Set inversion under functional uncertainties with joint meta-models

Reda El Amri, Céline Helbert, Miguel Munoz Zuniga, Clémentine Prieur, Delphine Sinoquet

Set inversion under functional uncertainties with joint meta-models *

Reda El Amri†, Céline Helbert‡, Miguel Munoz Zuniga§, Clémentine Prieur¶, and Delphine Sinoquet‖

Abstract. In this paper we propose an efficient sampling strategy to solve inversion problem under functional uncertainty. This approach aims to characterize region of a control space defined by exceedance above prescribed threshold. This study is motivated by an application on identifying the set of control parameters leading to meet the pollutant emission standards of a vehicle under driving profile uncertainties. In that context, the constrained response in the inversion problem is here formulated as the expectation over the functional random variable only known through a set of realizations and the unknown set is thus associated with the control variables. As often in industrial applications, this problem involves high-fidelity and time-consuming computational models. We thus proposed an approach that makes use of Gaussian Process meta-models built on the joint space of control and uncertain input variables. Specifically, we define a design criterion based on uncertainty in the excursion of the Gaussian Process and derive tractable expressions for the variance reduction in such a framework. Applications to analytical examples, followed by the automotive industrial test case show the accuracy and the efficiency brought by the proposed procedure.

Key words. Set inversion; Gaussian Process models; Data reduction; Functional uncertainties.

AMS subject classifications.

1. Introduction.

In recent years, engineers and scientists are increasingly relying on computer models as surrogate for physical experimentation generally too costly or impossible to execute ([BGL+12, CBG+14]). In particular, practitioners using these numerical simulations are not only interested in the response of their model for a given set of inputs (forward problem) but also in recovering the set of input values leading to a prescribed value or range for the output of interest. The problem of estimating such a set is called hereafter an inversion problem.

In our context, the numerical simulator modelling the system, denoted f, takes two types of input variables: a set of control variables $\mathbf{x} \in \mathbb{X}$, and a set of uncertain variables $\mathbf{v} \in \mathcal{V}$. Without considering any assumptions on the distribution of the uncertain variable $\mathbf{v}$, robust inversion consists in seeking the set of control variables $\mathbf{x} \in \mathbb{X}$ such that $\sup_{\mathbf{v}\in\mathcal{V}} f(\mathbf{x}, \mathbf{v})$ is smaller than a threshold c. Then, the difficulty of solving the robust inversion problem strongly depends on the uncertainty set $\mathcal{V}$. In our setting, $\mathcal{V}$ is a functional space, and we consider the inversion problem under uncertainty as a stochastic inversion problem, assuming that the uncertainty has a probabilistic description. Let $\mathbf{V}$ denote the associated random variable, valued in $\mathcal{V}$, modelling the uncertainty. In our framework, we are interested in recovering the set $\Gamma^* := \{\mathbf{x} \in \mathbb{X}\ ,\ g(\mathbf{x}) = \mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})] \leq c\}$, with $c \in \mathbb{R}$, and the functional random variable $\mathbf{V}$ is only known from a set of realizations. The expectation appearing in Γ^* has to be estimated. Moreover, the simulations are time consuming and thus the usual Monte Carlo method to estimate the expectation ought to be avoided.

Inversion problems have already been carried out in many applications, notably reliability engineering (see, e.g., [BGL+12], [CBG+14]), climatology (see, e.g., [BL15], [FS+13]) and many other fields. In the literature, one way to solve the problem is to adopt a sequential sampling strategy

*Submitted to the editors DATE.

Funding: This work was supported by IFPEN and the OQUAIDO chair.

†Formerly IFP Energies Nouvelles, Rueil-Malmaison, France (elamri.reda@yahoo.com)

‡ECL, ICJ, UMR 5208, Université de Lyon, 36 av. G. de Collongue, Ecully, France (celine.helbert@ec-lyon.fr).

§IFP Energies Nouvelles, Rueil-Malmaison, France (miguel.munoz-zuniga@ifpen.fr).

¶Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LJK, Grenoble, France (clementine.prieur@univ-grenoble-alpes.fr).

‖IFP Energies Nouvelles, Rueil-Malmaison, France (delphine.sinoquet@ifpen.fr).

based on a Gaussian Process (GP) emulator of $g : \mathbf{x} \mapsto \mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})]$. The underlying idea is that Gaussian Process emulators, which capture prior knowledge about the regularity of the unknown function, make it possible to assess the uncertainty about Γ^* given a set of evaluations of g. More specifically, for the estimation of an excursion set, these sequential strategies are closely related to the field of Bayesian global optimization (see, e.g., [CG13]). In the case of inversion problems, Stepwise Uncertainty Reduction (SUR) strategies based on set measures were introduced in [VB09]. More recently, a parallel implementation of these strategies has been proposed in [CBG+14] and applied to the recovery of an excursion set. Briefly, the strategy SUR gives sequentially the next location in the control space where to run the simulator in order to minimize an uncertainty measure of the excursion set ([EAHL+20]).

In the field of robust optimization where uncertainty comes from a real-valued (or vector-valued) random input, various methods exist and aim at optimizing the expectation taken with respect to the probability distribution of the random input (see [JLR13] or [WSN00]). These methods are based on the modelling of f by a Gaussian Process built in the joint space of deterministic and uncertain variables. Then a "projected" (integrated) Gaussian Process is defined by taking the expectation with respect to the probability distribution of the random input, leading to an approximation of the expected response g. Finally an adaptive design of experiments is proposed for optimizing the objective function g.

In the same spirit, we propose an original method to deal with a stochastic inversion problem with the aim of further reducing the number of simulations required. In this work f is approximated by a Gaussian Process model built on the joint space $\mathbb{X} \times \mathcal{V}$. For the iterative approximation of Γ^*, the sampling strategy in the joint space is based on two steps. Firstly a SUR approach is applied to the "projected" Gaussian Process to determine the next evaluation point $\mathbf{x}_{n+1} \in \mathbb{X}$. Secondly, in the uncertain space, the next function v_{n+1} is chosen such that the standard error of the "projected" process evaluated at $\mathbf{x}_{n+1}$ is minimized. Compared to methods based on an accurate estimation of the expectation and the construction of a surrogate of g ([EAHL+20]) our adaptive design of experiments, defined in the joint space, leads to further reduce the number of calls to the numerical simulator.

The article is structured as follows. Firstly, in Section 2, we recall the problem formulation and we extend the concept of Gaussian Process modelling to the case where the inputs contain a functional variable known through a finite set of realizations. In Section 3, we introduce a new adaptive sampling strategy to choose $(\mathbf{x}_n, \mathbf{v}_n)$. The whole algorithm for our robust inversion procedure is then detailed. In Section 4, our procedure is implemented on two analytical test cases (Sections 4.1 and 4.2), and the modelling assumptions are discussed (Section 4.3). Finally, our new procedure is tested on the industrial test case of a car pollution control system (Section 4.4).

2. Problem formulation.

We model the output of the industrial simulator by a function $f : \mathbb{X} \times \mathcal{V} \to \mathbb{R}$ with $\mathbb{X}$ the space of the control parameters a bounded subset of $\mathbb{R}^p$ and $\mathcal{V}$ the space of the functional uncertain input. We model the functional uncertain input by a random variable $\mathbf{V}$ valued in $\mathcal{V}$. We are interested in estimating the set

$$\Gamma^* = \{\mathbf{x} \in \mathbb{X} \ , \ g(\mathbf{x}) \leq c\}, \tag{2.1}$$

where $c \in \mathbb{R}$ is a threshold and $g : \mathbb{X} \to \mathbb{R}$ such that $g(\mathbf{x}) = \mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})]$. An additional constraint is that $\mathbf{V}$ is known through a finite set of realizations. The implication of this constraint will be specified in Section 3.3.
The proposed sequential strategy to approximate Γ^* involves two main ingredients introduced hereafter : functional data reduction to reduce the problem to a finite dimensional space and Gaussian

Process modelling in the joint space $Control \times Uncertain$.

2.1. Functional data reduction. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We assume that the random process $\mathbf{V}$ belongs to $\mathcal{H} = \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathcal{V})$ with

$$\mathcal{V} = \left\{ \mathbf{v} : [0,T] \to \mathbb{R}, ||\mathbf{v}|| = (< v, v >)^{1/2} = \left(\int_0^T \mathbf{v}(t)^2 \mathbf{dt} \right)^{1/2} < +\infty \right\}.$$

We assume that $\mathbf{V} \in \mathcal{H}$ has zero mean and continuous covariance function $C(t,s)$. Then

$$\mathbf{V}(t) = \sum_{i=1}^{\infty} U_i \psi_i(t), \quad t \in [0,T], \tag{2.2}$$

where $\{\psi_i\}_{i=1}^{\infty}$ is an orthonormal basis of eigenfunctions of the integral operator corresponding to C

$$\lambda_i \psi_i(t) = \int_0^T C(t,s)\psi_i(s)\mathrm{ds}, \tag{2.3}$$

and with $\{U_i\}_{i=1}^{\infty}$ denoting a set of uncorrelated random variables with zero mean and variance λ_i. Decomposition (2.2) is known as the Karhunen-Loève (KL) expansion of $\mathbf{V}$ ([LK10]). In the following we denote the truncated version of $\mathbf{V}$

$$\mathbf{V}_m(t) = \sum_{i=1}^{m} U_i \psi_i(t), \tag{2.4}$$

which represents, in the mean square error sense, the optimal m-term approximation of $\mathbf{V}$ ([LK10]).

