



HAL
open science

Feasible set estimation under functional uncertainty by Gaussian Process modelling

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

► **To cite this version:**

Mohamed Reda El Amri, Céline Helbert, Miguel Munoz Zuniga, Clémentine Prieur, Delphine Sinoquet. Feasible set estimation under functional uncertainty by Gaussian Process modelling. 2021. hal-02986558v5

HAL Id: hal-02986558

<https://ifp.hal.science/hal-02986558v5>

Preprint submitted on 20 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Feasible set estimation under functional uncertainty by Gaussian Process modelling

Mohamed Reda El Amri^{a,*}, Céline Helbert^b, Miguel Munoz Zuniga^c, Clémentine Prieur^d,
Delphine Sinoquet^e

^a*IFP Energies Nouvelles, Solaize, France (mohamed-reda.el-amri@ifpen.fr)*

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

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

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

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

Abstract

In this paper we deal with the estimation of a feasible set defined by an inequality constraint on the output of a time-consuming black-box simulator. We focus on the setting where the black-box simulator takes as inputs both a set of scalar controlled variables and a functional uncontrolled variable. We then place ourselves in a probabilistic framework, modelling the functional uncontrolled variable by a random process. The inequality constraint is formulated as the expectation of the output of the simulator conditional on the values taken by the set of controlled variables. We propose an original method to solve the above feasibility problem with a reduced number of evaluations of the costly simulator. A Gaussian Process model of the simulator is learned in the joint space of controlled and uncontrolled input variables, on the basis of a set of simulations which is enriched through a sequential procedure. This procedure aims to reduce the estimation error of the feasible set by evaluating the simulator on new points chosen sequentially in the joint input space according to specific enrichment criteria. It involves as a preliminary step the reduction of the dimension of the uncontrolled input space. A variation of this strategy is also proposed, which increases adaptively the dimension of the reduced space, leading to an improvement in terms of number of calls to the simulator. The procedure we propose is compared with other sampling procedures and another modelling approach on analytical examples. Finally our methodology is implemented on an automotive industrial application. For this application, the feasible set to be recovered is the set of values of controlled variables of a gas after-treatment device leading to the respect of pollutant emission standards of a vehicle under driving profile uncertainties.

Keywords: feasible set estimation; Gaussian Process model; dimension reduction; functional uncertainty.

1. Introduction

In recent years, engineers and scientists are increasingly relying on computer models as surrogates for physical experimentation generally too costly or impossible to execute ([1, 2]). 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

*Corresponding author

Email address: mohamed-reda.el-amri@ifpen.fr (Mohamed Reda El Amri)

6 values leading to a prescribed value or range for the output of interest. The problem of estimating
7 such a set is called hereafter feasible set estimation.

8 In our context, the numerical simulator, denoted f , takes two types of input variables: a set
9 of controlled variables $\mathbf{x} \in \mathbb{X}$, and a set of uncontrolled variables $\mathbf{v} \in \mathcal{V}$. Without considering
10 any assumptions on the set of uncontrolled variables \mathbf{v} , robust feasible set estimation consists in
11 seeking the set of controlled variables $\mathbf{x} \in \mathbb{X}$ such that $\sup_{\mathbf{v} \in \mathcal{V}} f(\mathbf{x}, \mathbf{v})$ is smaller than a threshold
12 c . Then, the difficulty of solving this estimation problem strongly depends on the set \mathcal{V} .

13 In our setting, the simulator f takes as inputs a set of scalar controlled variables and a functional
14 uncontrolled variable. Therefore we place ourselves in a probabilistic framework, modelling the
15 functional uncontrolled variable by a random process \mathbf{V} . We are interested in the estimation of
16 the feasible set defined as $\Gamma^* := \{\mathbf{x} \in \mathbb{X}, g(\mathbf{x}) = \mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})] \leq c\}$, with $c \in \mathbb{R}$. It is important to
17 note here that our study is driven by an industrial application on automotive depollution. More
18 precisely, we study an after-treatment device of diesel vehicles, depending on controlled variables,
19 in an uncertain environment corresponding to the uncontrolled driving profile. Knowledge on the
20 driving profile is provided through a finite set of realizations of moderate size (see Section 4.4 for
21 more details). In order to fit this setting, we make the assumption that the process \mathbf{V} is only
22 known through a finite set, denoted Ξ , of its realizations. Note that to estimate the expectation
23 $\mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})]$ appearing in the definition of Γ^* , a brute force Monte Carlo is out of reach as each
24 evaluation of the simulator is time consuming. Therefore we propose in the following a more
25 elaborate sampling strategy.

26 Feasible set estimation has already been carried out in many applications, notably reliability
27 engineering (see, e.g., [1], [2]), climatology (see, e.g., [3], [4]) and many other fields. One way to
28 tackle the problem is to adopt an active learning strategy based on a surrogate model of the simu-
29 lator output. Active learning strategies based on surrogates are quite common in the framework of
30 uncertainty quantification, to estimate a targeted quantity of interest built on the model output,
31 such as a probability of exceedance, a quantile, an optimum... Then two potentially complementary
32 learning strategies are possible: seek for precise surrogate prediction on the overall input space or
33 focus on sub-spaces important to the targeted quantity of interest. When possible, a compromise
34 between both strategies is often preferred. In the context of Gaussian process modelling, numerous
35 active learning methods have been developed tailored to different post processing of the surrogate
36 such as: meta-modelling [5, 6, 7, 8], optimization [9, 10, 11, 12], excursion/feasible set estimation
37 [2, 13, 14, 15], failure probability estimation [16, 17, 18], expectation estimation [19, 20, 21, 22],
38 each with dedicated criteria to enrich the Design of Experiments (DoE) and to stop the algorithm.
39 As a general rule, all criteria offer a compromise between exploration of the input space and ex-
40 ploitation of learned information important for the quantity of interest. A comparison of several
41 criteria is given in the recent review [23]. We can also notice that active learning is not restricted
42 to GP based surrogates [24] and has also been a hot topic in the broad field of machine learning
43 [25].

44 In the feasible set estimation problem of concern, one option is to implement an active DoE
45 learning strategy based on a Gaussian Process (GP) model for $g : \mathbf{x} \mapsto \mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})]$. The under-
46 lying idea is that Gaussian Process models, which capture prior knowledge about the regularity
47 of the unknown function, make it possible to assess the estimation error of Γ^* given a set of eval-
48 uations of g . More specifically, for the estimation of a feasible set, these sequential strategies are
49 closely related to the field of Bayesian global optimization (see, e.g., [26]). In the case of feasible
50 set estimation, specific Stepwise Uncertainty Reduction (SUR) strategies were introduced in [27].
51 More recently, a parallel implementation of these strategies has been proposed in [2] and applied
52 to the recovery of a feasible set. Briefly, the strategy SUR gives sequentially the next location in
53 the controlled space where to run the simulator in order to minimize a function (called uncertainty
54 function hereafter) measuring the estimation error of the feasible set.

55 In the field of robust optimization where uncertainty comes from a real-valued (or vector-valued)

56 random input, various methods exist and aim at optimizing the expectation taken with respect to
57 the probability distribution of the random input (see [28] or [29]). These methods are based on
58 the modelling of f by a Gaussian Process built in the joint space of controlled and uncontrolled
59 variables. Then a "projected" (integrated) Gaussian Process is defined by taking the expectation
60 with respect to the probability distribution of the random input, leading to an approximation of
61 the expected response g . Finally a sequential design of experiments is proposed for optimizing the
62 objective function g . In the same spirit, we propose an original method to solve a probabilistic
63 feasible set estimation problem with the aim of reducing at most the number of evaluations of the
64 simulator required. In this work f is approximated by a Gaussian Process model built on $\mathbb{X} \times \mathbb{R}^m$,
65 a finite-dimensional approximation of $\mathbb{X} \times \mathcal{V}$. The choice of the truncation argument m will be
66 discussed further. For the iterative approximation of Γ^* , the sampling strategy in the joint space
67 is based on two steps. Firstly a SUR approach is applied to the "projected" Gaussian Process
68 to determine the next evaluation point $\mathbf{x}_{n+1} \in \mathbb{X}$. Secondly, in the uncontrolled space, the next
69 realization v_{n+1} of the random process \mathbf{V} is chosen such that the standard error of the "projected"
70 process evaluated at \mathbf{x}_{n+1} is minimized.

71 As already mentioned, our procedure relies on a preliminary step which aims, for a truncation
72 argument m , at estimating the m -truncated Karhunen-Loève (KL) decomposition of the random
73 process \mathbf{V} from the finite set Ξ of realizations of \mathbf{V} . Building the sampling strategy on a GP model
74 built on the finite-dimensional space $\mathbb{X} \times \mathbb{R}^m$ certainly involves a loss of information. However,
75 we propose to mitigate this loss of information in the following manner: once the next point is
76 selected in $\mathbb{X} \times \mathbb{R}^m$, we evaluate the simulator on its closest neighbor in $\mathbb{X} \times \Xi$, where we recall that
77 Ξ is the available finite sample of realizations of \mathbf{V} (for more details, see the description of **step 8**
78 after the statement of Algorithm 1). With this latter process, we recover partial knowledge of the
79 full variability of the untruncated process \mathbf{V} , leading to a procedure which is robust with respect
80 to the truncation argument m . Also, we propose in Algorithm 2 a variation of our procedure, by
81 increasing adaptively the truncation argument m over the iterates.

82 Another procedure for solving the feasible set estimation problem described in this introduction
83 was proposed in [13]. The main difference with our work is that the sequential enrichment strategy
84 is defined on \mathbb{X} , the space of controlled variables and not on the joint space of controlled and
85 uncontrolled variables. Then, at each iteration, once the new point \mathbf{x}_{n+1} is selected, an accurate
86 approximation of $\mathbb{E}_{\mathbf{V}}[f(\mathbf{x}_{n+1}, \mathbf{V})]$ is computed via quantization. The procedure we propose in the
87 present work outperforms the one in [13] by allowing to reduce even more the number of evaluations
88 of the costly simulator (see Figure 7 Section 4.2).

89 The article is structured as follows. In Section 2, we recall the problem formulation and we
90 extend the concept of Gaussian Process modelling to the case where one of the inputs is a random
91 process known only through a finite set of realizations. In Section 3, we introduce a new sequential
92 sampling strategy targeted for robust feasible set estimation to choose the next point in the joint
93 space: $(\mathbf{x}_{n+1}, \mathbf{v}_{n+1})$ (Sections 3.1 and 3.2). In Section 3.3, together with numerical implementation
94 details, we summarize our strategy to tackle robust feasible set estimation in two algorithms: with
95 fixed KL parameter m (Algorithm 1) and its adaptive counterpart (Algorithm 2). The results of
96 these algorithms on two analytical test cases are presented in Section 4.1 to 4.3. In particular,
97 concerning the sampling enrichment in uncertain space, we compare our sampling strategy, based
98 on the standard deviation of the "projected" process evaluated at \mathbf{x}_{n+1} , with a uniform sampling
99 of v_{n+1} among the finite set of available realizations of the random process \mathbf{V} . We also compare
100 our procedure with the one introduced in [13] which combines the fitting of a Gaussian Process
101 model on the controlled space \mathbb{X} with a quantization estimation of the expectation of the output
102 conditional on the values taken by the controlled variables. In Section 4.4, our new procedure is
103 tested on the industrial application of a car pollution control system. Finally further discussion
104 on the modelling assumptions is postponed to Appendix A.

105 **2. Problem formulation**

106 We model the output of the industrial simulator by a function $f : \mathbb{X} \times \mathcal{V} \rightarrow \mathbb{R}$ with \mathbb{X} a bounded
 107 subset of \mathbb{R}^p being the controlled variable space and \mathcal{V} the functional space in which the random
 108 process \mathbf{V} , modelling uncertainties, takes its values. We are interested in estimating the feasible
 109 set

$$\Gamma^* = \{\mathbf{x} \in \mathbb{X}, g(\mathbf{x}) \leq c\}, \quad (1)$$

110 where $c \in \mathbb{R}$ is a threshold and $g : \mathbb{X} \rightarrow \mathbb{R}$ such that $g(\mathbf{x}) = \mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})]$. An additional constraint
 111 is that the random process \mathbf{V} is known only through a finite set of realizations, denoted by Ξ . The
 112 implication of this constraint will be specified in Section 3.3. The proposed sequential strategy to
 113 estimate Γ^* involves three main ingredients introduced hereafter: dimension reduction to reduce
 114 the random process \mathbf{V} to a m -dimensional random vector, Gaussian Process modelling in the joint
 115 space $\mathbb{X} \times \mathbb{R}^m$ and a wise selection of next point $(\mathbf{x}_{n+1}, \mathbf{u}_{n+1}) \in \mathbb{X} \times \mathbb{R}^m$ at which to evaluate the
 116 simulator. Although the Gaussian Process model is defined on the finite-dimensional truncated
 117 space $\mathbb{X} \times \mathbb{R}^m$, robustness with respect to truncation level m is mitigated through the sequential
 118 enrichment procedure of the design of experiments as for each selected point, the simulator is
 119 evaluated at point $(\mathbf{x}_{n+1}, \mathbf{v}_{n+1})$ where \mathbf{v}_{n+1} is the realization in Ξ corresponding to the truncated
 120 vector \mathbf{u}_{n+1} in a sense to be precised in the following subsections.

