



**HAL**  
open science

# Non-Parametric Measure Approximations for Constrained Multi-Objective Optimisation under Uncertainty

Mickael Rivier, Nassim Razaaly, Pietro Marco Congedo

► **To cite this version:**

Mickael Rivier, Nassim Razaaly, Pietro Marco Congedo. Non-Parametric Measure Approximations for Constrained Multi-Objective Optimisation under Uncertainty. *International Journal for Numerical Methods in Engineering*, 2023, 10.1002/nme.7403 . hal-03781832

**HAL Id: hal-03781832**

**<https://inria.hal.science/hal-03781832v1>**

Submitted on 20 Sep 2022

**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.



Distributed under a Creative Commons Attribution 4.0 International License

# Non-Parametric Measure Approximations for Constrained Multi-Objective Optimisation under Uncertainty

M. Rivier<sup>a,b</sup>, N. Razaaly<sup>a,c</sup>, P.M. Congedo<sup>a</sup>

<sup>a</sup>*Inria, Centre des Mathématiques Appliquées, Ecole Polytechnique, IPP, 91120 Palaiseau, France*

<sup>b</sup>*ArianeGroup, Le Haillan, 33185 Cedex, France*

<sup>c</sup>*Institut Pprime, ISAE-ENSMA, 2 Bd des Frères Lumière, 86360 Chasseneuil-du-Poitou, France*

---

## Abstract

In this paper, we propose non-parametric estimations of robustness and reliability measures approximation error, employed in the context of constrained multi-objective optimisation under uncertainty. These approximations with tunable accuracy permit to capture the Pareto front in a parsimonious way, and can be exploited within an adaptive refinement strategy.

First, we illustrate an efficient approach for obtaining joint representations of the robustness and reliability measures, allowing sharper discrimination of Pareto-optimal designs. A specific surrogate model of these objectives and constraints is then proposed to accelerate the optimisation process.

Secondly, we propose an adaptive refinement strategy, using these tunable accuracy approximations to drive the computational effort towards the computation of the optimal area. To this extent, an adapted Pareto dominance rule and Pareto optimal probability computation are formulated.

The performance of the proposed strategy is assessed on several analytical test-cases against classical approaches. We also illustrate the method on an engineering application, performing shape optimisation under uncertainty of an Organic Rankine Cycle turbine.

*Keywords:*

Optimisation under uncertainty, Gaussian processes, Robustness and reliability, Organic Rankine cycle

---

## 1. Introduction

There has been a growing interest in optimisation with uncertainty in recent years [1, 2, 3, 4]. We focus here specifically on cost-efficient and derivative-free strategies for tackling multiple objectives and constraints in this context. Optimisation Under Uncertainty (OUU) can be roughly classified under two categories, Reliability-Based Design Optimisation (RBDO), which deals with probabilistic and worst-case feasibility constraints, and Robust Design Optimisation (RDO), where the deterministic objectives are replaced with averaged or worst-case ones, possibly in a multi-objective context (*i.e.* Taguchi optimisation).

9 In the context of multi-objective optimisation [5, 6], a popular derivative-free optimisa-  
10 tion strategy for constrained multi-objective problems are Evolutionary Algorithms (EA).  
11 A comparison of multi-objective EA strategies is proposed in [7], in the bi- or tri-objective  
12 setting. When dealing with more objectives ( $\geq 4$ ), one should refer to the many-objective  
13 optimisation literature, notably reviewed in [8]. For more details, [9] gives a thorough de-  
14 scription of both the theoretical foundations and practical metaheuristics for application to  
15 engineering test cases.

16 In this work, we focus specifically on performing optimisation in the presence of uncer-  
17 tainties. Such a setting has received significant attention from the optimisation community,  
18 as can be seen in several reviews [10, 11, 12]. Notably, [10] proposed four types of uncertain-  
19 ties: (i) noise, which is an intrinsic bias on the evaluation of the fitness function, coming,  
20 for example, from imperfect sensors; (ii) robustness, which refers to specific parameters that  
21 may vary from their nominal value between the optimisation procedure and the real-world  
22 application; (iii) fitness approximation, when the true fitness value is not computable or very  
23 expensive, and less accurate but cheaper representation is used instead; (iv) time-varying  
24 fitness function, meaning that the optimum changes through time and that the optimisation  
25 algorithm should tackle this variability.