2.2. Gaussian Process modelling. We assume that $f(\mathbf{x}, \mathbf{v})$ is a realization of a Gaussian Process $Z_{(\mathbf{x},\mathbf{u})}$ defined on $\mathbb{X} \times \mathbb{R}^m$, where $\mathbf{u} = (< \mathbf{v}, \psi_1 >, \ldots, < \mathbf{v}, \psi_m >)^\top$.
Let m_Z be the mean function of $Z_{(\mathbf{x},\mathbf{u})}$ and k_Z its covariance function,

$$\begin{aligned} \mathbb{E}[Z_{(\mathbf{x},\mathbf{u})}] &= m_Z(\mathbf{x},\mathbf{u}), \\ \mathrm{Cov}(Z_{(\mathbf{x},\mathbf{u})}, Z_{(\mathbf{x}',\mathbf{u}')}) &= k_Z(\mathbf{x},\mathbf{u};\mathbf{x}',\mathbf{u}'). \end{aligned} \tag{2.5}$$

Let denote Z^n, the GP Z conditioned on the set of n observations (simulations) $\mathbf{Z}_n = \{f(\mathbf{x}_1,\mathbf{v}_1), \ldots, f(\mathbf{x}_n,\mathbf{v}_n)\}$ of Z at $\mathcal{X}_n \times \mathcal{U}_n = \{(\mathbf{x}_1,\mathbf{u}_1), \ldots, (\mathbf{x}_n,\mathbf{u}_n)\}$ where $\mathbf{u}_i = (< \mathbf{v}_i, \psi_1 > , \ldots, < \mathbf{v}_i, \psi_m >)^\top$

$$Z^n_{(\mathbf{x},\mathbf{u})} = [Z_{(\mathbf{x},\mathbf{u})} | Z_{\mathcal{X}_n \times \mathcal{U}_n} = \mathbf{Z}_n]. \tag{2.6}$$

The conditional expectation is

$$\mathbb{E}[Z^n_{(\mathbf{x},\mathbf{u})}] = m_Z(\mathbf{x},\mathbf{u}) + k_Z((\mathbf{x},\mathbf{u}); \mathcal{X}_n \times \mathcal{U}_n) k_Z(\mathcal{X}_n \times \mathcal{U}_n; \mathcal{X}_n \times \mathcal{U}_n)^{-1} (\mathbf{Z} - m_Z(\mathcal{X}_n \times \mathcal{U}_n)),$$

and the conditional covariance is

$$\begin{aligned} \mathrm{Cov}(Z^n_{(\mathbf{x},\mathbf{u})}, Z^n_{(\mathbf{x}',\mathbf{u}')}) &= k_Z((\mathbf{x},\mathbf{u});(\mathbf{x}',\mathbf{u}')) \\ &\quad - k_Z((\mathbf{x},\mathbf{u}); \mathcal{X}_n \times \mathcal{U}_n) k_Z(\mathcal{X}_n \times \mathcal{U}_n; \mathcal{X}_n \times \mathcal{U}_n)^{-1} k_Z(\mathcal{X}_n \times \mathcal{U}_n; (\mathbf{x}',\mathbf{u}')). \end{aligned}$$

It is important to note that the Gaussian Process $Z_{(\mathbf{x},\mathbf{u})}$ is defined on the finite-dimensional truncated space $\mathbb{X} \times \mathbb{R}^m$. A discussion about this model is proposed in Section 4.3.

2.3. Integrated Gaussian Process. Recall that $\Gamma^* = \{\mathbf{x} \in \mathbb{X}\ ,\ g(\mathbf{x}) = \mathbb{E}[f(\mathbf{x}, \mathbf{V})] \leq c\}$. Therefore, to model the function g, we introduce the integrated process

$$Y^n_{\mathbf{x}} = \mathbb{E}_{\mathbf{U}}[Z^n_{(\mathbf{x},\mathbf{U})}] = \int_{\mathbb{R}^m} Z^n_{(\mathbf{x},\mathbf{u})} d\rho(\mathbf{u}), \tag{2.7}$$

where $d\rho(u)$ is the probability distribution of $\mathbf{U} = (U_1, \ldots, U_m)^T$ introduced in (2.4). The process $Y^n_{\mathbf{x}}$ is a Gaussian Process ([JLR13]) fully characterized by its mean and covariance functions which are given by

$$\begin{aligned}\mathbb{E}[Y^n_{\mathbf{x}}] = &\int_{\mathbb{R}^m} m_Z(\mathbf{x}, \mathbf{u}) d\rho(u) + \\ &\int_{\mathbb{R}^m} k_Z((\mathbf{x}, \mathbf{u}); \mathcal{X}_n \times \mathcal{U}_n) k_Z(\mathcal{X}_n \times \mathcal{U}_n; \mathcal{X}_n \times \mathcal{U}_n)^{-1} (\mathbf{Z} - m_Z(\mathcal{X}_n \times \mathcal{U}_n)) d\rho(\mathbf{u}),\end{aligned} \tag{2.8}$$

and

$$\begin{aligned}\mathrm{Cov}(Y^n_{\mathbf{x}}, Y^n_{\mathbf{x}'}) = &\int_{\mathbb{R}^m} \int_{\mathbb{R}^m} k_Z((\mathbf{x}, \mathbf{u}); (\mathbf{x}', \mathbf{u}')) \\ &- k_Z((\mathbf{x}, \mathbf{u}); \mathcal{X}_n \times \mathcal{U}_n) k_Z(\mathcal{X}_n \times \mathcal{U}_n; \mathcal{X}_n \times \mathcal{U}_n)^{-1} k_Z(\mathcal{X}_n \times \mathcal{U}_n; (\mathbf{x}', \mathbf{u}')) d\rho(\mathbf{u}) d\rho(\mathbf{u}').\end{aligned} \tag{2.9}$$

3. Data driven infill strategy for stochastic inversion.

In this section we propose a two-step infill strategy in the joint space. The first step consists in choosing a point in the control space while the second one aims at enriching the design with a new point in the uncertain space.

3.1. Minimization of the Vorob'ev deviation: choice of next x. The objective of the first step is to wisely choose the points in the control space $\mathbb{X}$ in order to accurately estimate the set $\Gamma^* = \{\mathbf{x} \in \mathbb{X}\ ,\ g(\mathbf{x}) = \mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})] \leq c\}$. For this purpose, we consider in the inversion set the statistical model of the unobservable function g given by $Y^n_{\mathbf{x}}$ introduced in Section 2.3. Due to the stochastic nature of $(Y^n_{\mathbf{x}})_{\mathbf{x}\in\mathbb{X}}$, the associated excursion set,

$$\Gamma = \{\mathbf{x} \in \mathbb{X}\ ,\ Y^n_{\mathbf{x}} \leq c\} \tag{3.1}$$

is a well defined random closed set if $(Y^n_{\mathbf{x}})_{\mathbf{x}\in\mathbb{X}}$ has continuous sample paths ([Mol06] p.4, 23). Therefore, from now on, the considered random processes will be supposed separable ([Doo53], p.57), the mean m_Z continuous and the covariance function k_Z to be Matérn (5/2 or 3/2). Indeed, under these assumptions, we know that $(Z_{(\mathbf{x},\mathbf{u})})_{(\mathbf{x},\mathbf{u})\in\mathbb{X}\times\mathbb{R}^m}$ has continuous sample paths ([Pac03] p.44 table 2.1) and we can prove that the path continuity property remains valid for the integrated conditioned process $(Y^n_{\mathbf{x}})_{\mathbf{x}\in\mathbb{X}}$ by using the necessary criterion introduced in [Adl81] p.60 and presented in [Pac03] p.38 Eq.(2.9).

From the assumption that g is a realization of $Y^n_{\mathbf{x}}$, the true unknown set Γ^* can be seen as a realization of the random closed set Γ. The book of [Mol06] gives many possible definitions for the variance of a random closed set. In the present work we adapt the Stepwise Uncertainty Reduction (SUR) strategy introduced in [CG13] which aims at decreasing an uncertainty function defined as the Vorob'ev deviation ([Vor84, VL13]) of the random set.
More precisely the uncertainty function at step n is defined as

$$\mathcal{H}^{\text{uncert}}_n = \mathbb{E}[\mu(\Gamma \triangle Q_{n,\alpha^*_n}) \mid Z_{\mathcal{X}_n \times \mathcal{U}_n} = \mathbf{Z}_n],$$

where μ is the Lebesgue measure on $\mathbb{X}$, the Vorob'ev quantiles are given by $Q_{n,\alpha} = \{\mathbf{x} \in \mathbb{X}\ ,\ \mathbb{P}(Y^n_{\mathbf{x}} \leq c) \geq \alpha\}$, and the Vorob'ev expectation Q_{n,α^*_n} can be determined by tuning α to a level α^* such

that $\mu(Q_{n,\alpha_n^*}) = \mathbb{E}[\mu(\Gamma) \mid Z_{\mathcal{X}_n \times \mathcal{U}_n} = \mathbf{Z}_n]$.
Let

$$\mathcal{H}_{n+1}^{\text{uncert}}(\mathbf{x}) = \mathbb{E}[\mu(\Gamma \triangle Q_{n+1,\alpha_{n+1}^*}) \mid Z_{\mathcal{X}_n \times \mathcal{U}_n} = \mathbf{Z}_n, Y_{\mathbf{x}}^n].$$

The objective of the SUR strategy is thus to enrich the current design with a new point $\mathbf{x}_{n+1}$ satisfying

$$\begin{aligned} \mathbf{x}_{n+1} &\in \operatorname{argmin}_{\mathbf{x}\in\mathbb{X}} \mathbb{E}_{n,x}[\mathcal{H}_{n+1}^{\text{uncert}}(\mathbf{x})] \\ &:= \operatorname{argmin}_{\mathbf{x}\in\mathbb{X}} \mathcal{J}_n(\mathbf{x}), \end{aligned} \tag{3.2}$$

where $\mathbb{E}_{n,\mathbf{x}}$ denotes the expectation with respect to $Y_{\mathbf{x}}^n | Z_{\mathcal{X}_n \times \mathcal{U}_n} = \mathbf{Z}_n$ (for detailed formula and estimation of $\mathcal{J}_n(.)$ see [CG13]).
It remains now to enrich the design with a new point in the uncertain space.