121 *2.1. Random process finite dimensional representation*

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

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

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

$$\forall t \in [0, T], \mathbf{V}(t) = \sum_{i=1}^{\infty} U_i \psi_i(t), \quad (2)$$

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

$$\lambda_i \psi_i(t) = \int_0^T C(t, s) \psi_i(s) ds, \quad (3)$$

125 and with $\{U_i\}_{i=1}^{\infty}$ denoting a set of uncorrelated random variables with zero mean and variance λ_i .
 126 Decomposition (2) is known as the Karhunen-Loève (KL) expansion of \mathbf{V} ([30]). In the following
 127 we denote the truncated version of \mathbf{V} as \mathbf{V}_m :

$$\forall t \in [0, T], \mathbf{V}_m(t) = \sum_{i=1}^m U_i \psi_i(t), \quad (4)$$

128 which represents, in the mean square error sense, the optimal m -term approximation of \mathbf{V} ([30]).
 129 The value of the parameter m should be chosen such that the approximation is accurate enough.
 130 Its influence in practice is discussed in Section 4.2.

131 *2.2. Gaussian Process modelling*

132 We assume that f is a smooth function. In the following, we model f as a realization of a
 133 Gaussian Process $Z_{(\mathbf{x}, \mathbf{u})}$ defined on $\mathbb{X} \times \mathbb{R}^m$, where $\mathbf{u} = (\langle \mathbf{v}, \psi_1 \rangle, \dots, \langle \mathbf{v}, \psi_m \rangle)^\top$. Let m_Z be
 134 the mean function of $Z_{(\mathbf{x}, \mathbf{u})}$ and k_Z its covariance function,

$$\begin{aligned} \mathbb{E}[Z_{(\mathbf{x}, \mathbf{u})}] &= m_Z(\mathbf{x}, \mathbf{u}), \\ \text{Cov}(Z_{(\mathbf{x}, \mathbf{u})}, Z_{(\mathbf{x}', \mathbf{u}')} &= k_Z((\mathbf{x}, \mathbf{u}); (\mathbf{x}', \mathbf{u}')). \end{aligned} \quad (5)$$

135 Let us denote Z^n , the GP Z conditioned on the set of n observations $\mathbf{Z}_n =$
 136 $\{f(\mathbf{x}_1, \mathbf{v}_1), \dots, f(\mathbf{x}_n, \mathbf{v}_n)\}$ of Z at $(\mathcal{X}_n, \mathcal{U}_n) = \{(\mathbf{x}_1, \mathbf{u}_1), \dots, (\mathbf{x}_n, \mathbf{u}_n)\}$ where $\mathbf{u}_i = (\langle \mathbf{v}_i, \psi_1 \rangle$
 137 $, \dots, \langle \mathbf{v}_i, \psi_m \rangle)^\top$

$$Z_{(\mathbf{x}, \mathbf{u})}^n = [Z_{(\mathbf{x}, \mathbf{u})} | Z_{(\mathcal{X}_n, \mathcal{U}_n)} = \mathbf{Z}_n]. \quad (6)$$

The mean and covariance of Z^n are given by

$$\begin{aligned} \mathbb{E}[Z_{(\mathbf{x}, \mathbf{u})}^n] &= m_Z(\mathbf{x}, \mathbf{u}) + k_Z((\mathbf{x}, \mathbf{u}); (\mathcal{X}_n, \mathcal{U}_n)) \Sigma_{Z, n}^{-1} (\mathbf{Z} - m_Z(\mathcal{X}_n, \mathcal{U}_n)), \\ \text{Cov}(Z_{(\mathbf{x}, \mathbf{u})}^n, Z_{(\mathbf{x}', \mathbf{u}')}^n) &= k_Z((\mathbf{x}, \mathbf{u}); (\mathbf{x}', \mathbf{u}')) - k_Z((\mathbf{x}, \mathbf{u}); (\mathcal{X}_n, \mathcal{U}_n)) \Sigma_{Z, n}^{-1} k_Z((\mathcal{X}_n, \mathcal{U}_n); (\mathbf{x}', \mathbf{u}')). \end{aligned}$$

138 where $\Sigma_{Z, n} = k_Z((\mathcal{X}_n, \mathcal{U}_n); (\mathcal{X}_n, \mathcal{U}_n))$. The Gaussian Process $Z_{(\mathbf{x}, \mathbf{u})}$ is defined on the finite-
 139 dimensional truncated space $\mathbb{X} \times \mathbb{R}^m$. However, it is worth underlying that we recollect partial
 140 knowledge of the full variability of the untruncated process \mathbf{V} , despite the truncation of the KL
 141 expansion, by evaluating the simulator on design points in $\mathbb{X} \times \Xi$. A discussion about this model
 142 is proposed in Appendix A.

143 *2.3. Integrated Gaussian Process*

144 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
 145 introduce the integrated process

$$Y_{\mathbf{x}}^n = \mathbb{E}_{\mathbf{U}}[Z_{(\mathbf{x}, \mathbf{U})}^n] = \int_{\mathbb{R}^m} Z_{(\mathbf{x}, \mathbf{u})}^n d\rho(\mathbf{u}), \quad (7)$$

146 where ρ is the probability distribution of $\mathbf{U} = (U_1, \dots, U_m)^\top$ introduced in (4). The process $Y_{\mathbf{x}}^n$ is
 147 a Gaussian Process ([28]) fully characterized by its mean and covariance functions which are given
 148 by

$$\mathbb{E}[Y_{\mathbf{x}}^n] = \int_{\mathbb{R}^m} m_Z(\mathbf{x}, \mathbf{u}) + k_Z((\mathbf{x}, \mathbf{u}); (\mathcal{X}_n, \mathcal{U}_n)) \Sigma_{Z, n}^{-1} (\mathbf{Z} - m_Z(\mathcal{X}_n, \mathcal{U}_n)) d\rho(\mathbf{u}), \quad (8)$$

149 and

$$\text{Cov}(Y_{\mathbf{x}}^n, Y_{\mathbf{x}'}^n) = \iint_{\mathbb{R}^m} k_Z((\mathbf{x}, \mathbf{u}); (\mathbf{x}', \mathbf{u}')) - k_Z((\mathbf{x}, \mathbf{u}); (\mathcal{X}_n, \mathcal{U}_n)) \Sigma_{Z, n}^{-1} k_Z((\mathcal{X}_n, \mathcal{U}_n); (\mathbf{x}', \mathbf{u}')) d\rho(\mathbf{u}) d\rho(\mathbf{u}'). \quad (9)$$

150 **3. Data driven infill strategy for robust feasible set estimation**

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

154 *3.1. Minimization of the Vorob'ev deviation: choice of next \mathbf{x}*

155 The objective of the first step is to wisely choose the points in the controlled space \mathbb{X} in order
 156 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
 157 consider the statistical model of the non-observable function g given by $Y_{\mathbf{x}}^n$ introduced in Section
 158 2.3. In the following, we assume that the Gaussian Process $(Z_{(\mathbf{x}, \mathbf{u})})_{(\mathbf{x}, \mathbf{u}) \in \mathbb{X} \times \mathbb{R}^m}$ is separable with
 159 continuous mean function m_Z and Matérn (5/2 or 3/2) covariance function k_Z . Then the feasible
 160 set defined as $\Gamma = \{\mathbf{x} \in \mathbb{X}, Y_{\mathbf{x}}^n \leq c\}$ is a random closed set (see, e.g., [31] p.4, 23).

From the assumption that g is a realization of $Y_{\mathbf{x}}^n$, the true unknown set Γ^* can be seen as a
 realization of the random closed set Γ . The book of [31] gives many possible definitions for the
 variance of a random closed set. In the present work we focus on the Vorob'ev deviation ([32, 33])
 and we adapt the Stepwise Uncertainty Reduction (SUR) strategy introduced in [26] which aims
 at decreasing an uncertainty function defined as the Vorob'ev deviation ([32, 33]) of the random
 set. More precisely the uncertainty function at step n is defined as

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

where μ is the Lebesgue measure on \mathbb{X} , Δ the symmetric difference operator between two sets, the
 Vorob'ev quantiles are given by $Q_{n, \alpha} = \{\mathbf{x} \in \mathbb{X}, \mathbb{P}(Y_{\mathbf{x}}^n \leq c) \geq \alpha\}$, and the Vorob'ev expectation
 Q_{n, α_n^*} can be determined by tuning α to a level α^* such that $\mu(Q_{n, \alpha_n^*}) = \mathbb{E}[\mu(\Gamma) \mid Z_{(\mathcal{X}_n, \mathcal{U}_n)} = \mathbf{Z}_n]$.
 Let

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

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

$$\begin{aligned} \mathbf{x}_{n+1} &\in \operatorname{argmin}_{\mathbf{x} \in \mathbb{X}} \mathbb{E}_{n, \mathbf{x}}[\mathcal{H}_{n+1}^{\text{uncert}}(\mathbf{x})] \\ &:= \operatorname{argmin}_{\mathbf{x} \in \mathbb{X}} \mathcal{J}_n(\mathbf{x}), \end{aligned} \tag{10}$$

163 where $\mathbb{E}_{n, \mathbf{x}}$ denotes the expectation with respect to $Y_{\mathbf{x}}^n$. For the computation of $\mathcal{J}_n(\mathbf{x})$, we use
 164 the formula given in [34, Eq. (4.43) in Section 4.2] (see also [26, Section 3.1]). In this work, we
 165 add a single point at each iteration. In other words, we are seeking the point \mathbf{x}_{n+1} that minimizes
 166 $\mathcal{J}_n(\mathbf{x})$. However, it is worth noting that there exists a batch-mode sampling version, which can be
 167 beneficial when computations can be parallelized. Nevertheless, it should be acknowledged that
 168 the optimization process of $\mathcal{J}_n(\mathbf{x})$ becomes computationally expensive unless a greedy approach
 169 is implemented [35].

170
 171 The enrichment of the DoE consists in selecting a couple $(\mathbf{x}_{n+1}, \mathbf{u}_{n+1})$ in the joint space $\mathbb{X} \times \mathbb{R}^m$.
 172 \mathbf{x}_{n+1} has just been defined by (10), it remains now to choose a new point \mathbf{u}_{n+1} in the uncertain
 173 space.

174 *3.2. Minimization of the variance: choice of next \mathbf{u}*

175 The process Y^n approximates the expectation $\mathbb{E}_{\mathbf{V}}[f(\cdot, \mathbf{V})]$. It can be seen as a projection of
 176 Z^n from the joint space onto the controlled space. We propose to sample the point \mathbf{u}_{n+1} in the
 177 uncertain space in order to reduce at most the one-step-ahead variance at point \mathbf{x}_{n+1} , $\mathbb{V}\mathbb{A}\mathbb{R}(Y_{\mathbf{x}_{n+1}}^{n+1})$,
 178 whose expression is obtained from Eq. (9). More precisely,

$$\mathbf{u}_{n+1} = \operatorname{argmin}_{\mathbf{u} \in \mathbb{R}^m} \mathbb{V}\mathbb{A}\mathbb{R}(Y_{\mathbf{x}_{n+1}}^{n+1}), \tag{11}$$

179 with

$$\begin{aligned}
\text{VAR}(Y_{\mathbf{x}_{n+1}}^{n+1}) &= \vartheta(\tilde{\mathbf{u}}), \\
&= \iint_{\mathbb{R}^m} k_Z((\mathbf{x}_{n+1}, \mathbf{u}); (\mathbf{x}_{n+1}, \mathbf{u}')) d\rho(\mathbf{u}) d\rho(\mathbf{u}') \\
&\quad - \iint_{\mathbb{R}^m} k_Z((\mathbf{x}_{n+1}, \mathbf{u}); (\mathcal{X}_{n+1}, \mathcal{U}_{n+1})) \Sigma_{Z,n+1}^{-1} k_Z((\mathcal{X}_{n+1}, \mathcal{U}_{n+1}); (\mathbf{x}_{n+1}, \mathbf{u}')) d\rho(\mathbf{u}) d\rho(\mathbf{u}'),
\end{aligned} \tag{12}$$

180 where $\Sigma_{Z,n+1} = k_Z((\mathcal{X}_{n+1}, \mathcal{U}_{n+1}); (\mathcal{X}_{n+1}, \mathcal{U}_{n+1}))$ and $(\mathcal{X}_{n+1}, \mathcal{U}_{n+1}) = (\mathcal{X}_n, \mathcal{U}_n) \cup \{(\mathbf{x}_{n+1}, \tilde{\mathbf{u}})\}$.

181 3.3. Implementation

182 We present in this section two algorithms. Algorithm 1 is the global algorithm summarizing
183 the strategy we propose for robust feasible set estimation. The truncation argument is fixed once
184 for all in the algorithm. Algorithm 2 is a variation of Algorithm 1 with a data driven procedure
185 for increasing the truncation argument m over the iterations. This last algorithm allows to further
186 reduce the number of calls to the simulator.