26 Much work has been devoted to tackling noise within the optimisation process. This  
27 field is usually referred to as Noisy Optimisation (NO). In this context, some parsimonious  
28 techniques rely on the assumption that there exist either different levels of accuracy for  
29 computing the Quantities of Interest. Namely, multi-level techniques [13, 14] usually involve  
30 several mesh refinements for numerical simulation, and multi-fidelity approaches [15, 16]  
31 use different modeling strategies. Recent works have also proposed adaptive approaches for  
32 tackling problems where accuracy can be tuned online. This approach is particularly adapted  
33 when the fitness function is the limit of intermediate approximations of increasing fidelities,  
34 such as iterative solvers or Monte Carlo estimations. This concept has been proposed in [17]  
35 with online computation time allocation in a Bayesian Optimisation (BO) context. In this  
36 work, a similar online refinement strategy is employed to achieve high parsimony.

37 In a multi-objective setting, the presence of noise and cohabitation of different approx-  
38 imation accuracies raises the need for extending the classical Pareto dominance to an un-  
39 certain environment. The EA community has proposed techniques for problems featuring  
40 uniform [18] or Gaussian noises [19, 20, 21]. Interval and worst-case approaches are tackled  
41 in [22, 23, 24], within the classical NSGA-II and MOEA frameworks. Recent developments  
42 [25, 26, 27, 28] extends these computations to non-parametric noise distributions through  
43 the construction of histogram-based approximate densities, which permits fast computations  
44 of the probability of dominance. A straightforward extension of the Pareto dominance rule  
45 to interval uncertainty in the presence of constraints is proposed in [29] under the name of  
46 Bounding-Boxes. This extension has been exploited in several papers [30, 31] within an adap-  
47 tive strategy for performing multi-objective optimisation on objective functions computed  
48 with tunable accuracy.

49 In this paper, the core idea of the approach is to tune the accuracy of statistical objec-  
50 tives and constraints approximations to focus computational power on the most promising

51 designs. The first contribution is the computation of non-parametric representation of the  
52 robustness and reliability measures error, permitting to estimate the Pareto front in a par-  
53 simonious way. Specifically, we aim to avoid the restriction to independent uniform [31] or  
54 Gaussian [32, 33, 34, 35] distributions, which often rely on the so-called projected process  
55 technique. We propose to use empirical sampling-based distribution approximations, that  
56 should allow for more representative error estimations given the current information, and  
57 hopefully quicker convergence. An original approach is introduced for constructing these  
58 measure samples, which relies on a coupled space Gaussian Process (GP) surrogate model  
59 for drawing joint realisations of objectives and constraints over the whole design space. These  
60 approximated measures are treated as a random vector and are compared and ranked using  
61 the concept of probabilistic Pareto dominance [36, 26, 27]. The second main contribution  
62 of this work is an adaptive refinement strategy, where the accuracy is improved only on  
63 the most promising individuals. We assess the proposed approach on several (mono and bi-  
64 objective) optimisation problems to estimate its computational cost and accuracy in terms  
65 of average distance to the exact Pareto front. Finally, we tackle the problem of an Organic  
66 Rankine turbine shape optimisation with random operating conditions.

67 Section 2 illustrates the formulation of multi-objective optimisation problems under un-  
68 certainties treated here within a probabilistic setting. In Section 3, we introduce the nu-  
69 merical ingredients permitting to compute non-parametric measure approximations. Then,  
70 Section 4 illustrates the proposed general framework to solve the OUU problem. In Sec-  
71 tion 5, we analyse the performance of the proposed approach on some algebraic test cases  
72 by comparing it with known techniques in the literature. Finally, Section 6 is devoted to  
73 showing the applicability of the framework to complex engineering test cases with the shape  
74 optimisation of Organic Rankine cycles turbines.

## 75 2. Uncertainty-based optimisation problem

76 In this section, the context of this work is made explicit. The general formulation of  
77 the tackled uncertainty-based optimisation problem is given, and classical choices in terms  
78 of robustness and reliability measures are introduced. The proposed probabilistic setting  
79 is then given, which allows for comparing aleatory values in a constrained multi-objective  
80 context.

