordinate profiles and 0.19% and 5.08% for oblique profiles. The number of approximation points,  $k$ , is an important parameter for the approximation, the choice of  $k$  boils down to a trade-off between computational time and precision. For 1-dimensional profile extrema ( $p = 1$ ) the default choice implemented is  $k = 10\sqrt{d}$ . When the profile functions are smooth this heuristics often leads to good results, however it is generally better to compare different values for  $k$ . Cubic splines are implemented here only for  $p = 1$ , for  $p = 2$  we use kriging.

## 3 Uncertainty quantification for profile extrema

### 3.1 Profile extrema for expensive-to-evaluate functions

In the previous section we introduced profile extrema functions to visualize excursion sets of arbitrary functions. This method, however, requires performing potentially very high numbers of function evaluations which can be prohibitive in many computer experiments, as, for example, in the motivating application of Section [4](#). We therefore rely on a Gaus-



sian process emulator based on few evaluations of the expensive function. We consider a design of experiments (DoE) with  $n$  points  $X_n \in D^n$  and the corresponding function evaluations  $\mathbf{y}_n$  as described in Section 1. We select a prior GP  $(Z_{\mathbf{x}})_{\mathbf{x} \in D} \sim GP(\mu, \mathfrak{K})$  with mean  $\mu(\mathbf{x}) = \sum_{j=1}^m c_j h_j(\mathbf{x})$ , where  $\mathbf{c} = (c_1, \dots, c_m)$  is a vector of unknown coefficients and  $h_j$  are known basis functions. The covariance kernel is typically selected from a stationary parametric family, e.g. Matérn family, and its hyper-parameters are estimated with maximum likelihood. Here we use the R package `DiceKriging` (Roustant et al., 2012) for fitting and prediction. Given the data and hyper-parameters the posterior mean and covariance kernel can be computed with the following universal kriging (UK) equations.

$$\mu_n(\mathbf{x}) = \mathbf{h}(\mathbf{x})^T \hat{\mathbf{c}} + \mathfrak{K}(\mathbf{x}, X_n) \mathfrak{K}(X_n, X_n)^{-1} (\mathbf{y}_n - H \hat{\mathbf{c}}) \quad (7)$$

$$\mathfrak{K}_n(\mathbf{x}, \mathbf{x}') = \mathfrak{K}(\mathbf{x}, \mathbf{x}') - \mathfrak{K}(\mathbf{x}, X_n) \mathfrak{K}(X_n, X_n)^{-1} \mathfrak{K}(\mathbf{x}', X_n)^T + \quad (8)$$

$$\boldsymbol{\lambda}(\mathbf{x}, X_n) (H^T \mathfrak{K}(X_n, X_n)^{-1} H)^{-1} \boldsymbol{\lambda}(\mathbf{x}', X_n)^T, \text{ for } \mathbf{x}, \mathbf{x}' \in D \quad \text{with}$$

$$\boldsymbol{\lambda}(\mathbf{x}, X_n) = (\mathbf{h}(\mathbf{x})^T - \mathfrak{K}(\mathbf{x}, X_n) \mathfrak{K}(X_n, X_n)^{-1} H) \quad \text{and}$$

$$\hat{\mathbf{c}} = (H^T \mathfrak{K}(X_n, X_n)^{-1} H)^{-1} H^T \mathfrak{K}(X_n, X_n)^{-1} \mathbf{y}_n$$

where  $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \dots, h_m(\mathbf{x}))^T$ ,  $H = [h_j(\mathbf{x}_i)]_{i=1, \dots, n, j=1, \dots, m}$ ,  $\mathfrak{K}(\mathbf{x}, X_n) = (\mathfrak{K}(\mathbf{x}, \mathbf{x}_1), \dots, \mathfrak{K}(\mathbf{x}, \mathbf{x}_n))$ , for  $\mathbf{x} \in D$  and  $\mathfrak{K}(X_n, X_n) = (\mathfrak{K}(\mathbf{x}_i, \mathbf{x}_j))_{i,j=1, \dots, n} \in \mathbb{R}^{n \times n}$ . If  $m = 1$  and  $h_1 \equiv 1$  we have ordinary kriging (OK) equations.

Once we have fitted the GP, profile extrema functions can be directly computed with the posterior mean, leading to a visualization of a plug-in estimate  $\hat{\Gamma} = \{\mathbf{x} \in D : \mu_n(x) \geq \tau\}$  for  $\Gamma$ . The GP assumption allows us to also quantify the uncertainty on profile extrema functions, at fixed hyper-parameters. In particular, in the next section, we develop a Monte Carlo method that exploits approximate posterior realizations to obtain point-wise confidence intervals for  $P_{\mathbf{e}_i}^{\text{sup}} f$  and  $P_{\mathbf{e}_i}^{\text{inf}} f$ , for  $i = 1, \dots, d$ .

## 3.2 Approximation for posterior field realizations

Consider the GP  $(Z_{\mathbf{x}})_{\mathbf{x} \in D}$ , we denote with  $Z_{\mathbf{x}}(\omega)$  a realization of the process at  $\mathbf{x}$ , for  $\omega \in \Omega$ , the underlying probability space. We are interested in posterior realizations of  $Z_{\mathbf{x}}$  given  $Z_{X_n} = \mathbf{y}_n$ . A brute force approach would involve simulating  $s$  times the field over a space filling design on  $D$  and then evaluating the profile extrema function by taking the discrete extrema. This procedure however is very costly as it requires  $s$  exact simulations of the field over many points and its results strongly depend on the chosen discretization. Following Azzimonti et al. (2016) we use an approximating process in order to reduce the computational cost. We consider the sequence of points  $G = (\mathbf{g}_1, \dots, \mathbf{g}_\ell) \in D^\ell$  and the *approximating process of  $Z$  based on  $G$*

$$\tilde{Z}_{\mathbf{x}} = a(\mathbf{x}) + \mathbf{b}^T(\mathbf{x}) Z_G, \quad \mathbf{x} \in D, \quad (9)$$

where  $a : D \rightarrow \mathbb{R}$  is a trend function and  $\mathbf{b} : D \rightarrow \mathbb{R}^\ell$  is a continuous vector-valued function of deterministic weights, and  $Z_G = (Z_{\mathbf{g}_1}, \dots, Z_{\mathbf{g}_\ell})$  is the  $\ell$ -dimensional random vector given by the values of the original process  $Z$  at  $G$ . Here we select  $a(\mathbf{x}) = \mu(\mathbf{x})$ ,  $\mathbf{b}(\mathbf{x}) = \mathfrak{K}(G, G)^{-1} \mathfrak{K}(\mathbf{x}, G)^T$  where  $\mathfrak{K}(\mathbf{x}, G) = (\mathfrak{K}(\mathbf{x}, \mathbf{g}_1), \dots, \mathfrak{K}(\mathbf{x}, \mathbf{g}_\ell)) \in \mathbb{R}^{1 \times \ell}$  and  $\mathfrak{K}(G, G) =$

$[\mathcal{R}(\mathbf{g}_i, \mathbf{g}_j)]_{i,j=1,\dots,\ell} \in \mathbb{R}^{\ell \times \ell}$ . In what follows, the points  $G$  are called *pilot points*, borrowing the term from geostatistics (see, e.g. [Scheidt, 2006](#), Chapter 4.2), and here they are selected with Algorithm B from [Azzimonti et al. \(2016\)](#). The number of pilot points  $\ell$  can be empirically chosen by stopping when the optimum of Algorithm B's objective function stabilizes around a value. In moderate dimensions, usually  $\ell$  between 50 and 150 leads to distributions close to the true one. The bound introduced in section 3.4 and, in particular, the related tightness indicator, equation (13), can also be used to choose the number of pilot points. For a fixed  $\omega \in \Omega$ ,  $\tilde{Z}_{\mathbf{x}}(\omega)$  is a function of  $\mathbf{x}$  that has an analytical expression and only requires posterior simulations of the original field  $Z$  at  $G$  and linear operations with pre-calculated kriging weights. Moreover the gradient of  $\tilde{Z}_{\mathbf{x}}(\omega)$  with respect to  $\mathbf{x}$  is known, see Appendix B, therefore we compute profile extrema for each realization,  $P_{\Psi}^{\sup} \tilde{Z}_{\mathbf{x}}(\omega)$  and  $P_{\Psi}^{\inf} \tilde{Z}_{\mathbf{x}}(\omega)$ , with the algorithms introduced in Section 2.3.

### 3.3 Analytical example under uncertainty

Let us consider the analytical example illustrated in Section 2.2. We evaluate the function in Equation (5) only at  $n$  points and we approximate it on a dense grid  $100 \times 100$  with a GP. We consider two DoE with  $n = 20$  and  $n = 90$ , both obtained with (randomized) maximin Latin hypercube sampling (R package [DiceDesign](#), [Dupuy et al. \(2015\)](#)) and we choose a prior GP with constant mean and Matérn covariance kernel with  $\nu = 5/2$ . We estimate the covariance hyper-parameters with maximum likelihood and the constant mean value with the ordinary kriging formulae. The model with  $n = 20$  evaluations provides a rough approximation for  $f$ , while the model with  $n = 90$  is an accurate reconstruction. The  $Q^2$  criterion<sup>1</sup> computed over a dense grid of  $N_{test} = 10000$  test data is equal to 0.86 for the posterior mean with  $n = 20$  and 0.99 for  $n = 90$ . For both models, we consider  $\ell = 80$  pilot points and we compute the point-wise confidence intervals for the profile extrema on the mean with Monte Carlo simulations. Figure 5 shows the posterior GP mean obtained from  $n = 20$  evaluations of the function in Equation (5), the regions delimited by the oblique profile extrema on the mean along the directions  $\mathbf{v}_1, \mathbf{v}_2$  and its 90% point-wise confidence intervals for the threshold  $\tau = 0$ . Figure 6 shows the profile extrema for the mean, the profile extrema for  $s = 150$  approximate realizations of the posterior process and the empirical point-wise 90% confidence intervals (dark shaded tube, red). Figure 8 shows profile extrema for the mean on the model with  $n = 90$  observations. In this case the confidence intervals (dark shaded tube, red) are much tighter indicating a smaller uncertainty.

### 3.4 Bounds for the approximation

The approximating process  $\tilde{Z}$  does not provide proper probabilistic statements for  $P_{\Psi}^{\sup} Z, P_{\Psi}^{\inf} Z$  as it is based on the  $\ell$ -dimensional random vector  $Z_G$ . However since the distribution of the error process  $Z - \tilde{Z}$  can be expressed in closed form, we can control the approximating error with the following probabilistic bound. Proofs are in Appendix A.

---

<sup>1</sup> $Q^2 = 1 - \sum_{j=1}^{N_{test}} (\mu_n(x_j) - y_j^{test})^2 / (\sum_{j=1}^{N_{test}} (y_j^{test} - \overline{y^{test}})^2)$  with  $\overline{y^{test}} = 1/N_{test} \sum_{j=1}^{N_{test}} y_j^{test}$

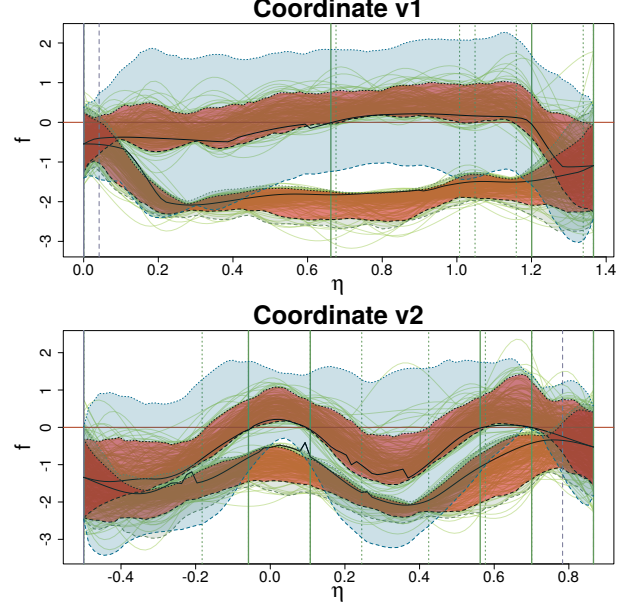
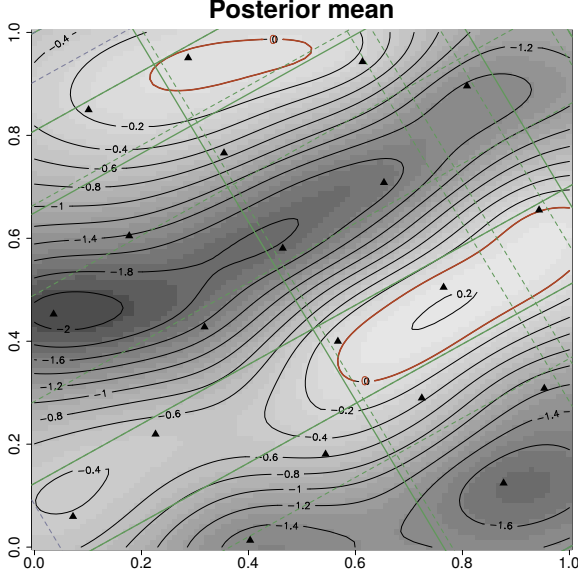


Figure 5: Posterior mean of GP trained on  $n = 20$  evaluations of function in equation (5). Excursion region ( $\tau = 0$ , shaded red), regions selected by the profile extrema on mean (solid) and on 95% – 5% quantiles (dashed).