187 The setting of our procedure is driven by our industrial application where the random process
188 \mathbf{V} is known only through a finite set of realizations $\Xi = \{\check{\mathbf{v}}_1, \dots, \check{\mathbf{v}}_N\}$. In this framework, points
189 a) and b) below detail the computation of KL decomposition and the minimization of the one-
190 step-ahead variance.

191 *a) Computational method for functional PCA.* We consider the empirical version of $C(s, t)$ defined
192 as $C^N(s, t) = \frac{1}{N} \sum_{i=1}^N \check{\mathbf{v}}_i(s) \check{\mathbf{v}}_i(t)$. The eigenvalue problem defined by Eq. (3) is then solved by
193 discretizing the trajectories $\{\check{\mathbf{v}}_i\}_{i=1, \dots, N}$ on $[0, T]$ and replacing C by C^N . Denoting by $\hat{\psi}_i$, $i =$
194 $1, \dots, m$, the m first estimated eigenfunctions, we define

$$195 \mathcal{G}_m = \{\check{\mathbf{u}}_1, \dots, \check{\mathbf{u}}_m\} \tag{13}$$

196 where $\check{\mathbf{u}}_i = (\langle \check{\mathbf{v}}_i, \hat{\psi}_1 \rangle, \dots, \langle \check{\mathbf{v}}_i, \hat{\psi}_m \rangle)^\top$.

197 *b) Minimization of the one-step-ahead variance.* Since \mathbf{V} is known through a finite set Ξ , Eq. (11)
198 is solved on the finite set \mathcal{G}_m .

199 We now state the global algorithm we propose for robust feasible set estimation (Algorithm 1)
200 and its variation with a data driven increase of the truncation argument (Algorithm 2).

201 3.3.1. A first algorithm for robust feasible set estimation

202 In this section we provide a global algorithm for the implementation of our methodology. Then
203 we comment some of its steps.

Algorithm 1 Robust feasible set estimation via joint space modelling

Require: The truncation argument m , the initial DoE of n_0 points $(\mathcal{X}_n, \mathcal{U}_n)$ in $\mathbb{X} \times \mathcal{G}_m$, and a maximal simulation budget

- 1: Set $n = n_0$.
 - 2: Calculate \mathbf{Z} the simulator responses at the design points $(\mathcal{X}_n, \mathcal{U}_n)$
 - 3: **while** $n \leq \text{budget}$ **do**
 - 4: Fit the GP model Z^n
 - 5: Induce the integrated GP $Y_{\mathbf{x}}^n$
 - 6: $\mathbf{x}_{n+1} \leftarrow$ sampling criterion \mathcal{J}_n
 - 7: $\mathbf{u}_{n+1} \leftarrow \text{argmin}_{\bar{\mathbf{u}} \in \mathcal{G}_m} \text{VAR}(Y_{\mathbf{x}_{n+1}}^{n+1})$
 - 8: 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}
 - 9: Update DoE : $(\mathcal{X}_{n+1}, \mathcal{U}_{n+1}) = (\mathcal{X}_n, \mathcal{U}_n) \cup \{(\mathbf{x}_{n+1}, \mathbf{u}_{n+1})\}$
 - 10: Update $\mathbf{Z} = \mathbf{Z} \cup \{f(\mathbf{x}_{n+1}, \mathbf{v}_{n+1})\}$
 - 11: Set $n = n + 1$
 - 12: **end while**
 - 13: Fit the final GP model Z^n
 - 14: Approximate Γ^* by the Vorob'ev expectation
-

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

208 **step 3** The algorithm stops when a maximal budget, in terms of number of model evaluations, is reached.

209 **step 4** The covariance kernel of the GP is chosen as a Matérn-5/2 covariance and we add a noise modelled
210 with a constant variance term. The homoscedastic modelling of the noise is discussed in Appendix A.
211 The mean function of the GP is modelled by a constant function. All types of parameters (mean,
212 correlation lengths, variance and noise) are estimated by maximum likelihood [36] at each iteration.

213 **step 5** In the framework where the uncertain vector \mathbf{U} is Gaussian as well as the covariance kernel, closed
214 form solutions of the integrals in (8) and (9) are given in [28]. In our framework, the integrals in (8)
215 and (9) are approximated by Monte Carlo using the sample \mathcal{G}_m .

216 **step 6** \mathbf{x}_{n+1} is obtained by solving (10) with a continuous global optimization algorithm: GENetic Opti-
217 mization Using Derivatives (GENOUD) [37].

218 **step 7** Once more the integrals in (12) are approximated by Monte Carlo. More details on the estimation
219 of (12) can be found in [28]. Here the minimization problem is solved by an exhaustive search on
220 the finite set \mathcal{G}_m defined in (13).

221 **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
222 Ξ corresponding to the truncated vector of coefficients \mathbf{u}_{n+1} . Note that evaluating the simulator at
223 a curve in the initial set of realizations whose coordinate in the uncertain space is \mathbf{u}_{n+1} and not a
224 projected curve on the basis composed with first eigenfunctions brings robustness with respect to
225 the truncation argument.

226 3.3.2. A variant of the proposed algorithm with a varying size of the reduced uncertain space

227 One limitation of our methodology is the prior choice of the truncation argument m . A nat-
228 ural choice is to fix this argument to a value m_0 guaranteeing that the m_0 first components in
229 the Karhunen-Loève decomposition explain a high prescribed percentage of the variability of the
230 random process V . But, depending on the complexity of the random process V , m_0 may be large
231 (more than 60 for the industrial application if we want to explain 95% of the random process

232 variability). If m_0 is large, the dimension of the GP space is also large. A potentially high num-
 233 ber of design points is then needed to produce an accurate response surface, which can lead to a
 234 prohibitive computational cost in terms of number of model evaluations. To overcome this issue,
 235 we introduce below an adaptive variant of our strategy. It consists in augmenting the uncertain
 236 space sequentially when needed. More precisely, a first Gaussian Process is defined in the $p + m$
 237 dimensional space, with m chosen small. Once the enrichment strategy (given by Algorithm 1)
 238 no longer provides information - associated to a stagnation in uncertainty reduction improvement
 239 and a potential rough or biased approximation of the feasible set - the dimension of the uncertain
 240 space is increased and the GP is updated in the $p + m + 1$ dimensional space. This alternative
 241 strategy is summarized by Algorithm 2:

Algorithm 2 Robust feasible set estimation via sequential joint space modelling

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

- 1: Set $n = n_0$.
 - 2: Calculate \mathbf{Z} the simulator responses at the design points $(\mathcal{X}_n, \mathcal{U}_n)$
 - 3: **while** $n \leq \text{budget}$ **do**
 - 4: $m \leftarrow \text{Update.Dimension}()$
 - 5: Fit the GP model Z^n
 - 6: Induce the integrated GP $Y_{\mathbf{x}}^n$
 - 7: $\mathbf{x}_{n+1} \leftarrow \text{sampling criterion } \mathcal{J}_n$
 - 8: $\mathbf{u}_{n+1} \leftarrow \arg \min_{\tilde{\mathbf{u}} \in \mathcal{G}_m} \text{VAR}(Y_{\mathbf{x}_{n+1}}^{n+1})$
 - 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 : $(\mathcal{X}_{n+1}, \mathcal{U}_{n+1}) = (\mathcal{X}_n, \mathcal{U}_n) \cup \{(\mathbf{x}_{n+1}, \mathbf{u}_{n+1})\}$
 - 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
-

242 In **step** 4 of Algorithm 2, the uncertain space dimension is updated based on a stagnation
 243 criterion of the Vorob'ev Deviation. More precisely, the dimension is updated from m to $m + 1$ if
 244 the following stopping criterion is verified

$$\forall 0 \leq j < n_0^{\text{SUR}}, e_{n-j}^{\text{SUR}} \leq \epsilon^{\text{SUR}}, \quad (14)$$

245 where $e_n^{\text{SUR}} = |\mathcal{H}_n^{\text{uncert}} - \mathcal{H}_{n-1}^{\text{uncert}}|$ is the absolute error between Vorob'ev deviations estimated
 246 for two successive iterations. The condition in Eq. (14) tests if the variation of the Vorob'ev
 247 deviation is smaller than a tolerance ϵ^{SUR} on n_0^{SUR} consecutive steps, with n_0^{SUR} to be tuned by
 248 the practitioner. For the tuning of parameters $(\epsilon^{\text{SUR}}, n_0^{\text{SUR}})$, we refer to the discussion in [13,
 249 Section 5]. Remark that the algorithm stops either when a maximal budget, in terms of number
 250 of model evaluations, or a maximal value for the truncation argument m is reached.

251 4. Numerical experiments

252 In the following, we first introduce in Section 4.1 two analytical examples on which we will
 253 test our strategy. Then in Section 4.2 we present the results obtained by implementing Algorithm
 254 1. In Section 4.3 we present the improvements we get by increasing adaptively the truncation
 255 argument in the KL decomposition over iterations (Algorithm 2). Finally Section 4.4 is devoted
 256 to the implementation of our strategy on an industrial application of automotive depollution.

257 *4.1. Two analytical examples - set-up*

258 To illustrate the behaviour of the proposed method, we consider two analytical examples. We
 259 suppose that a sample Ξ of size $N = 200$ realizations of the random process \mathbf{V} is available and
 260 its probability distribution is unknown. To highlight the robustness of our method regarding the
 261 random distribution of the uncertainties, we consider two types of random processes: max-stable
 262 process and the well-known brownian motion. Regarding the max-stable process, we consider the
 263 Schlather model with powered exponential correlation function, i.e., $\ell(h) = \exp\{-(h/\lambda)^\kappa\}$, where
 264 $\kappa = 1$ and $\lambda = 10$. This process is also known as the extremal Gaussian process [38]. The function
 265 `rmaxstab` from the R-package `SpatialExtremes` is used to generate a sample of realizations. As
 266 Algorithm 1 depends on the truncation argument m , different values are tested (see Table 1) to
 267 better understand the effect of the uncertain space dimension.

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

Table 1: Variance explained by the truncated KL decomposition according to m for two different random processes.

268 For all analytical examples, we consider a Gaussian Process prior $Z_{(\mathbf{x}, \mathbf{u})}$ with constant mean
 269 and Matérn covariance kernel with $\nu = 5/2$. Random Latin Hypercube Designs (RLHD) are used
 270 as initial DoEs in all the experiments. The number of points of the initial DoE is 20 for the first
 271 analytical example and 30 for the second one. RLHD induce variability in the behaviour of the
 272 algorithms. To account for this variability, the performance is averaged over 30 (respectively 10)
 273 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

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

274 where $\mathbf{x} \in \mathbb{X} = [-100, 100]^2$. The objective is to approximate the set $\Gamma^* = \{\mathbf{x} \in \mathbb{X}, g(\mathbf{x}) =$
 275 $\mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})] \leq 3500\}$ for the two different types of random processes (brownian motion and max-
 276 stable process).

Analytical example 2. For the second example we define a function that is not separable
 with respect to the controlled variables \mathbf{x} and the random process \mathbf{V} . The function involves the
 maximum and the minimum of $(V_t)_{t \geq 0}$, so catching the whole variability of \mathbf{V} becomes important.
 The function f is given by

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

277 where the controlled variables lie in $\mathbb{X} = [1.5, 5] \times [3.5, 5]$. The objective is to approximate the set
 278 $\Gamma^* = \{\mathbf{x} \in \mathbb{X}, g(\mathbf{x}) = \mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})] \leq c\}$, when $c = 1.2$ for the brownian motion and $c = 0.9$ for
 279 the max-stable process.

280 Note that for both examples the reference solution Γ^* is obtained from a 30×30 grid experiment,
 281 where at each grid point the expectation is empirically approximated using the whole sample Ξ . In
 282 the following, we measure the performance of the different strategies with the ratio of the volume of
 283 the symmetric difference between the reference set Γ^* and the estimated one Q_{n, α^*} to the volume
 284 of the reference set: $\mu(\Gamma^* \Delta Q_{n, \alpha^*}) / \mu(\Gamma^*)$ to which we will refer as the quality-ratio.