### 81 2.1. Optimisation

A deterministic constrained multi-objective optimisation problem can be written as fol-  
lows:

$$\begin{aligned}
& \text{minimise:} && \mathbf{f}(\mathbf{x}), \\
& \text{satisfying:} && \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \\
& \text{by changing:} && \mathbf{x} \in \mathcal{X},
\end{aligned} \tag{1}$$

82 with objectives  $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^{m_1}$  and constraints  $\mathbf{g}(\mathbf{x}) \in \mathbb{R}^{m_2}$ , and where the design vector  $\mathbf{x}$   
83 lives in  $\mathcal{X} \subset \mathbb{R}^n$ . Note that  $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$  stands for  $\forall i, g_i(\mathbf{x}) \leq 0$ .

84 The computation of  $\mathbf{f}$  and  $\mathbf{g}$  can also depend on a multitude of uncontrollable parameters  
 85 (*e.g.* environmental, material, geometrical, ...), here denoted by  $\boldsymbol{\xi} \in \Xi$ .

86 The most straightforward setting is to consider these uncertain parameters as aleatoric.  
 87 This work is limited to this scope and does not deal with epistemic uncertainties. In this  
 88 context,  $\boldsymbol{\xi}$ ,  $\mathbf{f}(\mathbf{x}, \boldsymbol{\xi})$  and  $\mathbf{g}(\mathbf{x}, \boldsymbol{\xi})$  are random variables and Equation (1) is not suitable any-  
 89 more.

90 It is classical to reformulate this problem into a measure-based Optimisation Under Un-  
 91 certainty (OUU) problem, where specific statistics of the random outputs  $\mathbf{f}(\mathbf{x}, \boldsymbol{\xi})$  and  $\mathbf{g}(\mathbf{x}, \boldsymbol{\xi})$   
 92 are plugged as objective and constraint functions.

## 93 2.2. Robustness and reliability measures

Writing  $\boldsymbol{\rho}_f$  and  $\boldsymbol{\rho}_g$  some statistical measures on the objective function  $\mathbf{f}$  and constraint  
 $\mathbf{g}$ , the OUU problem can be formulated as follows:

$$\begin{aligned} & \text{minimise: } \boldsymbol{\rho}_f(\mathbf{x}), \\ & \text{satisfying: } \boldsymbol{\rho}_g(\mathbf{x}) \leq \mathbf{0}, \\ & \text{by changing: } \mathbf{x} \in \mathcal{X}. \end{aligned} \tag{2}$$

94 As raised in the introduction, using statistical measures as objectives is usually referred  
 95 to as Robust Design Optimisation (RDO), and these statistics ( $\boldsymbol{\rho}_f$ ) are called robustness  
 96 measures. Similarly, Reliability-Based Design Optimisation (RBDO) deals with so-called  
 97 reliability measures  $\boldsymbol{\rho}_g$  as constraints.

In the following, the original functions  $\mathbf{f}$  and  $\mathbf{g}$  are gathered in a vector  $\mathbf{q} \in \mathbb{R}^{m_1+m_2}$ ,  
 that refers to Quantities of Interest (QoI). Namely,

$$\forall (\mathbf{x}, \boldsymbol{\xi}) \in \mathcal{X} \times \Xi, \mathbf{q}(\mathbf{x}, \boldsymbol{\xi}) = \begin{pmatrix} \mathbf{f}(\mathbf{x}, \boldsymbol{\xi}) \\ \mathbf{g}(\mathbf{x}, \boldsymbol{\xi}) \end{pmatrix}. \tag{3}$$

Similarly,  $\boldsymbol{\rho}_f$  and  $\boldsymbol{\rho}_g$  are gathered in a single statistical measures vector  $\boldsymbol{\rho} \in \mathbb{R}^{m_1+m_2}$ ,

$$\forall \mathbf{x} \in \mathcal{X}, \boldsymbol{\rho}(\mathbf{x}) = \begin{pmatrix} \boldsymbol{\rho}_f(\mathbf{x}) \\ \boldsymbol{\rho}_g(\mathbf{x}) \end{pmatrix}. \tag{4}$$