Figure 6: Profile extrema functions ( $n = 20$ ,  $\tau = 0$ ): mean (black, solid), breakpoints (vertical green; mean: solid, quantiles: dashed), quantiles (90%, black dashed, CI dark red tube), upper-lower bound (light blue tube).

**Theorem 1.** Consider a Gaussian process  $(Z_{\mathbf{x}})_{\mathbf{x} \in D} \sim GP(\mu, \mathfrak{K})$ , the approximating process  $\tilde{Z}$  of  $Z$  based on the points  $G$  (Equation (9)) and  $T \subset D$ , then for any  $u > \mu_T^{\tilde{\Delta}}$

$$Pr \left( \left| \sup_{\mathbf{x} \in T} Z_{\mathbf{x}} - \sup_{\mathbf{x} \in T} \tilde{Z}_{\mathbf{x}} \right| > u \right) \leq 2 \exp \left( - \frac{(u - \mu_T^{\tilde{\Delta}})^2}{2(\sigma_T^{\tilde{\Delta}})^2} \right), \quad (10)$$

where

$$\begin{aligned} \mu_T^{\tilde{\Delta}} &= \sup_{\mathbf{x} \in T} |\mu^{\tilde{\Delta}}(\mathbf{x})| \text{ and } (\sigma_T^{\tilde{\Delta}})^2 = \sup_{\mathbf{x} \in T} \mathfrak{K}^{\tilde{\Delta}}(\mathbf{x}, \mathbf{x}) \text{ with} \\ \mu^{\tilde{\Delta}}(\mathbf{x}) &= \mathbb{E}[Z_{\mathbf{x}} - \tilde{Z}_{\mathbf{x}}] = \mu(\mathbf{x}) - a(\mathbf{x}) - \mathbf{b}^T(\mathbf{x})\mu(G) \quad \mathbf{x} \in T \\ \mathfrak{K}^{\tilde{\Delta}}(\mathbf{x}, \mathbf{y}) &= \mathfrak{K}(\mathbf{x}, \mathbf{y}) - \mathfrak{K}(\mathbf{y}, G)\mathbf{b}(\mathbf{x}) - \mathfrak{K}(\mathbf{x}, G)\mathbf{b}(\mathbf{y}) + \mathbf{b}^T(\mathbf{x})\mathfrak{K}(G, G)\mathbf{b}(\mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in T \end{aligned} \quad (11)$$

If the approximating process  $\tilde{Z}$  is unbiased then  $\mu^{\tilde{\Delta}}(\mathbf{x}) \equiv 0$  and the inequality in (10) is valid for any  $u > 0$ .

**Corollary 1.** Consider the profile sup random functions defined as

$$P_{\Psi}^{\sup} \tilde{Z}(\eta) := \sup_{\{\mathbf{x} \in D: \Psi^T \mathbf{x} = \eta\}} \tilde{Z}_{\mathbf{x}} \text{ and } P_{\Psi}^{\sup} Z(\eta) := \sup_{\{\mathbf{x} \in D: \Psi^T \mathbf{x} = \eta\}} Z_{\mathbf{x}}, \quad \eta \in E_{\Psi},$$

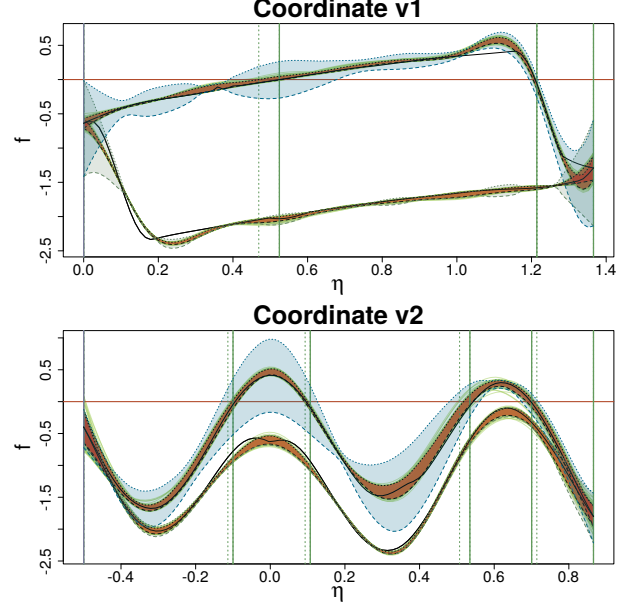
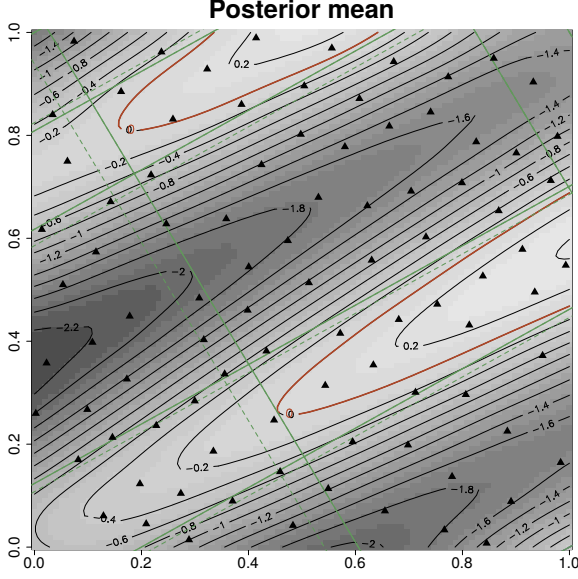


Figure 7: Posterior mean of GP trained on  $n = 90$  evaluations of function in equation (5). Excursion region ( $\tau = 0$ , shaded red), regions selected by the profile extrema on mean (solid) and on 95% – 5% quantiles (dashed).

Figure 8: Profile extrema functions ( $n = 90$ ,  $\tau = 0$ ): mean (black, solid), breakpoints (vertical green; mean: solid, quantiles: dashed), quantiles (90%, black dashed, CI dark red tube), upper-lower bound (light blue tube).

where  $\tilde{Z}$  is an unbiased approximate process in the form of Equation (9). For any  $\eta^* \in E_\Psi$  and  $\alpha \in (0, 1)$

$$\begin{aligned} Pr(P_\Psi^{\sup} Z(\eta^*) \leq u_\alpha^+) &\geq 1 - \alpha, & \text{with } u_\alpha^+ &= \tilde{u}_\beta^+ + \sqrt{2 \left( \sigma_T^{\tilde{\Delta}} \right)^2 \log \left( \frac{2}{\alpha - \beta} \right)} \text{ and} \\ Pr(P_\Psi^{\sup} Z(\eta^*) \geq u_\alpha^-) &\geq 1 - \alpha, & \text{with } u_\alpha^- &= \tilde{u}_\beta^- - \sqrt{2 \left( \sigma_T^{\tilde{\Delta}} \right)^2 \log \left( \frac{2}{\alpha - 2\beta} \right)} \end{aligned} \quad (12)$$

with  $(\sigma_T^{\tilde{\Delta}})^2$  as in (11),  $T = \{\mathbf{x} \in D : \Psi^T \mathbf{x} = \eta\}$ ,  $\alpha > 2\beta$  and  $\tilde{u}_\beta^+, \tilde{u}_\beta^-$  are the  $1-\beta, \beta$  quantiles for  $P_\Psi^{\sup} \tilde{Z}(\eta^*)$  respectively. Equations (12) imply  $Pr(P_\Psi^{\sup} Z(\eta^*) \in [u_\alpha^-, u_\alpha^+]) \geq 1 - 2\alpha$ .

In practice the quantiles  $\tilde{u}_\beta^\pm$  in equation (12) are estimated with sample quantiles  $\widehat{\tilde{u}}_\beta^\pm$  from the realizations of  $\tilde{Z}$  and  $u_\alpha^\pm$  are estimated by plugging-in  $\widehat{\tilde{u}}_\beta^\pm$  in equation (12).

Figures 6 and 8 show the conservative bound in equation (12) on  $P_i^{\sup} Z$  (skyblue, lightly shaded tube) and  $P_i^{\inf} Z$  (seagreen, lightly shaded),  $i = 1, 2$ , on the example presented in Section 3.3. The uncertainty is much smaller with  $n = 90$ , however the bound is still very conservative. An indicator for the bound tightness is the quantity  $(\sigma_T^{\tilde{\Delta}})^2(\eta, \ell)$ , for each  $\eta \in E_\Psi$ . Here we explicit the dependency of the approximation on  $\ell$ . In particular, we

study the integral of this map, i.e.

$$I(\sigma_T^{\tilde{\Delta}})^2(\ell) := \int_{E_{\Psi}} (\sigma_T^{\tilde{\Delta}})^2(\eta, \ell) d\eta \quad (13)$$

Appendix D shows a comparison of  $I(\sigma_T^{\tilde{\Delta}})^2$  from the example in figure 8, as a function of  $\ell$ . Figure 21, in particular, shows exponential decrease in  $I(\sigma_T^{\tilde{\Delta}})^2$  as  $\ell$  grows to 80.

### 3.5 Discussion on method’s parameters

Profile extrema functions for an expensive-to-evaluate function require the user to set several parameters summarized in Table 4, Appendix D. The DoE selected to train the GP model has the biggest influence on the profile extrema uncertainty. Figures 6 and 8 show that a GP model trained on more function evaluations, reduces the uncertainty on profiles. The bounds on profile extrema, equation (12), however, are also controlled by two parameters chosen by the user given a fixed DoE: the number of pilot points,  $\ell$ , and of GP realizations,  $s$ . The number of posterior GP realizations,  $s$ , controls the accuracy of the empirical quantiles  $\tilde{u}_{\beta}^{\pm}$ . More importantly, the number of pilot points  $\ell$  controls directly the tightness of the bounds in equation (12). This parameter plays a more prominent role in uncertainty quantification as more pilot points lead to a smaller  $\sigma_T^{\tilde{\Delta}}$  and a tighter bound, see, e.g., Figure 21, Appendix D.

## 4 Motivating application: coastal flooding

### 4.1 Motivation and study case

Coastal flooding models experienced recent progresses opening new research and applications perspectives. However, their computational cost ( $>$  hours) hinders their use when a large number of simulations is required for estimating the excursion region corresponding to the critical forcing conditions leading to inundation or when forecast is needed (Rohmer and Idier, 2012; Idier et al., 2013).

We focus here on coastal flooding induced by overflow and we consider the Boucholeurs area (French Atlantic coast, see Figure 9). This area is located close to La Rochelle and was flooded during the 2010 Xynthia storm event. This event was characterized by a high storm surge ( $> 1.5$  m at La Rochelle tide gauge) in phase with a high spring tide (Bertin et al., 2014). Here, we focus on these primary drivers (tide and storm surge) and on how they affect the resulting flooded surface ( $Y$ , in square meters).

#### 4.1.1 The forcing conditions

The offshore forcing conditions correspond to the tide and storm surge temporal evolution (see Figure 10a) are denoted with  $\mathbf{x} = (T, S, t_0, t_+, t_-)$ . They are parametrized as follows:

- the tide is simplified by a sinusoidal signal parametrised by its high tide level  $T \in [0.95 \text{ m}, 3.70 \text{ m}]$ , (see Figure 10a);



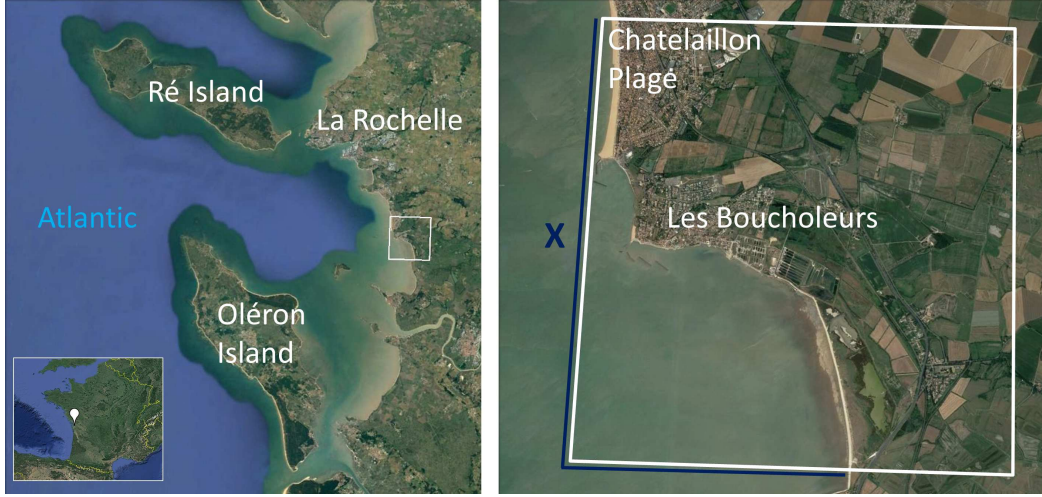


Figure 9: Study site location (left) and computational domain limits (right, in white) and location of the forcing conditions  $\mathbf{x}$  (right, in blue).

