



A SUR version of the Bichon criterion for excursion set estimation

Clément Duhamel, Céline Helbert, Miguel Munoz Zuniga, Clémentine Prieur, Delphine Sinoquet

► To cite this version:

Clément Duhamel, Céline Helbert, Miguel Munoz Zuniga, Clémentine Prieur, Delphine Sinoquet. A SUR version of the Bichon criterion for excursion set estimation. *Statistics and Computing*, 2023, 33 (2), pp.41. 10.1007/s11222-023-10208-4. hal-03780559v1

HAL Id: hal-03780559

<https://hal.science/hal-03780559v1>

Submitted on 19 Sep 2022 (v1), last revised 14 Mar 2023 (v2)

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.

A SUR version of the Bichon criterion for excursion set estimation

Clément Duhamel ¹, Céline Helbert ², Miguel Munoz Zuniga ³,
Clémentine Prieur ¹ and Delphine Sinoquet ³

September 19, 2022

Abstract

Many model inversion problems occur in industry. These problems consist in finding the set of parameter values such that a certain quantity of interest respects a constraint, for example remains below a threshold. In general, the quantity of interest is the output of a simulator, costly in computation time.

An effective way to solve this problem is to replace the simulator by a Gaussian process regression, with an experimental design enriched sequentially with a well chosen acquisition criterion. Different inversion-adapted criteria exist such as the Bichon (also known as Expected Feasibility Function) and deviation number criteria. There also exist a class of enrichment strategies (Stepwise Uncertainty Reduction) which select the next point by measuring the expected uncertainty reduction induced by its selection.

In this paper we propose a SUR version of the Bichon criterion. An explicit formulation of the criterion is given and test comparisons show good performances on classical test functions.

¹Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP*, LJK, 38000 Grenoble, France

* Institute of Engineering Univ. Grenoble Alpes

²Univ. Lyon, Ecole Centrale de Lyon, CNRS UMR 5208, Institut Camille Jordan, 69130 Ecully, France

³IFP Énergies Nouvelles, 92852 Reuil-Malmaison, France

1 Introduction

Nowadays, many industrial issues are related to a problem of excursion set estimation for instance, to find feasible solutions of complex optimal design problems. This problem consists in finding, for a given model, the set of input parameter values such that a quantity of interest defined from its outputs respects a constraint, for example remains below a threshold. In general, the quantity of interest is an output of a simulator, computationally expensive, which is often a black-box function. The problem is also known as an inversion problem (Chevalier [2013]). For example, the application to a vehicle pollution control system, allowing compliance with pollutant emission norms was studied in El Amri et al. [2020].

An effective way to solve this kind of problems is to replace the costly black-box function of interest by a surrogate model based on Gaussian process. This model is built thanks to a sequential Design of Experiments (DoE), whose points are chosen accordingly to the optimization of an acquisition criterion (see for example Bect et al. [2012] and Moustapha et al. [2021]). Acquisition criteria suitable for inversion include: the deviation number denoted U (Echard et al. [2011]), the Bichon criterion also known as Expected Feasibility Function (Bichon et al. [2008]), and the Ranjan criterion (Ranjan et al. [2008]). These two last criteria are adaptations of the classical optimization-oriented Expected Improvement criterion (Jones et al. [1998]) for excursion set estimation.

In addition, there is a more elaborate and in general more efficient class of criteria that anticipate the impact of adding new points to DoE: the SUR strategies, for Stepwise Uncertainty Reduction (Bect et al. [2012]). For example, SUR strategies based on the volume of the excursion set can be cited as particularly suitable for the inversion framework. It is shown in Bect et al. [2012] that those SUR criteria provide better performances compared to direct criteria.

Chevalier [2013] introduced a SUR strategy based on Vorob'ev random set theory (Molchanov [2005]), that goes beyond taking into account the volume of excursion sets like other SUR strategies. However, we have noticed that SUR Vorob'ev criterion (SUR Vorob'ev) is not robust enough in the sense that with a reasonable number of simulations it sometimes misses some of connected components of the set. Moreover, SUR Vorob'ev requires some approximations about the Vorob'ev threshold. Therefore we propose to tackle these issues by a SUR version of the Bichon criterion (SUR Bichon), which is easier to set up and more robust than SUR Vorob'ev. It should be noted that a SUR version of the U criterion could not have been envisaged due its nature and its dependence on the prediction mean.

This article is divided into three main sections. In section 2, the framework of excursion set estimation is recalled. Details on the construction of the sequential DoE based on Gaussian process regression are given as well as details on the Bichon acquisition criterion. Section 3 is dedicated to Stepwise Uncertainty Reduction (SUR) strategies. Our new criterion, a SUR version of the Bichon criterion is introduced in Section 4. Both theoretical and numerical aspects are discussed with respectively a simplified and easy-to-implement formulation and numerical tests of SUR Bichon performances, compared to those of SUR Vorob'ev for several analytical examples. Appendices present technical proofs, theoretical results on kriging and bases of Vorob'ev theory.

2 The framework for estimating excursion sets

2.1 Some reminders on Gaussian process regression

Let \mathbb{X} be a compact set of \mathbb{R}^d ($d \in \mathbb{N}^*$) and $g : \mathbb{X} \rightarrow \mathbb{R}$ a black-box function, whose analytical expression is unknown but which can be evaluated at any point of \mathbb{X} at a heavy computational cost. The objective of an excursion set estimation problem is to estimate the domain defined by

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

with T a fixed threshold, while limiting the number of costly evaluations of g .

Surrogate models, also known as metamodels, are approximations of the output of the simulator built from a sample of simulations and that are not expensive to evaluate. Therefore they can replace the original expensive simulator in a time-saving manner. Among surrogate models, Gaussian Process Regression (GPR) is very popular: g is considered as a realization of a Gaussian process (GP) ξ defined on a probabilistic space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e. $g(\mathbf{x}) = \xi(\mathbf{x}, \omega)$ for a given ω in Ω . This type of surrogate models gives, in addition to a prediction, an associated prediction error estimate.

More precisely, the process is written as the sum of a deterministic part and a stochastic part:

$$\xi(\mathbf{x}) := m(\mathbf{x}) + \mathbf{Z}(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{X} \quad (2)$$

with m the trend of ξ (deterministic part) and \mathbf{Z} a stationary GP, of zero mean, known covariance kernel $k : \mathbb{X}^2 \rightarrow \mathbb{R}$ and in particular variance function $\sigma^2(\mathbf{x}) := k(\mathbf{x}, \mathbf{x})$ for any \mathbf{x} in \mathbb{X} (stochastic part). The choice of the trend m is usually a linear combination of functions $\mathbf{f} := \{f_i\}_{i=1}^l$ with coefficients $\boldsymbol{\beta} := \{\beta_i\}_{i=1}^l$. The choice of the covariance kernel associated with the GP \mathbf{Z} is crucial since it determines the predictor regularity. Different choices are detailed in Section 4.2.

For an initial design of experiments $\chi_n := (\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$ belonging to \mathbb{X}^n , let's denote $g(\chi_n) := (g(\mathbf{x}^{(1)}), \dots, g(\mathbf{x}^{(n)}))^T$ the evaluations of g on χ_n and \mathcal{E}_n the event $\xi(\chi_n) = g(\chi_n)$, corresponding to n evaluations of ξ . $K := (k(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}))_{1 \leq i, j \leq n}$ is the covariance matrix on χ_n and $\mathbf{k}(\mathbf{x})$ the covariance vector between \mathbf{x} and χ_n defined by $\mathbf{k}(\mathbf{x}) := (k(\mathbf{x}, \mathbf{x}^{(1)}), \dots, k(\mathbf{x}, \mathbf{x}^{(n)}))^T$ for any \mathbf{x} in \mathbb{X} . We denote $\mathbf{f}(\mathbf{x}) := (f_1(\mathbf{x}), \dots, f_l(\mathbf{x}))^T$ the evaluation vector of \mathbf{f} on \mathbf{x} defining the trend and $\mathbb{F} \in \mathbb{R}^{n \times l}$ the matrix with $\mathbf{f}(\mathbf{x}^{(i)})^T$ as i^{th} row. The process ξ conditioned on the event \mathcal{E}_n is then still Gaussian (O'Hagan [1978]) with mean, variance and covariance respectively denoted m_n , σ_n^2 , and k_n given by

$$m_n(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} + \mathbf{k}(\mathbf{x})^T K^{-1} (g(\chi_n) - \mathbb{F} \boldsymbol{\beta}), \quad (3)$$

$$\sigma_n^2(\mathbf{x}) = \sigma^2(\mathbf{x}) - \mathbf{k}(\mathbf{x})^T K^{-1} \mathbf{k}(\mathbf{x}), \quad (4)$$

$$k_n(\mathbf{x}, \mathbf{x}') = k(\mathbf{x}, \mathbf{x}') - \mathbf{k}(\mathbf{x})^T K^{-1} \mathbf{k}(\mathbf{x}'). \quad (5)$$

We notice that the best linear unbiased predictor (BLUP) (with respect to mean quadratic error) is given by (3) with variance (4) and covariance (5).

When $\boldsymbol{\beta}$ is unknown and estimated by $\hat{\boldsymbol{\beta}} := (\mathbb{F}^T K^{-1} \mathbb{F})^{-1} \mathbb{F}^T K^{-1} g(\chi_n)$ the maximum likelihood estimator (MLE), formulas (3), (4) and (5) become

$$m_n(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \hat{\boldsymbol{\beta}} + \mathbf{k}(\mathbf{x})^T K^{-1} (g(\chi_n) - \mathbb{F} \hat{\boldsymbol{\beta}}), \quad (6)$$

$$\sigma_n^2(\mathbf{x}) = \sigma^2(\mathbf{x}) - \mathbf{k}(\mathbf{x})^T K^{-1} \mathbf{k}(\mathbf{x}) + (\mathbf{f}(\mathbf{x})^T - \mathbf{k}(\mathbf{x})^T K^{-1} \mathbb{F}) (\mathbb{F}^T K^{-1} \mathbb{F})^{-1} (\mathbf{f}(\mathbf{x})^T - \mathbf{k}(\mathbf{x})^T K^{-1} \mathbb{F})^T, \quad (7)$$

$$k_n(\mathbf{x}, \mathbf{x}') = k(\mathbf{x}, \mathbf{x}') - \mathbf{k}(\mathbf{x})^T K^{-1} \mathbf{k}(\mathbf{x}') + (\mathbf{f}(\mathbf{x})^T - \mathbf{k}(\mathbf{x})^T K^{-1} \mathbb{F}) (\mathbb{F}^T K^{-1} \mathbb{F})^{-1} (\mathbf{f}(\mathbf{x}')^T - \mathbf{k}(\mathbf{x}')^T K^{-1} \mathbb{F})^T. \quad (8)$$

Their interpretation as conditional expectation, covariance and variance is still possible in a bayesian context with a non informative prior distribution on the parameter $\boldsymbol{\beta}$ (Helbert et al. [2009]).

In order to save costly evaluations of function g , a sequential strategy of enrichment of the DoE is classically used (see e.g. Ginsbourger [2017]). Figure 1 provides a generic scheme of a sequential strategy. The stopping criteria are the budget of simulations and a threshold on the remaining uncertainty on the estimation of the excursion set.