3.2. Minimization of the variance: choice of next u. The process Y^n approximates the expectation $\mathbb{E}_{\mathbf{V}}[f(\cdot, \mathbf{V})]$. It can be seen as a projection of Z^n from the joint space onto the control space. We propose to sample the point $\mathbf{u}_{n+1}$ in the uncertain space in order to reduce at most the one-step-ahead variance at point $\mathbf{x}_{n+1}$, $\mathbb{VAR}(Y_{\mathbf{x}_{n+1}}^{n+1})$, whose expression is obtained from Eq.(2.9). More precisely,

$$\mathbf{u}_{n+1} = \operatorname{argmin}_{\tilde{\mathbf{u}}\in\mathbb{R}^m} \mathbb{VAR}(Y_{\mathbf{x}_{n+1}}^{n+1}), \tag{3.3}$$

with

$$\begin{aligned} \mathbb{VAR}(Y_{\mathbf{x}_{n+1}}^{n+1}) &= \vartheta(\tilde{\mathbf{u}}) \\ &= \int_{\mathbb{R}^m}\int_{\mathbb{R}^m} k_Z((\mathbf{x}_{n+1}, \mathbf{u}); (\mathbf{x}_{n+1}, \mathbf{u}'))d\rho(\mathbf{u})d\rho(\mathbf{u}') \\ &- \int_{\mathbb{R}^m}\int_{\mathbb{R}^m} k_Z((\mathbf{x}_{n+1}, \mathbf{u}); \mathcal{X}_{n+1} \times \mathcal{U}_{n+1}) \\ &k_Z(\mathcal{X}_{n+1} \times \mathcal{U}_{n+1}; \mathcal{X}_{n+1} \times \mathcal{U}_{n+1})^{-1} k_Z(\mathcal{X}_{n+1} \times \mathcal{U}_{n+1}; (\mathbf{x}_{n+1}, \mathbf{u}'))d\rho(\mathbf{u})d\rho(\mathbf{u}'), \end{aligned} \tag{3.4}$$

where $\mathcal{X}_{n+1} = (\mathcal{X}_n, \mathbf{x}_{n+1})$, $\mathcal{U}_{n+1} = (\mathcal{U}_n, \tilde{\mathbf{u}})$ and $\mathcal{X}_n$, $\mathcal{U}_n$ are the sample points in the control space, uncertain space at step n.

3.3. Implementation. The setting of our procedure is driven by our industrial application where the probability distribution of the uncertain variable V is known only through a finite set of realizations $\Xi = \{\breve{\mathbf{v}}_1, \ldots, \breve{\mathbf{v}}_N\}$.

Computational method for functional PCA. We consider the empirical version of $C(s,t)$ defined as $C^N(s,t) = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \breve{\mathbf{v}}_i(s)\breve{\mathbf{v}}_i(t)$. The eigenvalue problem defined by Eq. (2.3) is then solved by discretizing the trajectories $\{\breve{\mathbf{v}}_i\}_{i=1,\ldots,N}$ on $[0,T]$ and replacing C by C^N.
Denoting by $\hat{\psi}_i$, $i = 1, \ldots, m$, the estimated eigenfunctions, we define

$$\mathcal{G}_m = \{\breve{\mathbf{u}}_1, \ldots \breve{\mathbf{u}}_N\} \tag{3.5}$$

with $\breve{\mathbf{u}}_i = (< \breve{\mathbf{v}}_i, \hat{\psi}_1 >, \ldots, < \breve{\mathbf{v}}_i, \hat{\psi}_m >)^T$.

Minimization of the one-step-ahead variance. Since $\mathbf{V}$ is known through a finite set Ξ, Eq. (3.3) is solved on the finite set $\mathcal{G}_m$.

We now detail the implementation of our methodology. Let us first state the global algorithm and then comment some of its steps.

Algorithm 3.1 Stochastic inversion via joint space modelling

Require: The truncation argument m and the DoE of n points $\mathcal{X}_n \times \mathcal{U}_n$ in $(\mathbb{X}, \mathcal{G}_m)$

Set $n = n_0$.
Calculate $\mathbf{Z}$ the simulator responses at the design points $\mathcal{X}_n \times \mathcal{U}_n$
while $n \leq$ budget **do**
  Fit the GP model Z^n
  Induce the integrated GP $Y^n_{\mathbf{x}}$
  $\mathbf{x}_{n+1} \leftarrow$ sampling criterion $\mathcal{J}_n$
  $\mathbf{u}_{n+1} \leftarrow \operatorname{argmin}_{\tilde{\mathbf{u}} \in \mathcal{G}} \mathbb{VAR}(Y^{n+1}_{\mathbf{x}_{n+1}})$
  Simulation at $(\mathbf{x}_{n+1}, \mathbf{v}_{n+1})$, where $\mathbf{v}_{n+1} \in \Xi$ is the curve corresponding to $\mathbf{u}_{n+1}$
  Update DoE : $\mathcal{X}_{n+1} \times \mathcal{U}_{n+1} = \mathcal{X}_n \times \mathcal{U}_n \cup \{(\mathbf{x}_{n+1}, \mathbf{v}_{n+1})\}$
  Update $\mathbf{Z} = \mathbf{Z} \cup \{f(\mathbf{x}_{n+1}, \mathbf{v}_{n+1})\}$
  Set $n = n + 1$
end while
Fit the final GP model Z^n
Approximate Γ^* by the Vorob'ev expectation

step 1 Let $\mathbb{U}$ be the smallest m-rectangle containing $\mathcal{G}_m$, $\mathbb{U} = \prod_{i=1}^m [\min(< \Xi, \hat{\psi}_i >), \max(< \Xi, \hat{\psi}_i >)])$. For the initial DOE, we first build a Latin Hypercube Design of n points $\mathcal{X}_n \times \bar{\mathcal{U}}_n$ in the joint space $(\mathbb{X}, \mathbb{U})$. Then the set of points $\mathcal{U}_n$ is determined such that for $i = 1, ..., n$, $\mathbf{u}_i \in \mathcal{G}_m$ is the closest point from $\bar{\mathbf{u}}_i \in \bar{\mathcal{U}}_n$ (with respect to the euclidean norm in $\mathbb{R}^m$).

step 4 The covariance kernel of the GP is chosen as a sum of two terms: a Matèrn-5/2 covariance and a constant variance term modelling a homoscedastic noise. The homoscedastic modelling of the noise is discussed in Section 4.3. The mean function of the GP is modelled by a constant function. All types of parameters (mean, correlation lengths, variance and noise) are estimated by maximum likelihood [RGD12].

step 5 In the framework where the uncertain vector $\mathbf{U}$ is Gaussian as well as the covariance kernel, closed form solutions of the integrals in (2.8) and (2.9) are given in [JLR13]. In our framework, the integrals in (2.8) and (2.9) are approximated by Monte Carlo.

step 6 $\mathbf{x}_{n+1}$ is obtained by solving (3.2) with a continuous global optimization algorithm: GENetic Optimization Using Derivatives (GENOUD) [JS11].

step 7 Once more the integrals in (3.4) are approximated by Monte Carlo. More details on the estimation of (3.4) can be found in [JLR13]. Here the minimization problem is solved by an exhaustive search on the finite set $\mathcal{G}_m$ defined in (3.5).

step 8 The simulator is evaluated at point $(\mathbf{x}_{n+1}, \mathbf{v}_{n+1})$ where $\mathbf{v}_{n+1}$ is the curve of the initial set of curves Ξ corresponding to the truncated vector of coefficients $\mathbf{u}_{n+1}$.

Remark that Algorithm 3.1 depends on a prior choice of the truncation argument m. To overcome this, we propose another variant of this strategy. The approach consists in augmenting the uncertain space once convergence is established. More precisely, we start with a Gaussian Process defined in the $p + m$ dimensional space. Once the enrichment strategy (given by Algorithm 3.1) no longer provides information, the dimension of the uncertain space is increased and the GP is updated in the $p + m + 1$ dimensional space. It is important to underline that this approach does not require additional calls to the numerical simulator. This second strategy is summarized by Algorithm 3.2:

Algorithm 3.2 Stochastic inversion via sequential joint space modelling

Require: The initial truncation argument $m = 2$ and the DoE of n points $\mathcal{X}_n \times \mathcal{U}_n$ in $(\mathbb{X}, \mathcal{G}_m)$

1: Set $n = n_0$.
2: Calculate $\mathbf{Z}$ the simulator responses at the design points $\mathcal{X}_n \times \mathcal{U}_n$
3: **while** $n \leq$ budget **do**
4: $m \leftarrow$ Update.Dimension()
5: Fit the GP model Z^n
6: Induce the integrated GP $Y^n_{\mathbf{x}}$
7: $\mathbf{x}_{n+1} \leftarrow$ sampling criterion $\mathcal{J}_n$
8: $\mathbf{u}_{n+1} \leftarrow \texttt{arg}\min_{\tilde{\mathbf{u}} \in \mathcal{G}} \mathbb{VAR}(Y^{n+1}_{\mathbf{x}_{n+1}})$
9: Simulator response at $(\mathbf{x}_{n+1}, \mathbf{v}_{n+1})$, where $\mathbf{v}_{n+1} \in \Xi$ is the curve corresponding to $\mathbf{u}_{n+1}$
10: Update DoE : $\mathcal{X}_{n+1} \times \mathcal{U}_{n+1} = \mathcal{X}_n \times \mathcal{U}_n \cup \{(\mathbf{x}_{n+1}, \mathbf{v}_{n+1})\}$
11: Update $\mathbf{Z} = \mathbf{Z} \cup \{f(\mathbf{x}_{n+1}, \mathbf{v}_{n+1})\}$
12: Set $n = n + 1$
13: **end while**
14: Fit the GP model Z^n
15: Approximate Γ^* by the Vorob'ev expectation

In **step** 4 of Algorithm 3.2, the uncertain space dimension is updated based on a stagnation criterion of the Vorob'ev Deviation (see Eq.(26) in [EAHL+20]). If the criterion is verified then one dimension is added and thus $m = m + 1$.