- the surge signal is assumed to be described by a triangular model (see Figure 10a) using four parameters: the peak amplitude  $S \in [0.65 \text{ m}, 2.50 \text{ m}]$ , the phase difference  $t_0 \in [-6, 6]$  hours, between the surge peak and the high tide, the time duration of the raising part  $t_- \in [-12.0, -0.5]$  hours, and the falling part  $t_+ \in [0.5, 12.0]$  hours.

#### 4.1.2 The numerical model

The numerical modelling of the coastal flood relies on the MARS model (Lazure and Dumas, 2008). This finite-difference model solves the shallow-water equations and was originally designed to compute regional coastal hydrodynamics, e.g., tide and storm induced water level and currents. The MARS model has here been adapted to account for the specificities of local coastal flooding processes: hydraulic processes around connections like culverts and weirs, coastal defence breaching. This model has been implemented on the study site (white box in Figure 9) with a spatial resolution of 25 m and a total number of mesh cells of  $>39,000$ . The land cover effect on the flood is taken into account by using a spatially varying friction coefficient. The different hydraulic connections (e.g., the hydraulic culverts below the roads, dike, railway,) are taken into account in the modeling. The forcing conditions (time series deduced from the parameters  $\mathbf{x}$ , Figure 10a) are uniform over the open boundaries of the domain in blue on Figure 9. A single model run takes about 30-60 minutes of computation using a single CPU. For more details on the study site, the model set-up and validation, see Rohmer et al. (2018). It should be noted that when wave overtopping is dominant in the flooding processes, other types of models should be used (Le Roy et al., 2015), with computation times 2 orders of magnitude larger. An overflow case allows setting up statistical developments which will be useful also for the more expensive models. Figure 10b provides examples of the inundation depth ( $H$ ) computed in each cell for given forcing conditions  $\mathbf{x}$ , as well as the resulting flood surface value ( $Y$ ). We consider the threshold values  $\tau_1^{(Y)} = 1.2 \times 10^6 \text{ m}^2, \dots, \tau_4^{(Y)} = 6.5 \times 10^6 \text{ m}^2$  introduced in Figure 10b.

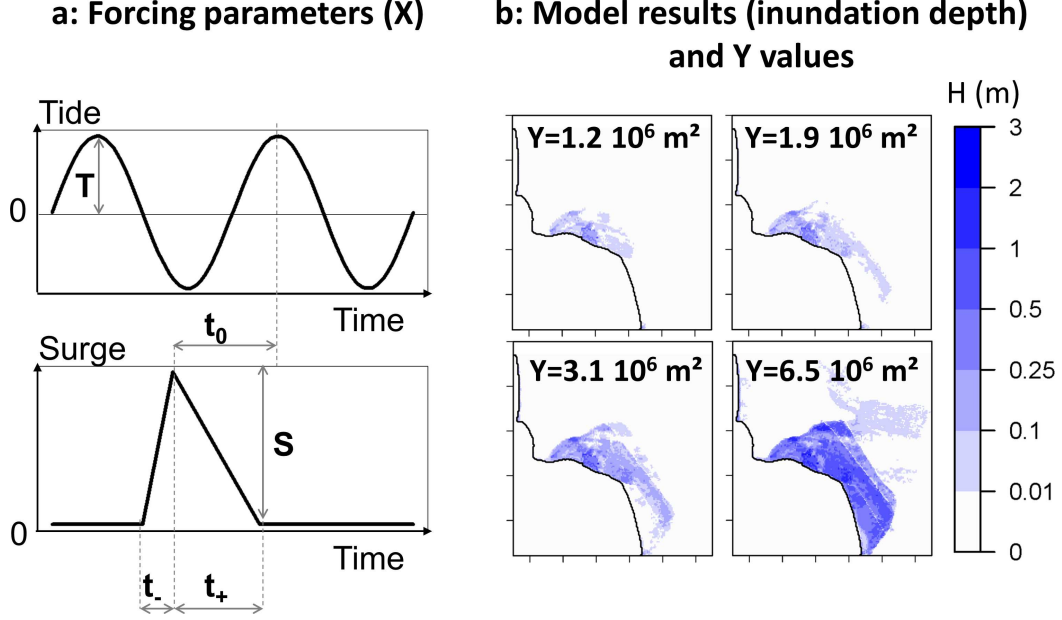


Figure 10: (a) Schematic representation of the tide and surge temporal signals and the different parameters describing them. (b) Maps of inland water height for given values of the  $\mathbf{x}$  parameters, and deduced  $Y$  value of flood surface (in square meter).

#### 4.1.3 Gaussian process model

We consider a rescaled input space  $D = [0, 1]^5$  and the function  $\check{f} : D \subset \mathbb{R}^5 \rightarrow \mathbb{R}_+$  with  $Y = \check{f}(\mathbf{x})$  for each  $\mathbf{x} = (x_1, \dots, x_5) = (T, S, t_o, t_+, t_-) \in D$ . In the remainder we keep the notation  $(T, S, t_o, t_+, t_-)$  for the rescaled input. We are interested in estimating

$$\Gamma = \{\mathbf{x} \in D : \check{f}(\mathbf{x}) \geq \tau\}, \quad \text{where } \tau = \tau_k^{(Y)}, \quad \text{for } k = 1, \dots, 4. \quad (14)$$

We consider the square root transformed output data  $f(\mathbf{x}) = \sqrt{\check{f}(\mathbf{x})} = \sqrt{Y}$  for  $\mathbf{x} \in D$  and the square root thresholds  $\tau_k = \sqrt{\tau_k^{(Y)}}$ . This transformation was chosen, after fitting the model on different scales for  $Y$ , because it provided the best cross-validation metrics.

We fix an initial DoE  $X_n \in D^n$ , with  $n = 200$  points obtained by evaluating the function  $f$  on the first 500 points of the 5-d Sobol' sequence and by selecting the first  $n = 200$  points leading to a flood of any magnitude. The evaluations  $f(X_n)$  at  $X_n$  are denoted with  $\mathbf{y}_n \in \mathbb{R}^n$ . We consider a GP model with a tensor product prior Matérn covariance kernel  $\nu = 3/2$  and prior mean of the form

$$\mu(\mathbf{x}) = c_0 + c_1 x_1 + c_2 x_2 + c_3 x_3^2 + c_4 x_4 + c_5 x_5 = c_0 + \sum_{j=1}^5 c_j h_j(\mathbf{x}). \quad (15)$$

The covariance kernel hyper-parameters are estimated with maximum likelihood from  $(X_n, \mathbf{y}_n)$  and the posterior mean and covariance kernel are obtained with Equations (7), (8). The GP and the basis functions  $h_j$  were selected using expert-based information achieving a  $Q^2 = 0.958$ . A comparison of different model fits is shown in Appendix C.



We estimate  $I$  from the posterior mean with

$$\hat{I}_{n,\tau_k} = \{\mathbf{x} \in D : \mu_n(\mathbf{x}) \geq \tau_k\}, \quad \text{for } k = 1, \dots, 4.$$

## 4.2 Procedure overview

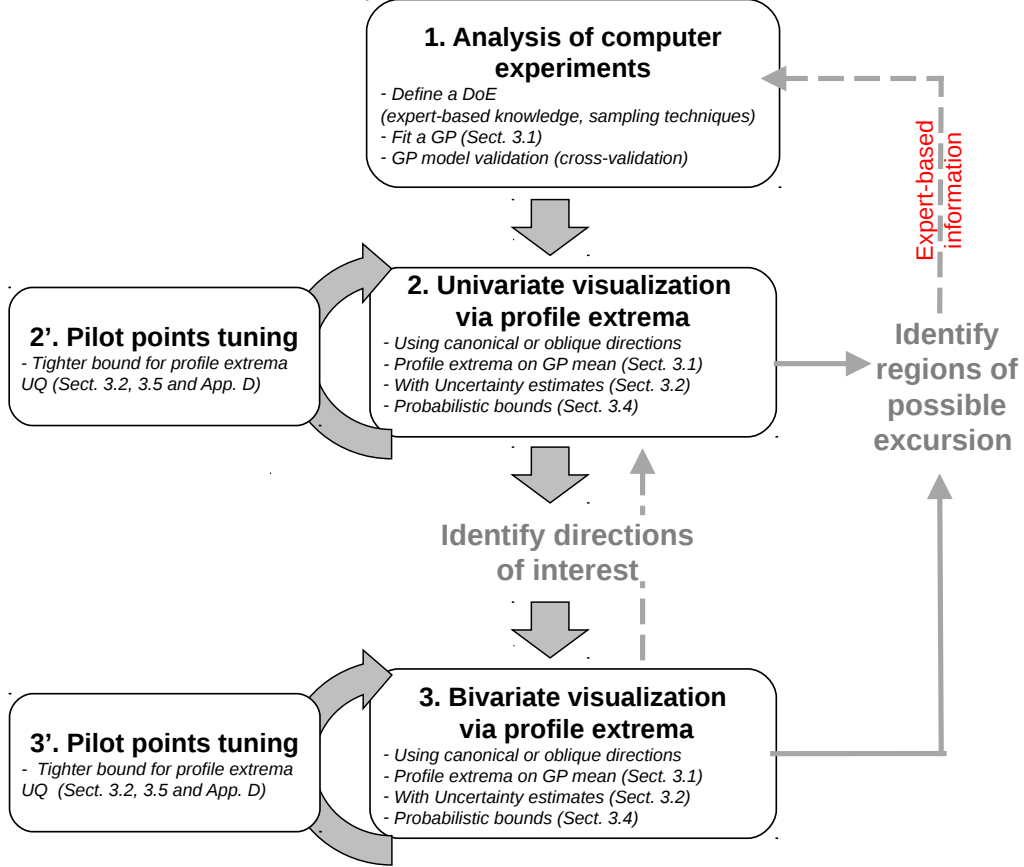


Figure 11: Flow chart for the procedure used in the test case.

In the following sections profile extrema are used to explore visually estimates for  $I$  and to quantify their uncertainty. The proposed procedure is summarized in Figure 11:

- step 1: design of experiment and emulation.** Select a DoE  $X_n$  ( $n = 200$ ) and run the MARS model (numerical simulator) to compute the flooded area  $Y_n$ ; Fit a GP (4.1.3) on  $X_n, Y_n$ , evaluate the emulation quality with leave-one-out-cross validation.
- step 2: univariate profile extrema.** Compute coordinate profile extrema (i.e. along the canonical directions) on the GP mean and give a first visual indication on the excursion set. Uncertainty quantification on profile extrema (Section 3) guides expert knowledge in identifying possible regions of excursion. A comparison of several profile

extrema plots with a different number of pilot points  $\ell$  increases the understanding of profile functions uncertainty without additional numerical simulator runs. The conclusions drawn from 1d profile extrema can be used to refine the DoE in step 1 and ultimately decrease the uncertainty on the excursion set. This first analysis, for example, shows that some offshore conditions do not influence the excursion. Oblique profile extrema could be used in this phase, see, e.g., section 2.2, if more informative directions are known in advance.

**step 3: bivariate profile extrema.** Explore combinations of input variables that lead to excursion with bivariate profile extrema functions. Orthogonal projections can potentially be used as shown in Section 2.2 on the analytical example. Similarly as for step 3, refinements of the DoE in the regions of interest can be performed.

## 4.3 Results on coastal flooding test case

### 4.3.1 Univariate profile extrema

We start the analysis of  $\hat{\Gamma}_{n,\tau_k}$ , for the thresholds  $\tau_1, \dots, \tau_4$ , with coordinate profile extrema on the posterior GP mean. The uncertainty is visualized with posterior quantiles of the profile extrema (Section 3.2) and with the upper and lower bound of equation (12).

Figure 12 shows the coordinate profile extrema for the posterior mean of the process based on the design with  $n = 200$  points described in Section 4.1.3, with the universal kriging prior mean defined in Equation (15) and the lowest and highest threshold  $\tau_1, \tau_4$ .