Among enrichment criteria, one can distinguish criteria that lead to an overall improvement of the model from goal-oriented criteria, which are adapted to particular frameworks such as optimization or inversion. The classical Mean Squared Error (MSE) criterion (Jin et al. [2002]), aims to select the point which has the highest prediction variance, as well as its integral versions IMSE and MMSE (Picheny et al. [2010]) standing for Integral and Maximum Mean Squared Error. Among direct goal oriented criteria, classical EI (Expected Improvement) from Jones et al. [1998], is adapted

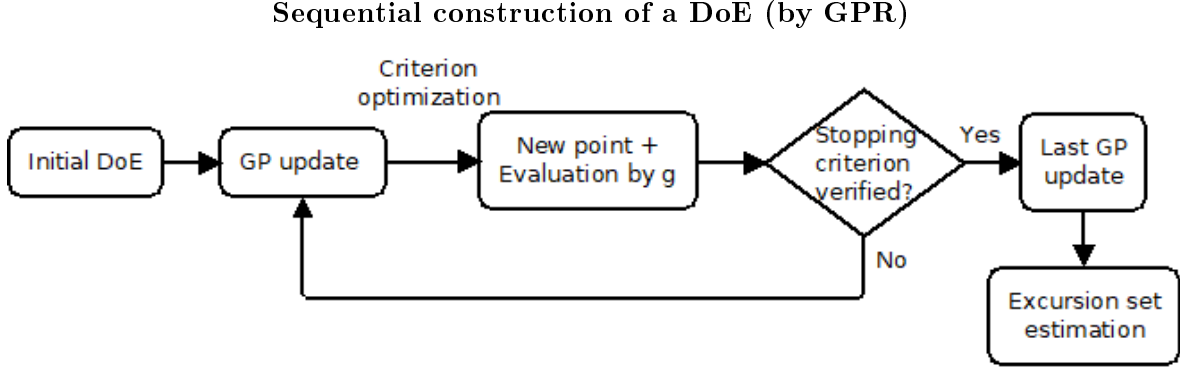


Figure 1: Diagram of the sequential DoE construction, coupled with GPR.

for optimization. Several inversion-adapted criteria can be cited: deviation number (Echard et al. [2011]), ratio of the distance of the model evaluations to the threshold to the kriging standard deviation; Bichon criterion also known as Expected Feasibility Function (EFF) (Bichon et al. [2008]), and the Ranjan criterion (Ranjan et al. [2008]). The first one is one of the simplest criterion for inversion.

2.2 Towards more exploration: Bichon criterion

Bichon criterion, originally presented in Bichon et al. [2008] is a direct goal-oriented criterion for the DoE enrichment. This criterion is an adaptation of the EI (Expected Improvement) criterion from Jones et al. [1998], used in the context of an optimization problem, to the inversion framework.

As a reminder, the original idea of the EI criterion is to select a point \mathbf{x} that allows an improvement of $g_{\min} - \xi(\mathbf{x})$ with g_{\min} the current observed minimum on the DoE, while taking into account the uncertainty of the surrogate model. The idea behind Bichon criterion is to adapt this strategy to the excursion set estimation framework by considering both the variability of the surrogate model and the potential improvement in the knowledge of the excursion set boundary. This is the exploration-exploitation compromise.

To introduce Bichon criterion, it is necessary to define, for a fixed \mathbf{x} belonging to \mathbb{X} , a random variable $\text{FF}(\mathbf{x})$ (Feasibility Function) defined by

$$\begin{aligned} \text{FF}(\mathbf{x}) &:= \varepsilon(\mathbf{x}) - \min \{ |T - \xi(\mathbf{x})|, \varepsilon(\mathbf{x}) \} \\ &= (\varepsilon(\mathbf{x}) - |T - \xi(\mathbf{x})|)^+, \end{aligned} \quad (9)$$

with $(\cdot)^+ := \max(\cdot, 0)$. This function represents the distance of the surrogate model to the bounds of the interval $[T - \varepsilon(\mathbf{x}), T + \varepsilon(\mathbf{x})]$ only if the surrogate model belongs to this interval and is 0 otherwise. In practice, the interval width $\varepsilon(\mathbf{x})$ is chosen proportionnal to the kriging standard deviation $\sigma_n(\mathbf{x})$, leading in particular to a null value of the criterion for the points already present in the DoE. An example is given in Figure 2. The feasibility function is drawn for one sample path of $\xi | \mathcal{E}_n$. Its maximization aims to select points close to the boundary of the excursion set estimate or points associated to high values of $\varepsilon(\mathbf{x})$.

Then, the average of FF over all sample paths, gives the Expected Feasibility Function (EFF)

$$\text{EFF}(\mathbf{x}) := \mathbb{E} \left[(\varepsilon(\mathbf{x}) - |T - \xi(\mathbf{x})|)^+ \mid \mathcal{E}_n \right]. \quad (10)$$

To interpret the criterion, it is possible to make an heuristic analogy with the theory of statistical tests (Dagnelie [1992]). Let \mathbf{x} be fixed, suppose that $\xi(\mathbf{x}) | \mathcal{E}_n \sim \mathcal{N}(m_n(\mathbf{x}), \sigma_n(\mathbf{x}))$ with $m_n(\mathbf{x})$

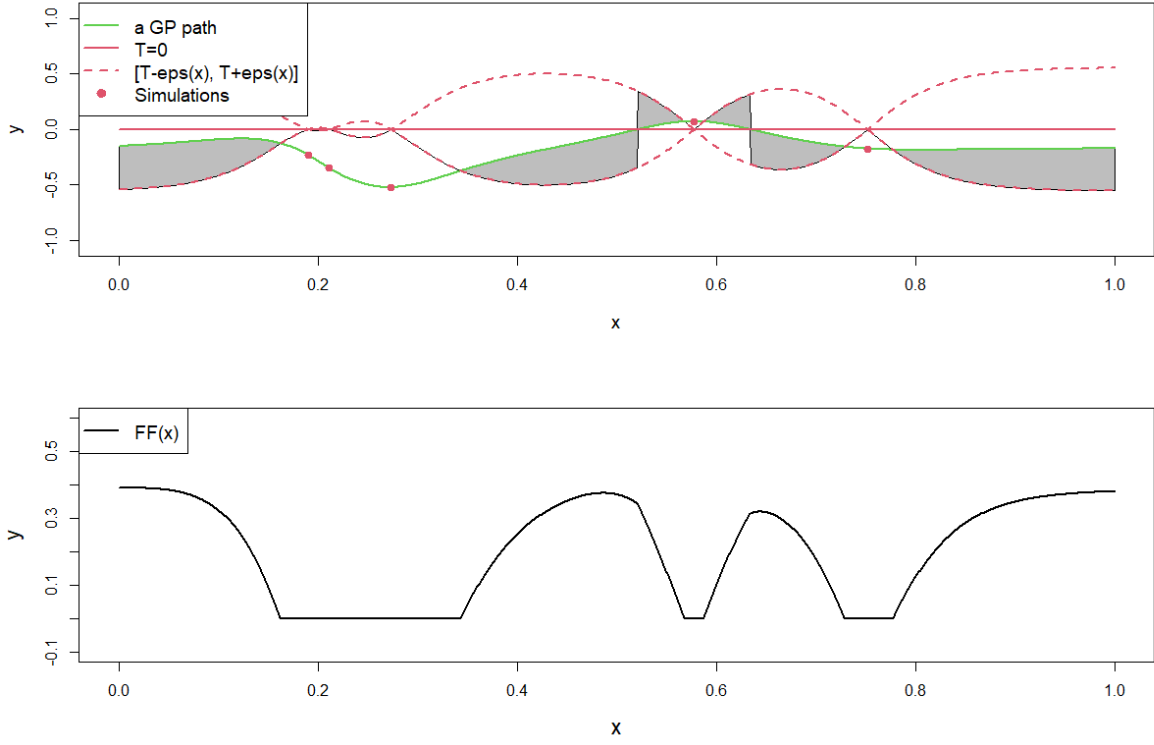


Figure 2: Representation of Feasibility Function (bottom) for a given example of a GP sample path conditioned on 5 evaluations of g function, a threshold T set to 0 and a given $\varepsilon(x)$ (top).

unknown and $\sigma_n(\mathbf{x}) > 0$ known and let us define the following statistical test

$$H_0 : m_n(\mathbf{x}) = T \quad \text{against} \quad H_1 : m_n(\mathbf{x}) \neq T \quad (11)$$

We choose $v_{\mathbf{x}} := \frac{\xi(\mathbf{x}) - T}{\sigma_n(\mathbf{x})} \Big| \mathcal{E}_n$ as the test statistic which follows standard normal distribution under H_0 . Consequently, if we want to refute Hypothesis H_0 at order α it is necessary that

$$|v_{\mathbf{x}}| > \kappa \quad (12)$$

with $\kappa := q_{1-\frac{\alpha}{2}}$ the quantile of order $1 - \frac{\alpha}{2}$ of the standard normal distribution.

However, what we wish to do is not to refute the hypothesis that $m_n(\mathbf{x}) = T$ but rather to select, among the \mathbf{x} for which the hypothesis H_0 is plausible i.e. $\kappa - |v_{\mathbf{x}}| > 0$, the \mathbf{x} for which the quantity $\kappa - |v_{\mathbf{x}}|$ is the largest in average. Multiplying $(\kappa - |v_{\mathbf{x}}|)^+$ by $\sigma_n(\mathbf{x})$ leads to Bichon criterion with $\varepsilon(\mathbf{x}) := \kappa \sigma_n(\mathbf{x})$, and has the effect to increase the exploration ability of the criterion.

Finally, an explicit formulation of the Bichon criterion (see Bect et al. [2012] for a proof) can be calculated based on the posterior kriging trend and variance, the threshold T and the width $\varepsilon(\mathbf{x})$

$$\begin{aligned} \text{EFF}(\mathbf{x}) = (m_n(\mathbf{x}) - T) & \left[2\phi\left(\frac{T - m_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) - \phi\left(\frac{T^- - m_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) - \phi\left(\frac{T^+ - m_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) \right] \\ & - \sigma_n(\mathbf{x}) \left[2\varphi\left(\frac{T - m_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) - \varphi\left(\frac{T^- - m_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) - \varphi\left(\frac{T^+ - m_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) \right] \\ & + \varepsilon(\mathbf{x}) \left[\phi\left(\frac{T^+ - m_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) - \phi\left(\frac{T^- - m_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) \right], \end{aligned} \quad (13)$$

with $T^\pm := T \pm \varepsilon(\mathbf{x})$, φ and Φ represent respectively the probability density and cumulative distribution functions of the reduced centered normal distribution. In practice, the enrichment of the DoE is done by maximizing the criterion given by Equation (13). The main objective of this work is to propose a SUR version of this direct goal-oriented criterion, in order to obtain a SUR method that is simpler to set up and more robust than SUR Vorob'ev.

3 SUR strategies based on Gaussian Process Regression

SUR strategies aim at maximizing the mean uncertainty reduction induced by new evaluations. Let us introduce an uncertainty measure \mathcal{H}_n , that depends on the event \mathcal{E}_n . From this uncertainty measure, the conditional measure $\mathcal{H}_{n+1}(\mathbf{x})$ is defined as the measure \mathcal{H}_n , conditionally on the addition of the point $(\mathbf{x}, \xi(\mathbf{x}))$ to the DoE. This conditional measure is a real random variable measurable with respect to $\xi(\mathbf{x})|\mathcal{E}_n$. The associated SUR strategy is then defined by

$$\mathbf{x}^{(n+1)} \in \operatorname{argmin}_{\mathbf{x} \in \mathbb{X}} \mathcal{J}_n(\mathbf{x}) \quad \text{with} \quad \mathcal{J}_n(\mathbf{x}) := \mathbb{E}[\mathcal{H}_{n+1}(\mathbf{x})]. \quad (14)$$

Note that the expectation in $\mathcal{J}_n(\mathbf{x})$ is relative to the distribution of $\xi(\mathbf{x})|\mathcal{E}_n$.