98 The most classical robustness and reliability measures for problem (2) are the following:

- 99 • Expectation:  $\rho(\mathbf{x}) = \mathbb{E}_{\boldsymbol{\xi}}[q(\mathbf{x}, \boldsymbol{\xi})]$ ,
- 100 • Variance:  $\rho(\mathbf{x}) = \mathbb{V}_{\boldsymbol{\xi}}[q(\mathbf{x}, \boldsymbol{\xi})]$ ,
- 101 • Worst case:  $\rho(\mathbf{x}) = \max_{\boldsymbol{\xi}}[q(\mathbf{x}, \boldsymbol{\xi})]$
- 102 • Quantile associated with probability  $p$ :  $\rho(\mathbf{x}) = q_{\boldsymbol{\xi}}^p[q(\mathbf{x}, \boldsymbol{\xi})]$  (also called Value at Risk,  
 103 VaR) ,
- 104 • Superquantile:  $\rho(\mathbf{x}) = \mathbb{E}_{\boldsymbol{\xi}}[q(\mathbf{x}, \boldsymbol{\xi}) \mid q(\mathbf{x}, \boldsymbol{\xi}) \geq q_{\boldsymbol{\xi}}^p[q(\mathbf{x}, \boldsymbol{\xi})]]$  (also called Expected Value  
 105 at Risk, EVaR).

106 Note that any combination of these may be used, and one has to choose the most relevant  
 107 for a given application.

108 In practice, robust objectives usually are either expectation or worst-case measures. The  
 109 latest can also be relaxed to quantile or superquantile measures.

110 As for reliability measures, they can impose hard constraint, by using worst-case mea-  
 111 sures, but are usually relaxed into probabilistic constraints, that ensures  $\mathbf{g}(\mathbf{x}, \boldsymbol{\xi}) \leq \mathbf{0}$  with  
 112 a given probability  $p$ . This can be readily obtained by using the  $p$ -quantile as a reliability  
 113 measure.

114 Given the distribution of the random input variable  $\boldsymbol{\xi}$ , and specific measures  $\boldsymbol{\rho}$ , prob-  
 115 lem (2) can be solved with classical techniques. However, if the unitary evaluation of  $\mathbf{q}$  is  
 116 expensive,  $\boldsymbol{\rho}$  becomes very tedious to compute. A compromise between the accuracy and  
 117 computational cost of  $\boldsymbol{\rho}$  must be chosen.

118 Specifically, we propose to consider these estimated measures as random variables, with  
 119 a given Probability Density Function (PDF) representing the associated inaccuracy. Next  
 120 section presents this probabilistic setting in details.

### 121 2.3. Probabilistic setting

122 In this work, we aim to tackle multi-objective optimisation problems ( $\leq 3$  objectives)  
 123 where the statistical measures  $\boldsymbol{\rho}$  used as objectives and constraint functions are not exactly  
 124 known and regarded as random fields  $\mathbf{P}$  over  $\mathcal{X}$ . In this context, the classical Pareto dom-  
 125 inance does not hold and a probabilistic one must be defined. Several propositions exist in  
 126 the literature [36, 26, 27], notably differing by the assumed distribution shape.

127 The classical Pareto dominance is described in the case of minimisation. For two vectors  
 128  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ ,

$$\begin{aligned}
 \mathbf{a} \succ \mathbf{b} \text{ (a dominates b)} &\iff \forall j \in \llbracket 1, m \rrbracket, a_j \leq b_j \quad \text{and} \\
 &\quad \exists j \in \llbracket 1, m \rrbracket, a_j < b_j, \\
 \mathbf{a} \succ\!\succ \mathbf{b} \text{ (a strictly dominates b)} &\iff \forall j \in \llbracket 1, m \rrbracket, a_j < b_j, \\
 \mathbf{a} \sim \mathbf{b} \text{ (a is indifferent to b)} &\iff \mathbf{a} \not\prec \mathbf{b} \text{ and } \mathbf{b} \not\prec \mathbf{a}.
 \end{aligned} \tag{5}$$