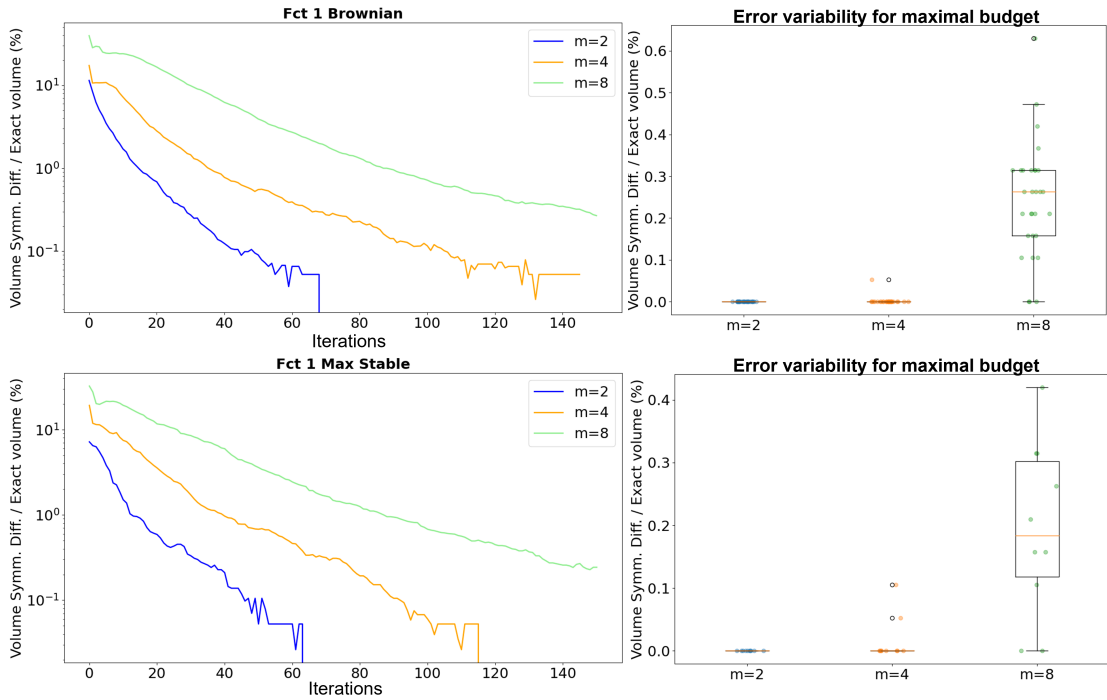


Figure 1: Analytical example 1 with brownian motion (top) and with max-stable process (bottom). Convergence of Algorithm 1 for $m = \{2, 4, 8\}$. Left: quality-ratio mean as function of the number of simulator calls in log scale. The mean is taken over the 30 independent runs of initial RLHD for brownian motion (respectively 10 for max-stable process). Right: quality-ratios associated with the random initial DoEs of 20 simulations, at the maximal budget of 170 simulations.

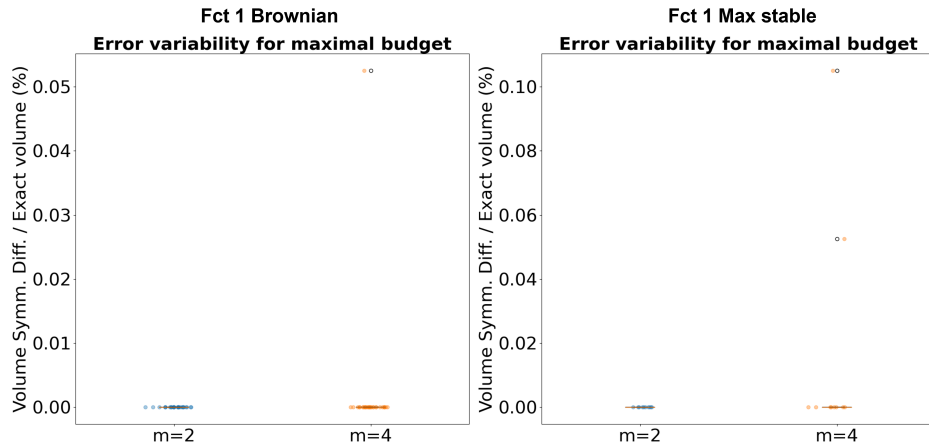


Figure 2: Zoom of Figure 1 : error variability for maximal simulation budget for $m = 2$ and $m = 4$.

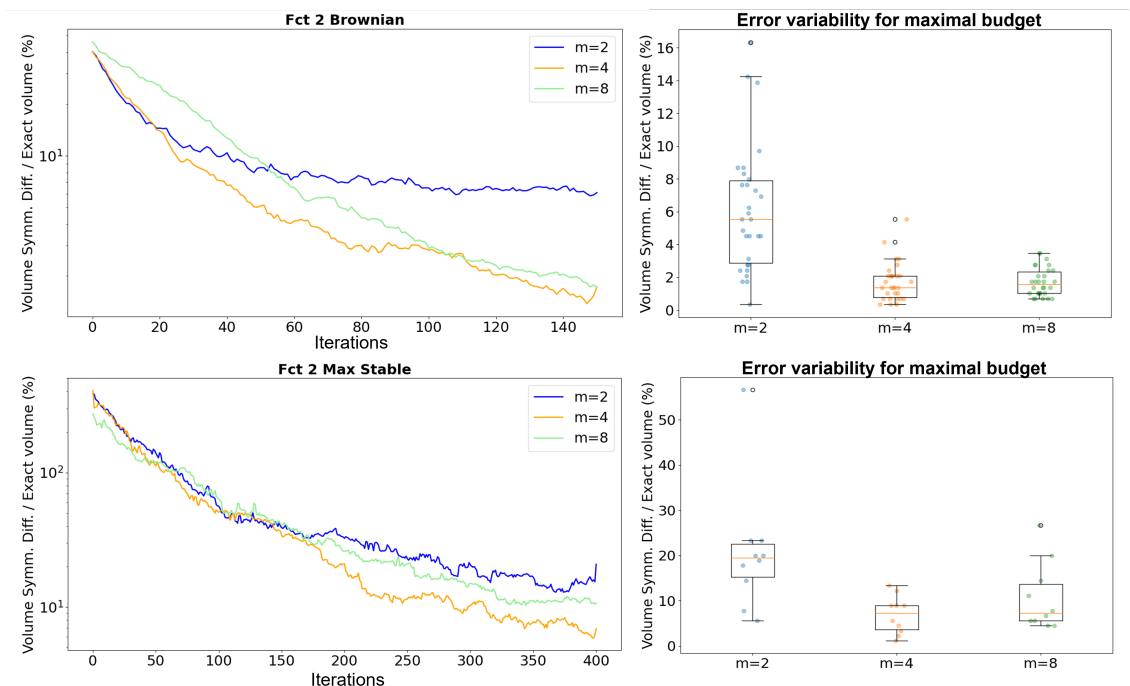


Figure 3: Analytical example 2 with brownian motion (top) and with max-stable process (bottom). Convergence of Algorithm 1 for $m = \{2, 4, 8\}$. Left: quality-ratio mean as function of the number of simulator calls in log scale. The mean is taken over the 30 independent runs of initial RLHD for brownian motion (respectively 10 for max-stable process). Right: quality-ratios associated with the random initial DoEs of 30 simulations, at the maximal budget of 180 simulations for brownian motion (respectively 430 for max-stable process).

286 In Figures 1 and 3, we show the evolution of the averaged quality-ratio with respect to the
 287 number of simulations involved in the Algorithm 1 on the two analytical examples with the two
 288 types of functional uncertainties (brownian and max-stable processes). The average is taken over
 289 the repeated runs of the complete approach corresponding to the 30 random initial designs (10 for
 290 the max-table process), and for 3 values of the truncation argument m .

291 For the first analytical example, the smaller values of m , the faster is the convergence. This
 292 observation can be explained by the fact that, in higher dimensional joined space (due to larger
 293 values of m), much more evaluation points are necessary to learn an accurate GP model (more
 294 hyper-parameters to determine). It is worth noting that even for 90% (for brownian motion) or
 295 58.8% (for max-stable process) of explained variance with $m = 2$ the proposed algorithm provides
 296 an efficient estimate of the true set Γ^* . Indeed, on stage 8 in Algorithm 1 the full curve $\mathbf{v}_{n+1} \in \Xi$
 297 associated to \mathbf{u}_{n+1} is recovered, such that the information lost after the dimension reduction is
 298 reduced, thereby further robustifying the method.

299 Regarding the second analytical example, the output depends on local behaviours of the stochas-
 300 tic process. The truncation argument $m = 2$ is too small to catch these dependencies, the function
 301 is sensitive to higher KL order. For the brownian motion, more than 95% of variance is explained
 302 with $m = 4$. It seems sufficient to obtain an accurate approximation of Γ^* . The improvement
 303 between $m = 2$ and $m = 4$ is noticeable. The improvement is not as important when the uncer-
 304 tainties are driven by a max-stable process since the percentage of explained variance increases
 305 slowly. Better results should be observed with $m = 8$. It is not the case because a higher dimension
 306 leads to difficulties in the estimation of the GP except by increasing consequently the number of
 307 observation points. Figure A.15 in Appendix shows the evolution of the feasible domain estimation

308 with respect to the iterations of Algorithm 1 for the second analytic case and the brownian motion,
 309 and for different truncation levels.

310 As shown in Figure 4, the higher the dimension of the GP space is, the longer the internal
 311 computations last. Moreover, the computational time needed to provide the next evaluation point
 312 increases with the number of simulator calls, and thus with the number of iterations, the cost of
 313 kriging being directly linked to the learning sample size. For example in the case of $m = 8$ (resp.
 314 $m = 2$), iteration 80 requires 203 (resp. 126) seconds to provide the next evaluation point whereas
 315 iteration 150 requires 275 (resp. 164) seconds.

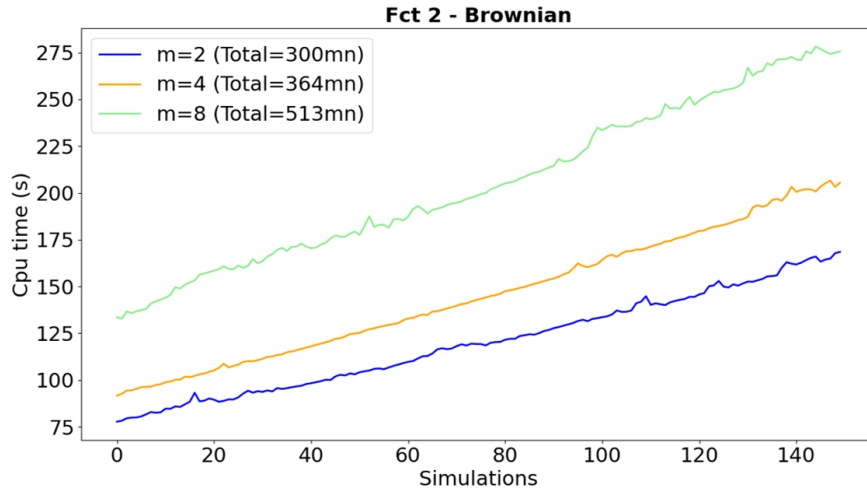


Figure 4: The computational time (sec.) needed to provide the next evaluation point as a function of iterations for the second analytic example with brownian motion. The values are averaged computational times for 5 runs of each strategy: $m = 2, 4, 8$.

316 To highlight the interest of the sampling criterion (11), we compared our approach to the
 317 one where \mathbf{u}_{n+1} is chosen accordingly to a uniform distribution. The results obtained for both
 318 analytical cases with the brownian motion are shown on Figure 5. We note that our criterion leads
 319 to a faster decrease of the quality-ratio and to a much smaller error variability, in comparison to
 320 a uniform sampling in the uncertain space. As shown in Figure 6, the points selected with the
 321 criterion based on (11) seem to concentrate in interest areas, in comparison to the points selected
 322 uniformly. Thus, the guided sampling in the uncertain space, that reduces the variance of the
 323 expectation estimation, leads to a faster convergence towards the feasible domain.

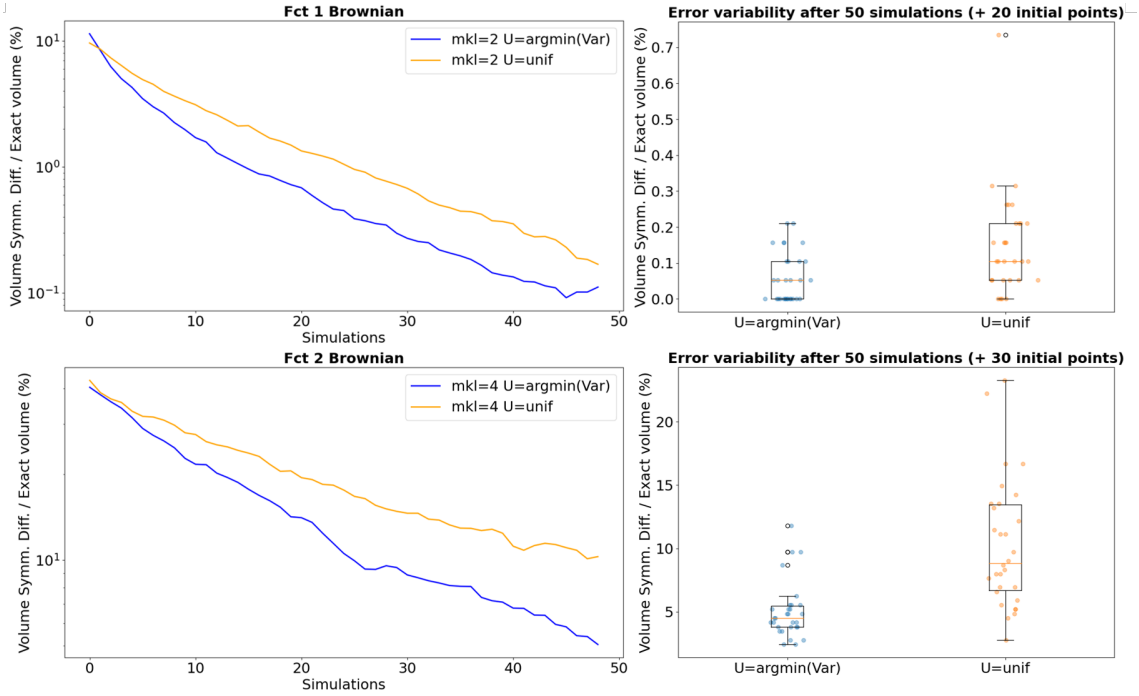


Figure 5: Analytical examples with the uncertainty modelled by a brownian motion. Decrease of the mean quality-ratio as function of the number of simulator calls (left) and boxplots for quality-ratios associated with the different random initial DoEs at the maximal simulation budget (right) for the first analytical case (top) and the second one (bottom), comparing criterion based on (11) with the uniform sampling in the uncertain space.