Since SUR strategies are characterized by their uncertainty function \mathcal{H}_n and therefore also by the corresponding sampling criterion \mathcal{J}_n , in the following, the abusive name of SUR sampling criterion will be used to designate \mathcal{J}_n . It is possible to easily generalize the SUR strategy definition to the case of simultaneous addition of q points (Chevalier [2013]). However in this article, only the special case of adding points one by one is presented. More details on SUR strategies and their origin from k -step lookahead strategies can be found in Bect et al. [2012].

The ξ -measurable random variable μ is defined by $\mu := \mathbb{P}_{\mathbb{X}}(\Gamma)$ with Γ the random set $\{\mathbf{x} \in \mathbb{X}, \xi(\mathbf{x}) \leq T\}$, $\mathbb{P}_{\mathbb{X}}$ a given σ -finite measure and $p_n(\mathbf{z}) := \mathbb{P}(\mathbf{z} \in \Gamma | \mathcal{E}_n)$ the coverage function of Γ knowing \mathcal{E}_n . Among the classical SUR sampling criteria based on GPR, we can quote for example the criterion using the variance of μ : $\mathcal{H}_n := \operatorname{Var}(\mu | \mathcal{E}_n)$ and one using the coverage function p_n : $\mathcal{H}_n := \int_{\mathbb{X}} p_n(\mathbf{z})(1 - p_n(\mathbf{z})) d\mathbb{P}_{\mathbb{X}}(\mathbf{z})$, both presented in Bect et al. [2012]. For the latter, the uncertainty is low when p_n is close to 0 or 1 and maximal when p_n is equal to 1/2. The conditional measure $\mathcal{H}_{n+1}(\mathbf{x})$ of these two criteria are respectively $\mathcal{H}_{n+1}(\mathbf{x}) := \operatorname{Var}(\mu | \xi(\mathbf{x}), \mathcal{E}_n)$ and $\mathcal{H}_{n+1}(\mathbf{x}) := \int_{\mathbb{X}} p_{n,\mathbf{x}}(\mathbf{z})(1 - p_{n,\mathbf{x}}(\mathbf{z})) d\mathbb{P}_{\mathbb{X}}(\mathbf{z})$ with $p_{n,\mathbf{x}}(\mathbf{z}) := \mathbb{P}(\mathbf{z} \in \Gamma | \xi(\mathbf{x}), \mathcal{E}_n)$.

More complex criteria requiring notions about the random set theory of Vorob'ev (Molchanov [2005]) are given in the literature. For example, the criterion associated with the Vorob'ev deviation, introduced in Chevalier [2013], can be cited. This latter will be used as a comparison reference in the numerical section. We refer to the appendix A for more details on Vorob'ev deviation (and more generally on Vorob'ev theory) and the associated SUR Vorob'ev method. Besides, recently a more conservative criterion also inspired by the Vorob'ev deviation, has been proposed in Azzimonti et al. [2021]. The latter is especially useful in industrial applications where the distinction between false positives and false negatives is critical, which is not the context of this paper.

Finally, SUR strategies are numerically more complex to implement, but are generally more efficient than direct goal oriented strategies (Bect et al. [2012] and Chevalier [2013]), for the same number of evaluations. A reduction in the numerical complexity of SUR strategies is frequently used through the use of kriging update formulas used in Chevalier [2013] (see Appendix B)

4 SUR Bichon criterion

4.1 Theoretical aspects

This section focuses on an adaptation of the Bichon criterion to a SUR strategy. The new SUR criterion we introduce can be implemented in the framework of GPR without any approximation

unlike SUR Vorob'ev (see Appendix A and Chevalier [2013]). Moreover, SUR Bichon corrects the lack of robustness observed when applying SUR Vorob'ev, at least on the test cases studied in Section 4.2.

Following the formalism of SUR strategies given in section 3, it is sufficient to define the uncertainty measure \mathcal{H}_n . This uncertainty measure is the mean of the Bichon criterion conditioned on \mathcal{E}_n , according to a given probability measure $\mathbb{P}_{\mathbb{X}}$:

$$\mathcal{H}_n := \int_{\mathbb{X}} \mathbb{E} \left[\left(\kappa \sigma_n(\mathbf{z}) - |T - \xi(\mathbf{z})| \right)^+ \middle| \mathcal{E}_n \right] d\mathbb{P}_{\mathbb{X}}(\mathbf{z}), \quad \text{with } \kappa > 0. \quad (15)$$

The corresponding conditional uncertainty measure is then defined by

$$\mathcal{H}_{n+1}(\mathbf{x}) := \int_{\mathbb{X}} \mathbb{E} \left[\left(\kappa \sigma_{n+1}(\mathbf{z}) - |T - \xi(\mathbf{z})| \right)^+ \middle| \xi(\mathbf{x}), \mathcal{E}_n \right] d\mathbb{P}_{\mathbb{X}}(\mathbf{z}), \quad \text{with } \kappa > 0, \quad (16)$$

with $\sigma_{n+1}(\mathbf{z})$ the kriging standard deviation given in Appendix B (Equation (31)) with $\mathbf{x}^{(n+1)} = \mathbf{x}$. $\mathcal{H}_{n+1}(\mathbf{x})$ is a real random variable $\xi(\mathbf{x})$ -measurable, conditioned on \mathcal{E}_n . In practice, the chosen probability measure $\mathbb{P}_{\mathbb{X}}$ is the Lebesgue measure restricted and normalized on \mathbb{X} and the coefficient κ is usually set to 1 like in the Bichon criterion (Bect et al. [2012]).

Lemma 1.

$$\mathcal{J}_n(\mathbf{x}) = \int_{\mathbb{X}} \mathbb{E} \left[\left(\kappa \sigma_{n+1}(\mathbf{z}) - |T - \xi(\mathbf{z})| \right)^+ \middle| \mathcal{E}_n \right] d\mathbb{P}_{\mathbb{X}}(\mathbf{z}). \quad (17)$$

Proof.

The integrand of the chosen uncertainty measure in (16) is a positive quantity, by positivity of the expectation. So, by reinjecting the expression of uncertainty (16) into (14), then applying Fubini–Tonelli theorem (thanks to σ -finite measures and positive integrand), we obtain

$$\mathcal{J}_n(\mathbf{x}) = \int_{\mathbb{X}} \mathbb{E} \left[\mathbb{E} \left[\left(\kappa \sigma_{n+1}(\mathbf{z}) - |T - \xi(\mathbf{z})| \right)^+ \middle| \xi(\mathbf{x}), \mathcal{E}_n \right] \right] d\mathbb{P}_{\mathbb{X}}(\mathbf{z}), \quad (18)$$

where the first expectation is relative to $\xi(\mathbf{x}) | \mathcal{E}_n$ and the second one is relative to $\xi(\mathbf{z})$ knowing $\xi(\mathbf{x}), \mathcal{E}_n$. In addition, in (18) it is possible to reduce the two expectations in one to obtain (17). \square

Finally, Proposition 1 below provides an explicit formula for the integrand of (17).

Proposition 1. *For all \mathbf{x}, \mathbf{z} belonging to \mathbb{X}^2 , we have*

$$\begin{aligned} & \mathbb{E} \left[\left(\kappa \sigma_{n+1}(\mathbf{z}) - |T - \xi(\mathbf{z})| \right)^+ \middle| \mathcal{E}_n \right] \\ &= (m_n(\mathbf{z}) - T) \left[2 \phi \left(\frac{T - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) - \phi \left(\frac{T^- - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) - \phi \left(\frac{T^+ - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) \right] \\ & \quad - \sigma_n(\mathbf{z}) \left[2 \varphi \left(\frac{T - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) - \varphi \left(\frac{T^- - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) - \varphi \left(\frac{T^+ - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) \right] \\ & \quad + \varepsilon_{\mathbf{x}}(\mathbf{z}) \left[\phi \left(\frac{T^+ - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) - \phi \left(\frac{T^- - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) \right] \end{aligned} \quad (19)$$

where $\varepsilon_{\mathbf{x}}(\mathbf{z}) := \kappa \sigma_{n+1}(\mathbf{z})$, $T^{\pm} := T \pm \varepsilon_{\mathbf{x}}(\mathbf{z})$, φ and ϕ denote the probability density and cumulative distribution functions of the reduced centered normal distribution, respectively.

The dependency in \mathbf{x} in Equation (19) is only given via $\varepsilon_{\mathbf{x}}(\mathbf{z}) = \kappa \sigma_{n+1}(\mathbf{z})$, therefore only via $\sigma_{n+1}(\mathbf{z})$, which is independent of the model evaluation on \mathbf{x} according to the kriging formulas. In practice, kriging update formulas (31) will be used to have a fast evaluation of $\varepsilon_{\mathbf{x}}(\mathbf{z})$. A proof of Proposition 1 is proposed in Appendix C.

4.2 Numerical experiments

The performances of SUR Bichon are illustrated on two analytical examples, and compared to SUR Vorob'ev performances. The DoE after n enrichment iterations with SUR Bichon (resp. SUR Vorob'ev) is noted $\chi_{1,n}$ (resp. $\chi_{2,n}$). The chosen test functions are the rescaled Branin function in dimension 2 and rescaled Hartmann function in dimension 6 (Picheny et al. [2013]). The choice of the threshold T for each of these functions is discussed later.

Several DoE enrichment strategies can be considered: an enrichment by SUR Bichon $\chi_{1,n}$ and one using SUR Vorob'ev $\chi_{2,n}$, both after n iterations. The criteria performance is evaluated through two different estimators: naive estimator $\hat{\Gamma}_1 := m_n^{-1}([-\infty, T])$ and Vorob'ev estimator noted $\hat{\Gamma}_2$, which corresponds to the Vorob'ev expectation (Appendix A). Performance of the different criteria are then compared after n iterations with the approximation error $\text{Err}(\hat{\Gamma}_i(\chi_{j,n})) := \mathbb{P}_{\mathbb{X}}(\hat{\Gamma}_i(\chi_{j,n}) \Delta \Gamma^*) / \mathbb{P}_{\mathbb{X}}(\Gamma^*)$, for $(i, j) \in \{1, 2\}^2$. This error measures the relative volume of the symmetric difference between estimator $\hat{\Gamma}_i(\chi_{j,n})$ and true excursion set Γ^* defined in (1).

Implementation choices

As mentioned earlier, the choices of trend m and covariance kernel k are fundamental. In the following, the trend is chosen as a single constant term m , see Roustant et al. [2012] for more details. A classical kernel product of type Matérn 5/2 is chosen:

$$k(\mathbf{x}, \mathbf{x}') := \text{Cov}(\mathbf{Z}(\mathbf{x}), \mathbf{Z}(\mathbf{x}')) = \sigma_c^2 \prod_{i=1}^d R_{\text{Matérn } 5/2}(h_i, \theta_i), \quad \forall (\mathbf{x}, \mathbf{x}') \in \mathbb{X}^2, \quad (20)$$

with a vector of parameters θ belonging to \mathbb{R}_+^d , estimated by maximizing the likelihood, $h_i = |\mathbf{x}_i - \mathbf{x}'_i|$, σ_c a fixed parameter and

$$R_{\text{Matérn } 5/2}(h_i, \theta_i) := \left(1 + \frac{\sqrt{5}|h_i|}{\theta_i} + \frac{5h_i^2}{3\theta_i^2}\right) \exp\left(-\frac{\sqrt{5}|h_i|}{\theta_i}\right). \quad (21)$$

This choice leads to trajectories of class C^2 (Paciorek [2003]).