129 In the presence of constraints, we follow the constrained Pareto dominance proposed in  
 130 [31]. For two vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ , with  $\mathbf{a}_f$  and  $\mathbf{b}_f$  the objective values and  $\mathbf{a}_g$  and  $\mathbf{b}_g$  the  
 131 constraint values,

$$\begin{aligned}
 \mathbf{a} \succ_c \mathbf{b} &\iff \mathbf{a} \in \mathcal{A} \text{ and } \mathbf{a}_f \succ \mathbf{b}_f \text{ or } \mathbf{b} \in \mathcal{F}, \\
 \mathbf{a} \succ\!\succ_c \mathbf{b} &\iff \mathbf{a} \in \mathcal{A} \text{ and } \mathbf{a}_f \succ\!\succ \mathbf{b}_f \text{ or } \mathbf{b} \in \mathcal{F}, \\
 \mathbf{a} \sim_c \mathbf{b} &\iff \mathbf{a} \not\prec_c \mathbf{b} \text{ and } \mathbf{b} \not\prec_c \mathbf{a},
 \end{aligned} \tag{6}$$

132 where  $\mathcal{A}$  is the admissible set, where all constraints are satisfied, and  $\mathcal{F}$  is the failure set,  
 133 where at least one constraint is not verified. Using this definition, two individuals that both  
 134 violate the constraints will mutually dominate each other. This allows to discard inefficient  
 135 individuals even when no satisfying individual has yet been found.

136 Here, a probabilistic Pareto dominance between two designs  $\mathbf{x}_1$  and  $\mathbf{x}_2$  hence consists in  
 137 computing the probability  $\mathbb{P}_{\mathbf{P}}[\mathbf{P}(\mathbf{x}_1) \succ_c \mathbf{P}(\mathbf{x}_2)]$  of a design to dominate the other. Recall  
 138 that  $\mathbf{P}$  is a random field, meaning that  $\mathbf{P}(\mathbf{x}_1)$  and  $\mathbf{P}(\mathbf{x}_2)$  are random variables. For better  
 139 generality and readability, these are denoted  $\mathbf{A}$  and  $\mathbf{B}$  in the following.

140 Formally, the unconstrained probability is defined as:

$$\mathbb{P}_{\mathbf{A},\mathbf{B}}[\mathbf{A} \succ \mathbf{B}] = \int_{\mathbb{R}_-^m} \phi_{\mathbf{A}-\mathbf{B}}(\mathbf{y})d\mathbf{y} = \Phi_{\mathbf{A}-\mathbf{B}}(\mathbf{0}), \quad (7)$$

141 where  $\phi_{\mathbf{A}-\mathbf{B}}$  is the Probability Density Function (PDF) of  $\mathbf{A}-\mathbf{B}$  and  $\Phi_{\mathbf{A}-\mathbf{B}}$  is the associated  
 142 Cumulative Density Function (CDF). An equivalent formula can be written using the joint  
 143 PDF  $\phi_{\mathbf{A},\mathbf{B}}$  of  $\mathbf{A}$  and  $\mathbf{B}$ :

$$\mathbb{P}_{\mathbf{A},\mathbf{B}}[\mathbf{A} \succ \mathbf{B}] = \iint_{\mathbf{a} \succ \mathbf{b}} \phi_{\mathbf{A},\mathbf{B}}(\mathbf{a}, \mathbf{b})d\mathbf{a}d\mathbf{b}. \quad (8)$$

144 As for the constrained case, definition (6) shows that for  $\mathbf{a}$  to dominate  $\mathbf{b}$ , either  $\mathbf{b}$  must  
 145 be in the failure domain, or  $\mathbf{a}$  must be in the admissible domain and dominate  $\mathbf{b}$  according  
 146 to the classical Pareto dominance rule. The integral in (8) can be split up depending on  $\mathbf{b}$   
 147 belonging to the failure or admissible domain, and an additional dominance constraint can  
 148 be added in the latest case. Hence, the domination probability reads:

$$\mathbb{P}_{\mathbf{A},\mathbf{B}}[\mathbf{A} \succ_c \mathbf{B}] = \int_{\mathbf{b} \in \mathcal{F}} \phi_{\mathbf{B}}(\mathbf{b})d\mathbf{b} + \iint_{\substack{\mathbf{a} \in \mathcal{A} \\ \mathbf{b} \in \mathcal{A} \\ \mathbf{a} \succ \mathbf{b}}} \phi_{\mathbf{A},\mathbf{B}}(\mathbf{a}, \mathbf{b})d\mathbf{a}d\mathbf{b}. \quad (9)$$

149 Note that this formulation naturally tackles the existence of dependency between the  
 150 random variables  $\mathbf{A}$  and  $\mathbf{B}$ , which greatly differs from the assumptions made in [26, 27].