Let us consider  $\tau_4 = 2549.5 = \sqrt{6.5e6}$ , the highest threshold in dark red. Coordinate profile extrema on the posterior mean tell us that if  $T < 0.57$ , there is no excursion independently of the other coordinates. The 90% point-wise confidence interval is based on  $s = 600$  approximate posterior realizations generated with  $\ell = 300$  pilot points. The point-wise confidence intervals identify a possible (90%) region of non-excursion  $\{\mathbf{x} \in D : T \in [0, \lambda] \text{ with } \lambda \in [0.52, 0.6]\}$ . If we consider the variable  $S$ , a possible region of non-excursion (above  $\tau_4$ ) is  $\{\mathbf{x} \in D : S \in [0, \lambda], \lambda \in [0.28, 0.43]\}$ . This region does not exist for low threshold values ( $< \tau_2$ ), thus indicating that small values of surge peak only play a role in moderate flooding events, i.e.  $< \tau_2$ . Similar assessments are available for the other coordinates and the other thresholds, see Table 3 in Appendix C for a summary. Note that the variables  $t_+, t_-$  do not bring information on the excursion regions as  $P_{\mathbf{e}_{4,5}}^{\text{inf}} \gamma_n(\eta)$  and  $P_{\mathbf{e}_{4,5}}^{\text{sup}} \gamma_n(\eta)$  are consistently below and above the thresholds  $\tau_1, \tau_4$  respectively. The bounds on the approximating process from equation (12), plotted as wide light blue tube, show that there is still uncertainty on this assessment. In fact, by accounting also for the approximation uncertainty, the possible region of non-excursion becomes  $\{\mathbf{x} \in D : T \in [0, 0.19]\}$ . The tightness of the bound is evaluated by looking the integrated variance of the difference, i.e. equation (13). We chose  $\ell = 300$  as the resulting integrated variance is small enough and at the same time the method is not computationally too expensive. For example the average integrated variance over all dimensions is 41.6% and 18.3% smaller than with  $\ell = 37$  for  $P^{\text{inf}}$  and  $P^{\text{sup}}$  respectively. On the other hand, the computational time for  $(\sigma_T^{\hat{\Delta}})$  grew from 302 seconds ( $\ell = 37$ ) to 1595 seconds for  $\ell = 300$ . More details in Appendix C.

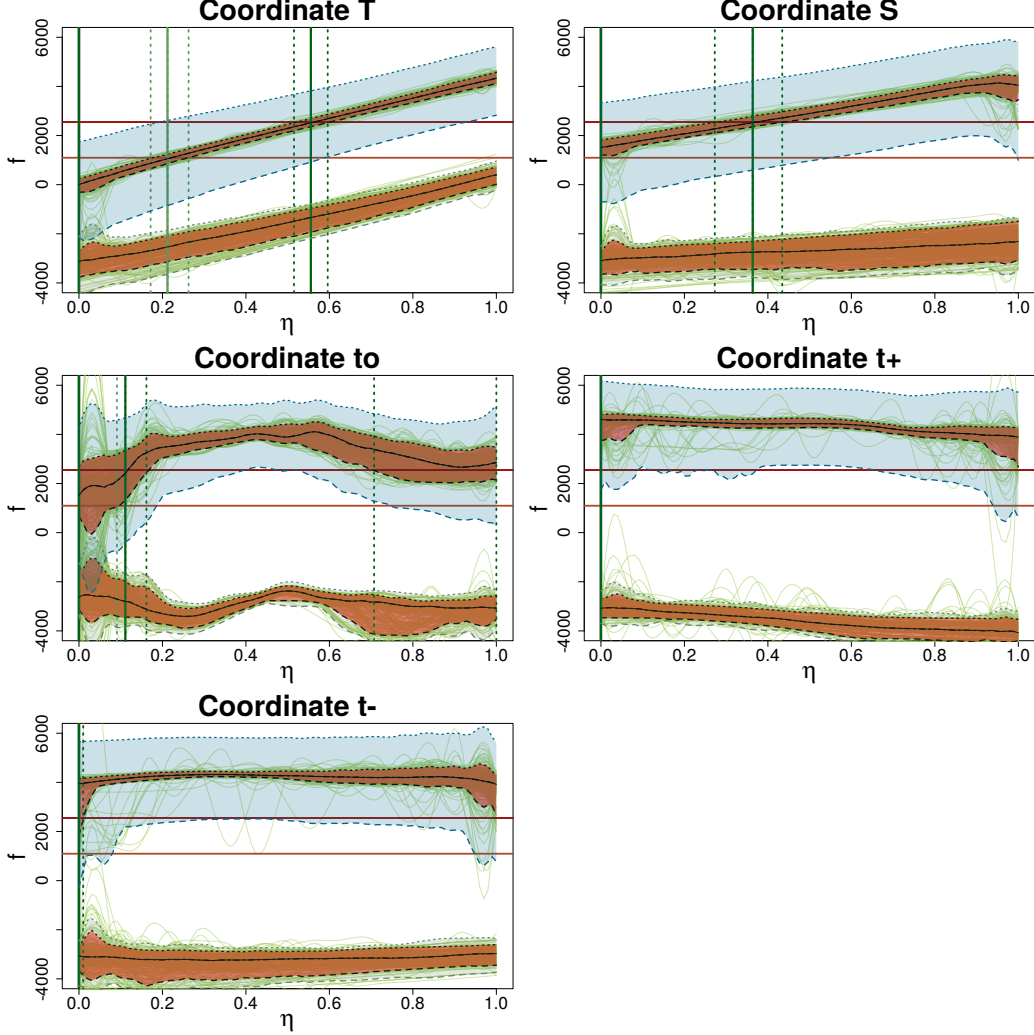


Figure 12: Coordinate profiles for 5-dimensional test case. Excursion thresholds (red, light  $\tau_1$ , dark  $\tau_4$ ), breakpoints  $P_i^{\text{sup}}/P_i^{\text{inf}}$  on mean (vertical green lines) and on 95%–5% quantiles (vertical green lines, dashed), CI dark red tube, bound range light blue tube.

#### 4.3.2 Bivariate profile extrema

We now focus on the combination of variables  $T, S$  and  $T, t_o$ , whose more prominent role on the excursion was outlined by the coordinate profile extrema, and we consider  $\tau = \tau_1$ . We explore which values of these combinations lead to excursion with bivariate profile extrema functions. For example, the profile extrema for  $T, S$  is obtained with

$$\Psi = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}^T.$$

We compute the empirical quantile maps and the bound with  $\ell = 300$  and  $s = 420$  approximate posterior realizations, see Section 3.2. Figures 13 (a,b,c) show the contour lines of  $P_\Psi^{\text{sup}}$  and  $P_\Psi^{\text{inf}}$  for the posterior GP mean based on the DoE described in Section 4.1.3. The

background colors indicate different heuristic measures of uncertainty: (a) the weighted inter-quantile range (i.e. the empirical inter-quantile range for profile extrema maps if the threshold is between the upper and lower quantile or zero, otherwise); (b) the upper-lower bound range (i.e. the difference between the upper and lower bound in (12) if  $\tau$  is between them, zero otherwise); (c) the standard deviation of the difference  $(\sigma_T^{\hat{\Delta}})(\eta)$ ,  $\eta \in E_{\Psi}$ .

Let us start by analyzing  $P_{\Psi}^{\text{sup}}$ ,  $P_{\Psi}^{\text{inf}}$  on the posterior mean, shown in the contour lines in Figures 13. The map on the left suggests that the region below the critical dashed red contour can be excluded, i.e. the region is outside the excursion. In a symmetric way,  $P_{\Psi}^{\text{inf}}$  (Figure 13, right) suggests that high values for  $T$  ( $> 0.8$ ) along with values for  $S$  in the range  $[0.05, 0.7]$  should lead to an excursion (i.e. flood). The bivariate profile extrema for  $T, t_o$  in Figure 19, instead, show (left) a non-flood region on the left of the critical contour and suggest (right) excursion in the region  $T > 0.95$ ,  $t_o \in [0.42, 0.68]$ . These conclusions are consistent with our physical knowledge, except for the excursion domain on  $T, S$  in Figure 13 (right). Indeed, an excursion domain bounded by maximum surge does not have a physical explanation because  $Y$  should increase with  $S$ , with no maximum bound of  $S$ .

The above analysis is based only on  $P_{\Psi}^{\text{sup}}$  and  $P_{\Psi}^{\text{inf}}$  for the posterior GP mean and, since the DoE is small and non-adaptive, there could be high uncertainty in some parts of the input domain. The indicators plotted as background colors allow us to quantify this uncertainty. Both the weighted inter-quantile range and the bound range for  $P_{\Psi}^{\text{inf}}$  in Figure 13 (right) are high the region of large  $T$  and  $S$ . Moreover, as shown in Figure 13(c), the values of  $(\sigma_T^{\hat{\Delta}})$  are not high in that region indicating that the uncertainty due to the approximating process is not very high. Those insights collectively suggest that more function evaluations should be added in the  $P_{\Psi}^{\text{inf}}$  uncertain region.

### 4.3.3 Summary of results

Coordinate profile extrema functions on the coastal flooding test case enabled: (i) to highlight the major role of the high tide level,  $T$ , whatever the considered thresholds, i.e., for small to large flooding events; (ii) to highlight the key role of the surge peak,  $S$ , only for moderate thresholds i.e., moderate flooding events; (iii) to highlight the moderate role of the phase difference  $t_o$  alone; (iv) to exclude a strong influence of  $t_-$  and  $t_+$  alone for the excursion of the response whatever the considered thresholds. Moreover by studying bivariate profile extrema we could: (i) exclude regions where  $T$  and  $S$  are simultaneously small (e.g.  $T < 0.4$  and  $S < 0.6$ ); (ii) highlight the role of phase difference  $t_o$  and tide  $T$  combined with a possible excursion in the region  $T > 0.95$ ,  $S \in [0.42, 0.68]$ . Furthermore, the comparison of approximation error indicators (Figure 13(c)) with uncertainty measures on profile extrema (Figure 13(b)) enabled us to nuance the results and to track the main uncertainty source. In the coastal flooding test case, the uncertainty unlikely stems from the approximating process, but rather from the lack of function evaluations.

## 5 Discussion

In this work we introduced a visualization technique, based on profile extrema functions, for excursion regions of expensive-to-evaluate functions. The main idea is to study the con-

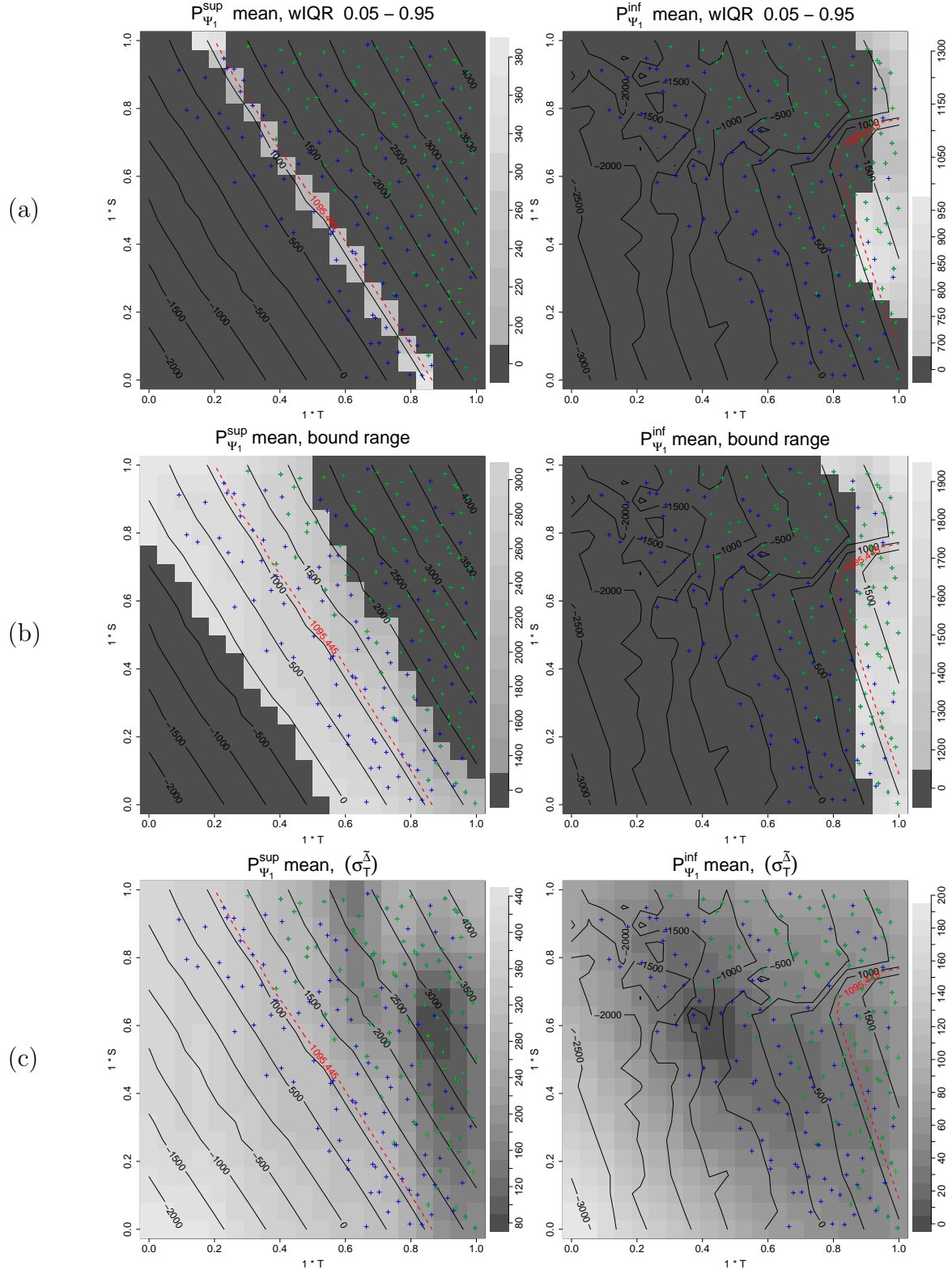


Figure 13: Bivariate  $(T, S)$  profile extrema mean (contour lines). Threshold  $\tau_1$  dashed red line,  $X_n$  locations (crosses, green above  $\tau_1$ ). Background: (a) weighted inter-quantile range (95% – 5%), (b) bound range (eq. (12)), (c) approximation error standard deviation.

strained optima of the functions on lower dimensional subspaces resulting from projections on lines or planes. By plotting profile extrema functions we can select regions of interest. If the function is expensive to evaluate, a GP model is used to emulate the function and we showed how profile extrema can be computed on the GP model. In this case the conclusions strongly depend on the chosen GP model and, as we show in Appendix C, the profile extrema uncertainty is affected by the modeling choices.