The implementation of SUR Bichon from formulas (17) and (19), is greatly inspired by the implementation of various SUR criteria in the package *KrigInv* (Chevalier et al. [2014]). In addition, the chosen measure $\mathbb{P}_{\mathbb{X}}$ is the Lebesgue measure restricted to \mathbb{X} , which is finite because \mathbb{X} is compact. In (17), the integration is performed using a Sobol' sequence with $n.points$ integration points (package *randtoolbox* Dutang and Savicky [2013]). The criterion is optimized with the genetic algorithm Genoud (with *pop.size* = 50d) (package *rgenoud*, Mebane Jr and Sekhon [2011]). Unless explicitly stated, κ in (19) is set to 1, initial DoEs are obtained by LHS optimized by maximizing minimal distances between the points (Latin Hypercube Sampling, Dupuy et al. [2015]) with size to be specified in the following, and $n.points$ is set to 10^4 .

The volume of Γ^* and $\hat{\Gamma}_i(\chi_{j,n}) \Delta \Gamma^*$ are approached using a quasi-Monte Carlo methods (Lemieux [2009]) with a Sobol' sequence of size 10^4 .

Performance tests on Branin-rescaled 2D function

The Branin-rescaled function, defined in Picheny et al. [2013] on $\mathbb{X} := [0, 1]^2$, is represented in Figure 3. The Γ^* excursion set is defined by the upper bound $T = 10$ on the function values, which leads to 3 disconnected areas and the volume of Γ^* represents 15.74% of the total volume of \mathbb{X} . The tests are performed on 100 different initial DoEs of size 10 and 20 iterations (1 simulation per iteration) are run for both SUR Bichon and SUR Vorob'ev.

The line plot representation on Figure 4 makes it possible to visualize the enrichment throughout the iterations for the chosen approximation error. SUR Bichon appears more robust than SUR

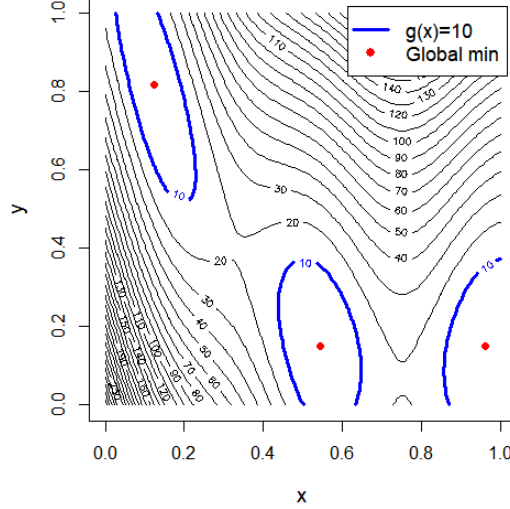


Figure 3: Representation of the Branin-rescaled function on \mathbb{X} .

Vorob'ev throughout the enrichment, no matter which estimator is chosen. Indeed, in the case of SUR Vorob'ev, several extreme cases are present with stagnation of the approximation error. These rare stagnations in the case of the SUR Vorob'ev are due to the late discovery of one of the three components of the excursion set as illustrated below on an example with Figure 5. The logarithmic scale of the graph seems to show a stagnation of the curve for these few cases, but it is in fact a faster progression towards a threshold value, due to the simplification of the problem to the two remaining components. Also, the approximation error with $\hat{\Gamma}_1$ seems more robust than with $\hat{\Gamma}_2$ (see Appendix D for more details), and this can be explained by the fact that estimators $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$ are respectively based on an extension of median and mean concepts (Appendix A).

We focus on one run of the enrichment with SUR Vorob'ev associated with one of the outliers of Figure 4. The associated $\hat{\Gamma}_1$ (resp. $\hat{\Gamma}_2$) estimators are represented as green full (resp. dotted) line on Figure 5, after 20 iterations. This figure shows that SUR Vorob'ev misses one of the three areas of the exact excursion set Γ^* presented in Figure 3. The use of SUR Bichon allows a better exploration of the design space, which here allows to detect the three areas of $\hat{\Gamma}^*$. We notice that there is very few difference between the two types of estimators and given the robustness of naive estimator $\hat{\Gamma}_1$ compared to $\hat{\Gamma}_2$ (Appendix D), naive estimator is kept for the remaining tests in dimension 2.

The characteristic statistical values of the empirical distribution for the error approximation with naïve $\hat{\Gamma}_1$ estimator are given in Table 1 for both SUR Bichon and SUR Vorob'ev. It can be seen that only for the quantile 5%, the results are slightly better for SUR Vorob'ev. In all other cases, SUR Bichon is more efficient than SUR Vorob'ev: this is due to the outliers of SUR Vorob'ev which have a strong influence on the characteristic values, for example mean or standard deviation.

To summarize, the study of the performances of SUR Bichon on Branin-rescaled function, showed that for $T = 10$ the sought excursion set with three connected components is better detected when using SUR Bichon. Indeed, the latter, unlike SUR Vorob'ev, makes it possible to avoid extreme cases for which one of the three connected components is completely missed.

Performance tests on Hartmann-rescaled 6D function

In this section the inversion results for the Hartmann-rescaled function from Picheny et al. [2013] on $\mathbb{X} := [0,1]^6$ are presented. The excursion set Γ^* is defined by the upper bound $T = -1.6$, which a volume that represents 15.45% of the total volume of \mathbb{X} . The tests are performed on 50 different initial DoEs of size 30 and 600 iterations of DoE enrichment are run for both SUR Bichon and SUR Vorob'ev.

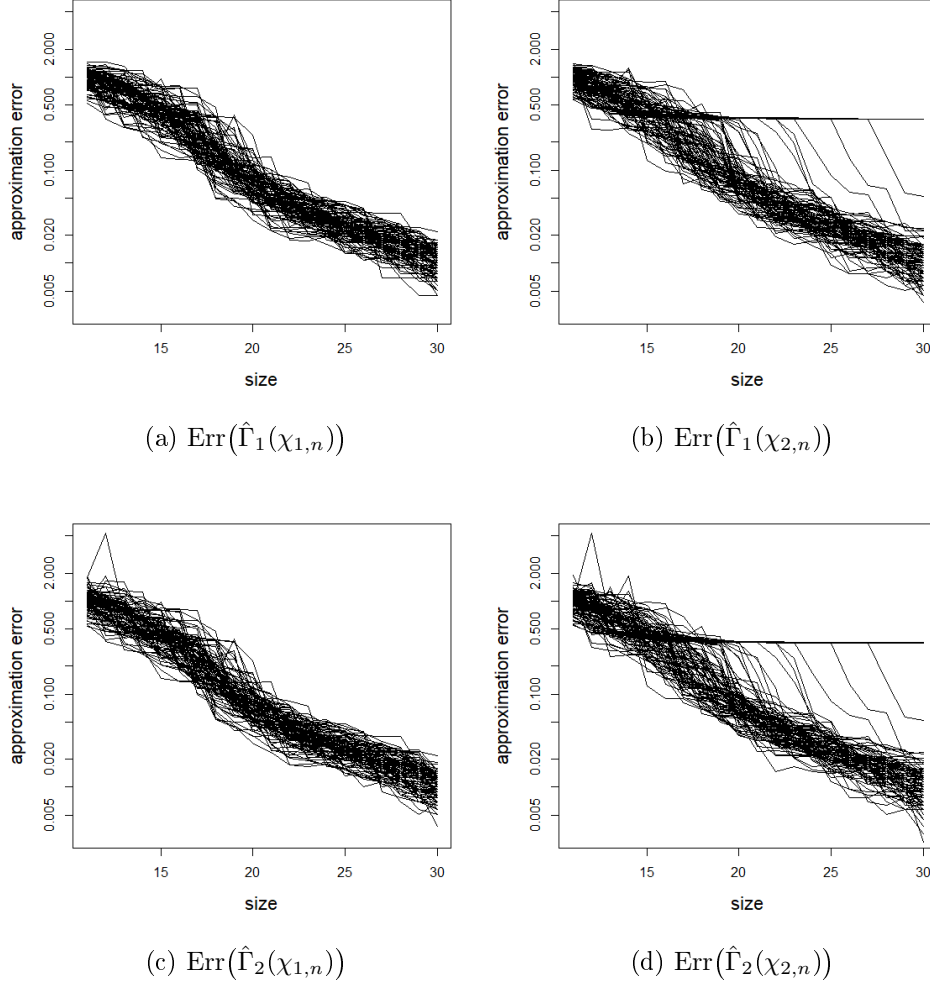


Figure 4: Line plots (with logarithmic scale) of the approximation error $\text{Err}(\hat{\Gamma}_i(\chi_{j,n}))$ for the different criteria during 20 iterations, for Branin-rescaled function inversion ($d = 2$) with $T = 10$, for 100 different initial DoEs of size 10 of type LHS Maximin, for $\kappa = 1$ and with $n.\text{points} = 10^4$. Left column: SUR Bichon with naive estimator (a) and Vorob'ev estimator (c). Right column: SUR Vorob'ev with naive estimator (b) and Vorob'ev estimator (d).

Crit. ($\times 100$)	ite = 10		ite = 20	
	SUR B.	SUR V.	SUR B.	SUR V.
Mean	7.82	12.35	1.09	3.18
Median	7.24	7.78	1.08	1.08
Quantile 5%	4.36	4.24	0.57	0.51
Quantile 95%	12.94	36.42	1.59	35.13
Standard Deviation	3.21	10.79	0.34	8.17
Interquartile Range	3.80	5.38	0.52	0.64

Table 1: Summary of empirical distributions of the $\text{Err}(\hat{\Gamma}_i(\chi_{j,n}))$ for the different criteria after 10 and 20 iterations, for Branin-rescaled function inversion ($d = 2$) with $T = 10$, for 100 different initial DoEs of size 10 of type LHS Maximin.

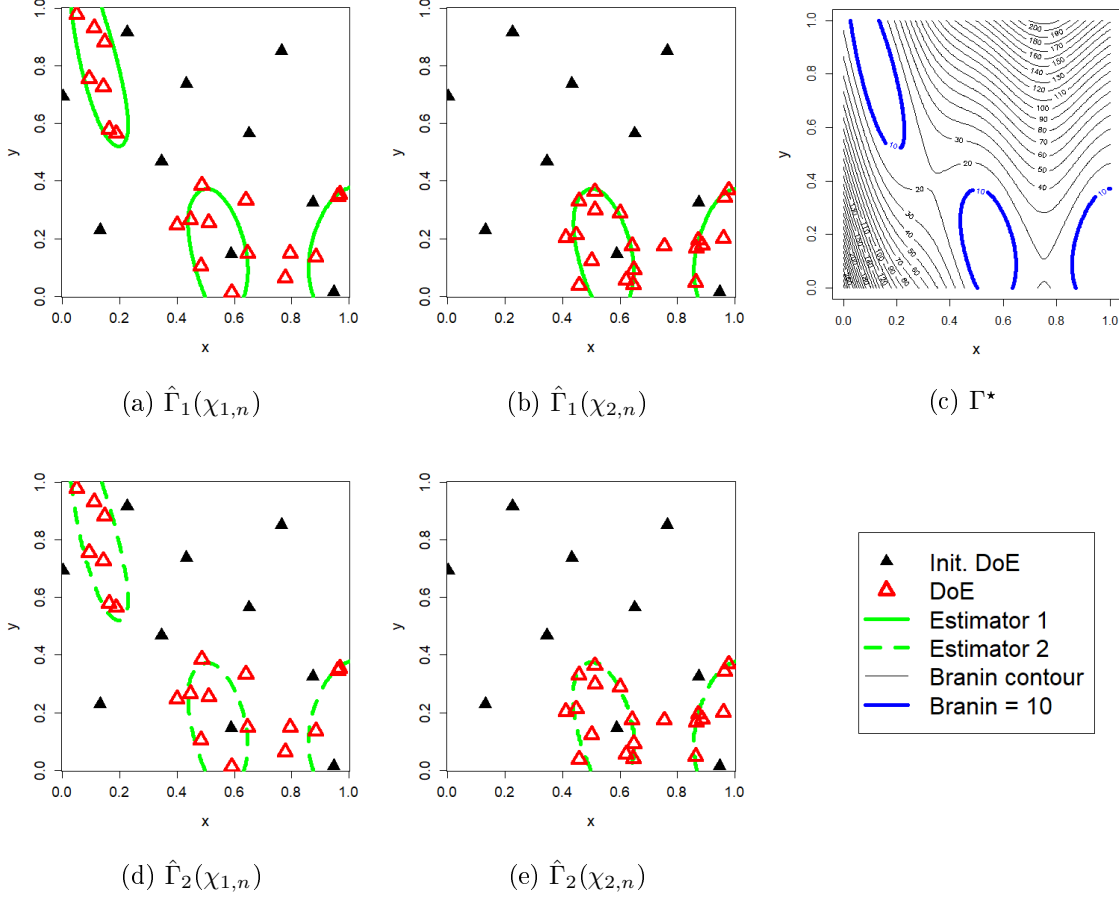


Figure 5: Representation of two $\hat{\Gamma}$ estimators for each of the two criteria after 40 iterations, in comparison with the true excursion set (top right), for Branin-rescaled function inversion ($d = 2$) with $T = 10$, for a particular initial DoE of size 10 where SUR Bichon outperforms SUR Vorob'ev. Left column: SUR Bichon with naive estimator (a) and Vorob'ev estimator (d). Middle column: SUR Vorob'ev with naive estimator (b) and Vorob'ev estimator (e). Right column: true excursion set and Branin-rescaled contour lines (c).