4. Numerical experiments.

4.1. Two analytical examples - set-up. To illustrate the behaviour of the proposed algorithm 3.1, we consider two analytical examples. We suppose that a sample Ξ of $N = 200$ realizations of the functional random variable $\mathbf{V}$ is available and its probability distribution is unknown. To highlight the robustness of our method regarding the random distribution of the uncertainties, we consider two types of functional random variables: Brownian motion and max-stable process. As Algorithm 3.1 depends on the truncation argument m, different values are tested (see Table.4.1) to better understand the effect of the uncertain space dimension.

m	2	4	8
$\mathbf{V}$: Brownian motion	90.1 %	95.2 %	97.6%
$\mathbf{V}$: Max-stable process	58.8 %	63.3 %	70%

Table 1

The explained variance of the functional data by the reduced variables in function of m for two types of uncertainties.

For the next two analytical examples, we consider a Gaussian Process prior $Z_{(\mathbf{x},\mathbf{u})}$ with constant mean and Matèrn covariance kernel with $\nu = 5/2$. Random Latin Hypercube Designs (RLHD) are used as initial DoE for the two algorithms. The number of points of the initial DoE is 20 for the first analytical example and 30 for the second one. The RLHD induces variability in the behaviour of the algorithms. To account for this variability in the tests, the performance of each method is averaged over 30 (respectively 10) independent runs for Brownian motion (respectively max-stable process).

Analytical example 1. We consider an additive function, sum of the two-dimensional Bo-

hachevsky function and a random term, defined as

$$f : (\mathbf{x}, \mathbf{V}) \mapsto \left(x_1^2 + 2x_2^2 - 0.3\cos(3\pi x_1) - 0.4\cos(4\pi x_2) + 0.7\right) + \int_0^T e^{\mathbf{V}_t} \mathrm{dt},$$

where $\mathbf{x} \in \mathbb{X} = [-100, 100]^2$. The objective is to construct the sets $\Gamma^* = \{\mathbf{x} \in \mathbb{X}\ ,\ g(\mathbf{x}) = \mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})] \leq 3500\}$ for the two different types of distribution of the random functional variable (Brownian motion and max-stable process).

Analytical example 2. For the second example we define a function that is not separable with respect to the control variables and uncertainties. The function involves the maximum and the minimum of the function $\mathbf{v}$, so catching the whole variability of $\mathbf{V}$ becomes important. The function f is given by

$$f : (\mathbf{x}, \mathbf{V}) \mapsto \max_t \mathbf{V}_t |0.1\cos(x_1 \max_t \mathbf{V}_t)\sin(x_2)(x_1 + x_2 \min_t \mathbf{V}_t)^2| \int_0^T (30 + \mathbf{V}_t)^{\frac{x_1 x_2}{20}} \mathrm{dt},$$

where the control variables lie in $\mathbb{X} = [1.5, 5] \times [3.5, 5]$. The objective is to construct the sets $\Gamma^* = \{\mathbf{x} \in \mathbb{X}\ ,\ g(\mathbf{x}) = \mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})] \leq c\}$, where $c = 1.2$ and $c = 0.9$ for the Brownian motion and the max-stable Process respectively.

To compare the performance of both algorithms, we use the ratio of the volume of the symmetric difference between the true set Γ^* and the estimated set Q_{n,α^*}: $\mu(\Gamma^* \triangle Q_{n,\alpha_n^*})/\mu(\Gamma^*)$.

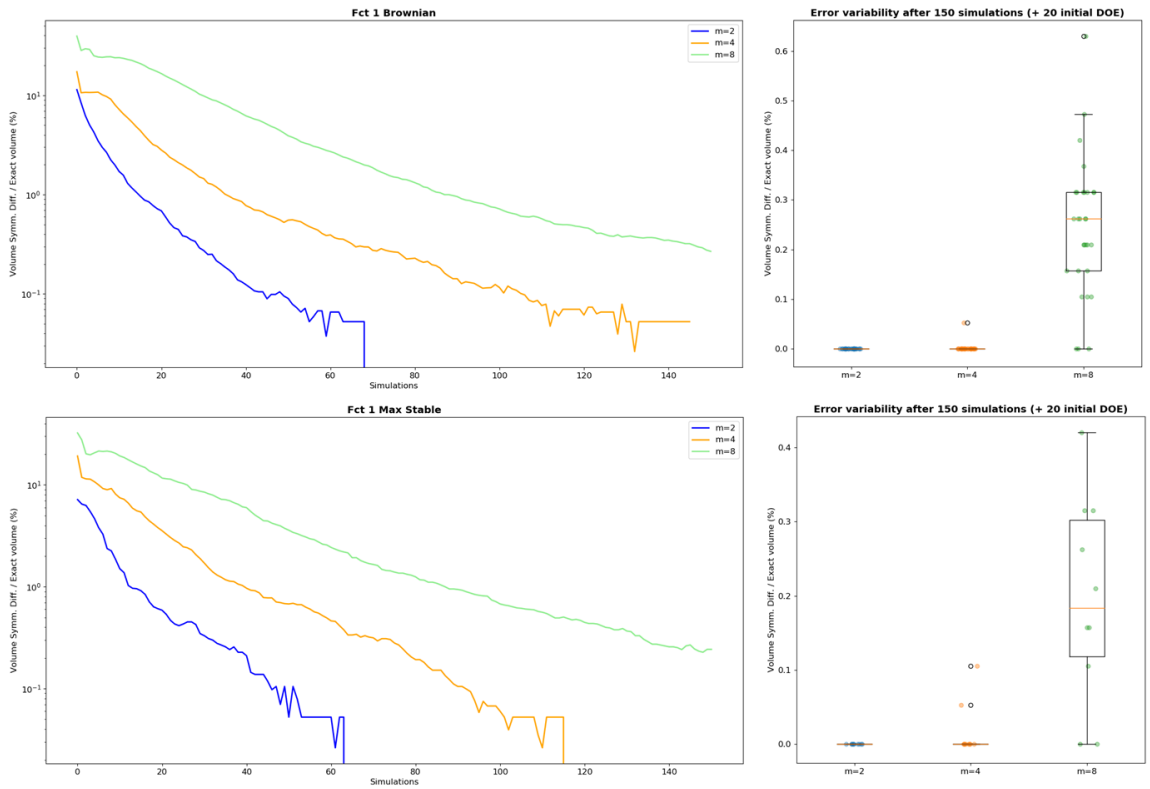

Figure 1. *Analytical example 1 with Brownian motion (top) and with max-stable process (bottom). Convergence of Algorithm 3.1 for $m = \{2, 4, 8\}$. Left: mean of the symmetric differences vs. number of simulator calls in log scale. The mean is taken over the independent runs of initial RLHD. Right: symmetric differences associated with the random initial DOEs at the maximal simulation budget.*