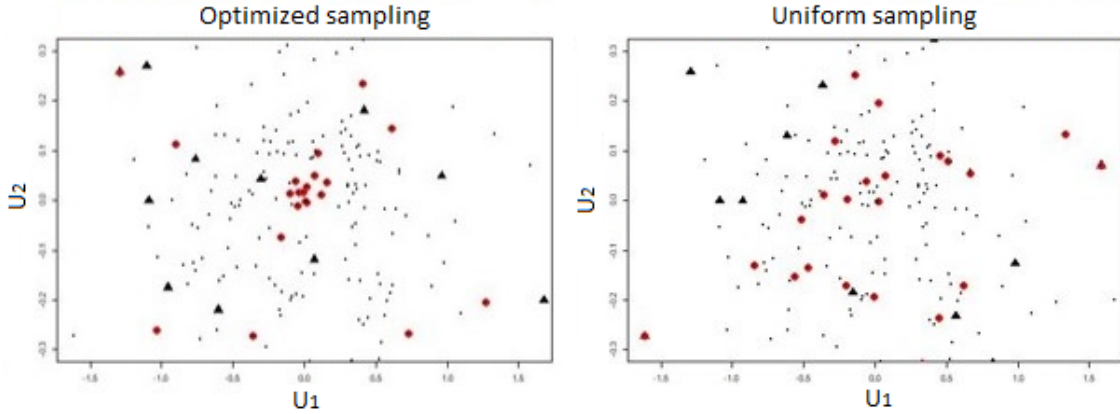


Figure 6: Analytical example 1 with the uncertainty modelled by a brownian motion. Black triangles correspond to the coefficients of the initial RLHD plotted in the uncertain truncated space ($m = 2$). Red points are the added points based on our criterion (left) and uniformly sampled (right) up to 50 simulations.

324 Finally we compare our joint modelling based method with the approach introduced in [13]
 325 which combines GP modelling in the controlled space with quantization to estimate the expectation
 326 in the uncontrolled space. Even without taking into account the costs induced by the initial designs
 327 in the controlled space (RLHD of size 9), the current approach based on joint GP modelling
 328 performs better regardless of the truncation argument m (Figure 7). Adding the costs induced

329 by the quantization on the initial sample points would disadvantage even more the approach
 330 introduced in [13] as the size of the quantizer is around 20 for each of the 9 points of the initial
 331 design, i.e. 180 simulator calls in average.

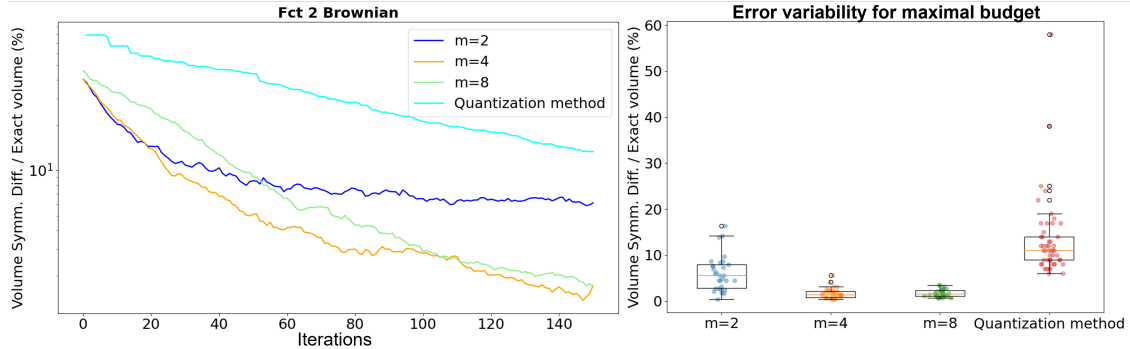


Figure 7: Comparison between the current approach based on joint GP modelling and the approach introduced in [13] combining GP modelling in the controlled space with quantization in the uncertain space. The costs induced by the initial designs (RLHD of size 9) are not counted. The uncertainty is modelled by a brownian motion.

332 *4.3. Evaluation of the adaptive strategy of Algorithm 2*

333 We now evaluate the adaptive strategy presented in Algorithm 2. We focus on the second
 334 analytical function of Section 4.1 which is the most complex one. A small initial value of m is
 335 chosen, $m = 2$, and it is then increased when the variation of the Vorob'ev deviation remains
 336 smaller than a given threshold $\epsilon = 0.005$ during $n_0 = 4$ consecutive iterations (see Eq. (14) in
 337 Section 3.3). It allows to increase the dimension of the KL reduced space only when it is necessary
 338 to obtain a better accuracy. As illustrated on Figure 8 it allows to save simulations and reduce
 339 computational time. The accuracy reached with this strategy is similar to the one obtained with
 340 the strategy with fixed $m = 8$ but with a gain of $\approx 12\%$ in terms of computational time (Figure
 341 9). We notice that the first iterations are performed with $m = 2$ and only the last iterations with
 342 $m = 8$.

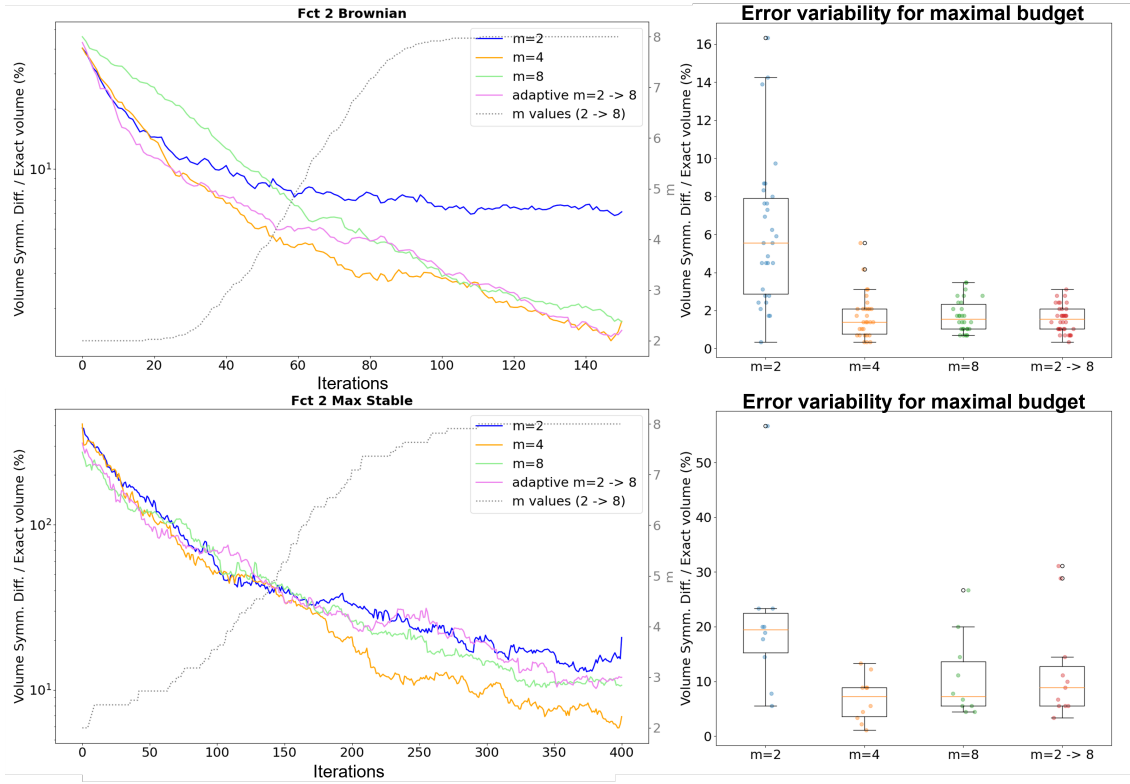


Figure 8: Analytical example 2 with brownian motion (top) and with max-stable process (bottom). Convergence of Algorithm 1 for $m = \{2, 4, 8\}$ and for adaptive choice of m value (Algorithm 2). Left: mean quality-ratio as function of the number of iterations in log scale. The dashed grey curve is the 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: boxplots for the quality-ratios associated with the different random initial DoEs at the maximal simulation budget.

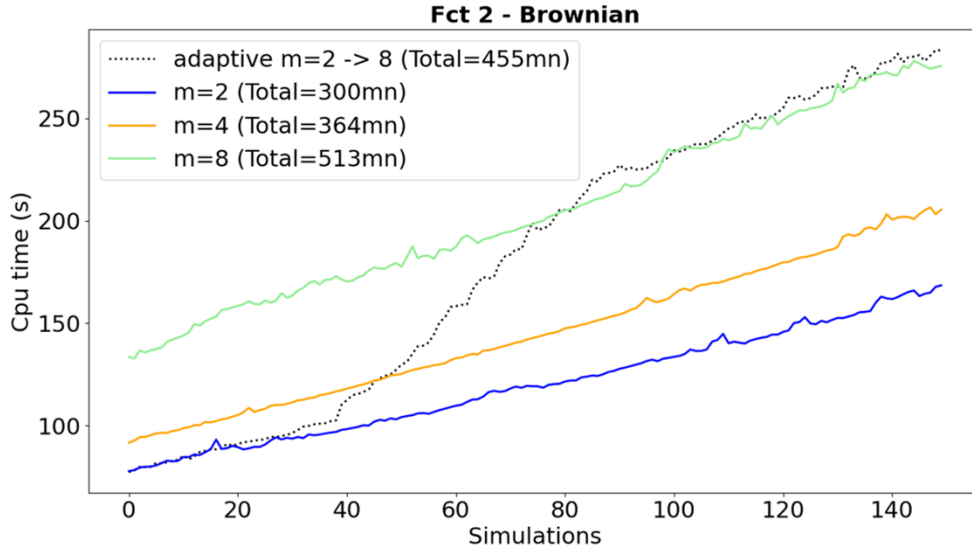


Figure 9: The computational time (sec.) needed to provide the next evaluation point as a function of iterations for the second analytic example with brownian motion. The values are averaged computational times for 5 runs of each strategy: $m = 2, 4, 8$ and adaptive choice of m value.

343 *4.4. Application to a pollution control system SCR*

344 In this section we test the proposed method on an automotive test case from IFPEN. The
 345 problem concerns an after-treatment device of diesel vehicles, called Selective Catalytic Reduction
 346 (SCR). This latter consists of a basic process of chemical reduction of nitrogen oxides (NO_x) to
 347 diatomic nitrogen (N₂) and water (H₂O) by the reaction of NO_x and ammonia NH₃. The reaction
 348 itself occurs in the SCR catalyst. Ammonia is provided by a liquid-reductant agent injected
 349 upstream of the SCR catalyst. The amount of ammonia introduced into the reactor is a critical
 350 quantity: overdosing causes undesirable ammonia slip (unreacted ammonia) downstream of the
 351 catalyst, whereas under-dosing causes insufficient NO_x reduction. In practice, ammonia slip is
 352 restricted to a prescribed threshold. We use an emission-oriented simulator developed by IFPEN,
 353 which models the vehicle, its engine and the exhaust after-treatment system. This latter takes
 354 as input the vehicle driving cycle profile and provides the time-series of corresponding exhaust
 355 emissions as output. A realistic SCR control law is used in this simulator (see [39] for more
 356 details). In this study, the two controlled variables are parameters of the SCR control law and lie
 357 in $\mathbb{X} = [0, 0.6]^2$. The random process describes the evolution of vehicle speed on $I = [0, 5400s]$ and
 358 is known through an available sample of 100 real driving cycles. A few samples are represented in
 359 Figure 10.