Line plots on Figure 6 show the approximation error along the iterations of the enrichment. Firstly, it can be observed that the improvement of the approximation error during the enrichment is slower than in dimension 2, which is consistent with the increasing difficulty of the problem in higher dimension. Moreover, we observe that SUR Bichon performs better than SUR Vorob'ev whatever the chosen estimator, with respect to the robustness to outliers but also on average, especially from the 300th iteration. In addition, it is clear from Figure 6 (a) and (c) that naive estimator $\hat{\Gamma}_1$ gives a more robust approximation error than $\hat{\Gamma}_2$, with SUR Bichon enrichment. We thus decide to present the numerical results in the following of the section for $\hat{\Gamma}_1$ only.

We focus on one of the outliers of Figure 6 (b) for which the enrichment strategy is based on SUR Vorob'ev. For an extreme case, we represent in pairwise projection the points of a Sobol' sequence of size 10^4 on \mathbb{X} belonging to $\hat{\Gamma}_1 \Delta \Gamma^*$, with $\hat{\Gamma}_1$ the estimator obtained after 600 iterations (see Figure 7). There are only 52 misclassified points observed for SUR Bichon, against 233 in the case of SUR Vorob'ev. Moreover, among the 233 points for the case SUR Vorob'ev, 176 correspond to $\hat{\Gamma} \setminus \Gamma^*$ (infeasible points that are predicted feasible), whereas the remaining misclassified points correspond to feasible points that were predicted unfeasible (as for 2D example). This allows to further illustrate the robustness of SUR Bichon compared to SUR Vorob'ev.

The characteristic statistical values of the empirical distribution for the approximation error are

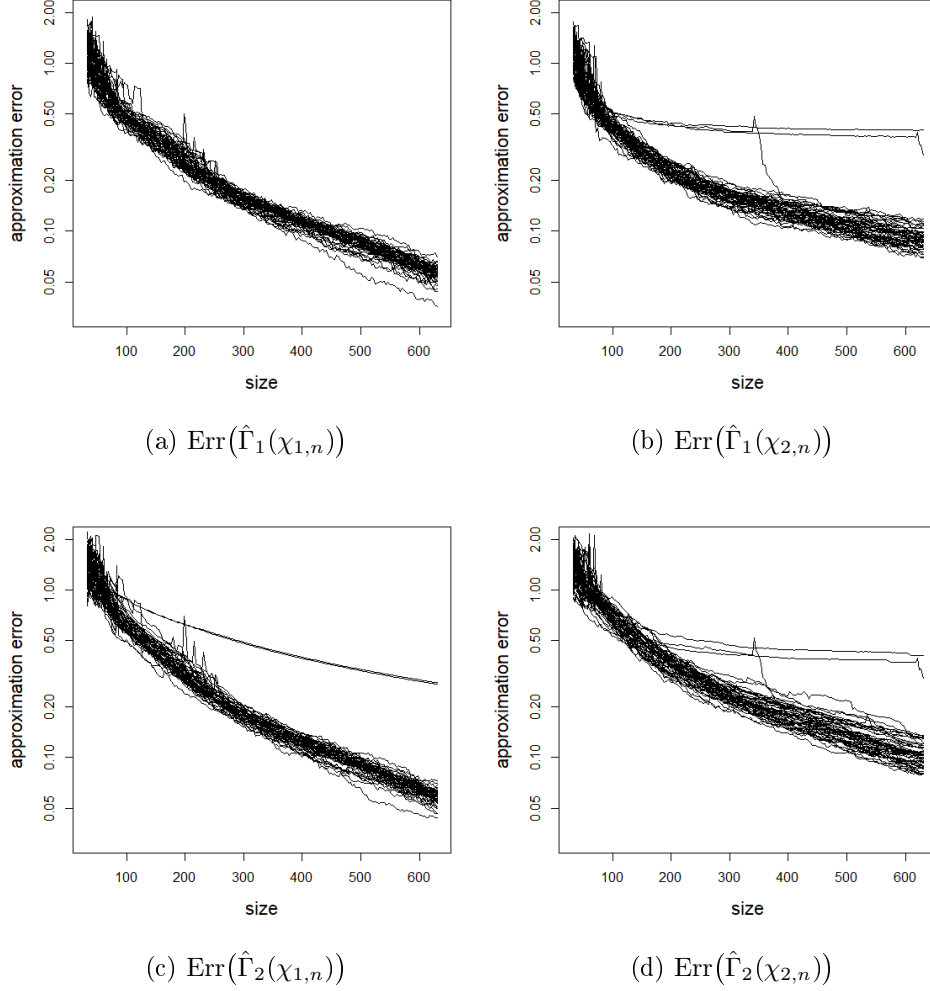


Figure 6: Line plots (with logarithmic scale) of the approximation error $\mathbb{P}_{\mathbf{x}}(\hat{\Gamma}_i(\chi_{j,n})\Delta\Gamma^*)/\mathbb{P}_{\mathbf{x}}(\Gamma^*)$ (with a Sobol' sequence of size 10^4) for the different criteria during 600 iterations, for Hartmann-rescaled function inversion ($d = 6$) with $T = -1.6$, for 50 different initial DoEs of size 30 of type LHS Maximin, for $\kappa = 1$ and with n.points= 10^4 . Left column: SUR Bichon with naive estimator (a) and Vorob'ev estimator (c). Right column: SUR Vorob'ev with naive estimator (b) and Vorob'ev estimator (d).

given in Figure 2 for both SUR Bichon and SUR Vorob'ev. It can be confirmed that except at the beginning where enrichment is not yet sufficient, SUR Bichon performs better for all indicators than SUR Vorob'ev, and not only in terms of robustness (see e.g. the quantile of order 5% or the median).

In summary, the tests on Hartmann-rescaled function in dimension 6 with $T = -1.6$, further highlight the robustness of SUR Bichon compared to SUR Vorob'ev, in dimension higher than 2. In addition to the robustness, it was also observed that in this 6 dimensional case, beyond about 300 iterations, SUR Bichon performs better than SUR Vorob'ev, even without considering the outliers.

Several hypotheses are proposed below to try to understand these differences. Firstly, the robustness of SUR Bichon can be explained by its exploratory capability forced by the kriging standard deviation in factor of the SUR Bichon formulation (see Section 2.2).

The second hypothesis is that the Vorob'ev expectation is strongly dependent on the stationarity assumption of the underlying Gaussian process (Appendix D), which has an influence on the enrichment, since the enrichment with SUR Vorob'ev is based on the Vorob'ev deviation (Appendix A), and then on the Vorob'ev expectation. Even if the two test cases studied seem to verify the stationarity hypothesis, this is never rigorously the case (unless the model is defined as a given

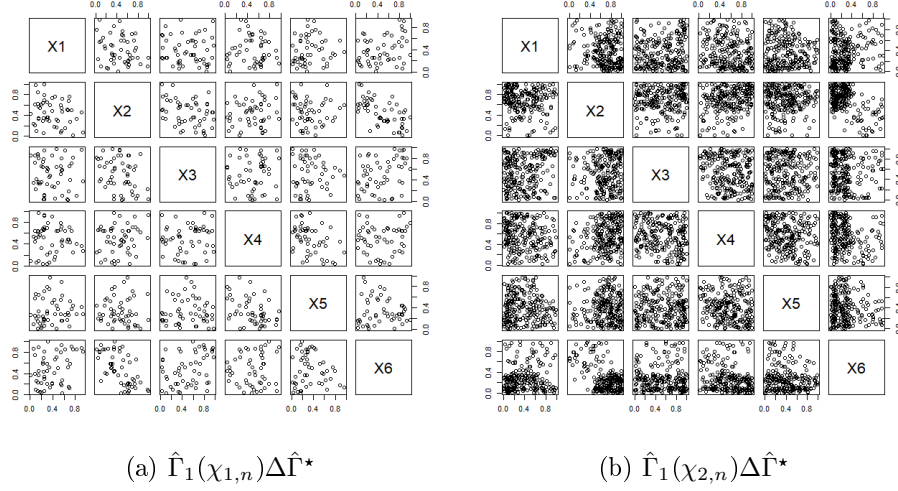


Figure 7: Pairwise projection plot of $\hat{\Gamma}_1 \Delta \Gamma^*$ for a 10^4 -Sobol' sequence, for the two criteria after 600 iterations, for Hartmann-rescaled function inversion ($d = 6$) with $T = -1.6$, for a particular initial DoE of size 30 where SUR Bichon outperforms on SUR Vorob'ev, for $\kappa = 1$ and for n.points= 10^4 . Left column: SUR Bichon. Right column: SUR Vorob'ev.

Crit. ($\times 100$)	ite = 99		ite = 300		ite = 600	
	SUR B.	SUR V.	SUR B.	SUR V.	SUR B.	SUR V.
Mean	40.42	35.05	13.92	16.55	5.61	9.94
Median	40.32	34.72	13.92	15.28	5.66	8.83
Quantile 5%	34.14	28.84	12.09	12.74	4.62	7.17
Quantile 95%	46.67	44.40	15.38	29.42	6.45	11.63
Standard Deviation	3.98	4.48	0.99	6.14	0.61	5.26
Interquartile Range	5.26	4.56	1.13	1.81	0.68	1.75

Table 2: Summary of empirical distributions of the approximation error $\mathbb{P}_{\mathbb{X}}(\hat{\Gamma} \Delta \Gamma^*) / \mathbb{P}_{\mathbb{X}}(\Gamma^*)$ (with a Sobol' sequence of size 10^4) for the different criteria after 99, 300 and 600 iterations, for Hartmann-rescaled function inversion ($d = 6$) with $T = -1.6$, for 50 different initial DoEs of size 30 of type LHS Maximin, for $\kappa = 1$ and with n.points= 10^4 .

realization of a stationary Gaussian process). The sensitivity of a criterion to this hypothesis can therefore have a relatively strong influence, especially when several points are successively added to the DoE, leading to an accumulation of errors.

The calculation of the Vorob'ev criterion is sensitive to the determination of the Vorob'ev threshold α^* (see Appendix A). However, we have verified that the determination of α^* was not the problem for the robustness of SUR Vorob'ev with verifying the numerical simplicity of the minimum search for the function $\alpha \mapsto |\mathbb{E}[\mathbb{P}_{\mathbb{X}}(\Gamma) | \mathcal{E}_n] - \mathbb{P}_{\mathbb{X}}(Q_\alpha)|$.