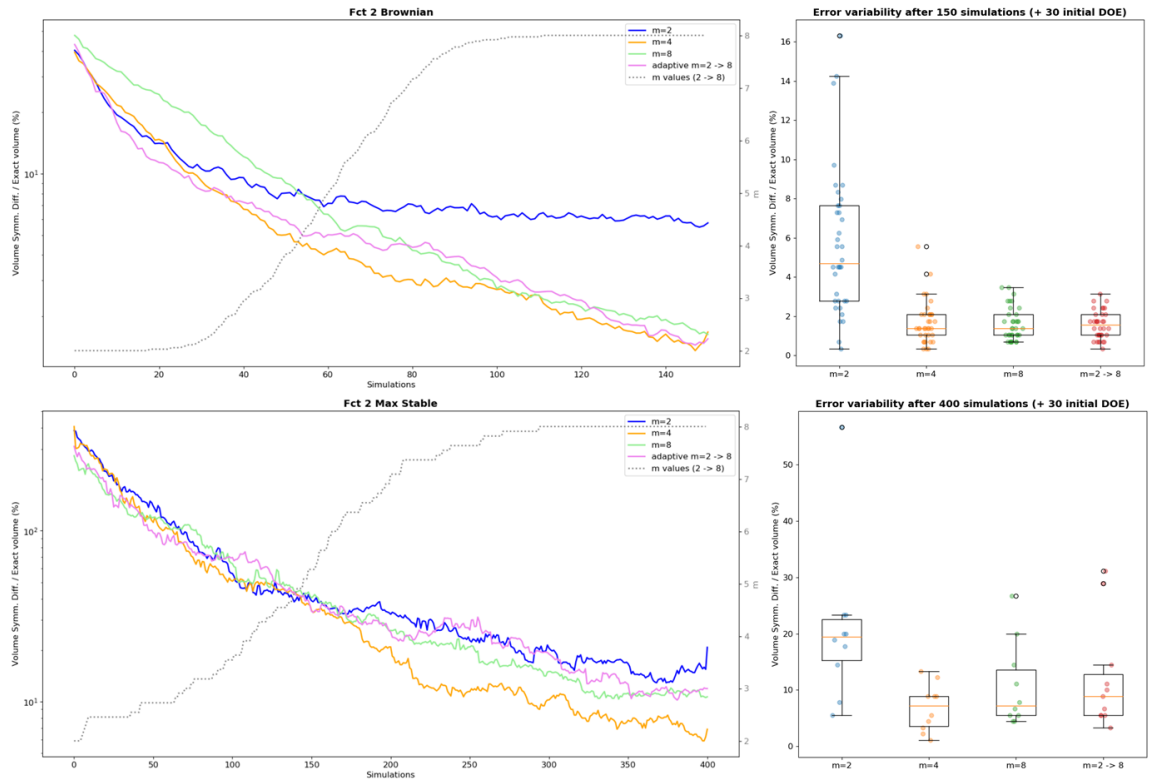

Figure 2. *Analytical example 2 with Brownian motion (top) and with max-stable process (bottom). Convergence of Algorithm 3.1 for $m = \{2, 4, 8\}$ and for adaptive choice of m value. Left: mean of the symmetric differences vs. number of simulator calls in log scale. The dashed grey curve is the mean of m values in the case of an adaptive choice of its value. The mean is taken over the independent runs of initial RLHD. Right: symmetric differences associated with the random initial DOEs at the maximal simulation budget.*

4.2. Two analytical examples - results. In Figures 1 and 2, we show the averaged convergence rates of Algorithm 3.1 on the two analytical examples with the two types of functional uncertainties (Brownian and max-stable processes). The average is taken over the repeated runs of the complete approach corresponding to the 30 random initial designs, and for 3 values of the truncation argument m.

For the first analytical example, the smaller values of m, the faster the convergence. This observation can be explained by the fact that, in higher dimensional joined space (due to larger values of m), much more evaluation points are necessary to learn an accurate GP model (more hyper parameters to determine). It is worth noting that even for 90% (for Brownian motion) or 58.8% (for max-stable process) of explained variance with $m = 2$ the proposed algorithm provides an efficient estimate of the true set Γ^*. Indeed, on stage 8 in Algorithm 3.1 the curve $\mathbf{v}_{n+1} \in \Xi$ associated to $\mathbf{u}_{n+1}$ is recovered, such that the information lost after the dimension reduction is reduced, thereby further robustifying the method.

For the second analytical example, the output depends on local behaviours of the stochastic process. The truncation argument $m = 2$ is too small to catch these dependencies, the function is sensitive to higher KL order. For the Brownian motion, more than 95% of variance is explained with $m = 4$. It seems sufficient to obtain an accurate approximation of Γ^*. The improvement between $m = 2$ and $m = 4$ is noticeable. The improvement is not as important when the uncertainties are driven by a max-stable process since the percentage of explained variance increases slowly. Better results should be observed with $m = 8$. It is not the case because a higher dimension leads to difficulties in the estimation of the GP except by increasing consequently the number of observation points.

In Figure 3 we can see the evolution of the feasible domain estimation with respect to the iterations of Algorithm 3.1 for the second analytic case and the Brownian motion, and for different truncation levels. From left to right we observe the increase of additional sampling points near the boundary with the iteration number.

As shown in Figure 4, the larger the dimension of the problem is, the larger the computational cost is. Moreover, the computational time needed to provide the next evaluation point increases with the number of simulator calls, and thus with the number of iterations, because of the cost of Kriging approximation directly linked with the learning sample size. For example at iteration 150, the run with $m = 8$ requires 275 seconds to perform the optimization and provide the next evaluation point whereas the one for iteration 80 requires 203 seconds. For $m = 2$, the computational time for iteration 150 is 164 seconds and 126 seconds for iteration 80.

From this observation, we propose to evaluate the strategy based on an adaptive choice of m presented in Algorithm 3.2: we start with a small value $m = 2$, and increase this number when the variation of the Vorob'ev deviation remains smaller than a given threshold ϵ (0.005) during l_0 consecutive iterations ($l_0 = 4$) (see Eq. (26) in [EAHL+20]). This adaptive strategy allows to increase the dimension of the KL reduced space only when it is necessary to obtain a better accuracy. It allows to save simulations and reduce computational time, as illustrated on the second analytical example with the Brownian motion on Figure 2 (top). The accuracy reached with this strategy is similar to the one obtained with the strategy with fixed $m = 8$ but with a gain of $\approx 12\%$ in terms of computational time (Figure 4). Only the last iterations are performed with $m = 8$ and the first 30 iterations are performed with $m = 2$ (see the dashed grey curve on Figure 2 (top left)).

4.3. Discussion on the GP model on the finite-dimensional truncated space. We discuss here the assumption stated in Section 2.2 that $f(\mathbf{x}, \mathbf{v})$ is a realization of a Gaussian Process $Z_{(\mathbf{x},\mathbf{u})}$ defined on the truncated space $\mathbb{X} \times \mathbb{R}^m$. It is worth underlying here that our aim was to reduce the simulation cost by considering a m-truncation of the KL expansion while accounting for our partial knowledge on the distribution of $\mathbf{V}$ through only a finite sample of realizations.

Let us consider two truncation arguments m and $L > m$, with L large enough to ensure that the part of variance explained by the KL terms indexed by $i > L$ is negligible.

For a given realization $\mathbf{v}$ of $\mathbf{V}$, let us introduce the notation $(\mathbf{u}, \tilde{\mathbf{u}}) \in \mathbb{R}^m \times \mathbb{R}^{L-m}$ where $\mathbf{u} = (< \mathbf{v}, \hat{\psi}_1 >, \ldots, < \mathbf{v}, \hat{\psi}_m >)^\top$ and $\tilde{\mathbf{u}} = (< \mathbf{v}, \hat{\psi}_{m+1} >, \ldots, < \mathbf{v}, \hat{\psi}_L >)^\top$.

In that setting $f(\mathbf{x}, \mathbf{V})$ can be expressed as

$$f(\mathbf{x}, \mathbf{V}) = f(\mathbf{x}, \hat{\mathbf{V}}_L) + \epsilon_T = f\big(\mathbf{x}, (\mathbf{U}, \tilde{\mathbf{U}})\hat{\mathbf{\Phi}}_\mathbf{L}\big) + \epsilon_T$$

where $\hat{\mathbf{V}}_L$ is the empirical version (estimated from C^N) of the KL approximation of $\mathbf{V}$ given by (2.4), $\hat{\mathbf{\Phi}}_L = (\hat{\psi}_1, ..., \hat{\psi}_L)^T$ and ϵ_T is the error associated to the KL truncation and empirical approximation, supposed small by construction.

Then, the best L^2-approximation of $f\big(\mathbf{x}, (\mathbf{U}, \tilde{\mathbf{U}})\hat{\mathbf{\Phi}}_L\big)$ by a measurable function of $\mathbf{U}$ only is the conditional expectation $\mathbb{E}_{\tilde{\mathbf{U}}}\big[f\big(\mathbf{x}, (\mathbf{U}, \tilde{\mathbf{U}})\hat{\mathbf{\Phi}}_L\big)|\mathbf{U}\big]$. We thus write:

$$f(\mathbf{x}, \mathbf{V}) = \mathbb{E}_{\tilde{\mathbf{U}}}\big[f\big(\mathbf{x}, (\mathbf{U}, \tilde{\mathbf{U}})\hat{\mathbf{\Phi}}_L\big)|\mathbf{U}\big] + \epsilon_P + \epsilon_T$$

with ϵ_P the L^2-projection error. We can further approximate the conditional expectation by

$$f\big(\mathbf{x}, (\mathbf{U}, \tilde{\mathbf{u}}(\mathbf{U}))\hat{\mathbf{\Phi}}_L\big) + \epsilon_E$$

where $\tilde{\mathbf{u}}(\mathbf{U})$ is one realization of $\tilde{\mathbf{U}}|\mathbf{U}$ and ϵ_E accounts for the expectation approximation. The latter approximation is motivated by the fact that, since $\mathbf{V}$ is only known through a finite sample, we only have access to one $\tilde{\mathbf{u}}(\mathbf{u})$ realization for each $\mathbf{u}$ corresponding to the sample $\mathbf{v}$ in the finite set Ξ. Thus we can write:

$$f(\mathbf{x}, \mathbf{V}) = f\big(\mathbf{x}, (\mathbf{U}, \tilde{\mathbf{u}}(\mathbf{U}))\hat{\mathbf{\Phi}}_L\big) + \epsilon \tag{4.1}$$