In the coastal flooding test case, as sketched in Figure 11, we selected directions of interest from coordinate profiles. For example, the profile sup along the direction of coordinate  $T$  is below the threshold of interest for some values. This indicates that  $T$  is a direction of interest for the excursion phenomena. This procedure was repeated for each coordinate. In test cases where canonical directions are not meaningful we could explore the excursion by looking at oblique directions as we did in the analytical example in Section 3.3.

The bivariate profile extrema maps in Figure 13 also suggest a principled way to develop adaptive design of experiments that sequentially reduce uncertainties. For example, we could select the next evaluations as the minimizers of the integrated (or maximum of)  $(\sigma_T^{\hat{A}})$ . Such criteria should be analytically tractable and could lead to adaptive designs similar to classic IMSE (MSE) strategies, Sacks et al. (1989). Alternative criteria could be obtained by minimizing bound-range uncertainties (Figure 13(b)), as they provide direct information on the excursion, however their tractability is still unclear.

Profile extrema functions require a continuously differentiable function on a compact domain. If the domain is not compact, then the optimization in the definition of profile extrema might not be well posed. In case of an open domain with a probability distribution on the inputs, profile extrema functions could be extended by using quantiles in place of the maximum (Roy and Notz, 2014). Profile extrema functions could be extended to non-linear subsets, however it is technically not straightforward and it might result in visualizations that are much harder to interpret. The overall approach developed here is a one-step procedure, and it could become part of an exploratory work flow. As shown in Appendix E, coordinate profile extrema, oblique and bivariate profiles can be combined to convey more information on the excursion set in simpler terms. A possible future extension could involve a treed procedure where the input space is restricted with constrained coordinate profile functions. Oblique coordinate profiles, i.e. profiles along non-canonical directions, require the user to choose which directions to explore. This choice could be driven by expert knowledge, such as in the motivating test case presented here. However, when such knowledge is not available, we could envisage a procedure following similar steps to projection pursuit (Cook et al., 1995) where we obtain the most informative direction.

## References

- Adler, R. J. and Taylor, J. E. (2007). *Random Fields and Geometry*. Springer, Boston.
- Asimov, D. (1985). The grand tour: A tool for viewing multidimensional data. *SIAM Journal on Scientific and Statistical Computing*, 6(1):128–143.

- Azaïs, J.-M. and Wschebor, M. (2009). *Level sets and extrema of random processes and fields*. Wiley Online Library.
- Azzimonti, D. (2016). *Contributions to Bayesian set estimation relying on random field priors*. PhD thesis, University of Bern.
- 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.
- Bayarri, M. J., Berger, J. O., Calder, E. S., Dalbey, K., Lunagomez, S., Patra, A. K., Pitman, E. B., Spiller, E. T., and Wolpert, R. L. (2009). Using statistical and computer models to quantify volcanic hazards. *Technometrics*, 51(4):402–413.
- Bertin, X., Li, K., Roland, A., Zhang, Y., Breilh, J., and Chaumillon, E. (2014). A modeling-based analysis of the flooding associated with Xynthia, central Bay of Biscay. *Coastal Engineering*, 94:80–89.
- Bingham, D., Ranjan, P., and Welch, W. (2014). Sequential design of computer experiments for optimization, estimating contours, and related objectives. In *Statistics in Action: A Canadian Outlook*, chapter 7, pages 109–124.
- Bolin, D. and Lindgren, F. (2015). Excursion and contour uncertainty regions for latent Gaussian models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77(1):85–106.
- Bonnans, J. F. and Shapiro, A. (1998). Optimization problems with perturbations: A guided tour. *SIAM Review*, 40(2):228–264.
- Byrd, R., Lu, P., Nocedal, J., and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208.
- Chevalier, C. (2013). *Fast uncertainty reduction strategies relying on Gaussian process models*. PhD thesis, University of Bern.
- Chevalier, C., Bect, J., Ginsbourger, D., Vazquez, E., Picheny, V., and Richet, Y. (2014). Fast kriging-based stepwise uncertainty reduction with application to the identification of an excursion set. *Technometrics*, 56(4):455–465.
- Cook, D., Buja, A., Cabrera, J., and Hurley, C. (1995). Grand tour and projection pursuit. *Journal of Computational and Graphical Statistics*, 4(3):155–172.
- Cook, D., Buja, A., Lee, E.-K., and Wickham, H. (2008). Grand tours, projection pursuit guided tours, and manual controls. In *Handbook of Data Visualization*, pages 295–314. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Danskin, J. M. (1967). *The Theory of Max-Min and its Application to Weapons Allocation Problems*. Springer Berlin Heidelberg, Berlin, Heidelberg.



- Dubourg, V., Sudret, B., and Deheeger, F. (2013). Metamodel-based importance sampling for structural reliability analysis. *Probabilistic Engineering Mechanics*, 33:47–57.
- Dupuy, D., Helbert, C., and Franco, J. (2015). DiceDesign and DiceEval: Two R packages for design and analysis of computer experiments. *Journal of Statistical Software*, 65(11):1–38.
- Edelsbrunner, H. and Harer, J. (2008). *Persistent homology – a survey*, volume 453, pages 257–282. American Mathematical Society, Providence, Rhode Island.
- Fukuda, K. (2007). *cddlib Reference Manual*. ETH Zürich, Institute for Operations Research, ver. 0.94 edition.
- Gerber, S., Bremer, P. T., Pascucci, V., and Whitaker, R. (2010). Visual exploration of high dimensional scalar functions. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1271–1280.
- Ginsbourger, D., Baccou, J., Chevalier, C., Perales, F., Garland, N., and Monerie, Y. (2014). Bayesian adaptive reconstruction of profile optima and optimizers. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):490–510.
- Idier, D., Rohmer, J., Bulteau, T., and Delvallée, E. (2013). Development of an inverse method for coastal risk management. *Natural Hazards and Earth System Sciences*, 13:999–1013.
- Inselberg, A. (1985). The plane with parallel coordinates. *The Visual Computer*, 1(2):69–91.
- Inselberg, A. (2009). *Parallel Coordinates: Visual Multidimensional Geometry and Its Applications*. Springer New York, New York, NY.
- Lazure, P. and Dumas, F. (2008). An external-internal mode coupling for a 3D hydrodynamical model for applications at regional scale (MARS). *Advances in Water Resources*, 31:233–250.
- Le Roy, S., Pedreros, R., Andrée, C., Paris, F., Lecacheux, S., Marche, F., and Vinchon, C. (2015). Coastal flooding of urban areas by overtopping: dynamic modelling application to the Johanna storm (2008) in Gâvres (France). *Natural Hazards and Earth System Sciences*, 15:2497–2510.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rohmer, J. and Idier, D. (2012). A meta-modelling strategy to identify the critical offshore conditions for coastal flooding. *Natural Hazards and Earth System Sciences*, 12(9):2943–2955.

- Rohmer, J., Idier, D., Paris, F., Pedreros, R., and Louisor, J. (2018). Casting light on forcing and breaching scenarios that lead to marine inundation: Combining numerical simulations with a random-forest classification approach. *Environmental Modelling & Software*, 104:64–80.
- Roustant, O., Ginsbourger, D., and Deville, Y. (2012). DiceKriging, DiceOptim: Two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization. *Journal of Statistical Software*, 51(1):1–55.
- Roy, S. and Notz, W. I. (2014). Estimating Percentiles in Computer Experiments: A Comparison of Sequential-Adaptive Designs and Fixed Designs. *Journal of Statistical Theory and Practice*, 8(1):12–29.
- Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989). Design and analysis of computer experiments. *Statist. Sci.*, 4(4):409–423.
- Santner, T. J., Williams, B. J., and Notz, W. I. (2003). *The Design and Analysis of Computer Experiments*. Springer-Verlag, New York.
- Scheidt, C. (2006). *Analyse statistique d’expériences simulées : Modélisation adaptative de réponses non-régulières par krigeage et plans d’expériences*. PhD thesis, Université Louis Pasteur - Strasbourg I.
- Schölkopf, B., Smola, A., and Müller, K.-R. (1997). Kernel principal component analysis. In *Artificial Neural Networks — ICANN’97: 7th International Conference Lausanne, Switzerland.*, pages 583–588. Springer Berlin Heidelberg, Berlin, Heidelberg.

## A Proofs

*Theorem 1.* Consider the GP regression set-up as described in Section 4.1.3. For ease of notation let us denote the posterior process as  $(Z_{\mathbf{x}})_{\mathbf{x} \in T} \sim GP(\mu, \mathfrak{K})$ , where we drop  $n$ , the number of observations as it is fixed in this section. Recall that the proposed approximate field  $(\tilde{Z}_{\mathbf{x}})_{\mathbf{x} \in T}$  is defined as follows

$$\tilde{Z}_{\mathbf{x}} = a(\mathbf{x}) + \mathbf{b}^T(\mathbf{x})Z_G \quad \mathbf{x} \in T,$$

where  $a : T \rightarrow \mathbb{R}$  is a continuous trend function,  $\mathbf{b} : T \rightarrow \mathbb{R}^\ell$  is a continuous vector-valued function of deterministic weights,  $G = (\mathbf{g}_1, \dots, \mathbf{g}_\ell) \in T^\ell$  is a fixed sequence of points in  $T$  and  $Z_G = (Z_{\mathbf{g}_1}, \dots, Z_{\mathbf{g}_\ell})^T$  is a  $\ell$ -dimensional random vector. Let us consider the difference process  $(\tilde{\Delta}_{\mathbf{x}})_{\mathbf{x} \in T}$ , defined as  $\tilde{\Delta}_{\mathbf{x}} := Z_{\mathbf{x}} - \tilde{Z}_{\mathbf{x}}$ , for each  $\mathbf{x} \in T$ . The mean function and covariance kernel of  $\tilde{\Delta}$  are

$$\begin{aligned} \mu^{\tilde{\Delta}}(\mathbf{x}) &= \mathbb{E}[Z_{\mathbf{x}} - \tilde{Z}_{\mathbf{x}}] = \mu(\mathbf{x}) - a(\mathbf{x}) - \mathbf{b}(\mathbf{x})^T \mu(G) \quad \mathbf{x} \in T \\ \mathfrak{K}^{\tilde{\Delta}}(\mathbf{x}, \mathbf{y}) &= \mathfrak{K}(\mathbf{x}, \mathbf{y}) - \mathfrak{K}(\mathbf{y}, G)\mathbf{b}(\mathbf{x}) - \mathfrak{K}(\mathbf{x}, G)\mathbf{b}(\mathbf{y}) + \mathbf{b}^T(\mathbf{x})\mathfrak{K}(G, G)\mathbf{b}(\mathbf{y}) \quad \mathbf{x}, \mathbf{y} \in T \end{aligned}$$

First of all notice that

$$Pr \left( \left| \sup_{\mathbf{x} \in T} Z_{\mathbf{x}} - \sup_{\mathbf{x} \in T} \tilde{Z}_{\mathbf{x}} \right| > u \right) \leq Pr \left( \sup_{\mathbf{x} \in T} |\tilde{\Delta}_{\mathbf{x}}| > u \right) \quad (16)$$

Let us now consider the centred process  $(\tilde{\Delta}_{\mathbf{x}}^C)_{\mathbf{x} \in T} := (\tilde{\Delta}_{\mathbf{x}} - \mu^{\tilde{\Delta}}(\mathbf{x}))_{\mathbf{x} \in T}$ . We have that

$$Pr \left( \sup_{\mathbf{x} \in T} |\tilde{\Delta}_{\mathbf{x}}^C| > u \right) \leq 2Pr \left( \sup_{\mathbf{x} \in T} \tilde{\Delta}_{\mathbf{x}}^C > u \right) \leq 2e^{-u^2/2(\sigma_T^{\tilde{\Delta}})^2}, \quad u > 0 \quad (17)$$

where  $(\sigma_T^{\tilde{\Delta}})^2 = \sup_{\mathbf{x} \in T} \mathbb{E}[(\tilde{\Delta}_{\mathbf{x}}^C)^2] = \sup_{\mathbf{x} \in T} \mathfrak{K}^{\tilde{\Delta}}(\mathbf{x}, \mathbf{x})$ . The first inequality follows from the symmetric distribution of the centered field  $\tilde{\Delta}_{\mathbf{x}}^C$  and the second is the Borell-TIS inequality, see, e.g., [Adler and Taylor \(2007\)](#), Chapter 2 for more detail.