151 We finally propose a *probabilistic constrained Pareto dominance* rule between two random  
 152 vectors:

$$\begin{aligned} \mathbf{A} \underset{0}{\succ}_c \mathbf{B} &\iff \mathbb{P}_{\mathbf{A},\mathbf{B}}[\mathbf{A} \succ_c \mathbf{B}] = 1, \\ \mathbf{A} \underset{0}{\succ\!\succ}_c \mathbf{B} &\iff \mathbb{P}_{\mathbf{A},\mathbf{B}}[\mathbf{A} \succ\!\succ_c \mathbf{B}] = 1, \\ \mathbf{A} \underset{0}{\sim}_c \mathbf{B} &\iff \mathbf{A} \not\underset{0}{\succ}_c \mathbf{B} \text{ and } \mathbf{B} \not\underset{0}{\succ}_c \mathbf{A}. \end{aligned} \quad (10)$$

153 This dominance rule generalises the *constrained boxed Pareto dominance* introduced in [31].

154 However, in the context of random variables with infinite support (*e.g.* Gaussian vari-  
 155 able), the dominance probability cannot reach exactly 1. It has been proposed in [26] to  
 156 relax this threshold. The  $\epsilon$ -relaxed *probabilistic constrained Pareto dominance* finally reads:

$$\begin{aligned} \text{Domination: } \mathbf{A} \underset{\epsilon}{\succ}_c \mathbf{B} &\iff \mathbb{P}_{\mathbf{A},\mathbf{B}}[\mathbf{A} \succ_c \mathbf{B}] \geq 1 - \epsilon, \\ \text{Strict domination: } \mathbf{A} \underset{\epsilon}{\succ\!\succ}_c \mathbf{B} &\iff \mathbb{P}_{\mathbf{A},\mathbf{B}}[\mathbf{A} \succ\!\succ_c \mathbf{B}] \geq 1 - \epsilon, \\ \text{Indifference: } \mathbf{A} \underset{\epsilon}{\sim}_c \mathbf{B} &\iff \mathbf{A} \not\underset{\epsilon}{\succ}_c \mathbf{B} \text{ and } \mathbf{B} \not\underset{\epsilon}{\succ}_c \mathbf{A}. \end{aligned} \quad (11)$$

157 The above formulation gives a general setting for comparing aleatory vectors, regardless  
 158 of their distribution shape. In practice,  $\epsilon$  is usually taken at  $10^{-2}$ . Next section proposes a  
 159 sampling-based technique to get an accurate representation of the random field  $\mathbf{P}$  over the  
 160 design space  $\mathcal{X}$ .

### 161 3. Non-parametric measure approximation

162 We aim to make as few assumptions as possible on the distribution of  $\mathbf{P}(\mathbf{x})$ , for  $\mathbf{x}$  in  
 163  $\mathcal{X}$ . To this extent, we only assume that realisations  $\mathbf{P}^{(i)}(\mathbf{x})$  can be drawn and exploit this  
 164 empirical set of samples to characterise  $\mathbf{P}(\mathbf{x})$ .

165 More precisely, because  $\mathbf{P}$  is a random field over the design space  $\mathcal{X}$ , we propose to draw  
 166 realisations  $\mathbf{P}^{(i)}$  of the whole random field. This consists in jointly drawing  $\mathbf{P}^{(i)}(\mathbf{x})$  on a set  
 167 of locations  $\mathbf{x}$ .

168 In this setting, dominance must only be verified between these joint samples. With a  
 169 number  $N_{samples}$  of realisations, the probability of dominance between two designs can be  
 170 empirically estimated as follows:

$$\mathbb{P}_{\mathbf{P}}[\mathbf{P}(\mathbf{x}_j) \not\prec_c \mathbf{P}(\mathbf{x}_k)] \simeq \frac{1}{N_{samples}} \sum_{i=1}^{N_{samples}} \mathbb{1}_{\{\mathbf{P}^{(i)}(\mathbf{x}_j) \not\prec_c \mathbf{P}^{(i)}(\mathbf{x}_k)\}}. \quad (12)$$

171 In practice, we tend to take  $N_{samples}$  between  $10^3$  and  $10^4$ .

172 The major task is to obtain realisations  $\mathbf{P}^{(i)}(\mathbf{x})$  of the robustness and reliability measures.  
 173 To this extent, we assume that a probabilistic surrogate can be constructed on the original  
 174 Quantities of Interest  $\mathbf{q}$  and draw realisations  $\mathbf{P}^{(i)}$  through measure estimations on functional  
 175 realisations  $\mathbf{q}^{(i)}(\mathbf{x}, \cdot)$ , simultaneously for every studied location  $\mathbf{x} \in \mathcal{X}$ . In this work, we use  
 176 Gaussian process surrogate models, that are built using the GPy python package [37].