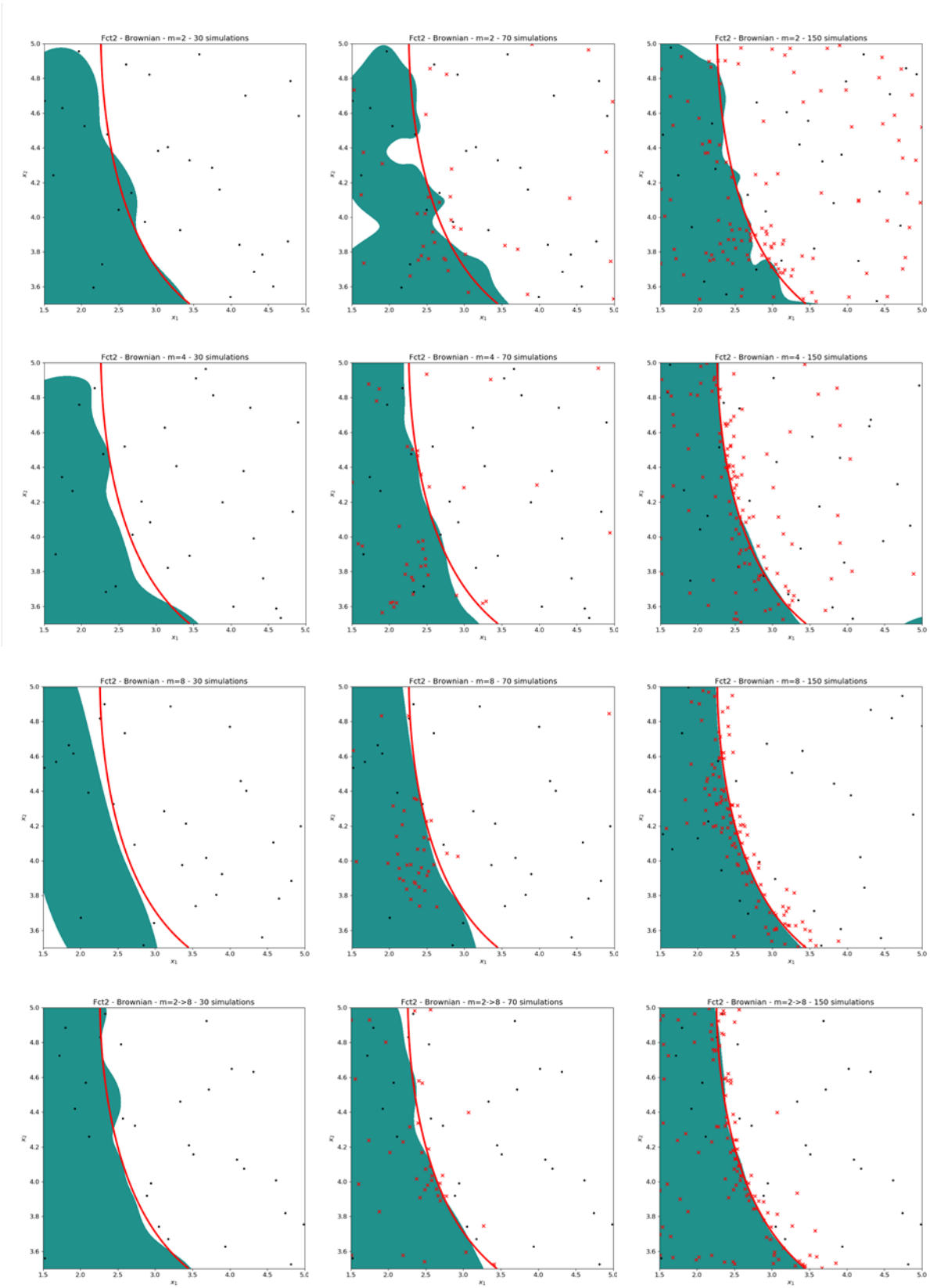

Figure 3. *Feasible domain estimation for function 2 in green and its boundary in red for 3 different iterations (30, 70 and 150 from left to right) and for the 3 values of $m = 2$, 4 and 8 and the adaptive choice of m value (from top to bottom). The black dots are the $\mathbf{x}$ coordinates of the points in the initial design of experiments, the red crosses are the additional points chosen by the algorithm.*

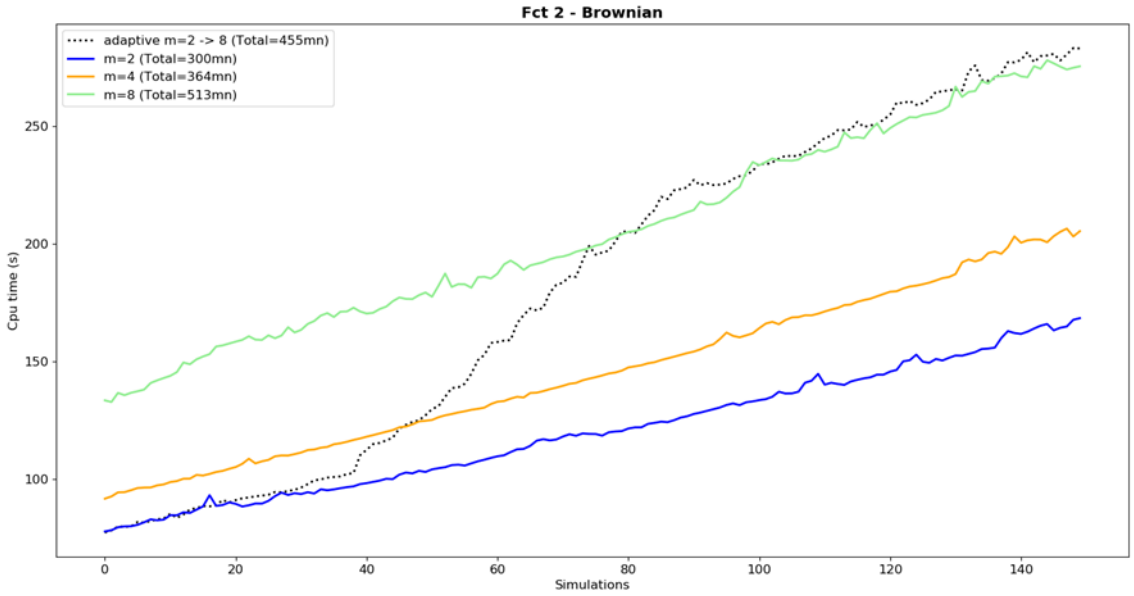

Figure 4. *The computational time (sec.) needed to provide the next evaluation point as a function of iterations for the second analytic example with Brownian motion. The values are average computational times for 5 runs of each strategy: $m = 2, 4, 8$ and adaptive choice of m value.*

with $\epsilon = \epsilon_T + \epsilon_P + \epsilon_E$. According to this last equation, the modelling assumption in Section 2.2 should include a noisy term. However, the estimation of this heteroscedastic noise comes with an extra estimation cost and as it can be seen in Figure 5, no significant model improvement is observed.

Indeed in Figure 5, for $m = 2$, we present the evolution of the symmetric difference for the noisy GP model $Z_{(\mathbf{x},\mathbf{u})}$ introduced from equation (4.1) when the noise ϵ is Gaussian and heteroscedastic with a variance function of $(\mathbf{x}, \mathbf{u})$:

$$\tau^2(\mathbf{x}, \mathbf{u}) = Var_{\tilde{U}}\left[f\left(\mathbf{x}, \left(\mathbf{u}, \tilde{\mathbf{U}}(\mathbf{u})\right)\hat{\mathbf{\Phi}}_L\right)\middle|\mathbf{U} = \mathbf{u}\right].$$

Moreover, supposing $\mathbf{V}$ Gaussian or "nearly Gaussian" then $\tilde{\mathbf{U}}$ can be considered independent of $\mathbf{U}$ and $\tau^2(\mathbf{x}, \mathbf{u})$ can be estimated by

$$\hat{\tau}^2(\mathbf{x}, \mathbf{u}) = \sum_{k=1}^{l} w_k \left[f\left(\mathbf{x}, \mathbf{V}_k^{Quant}\right) - \sum_{j=1}^{l} w_j f\left(\mathbf{x}, \mathbf{V}_j^{Quant}\right)\right]^2$$

where $l = 5$ and the $\mathbf{V}_k^{Quant}$ are greedy functional quantizers and w_k associated weights (see [EAHL+20] for more details). These quantizers are built from a set of N curves $\{\left(\mathbf{u}, \tilde{\mathbf{u}}_k\right)\hat{\mathbf{\Phi}}_L, k = 1, ..., N\}$ where $\tilde{\mathbf{u}}_k$ are independent samples of $\tilde{\mathbf{U}}$ which in practice are uniformly sampled in the finite set $\bar{\mathcal{G}} = \{\bar{\mathbf{u}}_1, ..., \bar{\mathbf{u}}_N\}$ where $\bar{\mathbf{u}}_i = (< \breve{v}_i, \hat{\psi}_{m+1} >, ..., < \breve{v}_i, \hat{\psi}_L >)$. Numerically we select 20 $(\mathbf{x}, \mathbf{u})$-points from the initial DoE set of size $n = 30$ and estimate the corresponding $\hat{\tau}^2$. To avoid further estimation of τ^2 at new locations (the remaining DoE points and during the infill strategy), we also build a second GP model of $\log(\hat{\tau}^2)$ based on the 20 initial estimations. Finally the noisy GP model Z is built using as noise variance $\exp\left(\hat{\log}(\hat{\tau}^2)\right)$. Overall we need additional $l \times 20 = 100$ costly evaluations of f to estimate the heteroscedastic noise.