5 Conclusion

In the framework of solving inversion problems using Gaussian Process Regression, we have proposed a new SUR criterion for DoE enrichment. Numerical simulations have demonstrated its good exploratory behavior, as far as its robustness from different points of view. Indeed, our new criterion

is robust with the stationarity assumption of the underlying Gaussian process. Moreover, it is robust to the geometry of the set to be retrieved, in particular in terms of number of connected components.

6 Acknowledgements

This research was conducted with the support of the consortium in Applied Mathematics CIRO-QUO, gathering partners in technological and academia in the development of advanced methods for Computer Experiments (<https://doi.org/10.5281/zenodo.6581216>). This work was also carried out as part of the “HPC/AI/HPDA Convergence for the Energy Transition” Inria/IFPEN joint research laboratory (High-performance computation / Artificial Intelligence / High-Performance Data Analysis). The authors would like to express their gratitude to Morgane Menz from IFP Énergies Nouvelles for her valuable help on numerical aspects and in particular for the study of the robustness of estimators with respect to the GP stationarity assumption. In addition, all the computations presented in this paper were performed using the GRICAD infrastructure (<https://gricad.univ-grenoble-alpes.fr>), which is supported by Grenoble research communities.

References

- D. Azzimonti, D. Ginsbourger, C. Chevalier, J. Bect, and Y. Richet. Adaptive design of experiments for conservative estimation of excursion sets. *Technometrics*, 63(1):13–26, 2021.
- J. Bect, D. Ginsbourger, L. Li, V. Picheny, and E. Vazquez. Sequential design of computer experiments for the estimation of a probability of failure. *Statistics and Computing*, 22(3):773–793, 2012.
- B. J. Bichon, M. S. Eldred, L. P. Swiler, S. Mahadevan, and J. M. McFarland. Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA journal*, 46(10):2459–2468, 2008.
- C. Chevalier. *Fast uncertainty reduction strategies relying on Gaussian process models*. PhD thesis, Universität Bern, 2013.
- C. Chevalier, V. Picheny, and D. Ginsbourger. Kriginv: An efficient and user-friendly implementation of batch-sequential inversion strategies based on kriging. *Computational statistics & data analysis*, 71:1021–1034, 2014.
- P. Dagnelie. *Statistique théorique et appliquée*. Presses agronomiques de Gembloux Belgium, 1992.
- D. Dupuy, C. Helbert, and J. Franco. DiceDesign and DiceEval: Two R packages for design and analysis of computer experiments. *Journal of Statistical Software*, 65(11), 2015.
- C. Dutang and P. Savicky. randtoolbox: Generating and testing random numbers. *R package*, 67, 2013.
- B. Echard, N. Gayton, and M. Lemaire. AK-MCS: an active learning reliability method combining kriging and Monte Carlo simulation. *Structural Safety*, 33(2):145–154, 2011.
- M. R. El Amri. *Analyse d’incertitudes et de robustesse pour les modèles à entrées et sorties fonctionnelles*. Theses, Université Grenoble Alpes, Apr. 2019. URL <https://tel.archives-ouvertes.fr/tel-02433324>.
- M. R. El Amri, C. Helbert, O. Lepreux, M. M. Zuniga, C. Prieur, and D. Sinoquet. Data-driven stochastic inversion via functional quantization. *Statistics and Computing*, 30(3):525–541, 2020.

- D. Ginsbourger. Sequential design of computer experiments. *Wiley StatsRef*, 99:1–11, 2017.
- C. Helbert, D. Dupuy, and L. Carraro. Assessment of uncertainty in computer experiments from universal to bayesian kriging. *Applied Stochastic Models in Business and Industry*, 25(2):99–113, 2009.
- R. Jin, W. Chen, and A. Sudjianto. On sequential sampling for global metamodeling in engineering design. In *International design engineering technical conferences and computers and information in engineering conference*, volume 36223, pages 539–548, 2002.
- D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492, 1998.
- C. Lemieux. Quasi–Monte Carlo constructions. In *Monte Carlo and Quasi-Monte Carlo Sampling*, pages 1–61. Springer, 2009.
- W. R. Mebane Jr and J. S. Sekhon. Genetic optimization using derivatives: the rgenoud package for R. *Journal of Statistical Software*, 42(1):1–26, 2011.
- I. S. Molchanov. *Theory of random sets*, volume 87. Springer, 2005.
- M. Moustapha, S. Marelli, and B. Sudret. A generalized framework for active learning reliability: survey and benchmark. *arXiv preprint arXiv:2106.01713*, 2021.
- A. O’Hagan. Curve fitting and optimal design for prediction. *Journal of the Royal Statistical Society: Series B (Methodological)*, 40(1):1–24, 1978.
- C. J. Paciorek. *Nonstationary Gaussian processes for regression and spatial modelling*. PhD thesis, Carnegie Mellon University, 2003.
- V. Picheny, D. Ginsbourger, O. Roustant, R. T. Haftka, and N.-H. Kim. Adaptive designs of experiments for accurate approximation of a target region. 2010.
- V. Picheny, T. Wagner, and D. Ginsbourger. A benchmark of kriging-based infill criteria for noisy optimization. *Structural and Multidisciplinary Optimization*, 48(3):607–626, 2013.
- P. Ranjan, D. Bingham, and G. Michailidis. Sequential experiment design for contour estimation from complex computer codes. *Technometrics*, 50(4):527–541, 2008.
- O. Roustant, D. Ginsbourger, and Y. Deville. Dicekriging, diceoptim: Two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization. 2012.

A Basics on Vorob’ev Theory and corresponding SUR strategy

A.1 Vorob’ev expectation

In this part, the notion of expectation for random closed sets in the sense of Vorob’ev is defined from Molchanov [2005]. The framework is a compact set $\mathbb{X} \subset \mathbb{R}^d$ and a random closed set Γ of \mathbb{X} . It is recalled that $\Gamma : \Omega \rightarrow \mathcal{C}$ is a random closed set if it is a measurable function on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in the set of all compacts of \mathbb{X} in the sense that:

$$\forall C \in \mathcal{C}, \{w \in \Omega, \Gamma(w) \cap C \neq \emptyset\} \in \mathcal{F}. \quad (22)$$

Let’s define the parametric family $\{Q_\alpha\}_{\alpha \in [0,1]}$ of Vorob’ev quantiles is defined by:

$$Q_\alpha := \{x \in \mathbb{X} : p(x) := \mathbb{P}(x \in \Gamma) \geq \alpha\}, \forall \alpha \in [0, 1]. \quad (23)$$

The elements of $\{Q_\alpha\}_{\alpha \in [0,1]}$ are called the Vorob'ev quantiles of the random closed set Γ and the function p is called the coverage function of Γ .

To define the expectation of the random closed set Γ , Molchanov [2005] comes back to the expectation of a real random variable: the measure of the Γ set $\mathbb{P}_\mathbb{X}(\Gamma)$. From the parametric family of Vorob'ev quantiles (see (23)), the expectation of Γ in the sense of Vorob'ev is then defined as the Vorob'ev quantile of measure equal (or the closest one higher) to the expectation of the measure of Γ . More precisely, the Vorob'ev expectation of a random closed set Γ is the set Q_{α^*} , where α^* is defined as the Vorob'ev threshold by

$$\forall \alpha > \alpha^*, \mathbb{P}_\mathbb{X}(Q_\alpha) < \mathbb{E}[\mathbb{P}_\mathbb{X}(\Gamma)] \leq \mathbb{P}_\mathbb{X}(Q_{\alpha^*}), \quad (24)$$

where $\mathbb{P}_\mathbb{X}$ denotes the Lebesgue measure on \mathbb{X} . α^* is called the Vorob'ev threshold.

Remark 1.

- Based on Equation (23)), the function $\alpha \mapsto \mathbb{P}_\mathbb{X}(Q_\alpha)$ is decreasing on $[0, 1]$.
- The uniqueness of α^* in the definition is easily checked. The existence of such α^* in the definition of Vorob'ev expectation is based on the decreasing and continuity to the left of the function $\alpha \mapsto \mathbb{P}_\mathbb{X}(Q_\alpha)$ which is itself guaranteed by the superior semi-continuity of the coverage function p (see Molchanov [2005] page 23).
- The continuity of the function $\alpha \mapsto \mathbb{P}_\mathbb{X}(Q_\alpha)$ ensures equality $\mathbb{P}_\mathbb{X}(Q_{\alpha^*}) = \mathbb{E}[\mathbb{P}_\mathbb{X}(\Gamma)]$ in the definition of Vorob'ev expectation.

In the particular case where Γ is given by $\Gamma := \{\mathbf{x} \in \mathbb{X}, \xi(\mathbf{x}) \leq T\}$ with ξ a stochastic process indexed by \mathbb{X} with continuous trajectories conditioned on the event \mathcal{E}_n corresponding to n evaluations of ξ and T a fixed threshold, Γ is a random closed set ([Molchanov, 2005, page 3]). A sufficient condition to obtain a stochastic process with continuous trajectories is to consider a separable Gaussian process with continuous mean and covariance kernel of type Matérn 3/2 or 5/2 ([Paciorek, 2003, pages 35 and 44]). Moreover, in this case, the function $\alpha \mapsto \mathbb{P}_\mathbb{X}(Q_\alpha)$ is continuous and so the equality $\mathbb{P}_\mathbb{X}(Q_{\alpha^*}) = \mathbb{E}[\mathbb{P}_\mathbb{X}(\Gamma) | \mathcal{E}_n]$ is verified. It is also important to notice that naive estimator $\hat{\Gamma}_1$ is almost surely equal to the median of Vorob'ev (quantile of order 1/2). Indeed, by noting ϕ the distribution function of the standard normal distribution,

$$\begin{aligned} \hat{\Gamma}_1 &= \{\mathbf{x} \in \mathbb{X}, m_n(\mathbf{x}) \leq T\} \\ &\stackrel{\text{a.s.}}{=} \left\{ \mathbf{x} \in \mathbb{X}, \frac{T - m_n(\mathbf{x})}{\sigma_n(\mathbf{x})} \geq 0 \text{ and } \sigma_n(\mathbf{x}) \neq 0 \right\} \\ &= \left\{ \mathbf{x} \in \mathbb{X}, \phi\left(\frac{T - m_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) \geq \phi(0) \text{ and } \sigma_n(\mathbf{x}) \neq 0 \right\} \quad \text{as } \phi \text{ increases} \\ &\stackrel{\text{a.s.}}{=} \left\{ \mathbf{x} \in \mathbb{X}, \mathbb{P}(\xi(\mathbf{x}) \leq T | \mathcal{E}_n) \geq \frac{1}{2} \right\} \\ &= Q_{\frac{1}{2}}. \end{aligned} \quad (25)$$

Repeating the previous calculation and replacing 1/2 by the Vorob'ev threshold α^* , we obtain :

$$Q_{\alpha^*} \stackrel{\text{a.s.}}{=} \{\mathbf{x} \in \mathbb{X}, m_n(\mathbf{x}) \leq T - \phi^{-1}(\alpha^*) \sigma_n(\mathbf{x})\} \quad (26)$$

A.2 Vorob'ev deviation

The introduction of the concept of Vorob'ev deviation is used to define uncertainty measure $\mathcal{H}_n(\mathbf{x})$ in a SUR strategy. Let's start by introducing the notion of distance between two random closed sets.