177 We first recall the main characteristics and formulas of Gaussian Process surrogate mod-  
 178 els. Then, an approximated approach for drawing these realisations of  $\mathbf{P}$  from a GP surrogate  
 179 of  $\mathbf{q}$  in the coupled space  $\mathcal{X} \times \Xi$  is proposed.

#### 180 3.1. Gaussian processes

181 This section gives a quick reminder of the GP setting, its interpretability and the main  
 182 formulas that will be used in the following. Gaussian process surrogate modelling, also called  
 183 Kriging, is a Bayesian kernel-based regression method [38].

184 The kernel setting, with the famous so-called kernel trick, allows to transpose most of  
 185 the classical parametric (*e.g.* linear) approaches to the non-parametric world. In practice,  
 186 basing on data points location  $\mathbf{x}_i$ , parametric features can be incorporated by replacing scalar  
 187 product between the different parameters  $\mathbf{x}_1^\top \mathbf{x}_2$  by a scalar product between chosen features  
 188  $\phi(\mathbf{x}_1)^\top \phi(\mathbf{x}_2)$ . Then, the kernel trick consists in choosing a kernel function  $k$ , and to use  
 189  $k(\mathbf{x}_1, \mathbf{x}_2)$  in place of the previous scalar product. In this way, the features are not explicitly  
 190 defined and can be of infinite number. Some kernels are called universal and engender a  
 191 space of functions that is dense in the space of continuous function.



192 The Bayesian setting, in addition, tracks the distribution of all parameters (of infinite  
 193 number with the kernel trick), and returns not only a predictive model of the Quantity of  
 194 Interest, but its whole distribution at each location  $\mathbf{x}$ . In the context of Gaussian process,  
 195 this distribution is fully determined by a predictive mean and a predictive variance for the  
 196 QoI.

In practice, Gaussian processes can also simply be viewed as an infinite dimensional  
 Gaussian distribution, which reduction to a finite set of points is also Gaussian. In this  
 setting, the kernel plays the role of covariance function between points and the predictive  
 mean and variance are the conditional mean and variances at each point  $\mathbf{x}$ . Given the  
 training data  $\{\mathbf{x}_{train}, \mathbf{y}_{train}\}$  where  $\mathbf{y}_{train} = f(\mathbf{x}_{train}) + \epsilon$  and  $\epsilon$  is a random Gaussian noise,  
 they read:

$$\begin{aligned}\boldsymbol{\mu}_{GP}(\mathbf{x}) &= \mathbf{k}_*^\top (K + \Delta)^{-1} \mathbf{y}_{train}, \\ \Sigma_{GP}(\mathbf{x}) &= \mathbf{k}_{**} - \mathbf{k}_*^\top (K + \Delta)^{-1} \mathbf{k}_*,\end{aligned}\tag{13}$$

197 where  $K$  is the covariance matrix between training points  $\mathbf{x}_{train}$ ,  $K = k(\mathbf{x}_{train}, \mathbf{x}_{train})$ ,  $\mathbf{k}_*$  is  
 198 the covariance  $k(\mathbf{x}_{train}, \mathbf{x})$  between the new points  $\mathbf{x}$  and the training points and  $\mathbf{k}_{**}$  is the  
 199 covariance between the new points,  $k(\mathbf{x}, \mathbf{x})$ . Matrix  $\Delta$  refers to the so-called input noise or  
 200 nugget effect. Usually, it regularises the problem with the same noise on each training point,  
 201  $\Delta = \sigma^2 I_N$ , with  $N$  the number of training data. However, in the heteroscedastic case, one  
 202 can readily assign a different noise to each training point, meaning  $\Delta = \text{diag}(\{\sigma_i^2\}_i)$ .