Since we have

$$\sup_{\mathbf{x} \in T} |\tilde{\Delta}_{\mathbf{x}}| \leq \sup_{\mathbf{x} \in T} |\tilde{\Delta}_{\mathbf{x}}^C| + \sup_{\mathbf{x} \in T} |\mu^{\tilde{\Delta}}(\mathbf{x})| \quad (18)$$

then following Equations (16), (17), (18), if  $u > \mu^{\tilde{\Delta}}(\mathbf{x})$

$$\begin{aligned} Pr \left( \left| \sup_{\mathbf{x} \in T} Z_{\mathbf{x}} - \sup_{\mathbf{x} \in T} \tilde{Z}_{\mathbf{x}} \right| > u \right) &\leq Pr \left( \sup_{\mathbf{x} \in T} |\tilde{\Delta}_{\mathbf{x}}^C| + \sup_{\mathbf{x} \in T} |\mu^{\tilde{\Delta}}(\mathbf{x})| > u \right) \\ &\leq Pr \left( \sup_{\mathbf{x} \in T} |\tilde{\Delta}_{\mathbf{x}}^C| > u - \sup_{\mathbf{x} \in T} |\mu^{\tilde{\Delta}}(\mathbf{x})| \right) \\ &\leq 2 \exp \left( - \frac{\left( u - \sup_{\mathbf{x} \in T} |\mu^{\tilde{\Delta}}(\mathbf{x})| \right)^2}{2(\sigma_T^{\tilde{\Delta}})^2} \right). \end{aligned}$$

If  $\tilde{Z}$  is an unbiased approximation for  $Z$ , then  $\mu^{\tilde{\Delta}}(\mathbf{x}) = 0$  and the inequality is valid for any  $u > 0$ .  $\square$

*Proof of Corollary 1.* Let us denote with  $A = P_{\Psi}^{\sup} Z(\eta^*)$  and  $B = P_{\Psi}^{\sup} \tilde{Z}(\eta^*)$ . By using Theorem 1 we have that for  $u > 0$

$$Pr(|A - B| \geq u) \leq 2e^{-\frac{(u - \mu_T^{\tilde{\Delta}})^2}{2(\sigma_T^{\tilde{\Delta}})^2}}$$

Moreover

$$\begin{aligned} Pr(A \leq u_{\alpha}^+) &= 1 - Pr(A \geq u_{\alpha}^+) = 1 - Pr(A - B + B \geq u_{\alpha}^+) \\ &= 1 - Pr(A - B + B \geq u_{\alpha}^+ \mid B \geq \tilde{u}_{\beta}^+) Pr(B \geq \tilde{u}_{\beta}^+) \\ &\quad - Pr(A - B + B \geq u_{\alpha}^+, B \leq \tilde{u}_{\beta}^+) \\ &\geq 1 - \beta - Pr(A - B + \tilde{u}_{\beta}^+ \geq u_{\alpha}^+, B \leq \tilde{u}_{\beta}^+) \\ &\geq 1 - \beta - Pr(|A - B| \geq u_{\alpha}^+ - \tilde{u}_{\beta}^+, B \leq \tilde{u}_{\beta}^+) \\ &\geq 1 - \beta - 2e^{-\frac{(u_{\alpha}^+ - \tilde{u}_{\beta}^+ - \mu_T^{\tilde{\Delta}})^2}{2(\sigma_T^{\tilde{\Delta}})^2}} \end{aligned}$$

By solving for  $u_\alpha^+$  in  $\beta + 2e^{-\frac{(u_\alpha^+ - \tilde{u}_\beta^+ - \mu_T^{\tilde{\Delta}})^2}{2(\sigma_T^{\tilde{\Delta}})^2}} = \alpha$  under the constraint  $u_\alpha^+ > \tilde{u}_\beta^+$ , we obtain (12). For the other side notice that

$$\begin{aligned} Pr(A \geq u_\alpha^-) &\geq Pr(B \geq \tilde{u}_\beta^-, |A - B| \leq \tilde{u}_\beta^- - u_\alpha^-) = 1 - Pr(B > \tilde{u}_\beta^-, |A - B| \geq \tilde{u}_\beta^- - u_\alpha^-) \\ &\quad - Pr(B \leq \tilde{u}_\beta^-, |A - B| \geq \tilde{u}_\beta^- - u_\alpha^-) - Pr(B \leq \tilde{u}_\beta^-, |A - B| \leq \tilde{u}_\beta^- - u_\alpha^-) \\ &\geq 1 - Pr(|A - B| \geq \tilde{u}_\beta^- - u_\alpha^-) - \beta - \beta \geq 1 - 2\beta - 2e^{-\frac{(u_\alpha^- - \tilde{u}_\beta^-)^2}{2(\sigma_T^{\tilde{\Delta}})^2}} \end{aligned}$$

By solving for  $u_\alpha^-$  in  $2\beta + 2e^{-\frac{(u_\alpha^- - \tilde{u}_\beta^-)^2}{2(\sigma_T^{\tilde{\Delta}})^2}} = \alpha$  under the constraint  $u_\alpha^- < \tilde{u}_\beta^-$ , we obtain (12).  $\square$

## B Gradient of $\tilde{Z}_x$ with respect to $\mathbf{x}$

We are interested in the approximating process for the posterior distribution of  $Z$  conditioned on  $(X_n, \mathbf{y}_n)$ . If  $a, \mathbf{b}$  are chosen as the posterior mean and the kriging weights respectively, then we can write  $\tilde{Z}$  as

$$\tilde{Z}_x = \Lambda(\mathbf{x})^T \begin{bmatrix} \mathbf{y}_n \\ Z_G \end{bmatrix} \quad \text{with}$$

$$\Lambda(\mathbf{x}) = K_{n+\ell}^{-1} (\mathfrak{K}_{n+\ell}(\mathbf{x}) + H_{n+\ell} K_H (\mathbf{h}(\mathbf{x}) - H_{n+\ell}^T K_{n+\ell}^{-1} \mathfrak{K}_{n+\ell}(\mathbf{x})))$$

$$\begin{aligned} \text{where } K_H &= (H_{n+\ell}^T K_{n+\ell}^{-1} H_{n+\ell})^{-1}, \quad K_{n+\ell} = \mathfrak{K}(A_{n+\ell}, A_{n+\ell}), \quad \mathfrak{K}_{n+\ell}(\mathbf{x}) = \mathfrak{K}(x, A_{n+\ell}), \\ H_{n+\ell} &= [h_j(A_{n+\ell})]_{j=1, \dots, m} \in \mathbb{R}^{(n+\ell) \times m} \quad \text{and} \quad A_{n+\ell} = [X_n^T, G^T]^T \in \mathbb{R}^{(n+\ell) \times d} \end{aligned}$$

Then  $\nabla_x \tilde{Z}_x = \nabla_x \Lambda(\mathbf{x})^T \begin{bmatrix} \mathbf{y}_n \\ Z_G \end{bmatrix}$  and it suffices to compute the gradient of  $\Lambda$  and

$$\begin{aligned} K_{n+\ell} \nabla_x \Lambda(\mathbf{x}) &= (\nabla_x \mathfrak{K}(\mathbf{x}, A_{n+\ell}) + H_{n+\ell} K_H (\nabla_x \mathbf{h}(\mathbf{x}) - H_{n+\ell}^T K_{n+\ell}^{-1} \nabla_x \mathfrak{K}(\mathbf{x}, A_{n+\ell}))) \\ \text{with } \nabla_x \mathfrak{K}(\mathbf{x}, A_{n+\ell}) &= [\nabla_x \mathfrak{K}(\mathbf{x}, a_1), \dots, \nabla_x \mathfrak{K}(\mathbf{x}, a_{n+\ell})]^T \in \mathbb{R}^{(n+\ell) \times d} \\ \text{and } \nabla_x \mathbf{h}(\mathbf{x}) &= [\nabla_x h_1(\mathbf{x}), \dots, \nabla_x h_m(\mathbf{x})] \in \mathbb{R}^{m \times d}. \end{aligned}$$

## C Full results on flooding test case

In this section, we report more details on the GP model used in the flooding test case in Section 4 and we show the profile extrema functions for all thresholds  $\tau_1, \dots, \tau_4$ .

In table 1 we compare different GP models according to two metrics:  $Q^2$  on leave-one-out predictions and log likelihood ( $\log Lik$ ). According to those metrics the prior mean function in equation 15 (universal kriging, UK) results in better fits than a constant prior mean (ordinary kriging, OK). On the other hand the difference between using a smoothness parameter  $\nu = 3/2$  or  $\nu = 5/2$  is very small. Here we chose the parameter  $\nu = 3/2$  as it leads to standardized model residuals that have a distribution closer to the normal one.

Table 1: Comparison of different GP models for flooding test case. OK: constant mean; UK: mean function as in equation 15. Best values in bold.

	Matern 3/2, OK	Matern 3/2, UK	Matern 5/2, OK	Matern 5/2, UK
$Q^2$	0.95	<b>0.96</b>	0.94	<b>0.96</b>
$\log Lik$	-1,406.02	<b>-1,363.73</b>	-1,410.71	-1,365.32

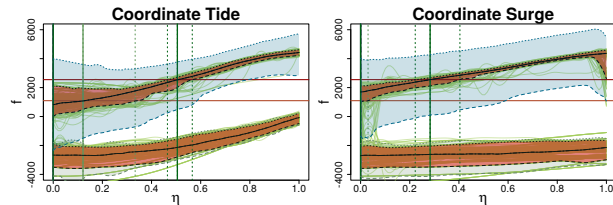


Figure 14: Coordinate profiles on GP model with Matérn  $\nu = 3/2$  and constant mean.

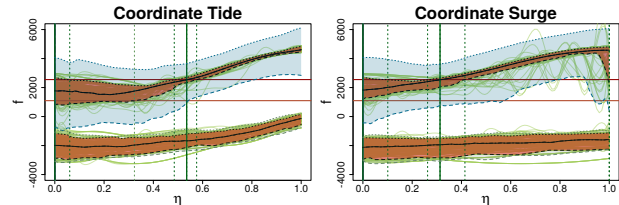


Figure 15: Coordinate profiles on GP model with Matérn  $\nu = 5/2$  and constant mean.

If the model fit is worse, then we obtain coordinate profiles with larger confidence bands and thus more uncertainty. We checked this assumption by computing coordinate profile extrema on each of the models in Table 1. Figures 14 and 15 show the first two coordinate profile plots for Matérn 3/2, OK and Matérn 5/2, OK. A quick glance already shows that the confidence bands are larger than in Figure 12. We further compared the integrated inter-quantile range for each model and we ranked them in increasing order. The average ranks over all coordinates for profile sup and profile inf are shown in Table 2. Note how the chosen model (Matérn 3/2 model with UK prior mean) has an average rank 1.8 for  $P^{\text{sup}}f$  and 1.4 for  $P^{\text{inf}}f$ .

Table 2: Average rank of integrated inter-quantile range for coordinate profile extrema.

	Matern 3/2, OK	Matern 3/2, UK	Matern 5/2, OK	Matern 5/2, UK
$P^{\text{sup}}f$	2.4	<b>1.8</b>	3	2.8
$P^{\text{inf}}f$	3.8	<b>1.4</b>	2.8	2

Let us now focus on the chosen model: Matérn  $\nu = 3/2$  with prior mean function as described in equation 15. Table 3 summarizes the intervals selected with the coordinate profile extrema functions. Figure 16 shows the coordinate profile extrema functions, the 5%, 95% quantiles, the boundary of the subsets selected (vertical lines) by posterior median and quantiles. The values of the bound with confidence level  $\alpha = 0.1$  are shown with the sky blue ( $P_i^{\text{sup}}Z$ ) and sea green ( $P_i^{\text{inf}}Z$ ) shaded regions. The bound on the profile inf function is almost overlapping with the 90% confidence intervals in red. The bound on the profile sup function instead provides an higher quantile and identifies as region of possible non excursion  $T \in [0, 0.19]$  above  $\tau_4$ . For the other coordinates the bound does not provide information on regions of non excursion.

Section 3.4 introduced the integrated variance of the difference, equation (13), as an indicator for the bound tightness. Figure 17 shows the value  $I(\sigma_T^{\hat{\Delta}})^2$  averaged over all

Table 3: Regions excluded with profile extrema functions on 5-d test case. Interval defined in the top line, for each threshold the interval for the boundary computed from the 90% approximate confidence intervals for the profiles is reported.