In Figure 5 we notice that compared to the noiseless model with $m = 2$, the noisy model achieves a faster symmetric difference volume reduction but the overcost, for the variance estimation, makes this approach interesting only for a large simulation budget: at least 130 simulations. For the Brownian case, on function 2, the noiseless models with higher m still perform better for a budget

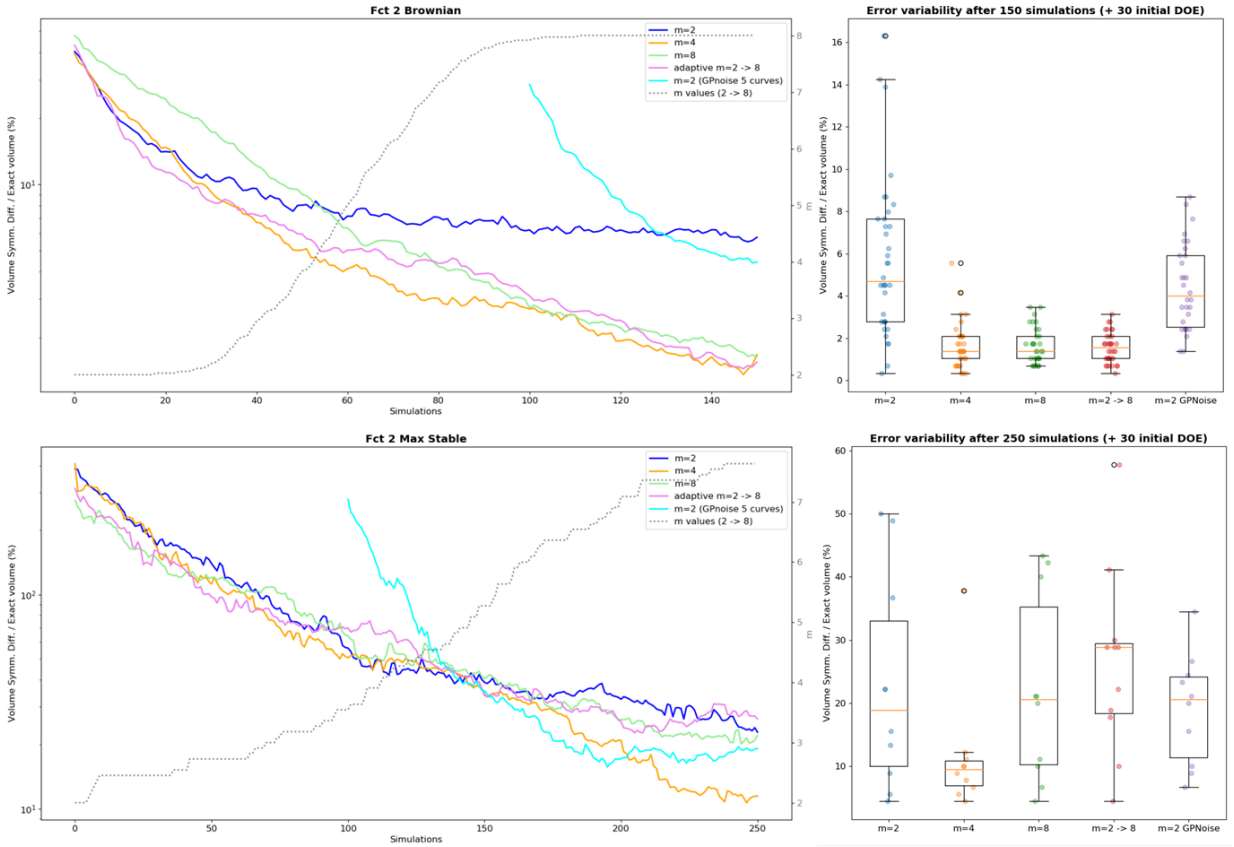

Figure 5. *Function 2 with Brownian (top) and max-stable processes (bottom) with a comparison with the heteroscedastic GP model. Convergence of Algorithm 3.1 for* $m = \{2, 4, 8\}$*, adaptive choice of* m *value. Left: mean of the symmetric difference vs. number of simulator calls. The mean is taken over the independent runs of initial RLHD. The additional curve (cyan) corresponds to* $m = 2$ *with the heteroscedastic model, it is translated to take into account the extra-cost of* 100 *simulations for the noise estimation. Right: symmetric differences associated with the random initial DOEs at the maximal simulation budget.*

up to 150 than the noisy one. A model with a small m, that is to say with a rough truncation error, involves a larger bias. Nevertheless, refining the heteroscedastic noise estimation should bring the method to a similar level but much further on the axis corresponding to the number of simulations. But on function 2 with a Max-stable process, the noisy model slightly outperforms the noiseless models ($m = 2, 4, 8$) when approaching the 150 simulations (Figure 5). We can understand this improvement by the fact that even with higher m a noiseless model does not make up for a wilder truncation error which is better approximated by a noisy model.

Note that it is possible to relax the Gaussian hypothesis on $\mathbf{V}$. In that case the same kind of heteroscedastic variance estimator could be used but would require an empirical estimation of the conditional distribution of $\tilde{\mathbf{U}}|\mathbf{U}$ which seems difficult in the context of our partial knowledge of $\mathbf{V}$ imposing on us to work on a finite predefined set $\mathcal{G} \cup \bar{\mathcal{G}}$.

4.4. Application to a pollution control system SCR. In this section we test the proposed method on an automotive test case from IFPEN. The problem concerns an after-treatment device of diesel vehicles, called Selective Catalytic Reduction (SCR). This latter consists of a basic process of chemical reduction of nitrogen oxides (NOx) to diatomic nitrogen (N2) and water (H2O) by the reaction of NOx and ammonia NH3. The reaction itself occurs in the SCR catalyst. Ammonia is provided by a liquid-reductant agent injected upstream of the SCR catalyst. The amount of ammonia introduced into the reactor is a critical quantity: overdosing causes undesirable ammonia slip

downstream of the catalyst, whereas under-dosing causes insufficient NOx reduction. In practice, ammonia slip is restricted to a prescribed threshold. We use an emission-oriented simulator developed by IFPEN, which models the vehicle, its engine and the exhaust after-treatment system. This latter takes as input the vehicle driving cycle profile and provides the time-series of corresponding exhaust emissions as output. A realistic SCR control law is used in this simulator. See [BCLP12] for an example of such a control law. In this study, the inputs are two control variables and a functional one considered as random. The control variables are parameters of the SCR control law. They set the targeted level of NH3 storage in the catalyst and then are indirectly related to the NH3 injected. They lie in $\mathbb{X} = [0, 0.6]^2$. The functional random variable describes the evolution of vehicle speed on $I = [0, 5400s]$ and is known through an available sample of 100 real driving cycles. Two samples are represented in Figure 6. In short, the ammonia emissions peak during a driving

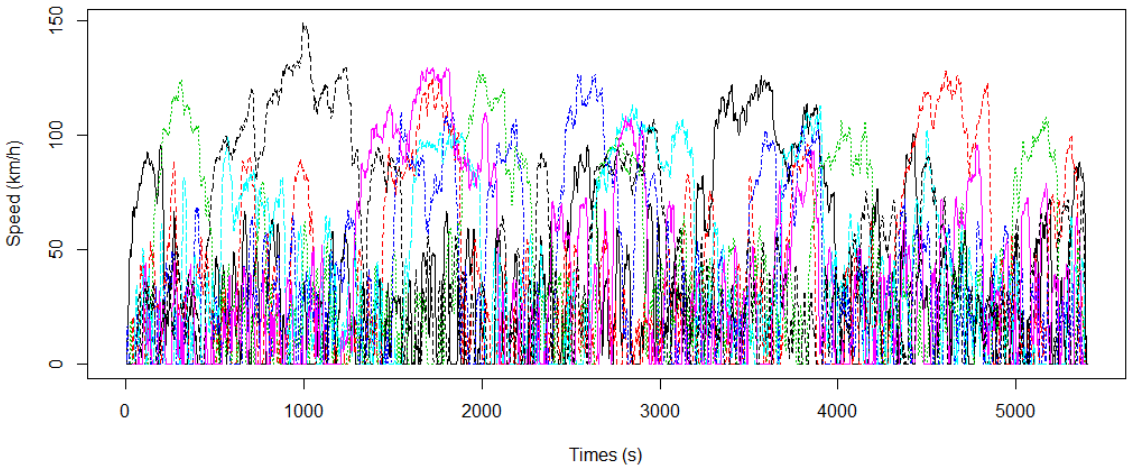

Figure 6. *Seven real-driving cycles extracted from the available sample of* 100 *cycles.*

cycle is modelled as a function

$$f : \begin{cases} \mathbb{X} \times \mathcal{V} & \to \mathbb{R} \\ (\mathbf{x}, \mathbf{V}) & \mapsto f(\mathbf{x}, \mathbf{V}) = \max_{t \in I} NH_3^{slip}(t) \end{cases} \tag{4.2}$$

We are interested in recovering the set $\Gamma^* = \{\mathbf{x} \in \mathbb{X}, g(\mathbf{x}) = \mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})] \leq c\}$, with $c = 30ppm$. Conducting this study on a full grid would consist on covering the space $[0, 0.6]^2$ with a fine mesh and evaluating the code 100 times at each point. Knowing that each simulation takes about two minutes, such study would require many hours of computational time, and thus using meta-models allows to tackle this computational issue.

As discussed in the previous subsection, we start by reducing the space dimension of the uncertain variable as described in Section 2.1 and fix the truncation argument to $m = 20$ in order to explain 80% of the variance. Thereafter, we consider a Gaussian Process prior $Z_{(\mathbf{x},\mathbf{u})}$, with constant mean function and Matérn covariance kernel with $\nu = 5/2$. The initial DoE consists of a $n = 5 \times (2 + 20) = 110$ points LHS design optimized with respect to the maximin criterion. The covariance kernel hyper-parameters are estimated by maximizing the likelihood. As for the analytical example, we proceed to add one point at each iteration of the SUR strategy.