The average distance $d_{\mathbb{P}_{\mathbb{X}}}$ with respect to a measure $\mathbb{P}_{\mathbb{X}}$ on all pairs of random closed sets included in \mathbb{X} is defined by: for all random closed sets Γ_1, Γ_2 defined on \mathbb{X} ,

$$d_{\mathbb{P}_{\mathbb{X}}}(\Gamma_1, \Gamma_2) := \mathbb{E}[\mathbb{P}_{\mathbb{X}}(\Gamma_1 \Delta \Gamma_2)] \quad (27)$$

where Δ is the random symmetric difference: $\forall \omega \in \Omega, \Gamma_1 \Delta \Gamma_2(\omega) := (\Gamma_1 \setminus \Gamma_2)(\omega) \cup (\Gamma_2 \setminus \Gamma_1)(\omega)$. In addition, the function $d_{\mathbb{P}_{\mathbb{X}}}$ checks the properties of a distance.

The following proposition Molchanov [2005] justifies the choice of the Vorob'ev quantile family to define the Vorob'ev expectation and also allows to define the Vorob'ev deviation. Moreover, when $\mathbb{E}[\mathbb{P}_{\mathbb{X}}(\Gamma)] = \mathbb{P}_{\mathbb{X}}(Q_{\alpha^*})$ (especially in the case of $\Gamma := \{\mathbf{x} \in \mathbf{X}, \xi(\mathbf{x}) \leq T\}$ with ξ a stochastic process indexed by \mathbb{X} with continuous trajectories), the condition $\alpha^* \geq \frac{1}{2}$ is no longer necessary (see [El Amri, 2019, page 28] for the proof)

Proposition 2. *Noting Q_{α^*} the Vorob'ev expectation of the random closed set Γ and assuming that $\alpha^* \geq \frac{1}{2}$, it results that: for any measurable set M included in \mathbb{X} such that $\mathbb{P}_{\mathbb{X}}(M) = \mathbb{E}[\mathbb{P}_{\mathbb{X}}(\Gamma)]$,*

$$d_{\mathbb{P}_{\mathbb{X}}}(\Gamma, Q_{\alpha^*}) \leq d_{\mathbb{P}_{\mathbb{X}}}(\Gamma, M) \quad (28)$$

The Vorob'ev deviation of the random set Γ is defined as the quantity $d_{\mathbb{P}_{\mathbb{X}}}(\Gamma, Q_{\alpha^*})$. The Vorob'ev deviation quantifies the variability of the random closed set Γ relative to its Vorob'ev expectation.

A.3 SUR Vorob'ev criterion

Once the basic elements of Vorob'ev theory are introduced, the associated SUR strategy is simply defined from the definition of SUR strategies via Equation (14) by taking:

$$\mathcal{H}_n := \mathbb{E}[\mathbb{P}_{\mathbb{X}}(\Gamma \Delta Q_{n, \alpha_n^*}) | \mathcal{E}_n] \quad \text{and} \quad \mathcal{H}_{n+1}(\mathbf{x}) := \mathbb{E}[\mathbb{P}_{\mathbb{X}}(\Gamma \Delta Q_{n+1, \alpha_{n+1}^*}) | \xi(\mathbf{x}), \mathcal{E}_n] \quad (29)$$

where Q_{n, α_n^*} denotes the Vorob'ev expectation conditioned on \mathcal{E}_n and Q_{n+1, α_{n+1}^*} the Vorob'ev expectation conditioned on \mathcal{E}_n and the addition of the point $(\mathbf{x}, \xi(\mathbf{x}))$ to the DoE. The idea behind (29) is to take as an uncertainty function, the variation with respect to the Vorob'ev expectation of the random closed set Γ , with $\Gamma := \{\mathbf{x} \in \mathbf{X}, \xi(\mathbf{x}) \leq T\}$. With the assumption that $\alpha_{n+1}^* = \alpha_n^*$ and by reinjecting the quantity $\mathcal{H}_{n+1}(\mathbf{x})$ of (29) in the \mathcal{J}_n criterion (14), it is possible to find a simplified formulation involving only an integral of a simple quantity Chevalier [2013]. This quantity is dependent on the cumulative distribution functions of the standard normal distribution and the bivariate centered normal distribution with given covariance matrix. Such a formulation then allows less time consuming computations and therefore is implemented in the package *KrigInv* (Chevalier et al. [2014]).

B Kriging update formulas

In the context of SUR strategies, the quantity $\mathcal{J}_n(\mathbf{x})$ in Equation (14) for a fixed \mathbf{x} is usually simplified thanks to formulas of kriging and conditionally on the point $(\mathbf{x}, \xi(\mathbf{x}))$ added to the DoE, and more particularly using the kriging standard deviation. Indeed, contrary to the trend, the kriging standard deviation does not depend on surrogate model observations. For instance the recurrent formula, used in Chevalier [2013] is efficient for calculating kriging model in the context of universal kriging and when the kriging parameters β and θ do not need to be reestimated. These kriging update formulas are given for all \mathbf{y}, \mathbf{y}' in \mathbb{X}^2 by

$$m_{n+1}(\mathbf{y}) = m_n(\mathbf{y}) + k_n(\mathbf{y}, \mathbf{x}^{(n+1)}) k_n(\mathbf{x}^{(n+1)}, \mathbf{x}^{(n+1)})^{-1} \left(g(\mathbf{x}^{(n+1)}) - m_n(\mathbf{x}^{(n+1)}) \right), \quad (30)$$

$$\sigma_{n+1}^2(\mathbf{y}) = \sigma_n^2(\mathbf{y}) - k_n^2(\mathbf{y}, \mathbf{x}^{(n+1)}) \sigma_n^2(\mathbf{x}^{(n+1)})^{-1}, \quad (31)$$

$$k_{n+1}(\mathbf{y}, \mathbf{y}') = k_n(\mathbf{y}, \mathbf{y}') - k_n(\mathbf{y}, \mathbf{x}^{(n+1)})k_n(\mathbf{x}^{(n+1)}, \mathbf{x}^{(n+1)})^{-1}k_n(\mathbf{y}', \mathbf{x}^{(n+1)}), \quad (32)$$

with $\mathbf{x}^{(n+1)}$ the $n + 1^{\text{th}}$ observation point.

As for SUR strategies, it is possible to generalize these formulas in the case of simultaneous additions of q points (Chevalier [2013]). The advantage of these formulas is that the expressions of m_n , σ_n , and k_n are reused to reduce computational time. It is particularly useful in SUR strategies where many evaluations of the kriging formulas may be required for the numerous evaluations of the sampling criterion (in the context of its minimization (14)). Finally, it can be shown that these kriging formulas still coincide with the Gaussian process conditional formulas in the context of universal kriging (see Appendix A of Chevalier [2013] for a proof).

C Proof of the explicit formula for SUR Bichon

The interest of this appendix is to propose a demonstration of Proposition 1, allowing to give an explicit expression of SUR Bichon. Let's start by stating and proving an intermediate lemma.

Lemma 2. *Let N be a standard Gaussian random variable and $(a, b) \in \mathbb{R}^2$ such that $a < b$, then:*

$$\mathbb{E}[N\mathbf{1}_{[a,b]}(N)] = -\varphi(b) + \varphi(a) \quad (33)$$

where φ is the probability density function of the standard normal distribution.

Proof.

$$\mathbb{E}[N\mathbf{1}_{[a,b]}(N)] = \frac{1}{\sqrt{2\pi}} \int_a^b te^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \left[-e^{-\frac{t^2}{2}} \right]_a^b = -\varphi(b) + \varphi(a) \quad (34)$$

□

Proposition 1. *For all \mathbf{x}, \mathbf{z} belonging to \mathbb{X}^2 , we have:*

$$\begin{aligned} & \mathbb{E} \left[\left(\kappa\sigma_{n+1}(\mathbf{z}) - |T - \xi(\mathbf{z})| \right)^+ \middle| \mathcal{E}_n \right] \\ &= (m_n(\mathbf{z}) - T) \left[2\phi\left(\frac{T - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})}\right) - \phi\left(\frac{T^- - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})}\right) - \phi\left(\frac{T^+ - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})}\right) \right] \\ & \quad - \sigma_n(\mathbf{z}) \left[2\varphi\left(\frac{T - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})}\right) - \varphi\left(\frac{T^- - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})}\right) - \varphi\left(\frac{T^+ - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})}\right) \right] \\ & \quad + \varepsilon_{\mathbf{x}}(\mathbf{z}) \left[\phi\left(\frac{T^+ - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})}\right) - \phi\left(\frac{T^- - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})}\right) \right] \end{aligned} \quad (35)$$

where $\varepsilon_{\mathbf{x}}(\mathbf{y}) := \kappa\sigma_{n+1}(\mathbf{z})$, $T^\pm := T \pm \varepsilon_{\mathbf{x}}(\mathbf{z})$, φ and ϕ denote the probability density and cumulative distribution functions of the standard normal distribution, respectively.

Proof.

For the proof only and for the sake of lightening the notations, we note \mathbb{E}_n the conditional expectation $\mathbb{E}[\cdot | \mathcal{E}_n]$. In addition, the expression to be calculated is separated into three terms that are calculated separately. Specifically:

$$\begin{aligned} & \mathbb{E}_n \left[\left(\kappa\sigma_{n+1}(\mathbf{z}) - |T - \xi(\mathbf{z})| \right)^+ \right] = \mathbb{E}_n \left[\left(\varepsilon_{\mathbf{x}}(\mathbf{z}) - |T - \xi(\mathbf{z})| \right) \mathbf{1}_{[T^-, T^+]}(\xi(\mathbf{z})) \right] \\ &= \mathbb{E}_n \left[\varepsilon_{\mathbf{x}}(\mathbf{z}) \mathbf{1}_{[T^-, T^+]}(\xi(\mathbf{z})) \right] - \mathbb{E}_n \left[(T - m_n(\mathbf{z}) + m_n(\mathbf{z}) - \xi(\mathbf{z})) \mathbf{1}_{[T^-, T]}(\xi(\mathbf{z})) \right] \\ & \quad + \mathbb{E}_n \left[(T - m_n(\mathbf{z}) + m_n(\mathbf{z}) - \xi(\mathbf{z})) \mathbf{1}_{[T, T^+]}(\xi(\mathbf{z})) \right] \end{aligned}$$

$$\begin{aligned}
 &= \underbrace{\mathbb{E}_n \left[\varepsilon_{\mathbf{x}}(\mathbf{z}) \mathbf{1}_{[T^-, T^+]}(\xi(\mathbf{z})) \right]}_{\textcircled{1}} + \underbrace{(T - m_n(\mathbf{z})) \left[\mathbb{E}_n \left[\mathbf{1}_{[T, T^+]}(\xi(\mathbf{z})) \right] - \mathbb{E}_n \left[\mathbf{1}_{[T^-, T]}(\xi(\mathbf{z})) \right] \right]}_{\textcircled{2}} \\
 &\quad - \underbrace{\left[\mathbb{E}_n \left[(m_n(\mathbf{z}) - \xi(\mathbf{z})) \mathbf{1}_{[T^-, T]}(\xi(\mathbf{z})) \right] - \mathbb{E}_n \left[(m_n(\mathbf{z}) - \xi(\mathbf{z})) \mathbf{1}_{[T, T^+]}(\xi(\mathbf{z})) \right] \right]}_{\textcircled{3}} \quad (36)
 \end{aligned}$$