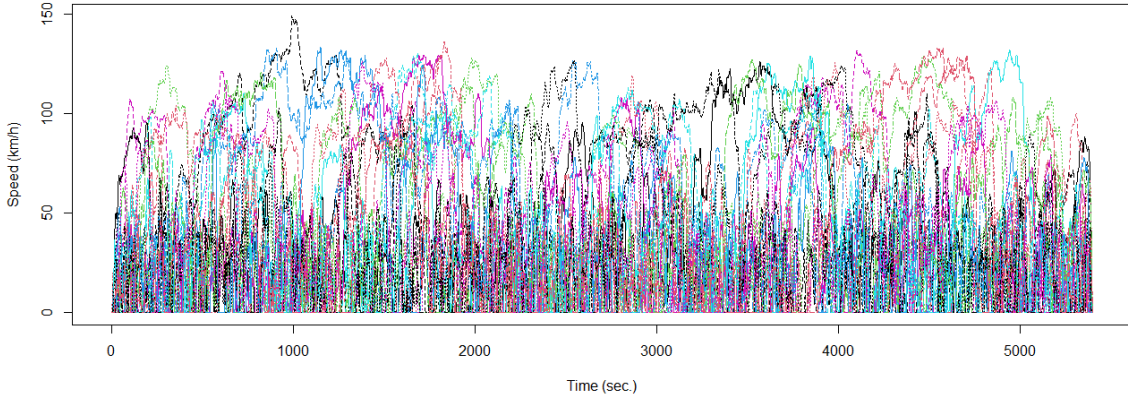


Figure 10: Seven real-driving cycles extracted from the available sample of size 100.

360 In short, the ammonia emissions peak during a driving cycle is modelled as a function

$$f : \begin{cases} \mathbb{X} \times \mathcal{V} & \rightarrow \mathbb{R} \\ (\mathbf{x}, \mathbf{V}) & \mapsto f(\mathbf{x}, \mathbf{V}) = \max_{t \in I} \text{NH}_3(t) \end{cases} \quad (15)$$

361 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 $c = 30ppm$.
 362 Conducting this study on a full grid would consist in covering the space $[0, 0.6]^2$ with a fine mesh
 363 and evaluating the code 100 times at each point. Knowing that each simulation takes about two
 364 minutes, such study would require many hours of computational time, and thus using meta-models
 365 allows to tackle this computational issue.

Truncation argument	$m = 15$	$m = 17$	$m = 19$	$m = 21$	$m = 23$	$m = 25$	$m = 64$
Explained variance	63.11%	66.64%	69.64%	72.45%	74.86%	76.87%	95.32%

Table 2: SCR pollution control system. The Variance explained by the KL decomposition according to m for the random cycle process \mathbf{V} .

366 In this industrial case, the explained variance grows very slowly as shown in the Table 2.
 367 Therefore, to represent at best the driving cycle \mathbf{V} one should consider a high-dimensional threshold
 368 ($m = 64$). In order to avoid the curse of dimensionality involved by a large value of m , we adopt
 369 the adaptive strategy introduced in section 3.3.2 and tested in subsection 4.3 on a toy problem.
 370 More precisely, we start modelling the numerical simulator in a reduced uncertain space by taking
 371 $m = 15$. We increase the dimension of the uncertain space as soon as the condition in Eq. (14)
 372 is satisfied. We set the condition parameters at $\epsilon^{\text{SUR}} = 5 \times 10^{-3}$ and $n_0^{\text{SUR}} = 4$. The maximal
 373 budget is fixed to 1000 iterations of the algorithm while the maximal truncation argument is fixed
 374 to $m_0 = 64$. We consider a Gaussian Process prior $Z_{(\mathbf{x}, \mathbf{u})}$, with constant mean function and Matérn
 375 covariance kernel with $\nu = 5/2$. The initial DoE consists of a $n = 60$ points LHS design optimized
 376 with respect to the maximin criterion. The covariance kernel hyper-parameters are re-estimated
 377 by maximizing the likelihood at each iteration. Figure 11 shows the so-called excursion probability
 378 function defined by $\mathbf{x} \mapsto \mathbb{P}(Y_{\mathbf{x}}^n \leq c)$, with $Y_{\mathbf{x}}^n$ the integrated Gaussian Process $Y_{\mathbf{x}}$ conditionally
 379 to the n available observations. The initial estimate of Γ^* is given by the green set with blue
 380 boundary. In the following and as for the analytical examples, we proceed to add one point at
 381 each iteration of the SUR strategy.

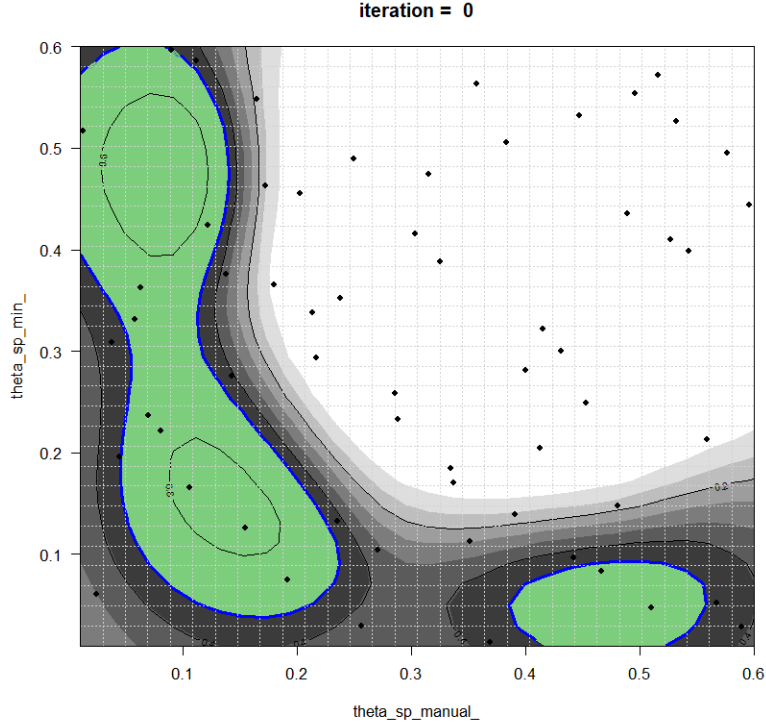


Figure 11: SCR pollution control system. The initial DoE (black triangles) and the initial estimate set (green). The contour plot in grey represents the excursion probability: darker corresponding to higher probability for the integrated process to be under the threshold.

382 Figure 12 shows the evolution of Vorobev deviation and of the dimension during the itera-
 383 tions of the algorithm. Initially, we reduce the Vorobev deviation based on GP modelling in a
 384 17-dimensional joint space. At iteration 754, as the algorithm detects a stagnation of Vorobev
 385 deviation in the current space, the dimension m is increased to $m + 1$. We namely interpret a
 386 stagnation of Vorobev deviation in the current space as a convergence to a biased estimation of the
 387 feasible set of interest. Each time such a stagnation (measured with parameters ϵ^{SUR} and n_0^{SUR})
 388 is observed, the current dimension is increased. For this application, the algorithm stops when
 389 the maximal budget of 1000 iterations is reached. The truncation argument has then reached the
 390 value $m = 24$, corresponding to approximately 75% of the variability of the driving cycle explained
 391 (see Table 2). The estimated set at different iterations is given in Figure A.16 in Appendix. From
 392 Figure 13, we note that the SUR algorithm heavily visits the boundary region of Γ^* and explore
 393 also other potentially interesting regions. Actually, after 1000 iterations the whole domain \mathbb{X} has
 394 an excursion probability close to either 0 or 1.

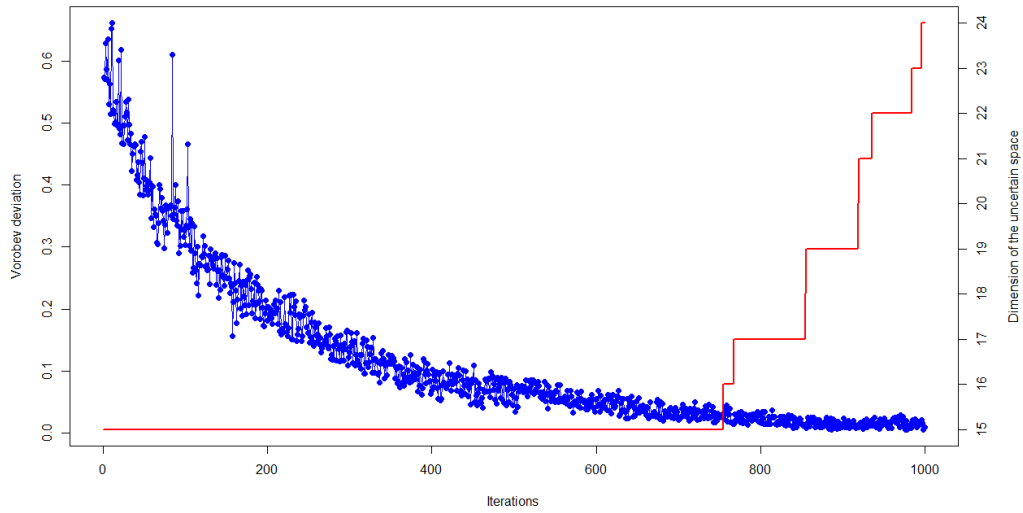


Figure 12: SCR pollution control system. The Vorob'ev deviation and the uncertain space dimension in function of the number of simulations.

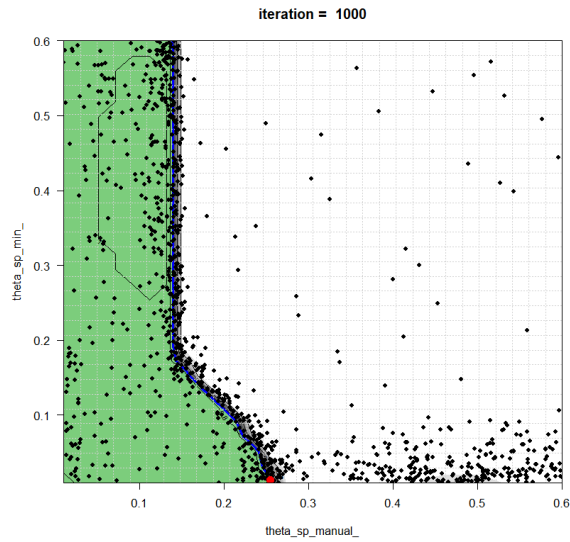


Figure 13: SCR pollution control system. The coverage probability function, the last proposed point by the algorithm (red point) and the estimate set (green set) after 1000 added points (black points). The contour plot in grey represents the excursion probability.

395 **5. Conclusion**

396 The aim of this paper is to propose a new feasible set estimation procedure for an automotive
 397 control system in presence of functional uncontrolled variables modelled by a random process. Our
 398 procedure outperforms the one recently introduced in [13], as it requires less evaluations of the
 399 high-fidelity and expensive-to-evaluate model used to simulate the behaviour of the automotive
 400 control system to achieve a similar accuracy.

401 Our procedure is based on a new enrichment strategy whose main ingredient is to fit a Gaussian
402 Process model to the initial expensive-to-evaluate code in the joint input space of controlled and
403 uncontrolled variables. In our framework, the uncontrolled variable is modelled by a random
404 process \mathbf{V} to which a preliminary step of dimension reduction is applied. Moreover our knowledge
405 of \mathbf{V} is limited to a finite set of realizations Ξ . Note that our approach guarantees robustness with
406 respect to the order of reduction. Indeed, partial knowledge of the full variability of \mathbf{V} is recovered
407 by evaluating the simulator on design points in $\mathbb{X} \times \Xi$. More precisely, at each step of the enrichment
408 procedure, the simulator is evaluated in the point in $\mathbb{X} \times \Xi$ whose projection corresponds to the
409 point obtained by optimizing the criterion in the truncated joint input space. The enrichment
410 procedure guides the sampling toward informative regions for the feasible set, allowing by the way
411 an accurate estimation with less evaluations of the expensive-to-evaluate code. We also propose
412 a variation of our strategy, in which the order of reduction of \mathbf{V} is increased adaptively. This
413 approach consists in increasing the order of reduction sequentially, only when necessary, leading
414 to even more computational savings.

415 Two bi-dimensional analytical examples are considered for which a reference solution can be
416 computed. Then our procedure is validated by computing the quality-ratio defined as the ratio
417 of the volume of the symmetric difference between the reference set and the estimated one to the
418 volume of the true set. On these analytical examples, our procedure outperforms the one in [13] by
419 achieving similar accuracy with much less evaluations of the expensive-to-evaluate code. Finally,
420 we apply our procedure to an industrial problem related to the pollution control system of an
421 automotive. A feasible set solution is found within a reasonable number of simulations.

422 The paper focuses on a formulation of the excursion set involving a unique constraint and where
423 the uncertainty is summarized via its expectation. Nevertheless, as perspective, other reliability
424 measures may also be of great interest. For example, one may be interested in ensuring a certain
425 level of reliability with high probability or in considering multiple constraints, e.g., on the mean
426 and the variance.