	$T \in [0, \lambda_{T,k}]$	$S \in [0, \lambda_{S,k}]$	$t_o \in [0, \lambda_{t_o,k}^a] \cup [\lambda_{t_o,k}^b, 1]$	$t_+$	$t_-$
$\tau_1$	$\lambda_{T,1} \in [0.17, 0.26]$	—	$\lambda_{t_o,1}^a = 0, \lambda_{t_o,1}^b = 1$	—	—
$\tau_2$	$\lambda_{T,2} \in [0.24, 0.32]$	$\lambda_{S,2} \in [0, 0.06]$	$\lambda_{t_o,2}^a \in [0, 0.11], \lambda_{t_o,2}^b = 1$	—	—
$\tau_3$	$\lambda_{T,3} \in [0.33, 0.40]$	$\lambda_{S,2} \in [0, 0.14]$	$\lambda_{t_o,3}^a \in [0, 0.12], \lambda_{t_o,3}^b = 1$	—	—
$\tau_4$	$\lambda_{T,4} \in [0.52, 0.60]$	$\lambda_{S,2} \in [0.28, 0.43]$	$\lambda_{t_o,4}^a \in [0, 0.16], \lambda_{t_o,4}^b = 1$	—	—

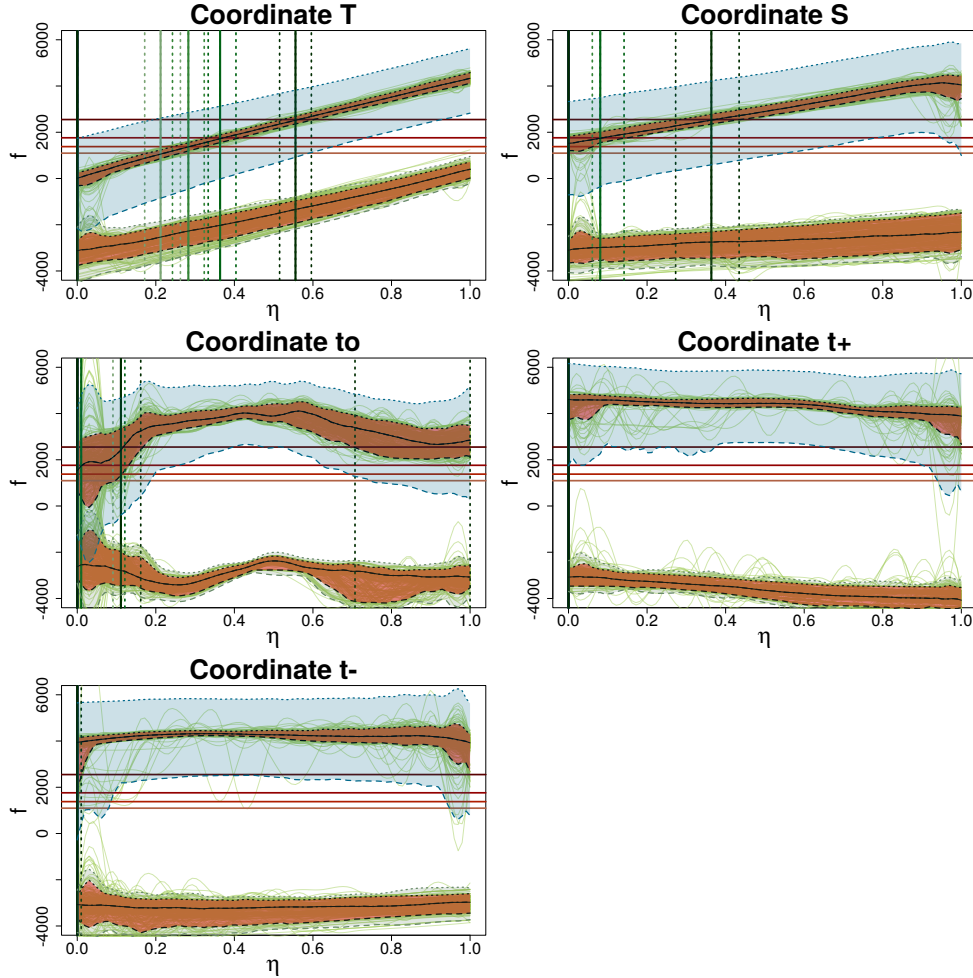


Figure 16: Coordinate profiles for 5-dimensional test case. Thresholds (red, light to dark for  $\tau_1, \dots, \tau_4$ ), regions selected by profile extrema on mean (vertical green lines) and regions selected by profile extrema on 95% – 5% quantiles (vertical green lines, dashed). Bound on the confidence region ( $\alpha = 0.05$ ) for  $P_i^{\text{sup}}Z$  (sky blue, shaded) and for  $P_i^{\text{inf}}Z$  (sea green, shaded).

dimensions for  $P^{\text{sup}}$  and  $P^{\text{inf}}$  versus the number of pilot points  $\ell$ . Note the “diminishing

returns” as  $\ell$  increases. In particular, here we chose  $\ell = 300$  for computational reasons.

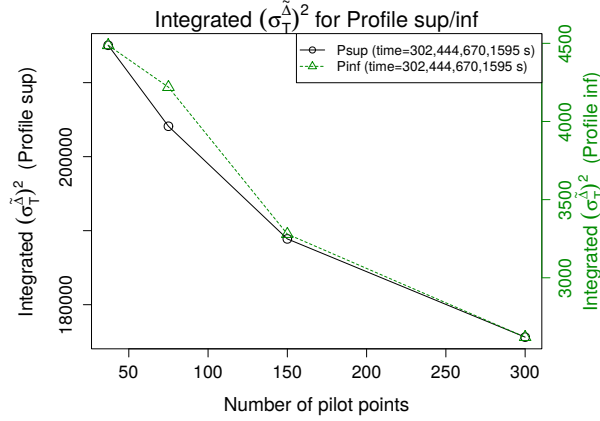


Figure 17: Average integrated variance of the difference, equation (13), for profile inf (dashed, triangles) and profile sup (solid, circles) versus  $\ell$  (37, 75, 150, 300). Computational time for  $(\sigma_T^{\Delta})^2$  in legend. Profile coordinates, 5-d test case.

Figure 18 shows the cuts at  $T = 0, 0.5, 1$  of the bivariate profile functions on the mean shown in Figure 13(b) with their respective upper and lower bounds. Such plots can be used to zoom in particular regions of the bivariate profiles. They do not plot more information than bivariate profile maps, however they are simpler to read. For example, Figure 18 shows that  $\{\mathbf{x} \in D : T = 0 \text{ and } S < 0.79\}$  is a non-excursion region with high probability.

Figure 19 shows the bivariate profile extrema for the variables  $T, t_o$ . In particular the left-most map allows us to exclude all input regions on the left of the red dotted curve. In a symmetric way, the right-most plot tells us that values for  $T, t_o$  in the region on the right-hand side of the red dotted line lead to an excursion. Also in this case the uncertainty quantification adds more insights on the result. The profile inf shows much uncertainty around the threshold, therefore while region of interest could be selected the uncertainty must be accounted for. The profile sup also tells us that close to the threshold

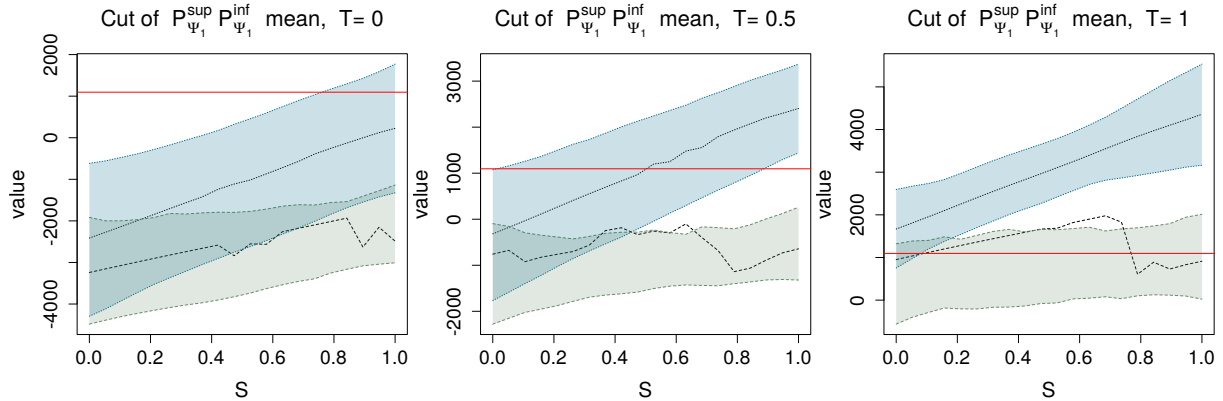


Figure 18: Cut with  $T = 0, 0.5, 1$  of  $P_{\psi_1}^{\text{sup}}, P_{\psi_1}^{\text{inf}}$  on the profile mean, see also Figure 13.



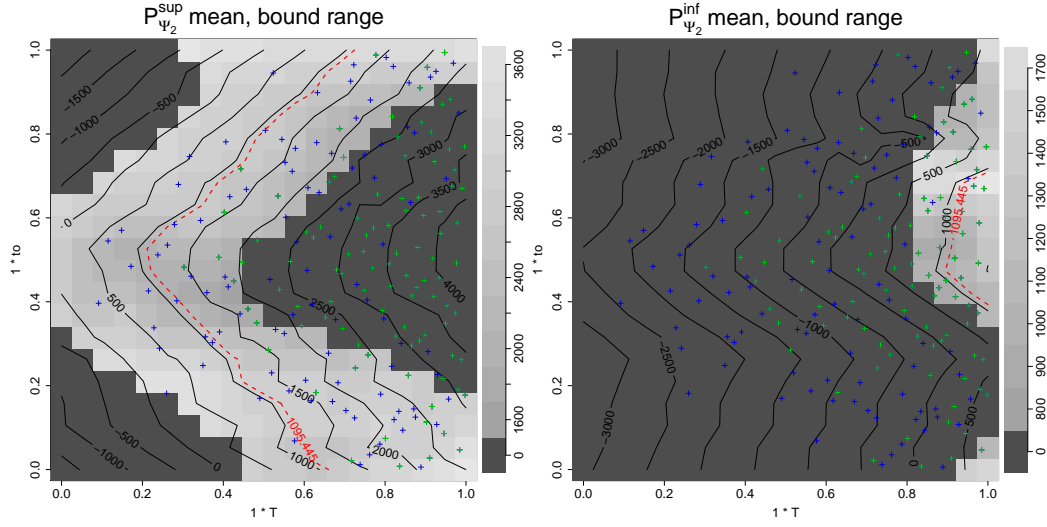


Figure 19: Bivariate profile extrema for 5-dimensional test case with  $T, t_o$ . Excursion threshold  $\tau_1$  in red, bivariate profile mean values in contour lines,  $X_n$  locations as crosses (green if above  $\tau_1$ ), background color denotes the upper-lower bound range (equation (12)).

the uncertainty is still high, however the top and bottom left triangles are not in the excursion with very little uncertainty. A very conservative assessment should select only this input region as non-excursion. Figure 20 shows the cuts at  $T = 0, 0.5, 1$  of these bivariate profile functions. As in the other case such plots confirm information shown in bivariate maps. For example, we can clearly see that for  $T = 0$  and  $t_o \in [0, 0.37] \cup [0.63, 1]$  there is no excursion with high probability.

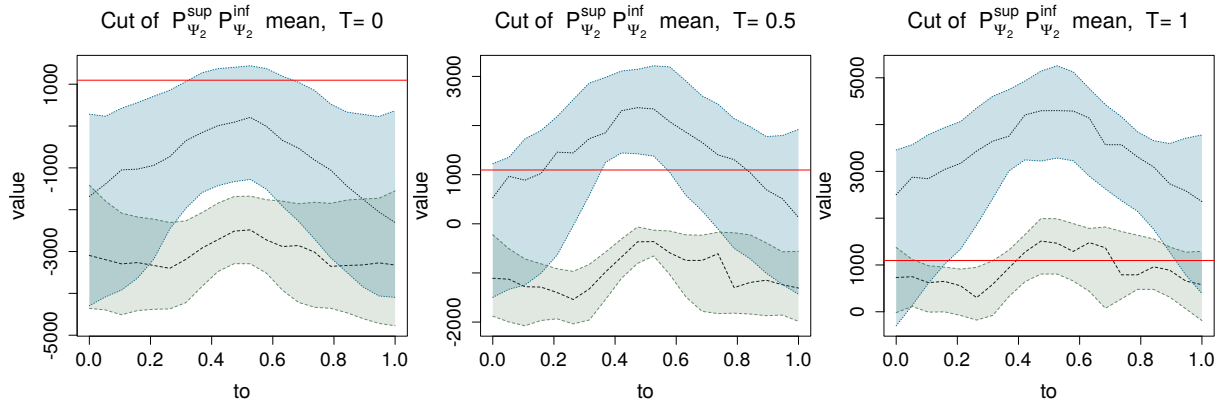


Figure 20: Cut with  $T = 0, 0.5, 1$  of  $P_{\Psi_2}^{\sup}, P_{\Psi_2}^{\inf}$  on the profile mean, see also Figure 19.

Table 4: Summary of profile extrema parameters.