Figure 7 shows the coverage probability function defined by the integrated Gaussian Process $Y_{\mathbf{x}}$ conditionally to the n available observations. The initial estimate of Γ^* is given by the green set of blue boundary. From Figure 8, we note that, for each additional point, the new observed response affects the estimation of the excursion set and its uncertainty. Thus, the Vorob'ev deviation generally decreases in function of the iterations. SUR algorithm heavily visits the boundary region

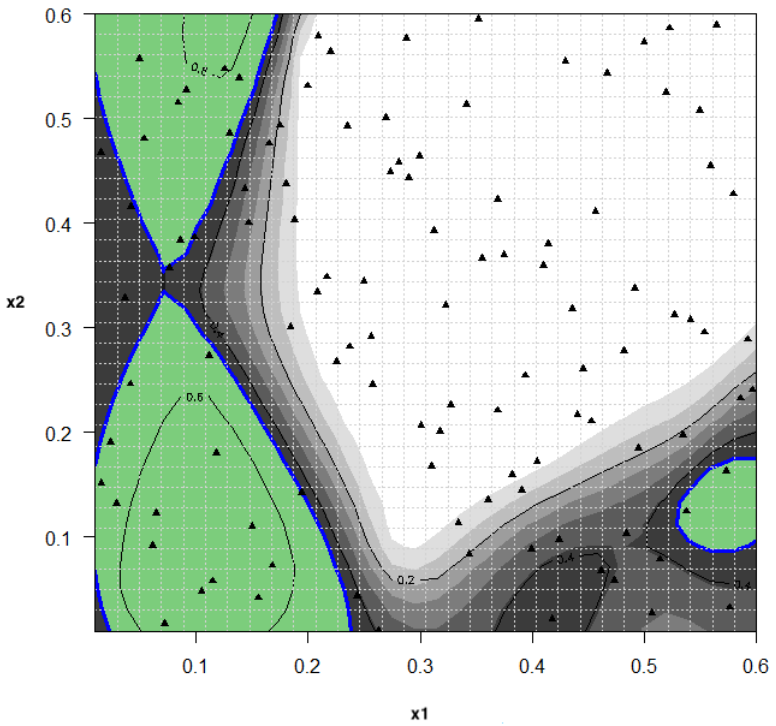

Figure 7. *SCR pollution control system. The initial DoE (black triangles) and the initial estimate set (green). The contour plot in grey represents the excursion probability.*

of Γ^* and explore also other potentially interesting regions. Actually, after 400 iterations (510 evaluations) the whole domain $\mathbb{X}$ has an excursion probability close to either 0 or 1.

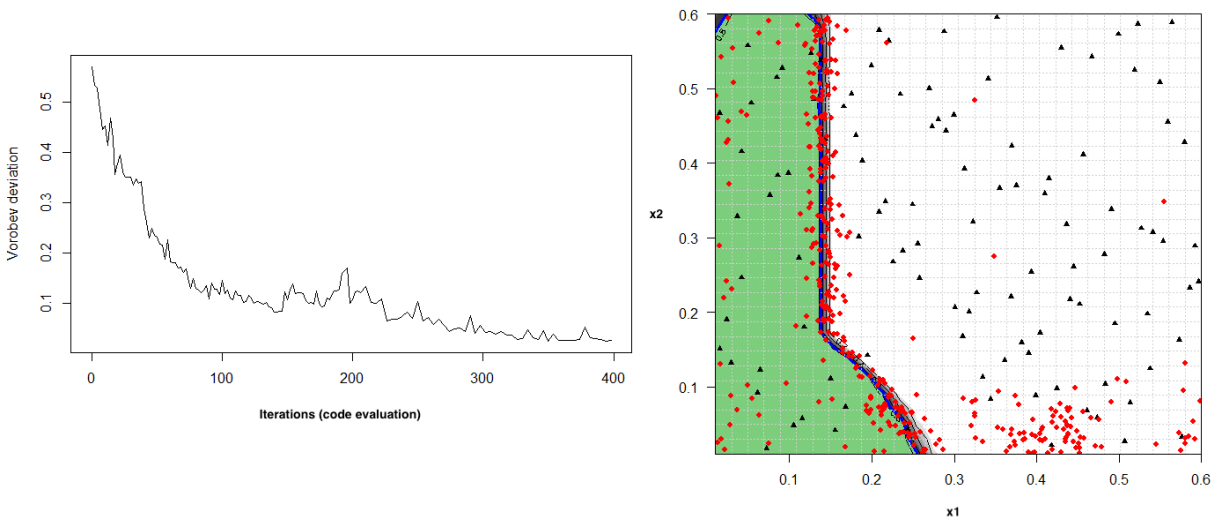

Figure 8. *SCR pollution control system. The Vorob'ev deviation in function of the number of simulations (left). The coverage probability function, the initial DoE (black triangles) and the estimate set (green set) after* 400 *added points (red). The contour plot in grey represents the excursion probability.*

5. Conclusion. The aim of this paper is to propose an excursion set inversion procedure for control system in an uncertain environment. Furthermore, control systems whose behaviour is simulated by high-fidelity and expensive-to-evaluate models are considered. Gaussian Process modelling approaches are therefore introduced as computationally costless approximations of the outputs of

the simulator.

The proposed strategy minimizes the uncertainties on the excursion set of the simulator output by, first, creating a Gaussian Process model in the joint space of deterministic and uncertain input variables. The vector-valued random variables result from a dimension reduction of the functional input variable. Then another "projected" Gaussian Process is built to represent the mean of the quantity of interest (output of the simulator). Enrichment of the design of experiments is performed in the joint space. This allows us to direct the experimental design points toward regions of the space that decrease significantly the uncertainties on the excursion set while limiting the number of simulation cost.

Two bi-dimensional examples based on analytic expressions are considered to validate the proposed procedure. This allows us to validate the proposed method with comparison with exact solutions. The application of the proposed procedure shows increased efficiency as the number of calls to the complex simulator is reduced. Finally, we apply the methodology to an industrial problem related to the pollution control system of an automotive. An excursion set solution is found within a reasonable number of simulations.

The paper focuses on the expectation while other reliability measures may also be of great importance. For example, one may interested in ensuring a certain level of reliability with a high probability or satisfying multiple constraints, e.g., on the mean and the variance.

REFERENCES

[Adl81] Robert J. Adler. *The Geometry of Random Fields*. John Wiley & Sons, Chichester, 1981.

[BCLP12] Anthony Bonfils, Yann Creff, Olivier Lepreux, and Nicolas Petit. Closed-loop control of a scr system using a nox sensor cross-sensitive to nh3. *IFAC Proceedings Volumes*, 45(15):738–743, 2012.

[BGL+12] Julien Bect, David Ginsbourger, Ling Li, Victor Picheny, and Emmanuel Vazquez. Sequential design of computer experiments for the estimation of a probability of failure. *Statistics and Computing*, 22(3):773–793, 2012.

[BL15] David Bolin and Finn Lindgren. Excursion and contour uncertainty regions for latent gaussian models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77(1):85–106, 2015.

[CBG+14] Clément Chevalier, Julien Bect, David Ginsbourger, Emmanuel Vazquez, Victor Picheny, and Yann Richet. Fast parallel kriging-based stepwise uncertainty reduction with application to the identification of an excursion set. *Technometrics*, 56(4):455–465, 2014.

[CG13] Clément Chevalier and David Ginsbourger. Fast computation of the multi-points expected improvement with applications in batch selection. In *International Conference on Learning and Intelligent Optimization*, pages 59–69. Springer, 2013.

[Doo53] Joseph Leo Doob. *Stochastic processes*. New York: John Wiley & Sons, 1953.

[EAHL+20] Mohamed Reda El Amri, Céline Helbert, Olivier Lepreux, Miguel Munoz Zuniga, Clémentine Prieur, and Delphine Sinoquet. Data-driven stochastic inversion via functional quantization. *Statistics and Computing*, 30(3):525–541, 2020.

[FS+13] Joshua P French, Stephan R Sain, et al. Spatio-temporal exceedance locations and confidence regions. *The Annals of Applied Statistics*, 7(3):1421–1449, 2013.

[JLR13] Janis Janusevskis and Rodolphe Le Riche. Simultaneous kriging-based estimation and optimization of mean response. *Journal of Global Optimization*, 55(2):313–336, 2013.

[JS11] Walter Mebane Jr. and Jasjeet Sekhon. Genetic optimization using derivatives: The rgenoud package for r. *Journal of Statistical Software, Articles*, 42(11):1–26, 2011.

[LK10] Olivier Le Maître and Omar M. Knio. *Spectral Methods for Uncertainty Quantification*. Scientific Computation. Springer, Dordrecht, 2010.

[Mol06] Ilya Molchanov. *Theory of random sets*. Springer Science & Business Media, 2006.

[Pac03] Christopher Joseph Paciorek. *Nonstationary Gaussian processes for regression and spatial modelling*. PhD thesis, Department of Statistics, Carnegie Mellon, University Pittsburgh Pennsylvania 15213, 2003.

[RGD12] Olivier Roustant, David Ginsbourger, and Yves Deville. 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, 2012.

[VB09] Emmanuel Vazquez and Julien Bect. A sequential bayesian algorithm to estimate a probability of failure. *IFAC Proceedings Volumes*, 42(10):546–550, 2009.

[VL13] Oleg Yu. Vorobyev and Natalia A. Lukyanova. A mean probability event for a set of events. Mpra

paper, University Library of Munich, Germany, 2013.

[Vor84] O Yu Vorob'ev. Srednemernoje modelirovanie (mean-measure modelling), 1984.

[WSN00] Brian J Williams, Thomas J Santner, and William I Notz. Sequential design of computer experiments to minimize integrated response functions. *Statistica Sinica*, pages 1133–1152, 2000.