427 **References**

- 428 [1] Julien Bect, David Ginsbourger, Ling Li, Victor Picheny, and Emmanuel Vazquez. Sequential
429 design of computer experiments for the estimation of a probability of failure. Statistics and
430 Computing, 22(3):773–793, 2012.
- 431 [2] Clément Chevalier, Julien Bect, David Ginsbourger, Emmanuel Vazquez, Victor Picheny, and
432 Yann Richet. Fast parallel kriging-based stepwise uncertainty reduction with application to
433 the identification of an excursion set. Technometrics, 56(4):455–465, 2014.
- 434 [3] David Bolin and Finn Lindgren. Excursion and contour uncertainty regions for latent gaussian
435 models. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 77(1):
436 85–106, 2015.
- 437 [4] Joshua P. French, Stephan R. Sain, et al. Spatio-temporal exceedance locations and confidence
438 regions. The Annals of Applied Statistics, 7(3):1421–1449, 2013.
- 439 [5] Loic Le Gratiet and Claire Cannamela. Kriging-based sequential design strategies using fast
440 cross-validation techniques with extensions to multi-fidelity computer codes. working paper
441 or preprint, October 2012. URL <https://hal.archives-ouvertes.fr/hal-00744432>.
- 442 [6] Robert B. Gramacy and Daniel W. Apley. Local gaussian process approximation for large
443 computer experiments. Journal of Computational and Graphical Statistics, 24(2):561–578,
444 2015. doi: 10.1080/10618600.2014.914442.
- 445 [7] Haitao Liu, Jianfai Cai, and Yew-Soon Ong. An adaptive sampling approach for kriging
446 metamodeling by maximizing expected prediction error. Computers & Chemical Engineering,
447 106:171 – 182, 2017. ISSN 0098-1354. doi: [https://doi.org/10.1016/j.compchemeng.2017.](https://doi.org/10.1016/j.compchemeng.2017.05.025)
448 05.025. URL <http://www.sciencedirect.com/science/article/pii/S009813541730234X>.
449 ESCAPE-26.
- 450 [8] Aikaterini P. Kyprioti, Jize Zhang, and Alexandros A. Taflanidis. Adaptive design of ex-
451 periments for global kriging metamodeling through cross-validation information. Structural
452 and Multidisciplinary Optimization, 62(3):1135–1157, Sep 2020. ISSN 1615-1488. doi:
453 10.1007/s00158-020-02543-1. URL <https://doi.org/10.1007/s00158-020-02543-1>.
- 454 [9] Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient global optimization of
455 expensive black-box functions. J. Global Optim., 13(4):455–492, 1998.
- 456 [10] Chunna Li and Qifeng Pan. Adaptive optimization methodology based on kriging modeling
457 and a trust region method. Chinese Journal of Aeronautics, 32(2):281–295, 2019. ISSN 1000-
458 9361. doi: <https://doi.org/10.1016/j.cja.2018.11.012>. URL [https://www.sciencedirect.](https://www.sciencedirect.com/science/article/pii/S100093611930041X)
459 [com/science/article/pii/S100093611930041X](https://www.sciencedirect.com/science/article/pii/S100093611930041X).
- 460 [11] Youssef Diouane, Victor Picheny, Rodolphe Le Riche, and Alexandre S. Di Perrotolo. Trego:
461 a trust-region framework for efficient global optimization. Journal of Global Optimization,
462 86(1):1–23, May 2023. ISSN 1573-2916. doi: 10.1007/s10898-022-01245-w. URL <https://doi.org/10.1007/s10898-022-01245-w>.
- 464 [12] Roman Garnett. Bayesian Optimization. Cambridge University Press, 2023. doi: 10.1017/
465 9781108348973.
- 466 [13] Mohamed Reda El Amri, Céline Helbert, Olivier Lepreux, Miguel Munoz Zuniga, Clémentine
467 Prieur, and Delphine Sinoquet. Data-driven stochastic inversion via functional quantiza-
468 tion. Statistics and Computing, 30(3):525–541, 2020. doi: 10.1007/s11222-019-09888-8. URL
469 <https://doi.org/10.1007/s11222-019-09888-8>.

- 470 [14] Dario Azzimonti, David Ginsbourger, Clément Chevalier, Julien Bect, and Yann Richet. Adaptive
471 design of experiments for conservative estimation of excursion sets. Technometrics, 63
472 (1):13–26, 2021. doi: 10.1080/00401706.2019.1693427.
- 473 [15] Clément Duhamel, Céline Helbert, Miguel Munoz Zuniga, Clémentine Prieur, and Delphine
474 Sinoquet. A SUR version of the Bichon criterion for excursion set estimation. Statistics and
475 Computing, 33(2):41, February 2023. ISSN 1573-1375. doi: 10.1007/s11222-023-10208-4.
476 URL <https://doi.org/10.1007/s11222-023-10208-4>.
- 477 [16] Julien Bect, David Ginsbourger, L. Li, Victor Picheny, and Emmanuel Vazquez. Sequential
478 design of computer experiments for the estimation of a probability of failure. Statistics and
479 Computing, 22:773–793, 2012.
- 480 [17] Nicolas Lelièvre, Pierre Beaurepaire, Cécile Mattrand, and Nicolas Gayton. Ak-mcsi: A
481 kriging-based method to deal with small failure probabilities and time-consuming models.
482 Structural Safety, 73:1–11, 2018.
- 483 [18] Alexis Cousin, Josselin Garnier, Martin Guiton, and Miguel Munoz Zuniga. A two-step proce-
484 dure for time-dependent reliability-based design optimization involving piece-wise stationary
485 gaussian processes. Structural and Multidisciplinary Optimization, 65(120), 2022.
- 486 [19] Ferenc Huszár and David Duvenaud. Optimally-weighted herding is bayesian quadrature.
487 Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence, page
488 377–386, 2012.
- 489 [20] François-Xavier Briol, Chris J. Oates, Marc Girolami, Michael A. Osborne, and Dino Sejdin-
490 novic. Rejoinder: Probabilistic Integration: A Role in Statistical Computation? Statistical
491 Science, 34(1):38–42, 2019. ISSN 0883-4237. URL [https://www.jstor.org/stable/
492 26771030](https://www.jstor.org/stable/26771030). Publisher: Institute of Mathematical Statistics.
- 493 [21] Motonobu Kanagawa and Philipp Hennig. Convergence Guarantees for Adaptive Bayesian
494 Quadrature Methods. In Advances in Neural Information Processing Systems, volume 32.
495 Curran Associates, Inc., 2019. URL [https://proceedings.neurips.cc/paper/2019/hash/
496 165a59f7cf3b5c4396ba65953d679f17-Abstract.html](https://proceedings.neurips.cc/paper/2019/hash/165a59f7cf3b5c4396ba65953d679f17-Abstract.html).
- 497 [22] Luc Pronzato and Anatoly Zhigljavsky. Bayesian Quadrature, Energy Minimization, and
498 Space-Filling Design. SIAM/ASA Journal on Uncertainty Quantification, 8(3):959–1011,
499 January 2020. doi: 10.1137/18M1210332. URL [https://epubs.siam.org/doi/10.1137/
500 18M1210332](https://epubs.siam.org/doi/10.1137/18M1210332). Publisher: Society for Industrial and Applied Mathematics.
- 501 [23] Jan N. Fuhg, Amélie Fau, and Udo Nackenhorst. State-of-the-art and comparative review of
502 adaptive sampling methods for kriging. Archives of Computational Methods in Engineering,
503 28:2689–2747, 2021.
- 504 [24] Stefano Marelli and Bruno Sudret. An active-learning algorithm that combines sparse poly-
505 nomial chaos expansions and bootstrap for structural reliability analysis. Structural Safety,
506 75:67–74, 2018.
- 507 [25] Punit Kumar and Atul Gupta. Active learning query strategies for classification, regression,
508 and clustering: A survey. Journal of Computer Science and Technology, 35(4):913–945, Jul
509 2020. ISSN 1860-4749. doi: 10.1007/s11390-020-9487-4. URL [https://doi.org/10.1007/
510 s11390-020-9487-4](https://doi.org/10.1007/s11390-020-9487-4).
- 511 [26] Clément Chevalier and David Ginsbourger. Fast computation of the multi-points expected
512 improvement with applications in batch selection. In International Conference on Learning
513 and Intelligent Optimization, pages 59–69. Springer, 2013.

- 514 [27] Emmanuel Vazquez and Julien Bect. A sequential bayesian algorithm to estimate a probability
515 of failure. IFAC Proceedings Volumes, 42(10):546–550, 2009.
- 516 [28] Janis Janusevskis and Rodolphe Le Riche. Simultaneous kriging-based estimation and opti-
517 mization of mean response. Journal of Global Optimization, 55(2):313–336, 2013.
- 518 [29] Brian J Williams, Thomas J Santner, and William I Notz. Sequential design of computer
519 experiments to minimize integrated response functions. Statistica Sinica, pages 1133–1152,
520 2000.
- 521 [30] Olivier Le Maître and Omar M. Knio. Spectral Methods for Uncertainty Quantification. Sci-
522 entific Computation. Springer, Dordrecht, 2010. doi: 10.1007/978-90-481-3520-2.
- 523 [31] Ilya Molchanov. Theory of random sets. Springer Science & Business Media, 2006.
- 524 [32] O Yu Vorob'ev. Srednemernoje modelirovanie (mean-measure modelling), 1984.
- 525 [33] Oleg Yu. Vorobyev and Natalia A. Lukyanova. A mean probability event for a set of events.
526 Mpra paper, University Library of Munich, Germany, 2013. URL [https://EconPapers.
527 repec.org/RePEc:pra:mprapa:48101](https://EconPapers.repec.org/RePEc:pra:mprapa:48101).
- 528 [34] Clément Chevalier. Fast uncertainty reduction strategies relying on Gaussian process models.
529 Theses, Universität Bern, September 2013. URL [https://theses.hal.science/
530 tel-00879082](https://theses.hal.science/tel-00879082).
- 531 [35] Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-optimal sensor placements in gaussian
532 processes: Theory, efficient algorithms and empirical studies. J. Mach. Learn. Res., 9:235–284,
533 jun 2008. ISSN 1532-4435.
- 534 [36] Olivier Roustant, David Ginsbourger, and Yves Deville. Dicekriging, diceoptim: Two r pack-
535 ages for the analysis of computer experiments by kriging-based metamodeling and optimiza-
536 tion. Journal of Statistical Software, 51(1):1–55, 2012. URL [http://www.jstatsoft.org/
537 v51/i01/](http://www.jstatsoft.org/v51/i01/).
- 538 [37] Walter Mebane Jr. and Jasjeet Sekhon. Genetic optimization using derivatives: The rgenoud
539 package for r. Journal of Statistical Software, 42(11):1–26, 2011. ISSN 1548-7660. doi:
540 10.18637/jss.v042.i11. URL <https://www.jstatsoft.org/v042/i11>.
- 541 [38] Martin Schlather. Models for stationary max-stable random fields. Extremes, 5(1):33–44,
542 2002.
- 543 [39] Anthony Bonfils, Yann Creff, Olivier Lepreux, and Nicolas Petit. Closed-loop control of a scr
544 system using a nox sensor cross-sensitive to nh3. IFAC Proceedings Volumes, 45(15):738–743,
545 2012.

546 **Appendix A. 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 $\mathbb{X} \times \mathbb{R}^m$. Considering a m -truncation of the random process KL expansion, we reduce the hyperspace on which the GP is defined. 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 $\mathbf{u} = (\langle \mathbf{v}, \hat{\psi}_1 \rangle, \dots, \langle \mathbf{v}, \hat{\psi}_m \rangle)^\top$ and $\tilde{\mathbf{u}} = (\langle \mathbf{v}, \hat{\psi}_{m+1} \rangle, \dots, \langle \mathbf{v}, \hat{\psi}_L \rangle)^\top$. In that setting $f(\mathbf{x}, \mathbf{V})$ can be expressed as

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

547 where $\hat{\mathbf{V}}_L$ is the empirical version (estimated from C^N) of the KL approximation of \mathbf{V} given by
 548 (4) (replacing m by L), $\hat{\Phi}_L = (\hat{\psi}_1, \dots, \hat{\psi}_L)^\top$ and ϵ_T is the error associated to the KL truncation
 549 and empirical approximation, supposed small by construction.

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

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

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

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

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

$$f(\mathbf{x}, \mathbf{V}) = f(\mathbf{x}, (\mathbf{U}, \tilde{\mathbf{u}}(\mathbf{U}))\hat{\Phi}_L) + \epsilon \tag{A.1}$$

with $\epsilon = \epsilon_T + \epsilon_P + \epsilon_E$. According to this last equation, the modelling assumption in Section 2.2 should include a noise term. However, the estimation of this heteroscedastic noise comes with an extra estimation cost and as it can be seen in Figure A.14, no significant model improvement is observed. Indeed in Figure A.14, for $m = 2$, we present the evolution of the symmetric difference for the noisy GP model $Z_{(\mathbf{x}, \mathbf{u})}$ introduced from equation (A.1) when the noise ϵ is Gaussian and heteroscedastic with a variance function of (\mathbf{x}, \mathbf{u}) :

$$\tau^2(\mathbf{x}, \mathbf{u}) = \text{Var}_{\tilde{\mathbf{U}}}[f(\mathbf{x}, (\mathbf{u}, \tilde{\mathbf{U}}(\mathbf{u}))\hat{\Phi}_L)|\mathbf{U} = \mathbf{u}].$$

Moreover, supposing \mathbf{V} Gaussian or "nearly Gaussian", that is assuming that $\tilde{\mathbf{U}}$ can be considered in first approximation as independent of \mathbf{U} , then $\tau^2(\mathbf{x}, \mathbf{u})$ can be estimated by

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

554 where $l = 5$ and the \mathbf{V}_k^{Quant} are greedy functional quantizers and w_k associated weights (see [13]
 555 for more details). These quantizers are built from a set of N curves $\{(\mathbf{u}, \tilde{\mathbf{u}}_k)\hat{\Phi}_L, k = 1, \dots, N\}$
 556 where $\tilde{\mathbf{u}}_k$ are independent samples of $\tilde{\mathbf{U}}$ which in practice are uniformly sampled in the finite set
 557 $\bar{\mathcal{G}}_{m,L} = \{\tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_N\}$ where $\tilde{\mathbf{u}}_i = (\langle \tilde{v}_i, \hat{\psi}_{m+1} \rangle, \dots, \langle \tilde{v}_i, \hat{\psi}_L \rangle)$. Numerically we select 20 (\mathbf{x}, \mathbf{u}) -
 558 points from the initial DoE set of size $n = 30$ and estimate the corresponding $\hat{\tau}^2$. To avoid further
 559 estimation of τ^2 at new locations (the remaining DoE points and during the infill strategy), we
 560 build a second GP model of $\log(\hat{\tau}^2)$ based on the 20 initial estimations. Finally the noisy GP
 561 model Z is built using as noise variance $\exp(\hat{\log}(\hat{\tau}^2))$. Overall we need additional $l \times 20 = 100$
 562 costly evaluations of f to estimate the heteroscedastic noise.