The calculation of the three terms separately is as follows:

$$\begin{aligned}
 \textcircled{1} &:= \mathbb{E}_n \left[\varepsilon_{\mathbf{x}}(\mathbf{z}) \mathbf{1}_{[T^-, T^+]}(\xi(\mathbf{z})) \right] \\
 &= \varepsilon_{\mathbf{x}}(\mathbf{z}) \mathbb{E}_n \left[\mathbf{1}_{\left[\frac{T^- - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})}, \frac{T^+ - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right]} \left(\frac{\xi(\mathbf{z}) - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) \right] \\
 &= \varepsilon_{\mathbf{x}}(\mathbf{z}) \left[\phi \left(\frac{T^+ - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) - \phi \left(\frac{T^- - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) \right] \quad (37)
 \end{aligned}$$

$$\begin{aligned}
 \textcircled{2} &:= (T - m_n(\mathbf{z})) \left[\mathbb{E}_n \left[\mathbf{1}_{[T, T^+]}(\xi(\mathbf{z})) \right] - \mathbb{E}_n \left[\mathbf{1}_{[T^-, T]}(\xi(\mathbf{z})) \right] \right] \\
 &= (m_n(\mathbf{z}) - T) \left[\mathbb{E}_n \left[\mathbf{1}_{[T^-, T]}(\xi(\mathbf{z})) \right] - \mathbb{E}_n \left[\mathbf{1}_{[T, T^+]}(\xi(\mathbf{z})) \right] \right] \\
 &= (m_n(\mathbf{z}) - T) \left[2 \phi \left(\frac{T - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) - \phi \left(\frac{T^- - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) - \phi \left(\frac{T^+ - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) \right] \quad (38)
 \end{aligned}$$

For the calculation of $\textcircled{3}$, the use of the lemma 2 is necessary, using the notations of it.

$$\begin{aligned}
 \textcircled{3} &:= - \left[\mathbb{E}_n \left[(m_n(\mathbf{z}) - \xi(\mathbf{z})) \mathbf{1}_{[T^-, T]}(\xi(\mathbf{z})) \right] - \mathbb{E}_n \left[(m_n(\mathbf{y}) - \xi(\mathbf{z})) \mathbf{1}_{[T, T^+]}(\xi(\mathbf{z})) \right] \right] \\
 &= \sigma_n(\mathbf{z}) \left[\mathbb{E}_n \left[\left(\frac{\xi(\mathbf{z}) - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) \mathbf{1}_{\left[\frac{T^- - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})}, \frac{T - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right]} \left(\frac{\xi(\mathbf{z}) - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) \right] \right. \\
 &\quad \left. - \mathbb{E}_n \left[\left(\frac{\xi(\mathbf{z}) - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) \mathbf{1}_{\left[\frac{T - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})}, \frac{T^+ - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right]} \left(\frac{\xi(\mathbf{z}) - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) \right] \right] \\
 &= \sigma_n(\mathbf{z}) \left[\mathbb{E} \left[N \mathbf{1}_{\left[\frac{T^- - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})}, \frac{T - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right]}(N) \right] - \mathbb{E} \left[N \mathbf{1}_{\left[\frac{T - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})}, \frac{T^+ - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right]}(N) \right] \right] \\
 &= -\sigma_n(\mathbf{z}) \left[2 \varphi \left(\frac{T - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) - \varphi \left(\frac{T^- - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) - \varphi \left(\frac{T^+ - m_n(\mathbf{z})}{\sigma_n(\mathbf{z})} \right) \right] \quad (39)
 \end{aligned}$$

The expected result is then obtained by reinjecting the expressions of $\textcircled{1}$, $\textcircled{2}$, and $\textcircled{3}$ obtained in Equations (37) to (39) in Equation (36). \square

Remark 2.

• In this proof of the proposition 1, the fact that $m_n(\mathbf{z})$, $\sigma_n(\mathbf{z})$, $\varepsilon_{\mathbf{x}}(\mathbf{z})$ and T^\pm are constant with respect to $\xi(\mathbf{z})$ is implicitly used, in particular to output $m_n(\mathbf{z})$, $\sigma_n(\mathbf{z})$ and $\varepsilon_{\mathbf{x}}(\mathbf{z})$ of the conditional expectation \mathbb{E}_n , but also for the renormalization of $\xi(\mathbf{z})$ and the transition to the density probability and cumulative distribution functions of the reduced centered normal distribution.

D Robustness of estimators with respect to the GP stationarity assumption.

A certain instability of the approximation error was observed at the beginning of the enrichment in Figures 4 and 6 in the case of estimator $\hat{\Gamma}_2$ corresponding to Vorob'ev expectation, in comparison to $\hat{\Gamma}_1$ naïve estimator. This robustness of naïve estimator $\hat{\Gamma}_1$ compared to $\hat{\Gamma}_2$ can be explained by the fact that naïve estimator corresponds to the median of Vorob'ev (Appendix A, equation (25)). Indeed, even if the notions of expectation and median of Vorob'ev are not similar to the classical ones, the property of minimising the first absolute central moment of the median is preserved when extending the notion of median to the framework of Vorob'ev random sets (Molchanov [2005] page 178). It is also possible to "read" the lack of robustness on Equation (26) of Appendix A which is recalled below:

$$Q_{\alpha^*} \stackrel{\text{a.s.}}{=} \{ \mathbf{x} \in \mathbb{X}, m_n(\mathbf{x}) \leq T - \phi^{-1}(\alpha^*) \sigma_n(\mathbf{x}) \}.$$

Knowing the strong dependence of the σ_n term on the stationarity condition of the process, it is straightforward that the term $\phi^{-1}(\alpha^*) \sigma_n(\mathbf{x})$ plays an important role in the non-robustness of $\hat{\Gamma}_2 = Q_{\alpha^*}$ estimator, compared to naïve estimator $\hat{\Gamma}_1$ where this term $\phi^{-1}(\alpha^*) \sigma_n(\mathbf{x})$ does not appear.

To illustrate this robustness issue, we define the Loggruy function in dimension 2 as follows:

$$\forall \mathbf{x} \in \mathbb{R}^2, \text{Loggruy}(\mathbf{x}) = 10 \log_{10} \left(3.6 + 10^4 \left(\sum_{i=1}^2 (\mathbf{x}_i - \mathbf{a}_i)^2 - r^2 \right) \times \cdots \times \left(\sum_{i=1}^2 (\mathbf{x}_i - \mathbf{e}_i)^2 - r^2 \right) \right) \quad (40)$$

with $(\mathbf{a}_i)_i = \begin{pmatrix} 0.153 \\ 0.939 \end{pmatrix}$; $(\mathbf{b}_i)_i = \begin{pmatrix} 0.854 \\ 0.814 \end{pmatrix}$; $(\mathbf{c}_i)_i = \begin{pmatrix} 0.510 \\ 0.621 \end{pmatrix}$; $(\mathbf{d}_i)_i = \begin{pmatrix} 0.207 \\ 0.386 \end{pmatrix}$; $(\mathbf{e}_i)_i = \begin{pmatrix} 0.815 \\ 0.146 \end{pmatrix}$
and $r := 0.07$.

The threshold chosen is $T = 10 \log_{10}(3.6)$ and the corresponding Γ^* excursion set (Equation (1)) is composed of 5 disconnected components. This function is particularly interesting in our context, since it has strong gradients at the edges of the domain and weaker gradients in the middle, where the different zones of the Γ^* excursion set are located (Figure 8c). This means that the stationarity assumption of the kriging metamodel cannot be verified.

Figure 8 represents in the case of an enrichment of 100 and 200 points of SUR Vorob'ev, from an initial DoE of size 10, the contour lines of the coverage probability p_n and the kriging mean m_n . Estimators $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$ are also represented and compared to Γ^* . An irregularity is observed for the contour lines of p_n either after 100 or 200 iterations. But, m_n is relatively accurate even after 100 iterations. $\hat{\Gamma}_2$ estimator is then less performant than $\hat{\Gamma}_1$ estimator.

In summary, $\hat{\Gamma}_2$ estimator related to Vorob'ev theory is more dependent on the stationarity hypothesis than $\hat{\Gamma}_1$ naïve estimator, and this is essentially explained by the construction of Vorob'ev expectation which is more sensitive to the kriging standard deviation (Equation (26)). Consequently, the approximation error $\text{Err}(\hat{\Gamma}_2(\chi_{j,n}))$ is more dependent on the stationarity hypothesis than the naïve approximation error $\text{Err}(\hat{\Gamma}_1(\chi_{j,n}))$.

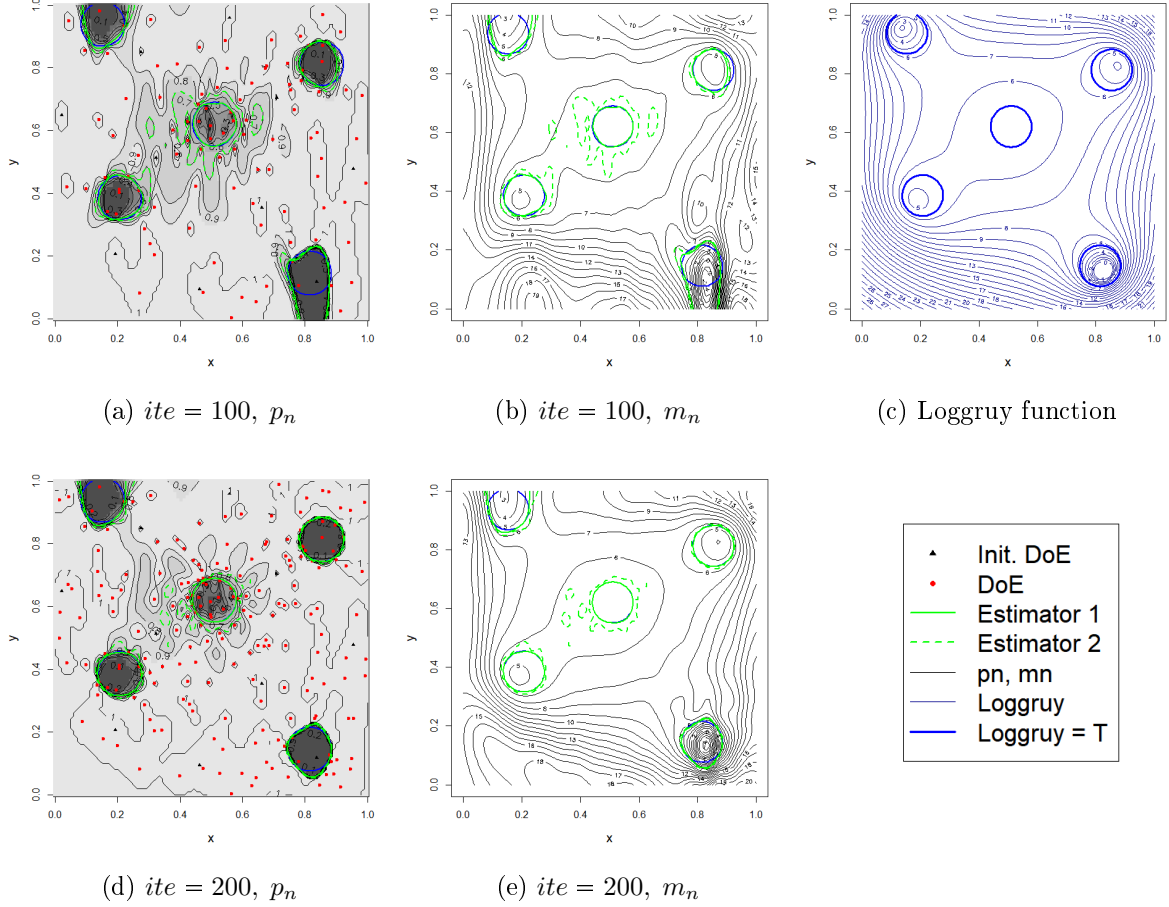


Figure 8: At left and center, representation of the two $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$ estimators after 100 iterations (first line) and 200 iterations (second line) for SUR Vorob'ev, and with the contour lines of p_n (first column) and m_n (second column), in comparison with the true excursion set (top right), for Loggruy function inversion ($d = 2$) with $T = 10 \log_{10}(3.6)$, for a particular initial DoE of size 10.