Symbol	Meaning	Section
$n$	Number of expensive computer experiment simulations	1
$d$	Dimension of input space	1
$p$	Dimension of profile extrema projection, usually 1 or 2	2
$k$	Number of points used for deterministic approximation of $P^{\text{sup}}, P^{\text{inf}}$	2.4
$\ell$	Number of pilot points for GP approximation in (9)	3.2
$s$	Number of posterior approximate GP realizations	3.2

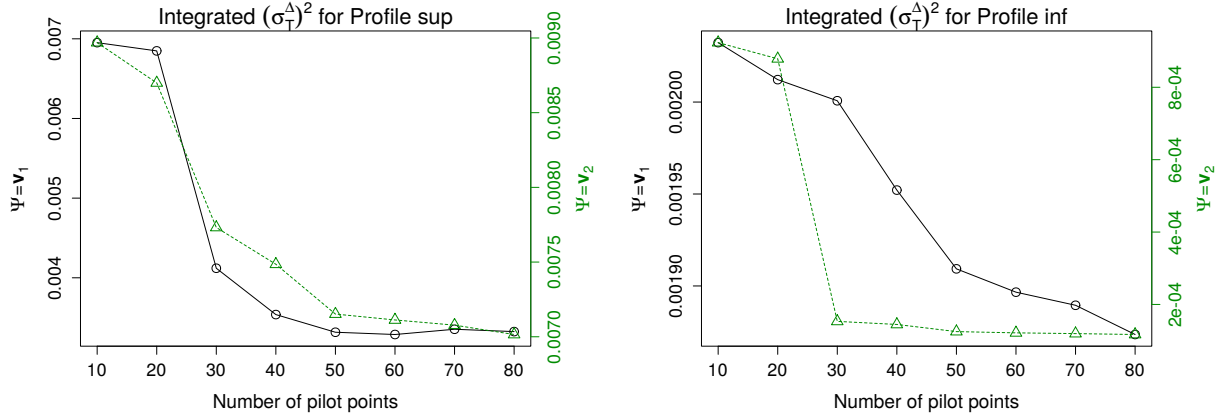


Figure 21: Values of  $I(\sigma_T^{\Delta})^2(\ell)$ , Eq. (13), for  $\Psi = \mathbf{v}_1$  (solid, circles) and  $\Psi = \mathbf{v}_2$  (dashed, triangles) as functions of  $\ell$ . Profile sup (left) and Profile inf (right).

## D An indicator of tightness for the bound

The bounds for the uncertainty on  $P_{\Psi}^{\text{sup}}Z$  are computed with equation (12). For a given  $\ell$ , the tightness of this approximation is controlled by  $(\sigma_T^{\Delta})^2(\eta, \ell)$ , i.e. the sup of the difference variance. This is a non-negative function of  $\eta$ , as the difference depends on where we evaluate  $P_{\Psi}^{\text{sup}}Z(\eta)$ . The integrated variance  $I(\sigma_T^{\Delta})^2(\ell)$ , Equation (13), is a summary of this quantity. Since  $(\sigma_T^{\Delta})^2(\eta, \ell)$  is non-negative for each  $\eta \in E_{\Psi}$ , a smaller integral implies less variability and a tighter bound. We can control  $I(\sigma_T^{\Delta})^2(\ell)$  with the number of pilot points  $\ell$  chosen for the approximate realizations. Figure 21 shows the estimated integral  $I(\sigma_T^{\Delta})^2(\ell) = \int_{E_{\mathbf{v}_i}} (\sigma_T^{\Delta})^2(\eta, \ell) d\eta$ ,  $i = 1, 2$  for the synthetic example introduced in Section 3.3 with  $n = 90$  as a function of  $\ell$ . More pilot points lead to a smaller integral for each coordinate and for both profile extrema, however the rate of decrease is test case dependent.

## E Uni and bivariate profiles, a combined approach

In this section we consider the following function

$$\gamma(\mathbf{x}) = \sin(a\mathbf{v}_1^T \mathbf{x} + b) + \cos(c\mathbf{v}_2^T \mathbf{x} + d) + \sin(e\mathbf{v}_3^T \mathbf{x} + f) - 1.5 \quad \mathbf{x} \in [0, 1]^3, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3 \quad (19)$$

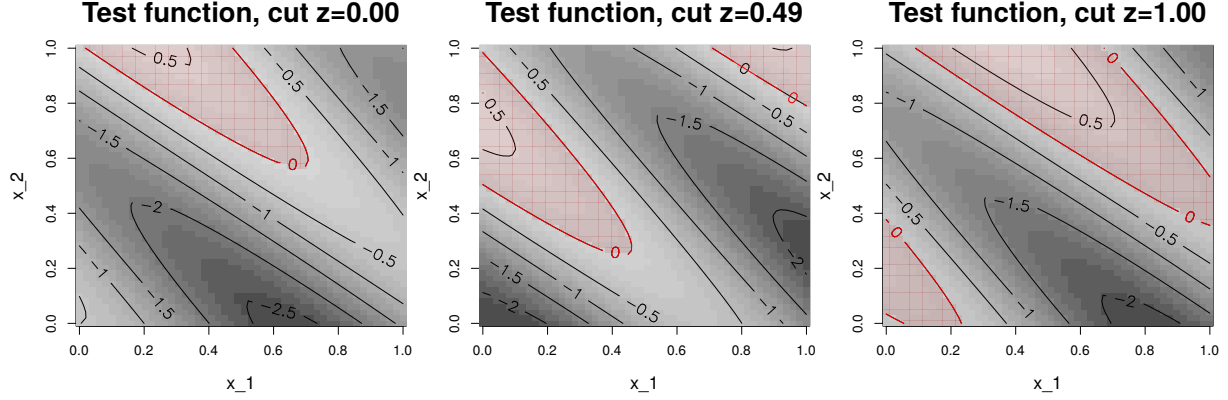


Figure 22: Cuts of the function in Equation (19).

where  $a, b, c, d, e, f \in \mathbb{R}$  and

$$\mathbf{v}_1 = \begin{bmatrix} \sin(\theta) \cos(\phi) \\ \sin(\theta) \sin(\phi) \\ \cos(\theta) \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} \cos(\theta) \cos(\phi) \\ \cos(\theta) \sin(\phi) \\ -\sin(\theta) \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -\sin(\phi) \\ \cos(\phi) \\ 0 \end{bmatrix}$$

with  $\theta = \pi/4$ ,  $\phi = \pi/4$ . We fix  $[a, b, c, d, e, f] = [1, 0, 10, 0, 1, 0]$  and we study the excursion set above  $\tau = 0$ . Figure 22 shows three cuts corresponding to the planes  $z = 0, 0.49, 1$ .

Since the excursion set sits along directions which are oblique with respect to the canonical coordinates, the coordinate profile extrema do not provide information on this set. Figure 23 shows that  $P_i^{\text{sup}} f, i = 1, 2, 3$  are always above the threshold and  $P_i^{\text{inf}} f, i = 1, 2, 3$  are always below the threshold, thus we cannot exclude any region. In such situations we propose two approaches: compute oblique profiles according to more informative directions or compute bivariate coordinate profiles. Figure 24 show the oblique profile extrema along the generating directions  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ . This plot excludes part of the input space by considering the planes delimited by the interceptions between the profile sup and the threshold, e.g. all points between the planes  $\mathbf{v}_1 \mathbf{x} = 0$  and  $\mathbf{v}_1 \mathbf{x} = 0.31$ .

The second option involves computing the profile extrema with projections on 2 dimensional subspaces. Here we consider all combinations of bivariate projections on canonical axes. Figures 25, 26 show the maps obtained for the  $P_\Psi^{\text{sup}} f, P_\Psi^{\text{inf}} f$  for the three matrices build by combining the canonical directions. Figure 25 in particular highlights darker shaded (red) regions that are not part of the excursion set. Consider  $P_{\Psi_1}^{\text{sup}} \gamma$ , for all  $x_1, x_2$  in the shaded region the profile tells us that there is no excursion. For example if  $x_1 \in [0.115, 1]$  and  $x_2 = 0.2$ , the segment highlighted in the top left plot of Figure 25, there is no excursion. In this case,  $P_\Psi^{\text{inf}} f$ , Figure 26, does not allow us to select any excursion region.

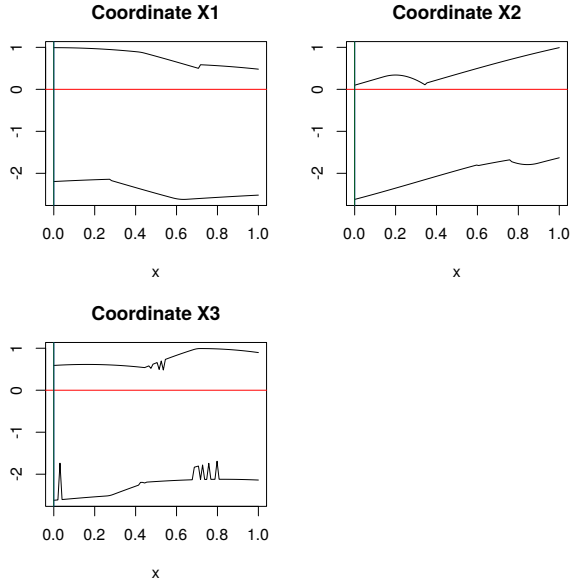


Figure 23: Coordinate profile extrema for the function in Equation (19).

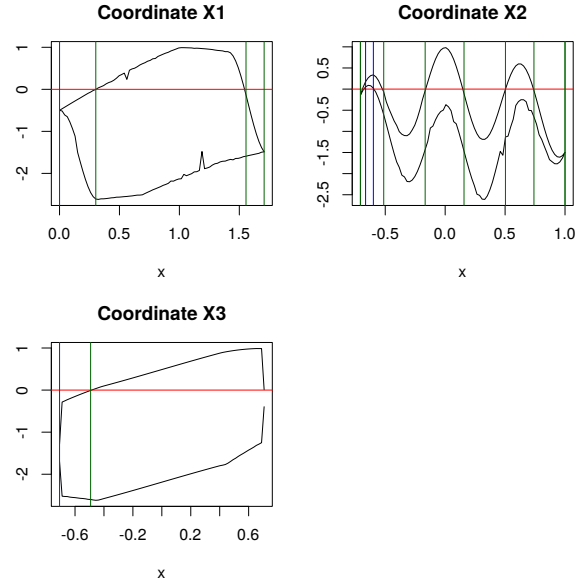


Figure 24: Profile extrema for the function in Equation (19) along the directions  $\mathbf{v}_1$ ,  $\mathbf{v}_2$ ,  $\mathbf{v}_3$ .

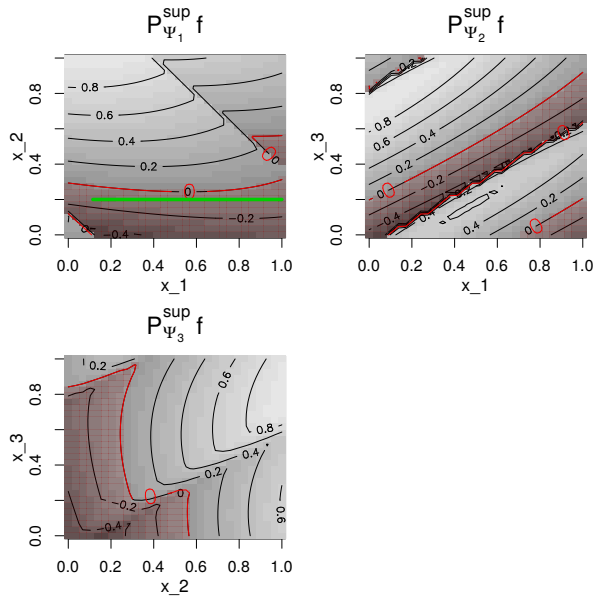


Figure 25: Bivariate profile sup for the function in Equation (19) along projections on combinations of three canonical axes.

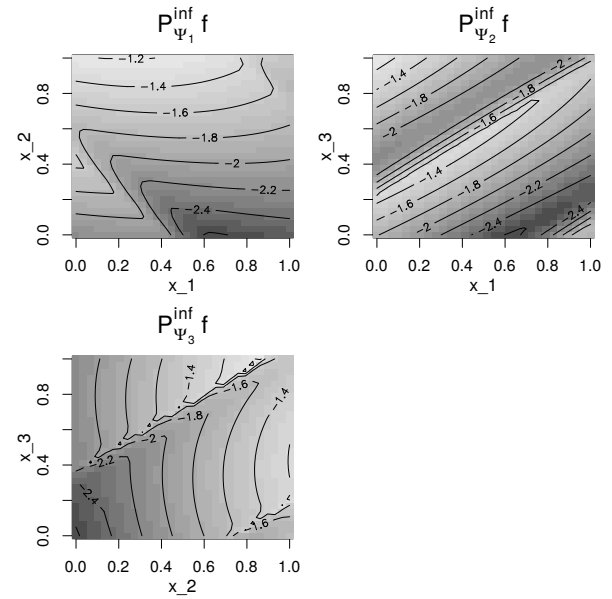


Figure 26: Bivariate profile inf for the function in Equation (19) along projections on combinations of three canonical axes.