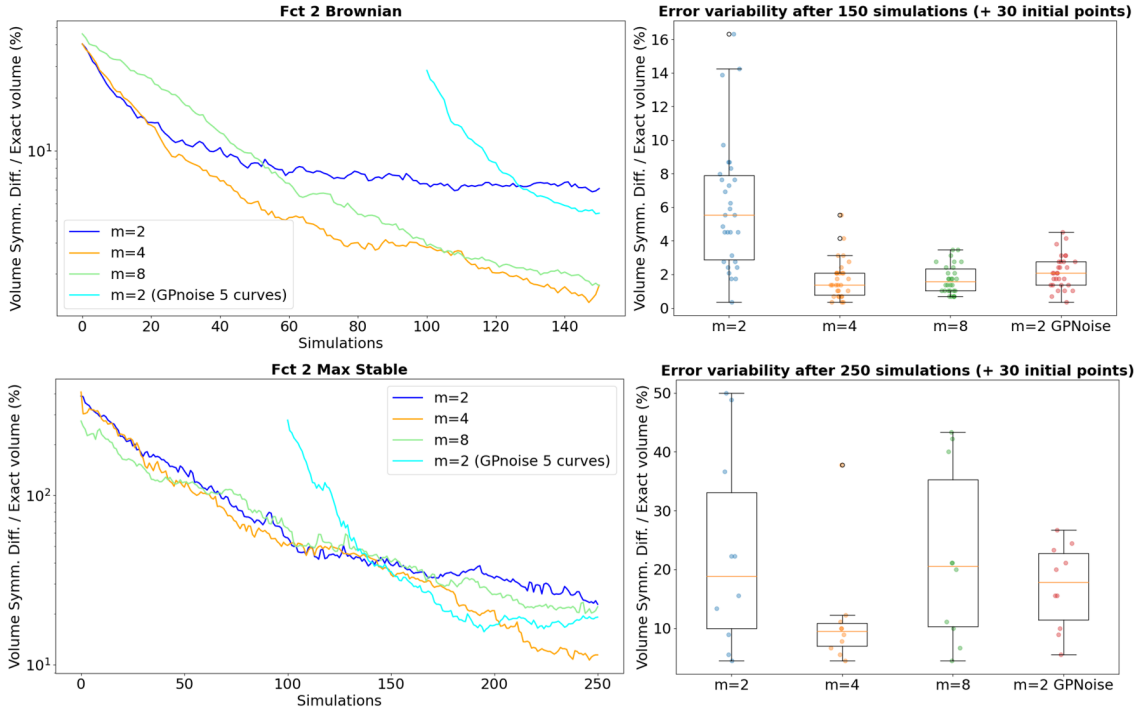


Figure A.14: Function 2 with brownian (top) and max-stable processes (bottom) with a comparison with the heteroscedastic GP model. Convergence of Algorithm 1 for $m = \{2, 4, 8\}$. Left: mean of the symmetric difference vs. number of simulator calls. The mean is taken over the independent runs of initial 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.

563 In Figure A.14 we notice that compared to the homoscedastic model with $m = 2$, the model
564 with heteroscedastic noise achieves a faster symmetric difference volume reduction but the overcost,
565 for the variance estimation, makes this approach interesting only for a large simulation budget:
566 at least 130 simulations. For the brownian case, on function 2, the homoscedastic models with
567 higher m still perform better for a budget up to 150 than the heteroscedastic one. A model with a
568 small m , that is to say with a rough truncation error, involves a larger bias. Nevertheless, refining
569 the heteroscedastic noise estimation should bring the method to a similar level but much further
570 on the axis corresponding to the number of simulations. But on function 2 with a max-stable
571 process, the heteroscedastic model slightly outperforms the homoscedastic models ($m = 2, 4, 8$)
572 when approaching the 150 simulations (Figure A.14). We can understand this improvement by
573 the fact that even with higher m a homoscedastic model does not make up for a wider truncation
574 error which is better approximated by a heteroscedastic model. Note that it is possible to relax
575 the "nearly Gaussian" hypothesis on \mathbf{V} . In that case the same kind of heteroscedastic variance
576 estimator could be used but would require an empirical estimation of the conditional distribution
577 of $\tilde{\mathbf{U}}|\mathbf{U}$ which seems difficult in the context of our partial knowledge of \mathbf{V} imposing on us to work
578 on finite predefined sets \mathcal{G} and $\tilde{\mathcal{G}}_{m,L}$.

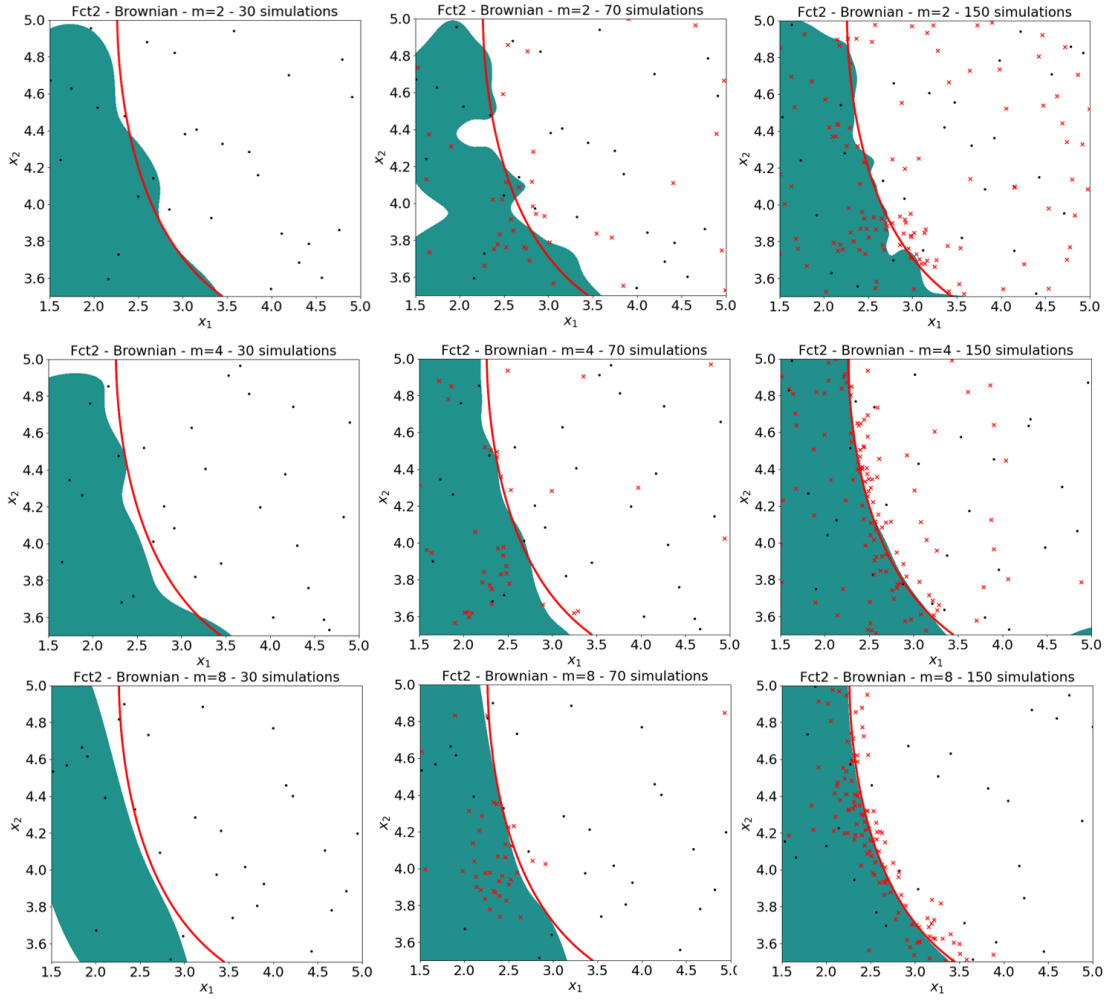


Figure A.15: Feasible domain estimation for analytical example 2 with brownian motion in green and its boundary in red for 3 different iterations (30, 70 and 150 from left to right) and for the 3 values of $m = 2, 4$ and 8 (from top to bottom). The black dots are the x coordinates of the points in the initial design of experiments, the red crosses are the additional points chosen by the algorithm.

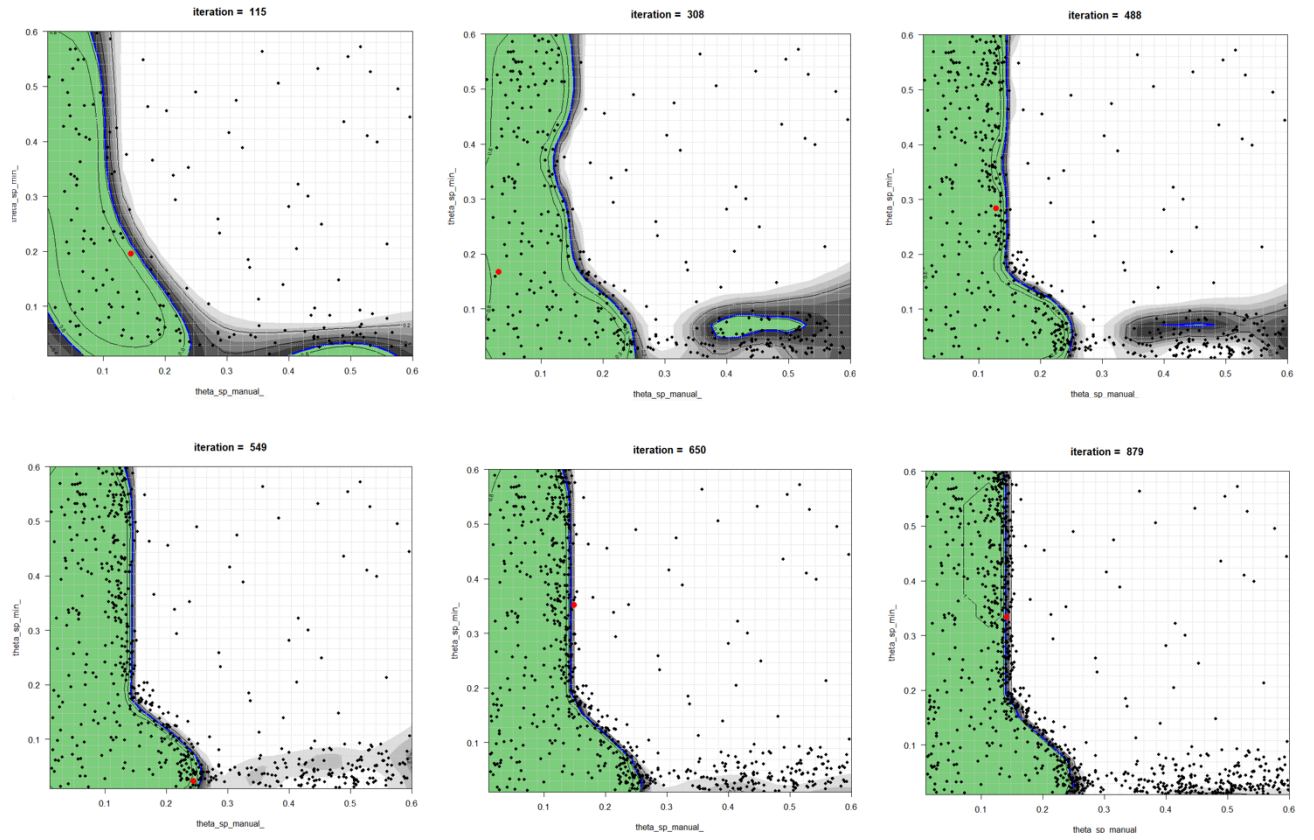


Figure A.16: SCR pollution control system. The estimated feasible domain at 6 different iterations.