203 In this paper, hyperparameters of the kernel function are optimised by Maximum Like-  
 204 lihood. Because Automatic Relevance Determination (ARD) is exploited, a lengthscale  
 205 parameter is assigned to each dimension. This heavily limits the use of these models when  
 206 the problem features more than tens of dimensions.

Finally, one can draw functional realisations of a GP surrogate model, discretised on a  
 finite set of points. A GP is a Gaussian random vector, so a realisation  $\mathbf{f}^{(i)} \sim \mathcal{N}(\boldsymbol{\mu}_{GP}, \Sigma_{GP})$   
 is drawn as follows:

$$\mathbf{f}^{(i)} = \boldsymbol{\mu}_{GP} + L_{GP} Y^{(i)}, \quad Y \sim \mathcal{N}(\mathbf{0}, I_N),\tag{14}$$

207 with  $L_{GP}$  referring to the Cholesky decomposition of the covariance matrix, so that  $\Sigma_{GP} =$   
 208  $L_{GP} L_{GP}^\top$ .

### 209 3.2. Direct measure sampling in coupled space

210 We assume that a GP surrogate model can be built on  $\mathbf{q}$  in the coupled design/uncertainty  
 211 space. We here simply build a scalar GP for each output dimension. Based on training data  
 212  $\{\mathbf{q}(\mathbf{x}_i, \boldsymbol{\xi}_i)\}_i$ , the surrogate  $\hat{\mathbf{q}}$  is constructed and yields an estimation of  $\mathbf{q}$  at any location  
 213  $(\mathbf{x}, \boldsymbol{\xi})$  in the coupled space  $\mathcal{X} \times \Xi$ .

This surrogate model can be exploited to obtain an estimation of the robustness and  
 reliability measure through Monte Carlo integration. This integration is directly performed  
 on the predictive mean  $\hat{\mathbf{q}}$  of the GP. For example, with the expectation measure  $\boldsymbol{\rho}(\mathbf{x}) =$

$\mathbb{E}_{\boldsymbol{\xi}}[\mathbf{q}(\mathbf{x}, \boldsymbol{\xi})]$ , one can use the estimator :

$$\tilde{\boldsymbol{\rho}}(\mathbf{x}) = \frac{1}{N_{\boldsymbol{\xi}}} \sum_{j=1}^{N_{\boldsymbol{\xi}}} \hat{\mathbf{q}}(\mathbf{x}, \boldsymbol{\xi}_j). \quad (15)$$

214 More interestingly, one can draw realisations of the GP built on  $\mathbf{q}$ , which directly trans-  
 215 lates into realisations of  $\mathbf{P}$  by integrating over the uncertain dimension.

For example, in the case of the expectation measure  $\boldsymbol{\rho}(\mathbf{x}) = \mathbb{E}_{\boldsymbol{\xi}}[\mathbf{q}(\mathbf{x}, \boldsymbol{\xi})]$ , by evaluating  $\mathbf{q}$  with  $N_{\boldsymbol{\xi}}$  locations in the uncertain space at  $N_x$  different designs, the joint realisations of  $\mathbf{P}$  over all these designs write as follows:

$$\forall k \in \llbracket 1, N_x \rrbracket, \quad \mathbf{P}^{(i)}(\mathbf{x}_k) = \frac{1}{N_{\boldsymbol{\xi}}} \sum_{j=1}^{N_{\boldsymbol{\xi}}} \mathbf{q}^{(i)}(\mathbf{x}_k, \boldsymbol{\xi}_j), \quad (16)$$

216 drawn simultaneously on a realisation  $\mathbf{q}^{(i)}$  at all designs.

217 The same can be carried out with empirical estimators of other robustness and reliability  
 218 measures.

219 However, one can notice that  $\mathbf{q}^{(i)}$  is drawn at  $N_x \times N_{\boldsymbol{\xi}}$  locations, which can quickly reach  
 220 a value of more than  $10^5$ . Looking back at Equation (14), it reveals that Cholesky decom-  
 221 position should hence be performed on a matrix of size  $10^5 \times 10^5$ . The computational cost  
 222 associated with such a task is prohibitive. This problem is quite common in the literature,  
 223 and several approximations techniques have been proposed, such as a circulant embedding  
 224 of the covariance matrix [39], regularisation [40] or the use of a Gibbs sampler for Gaussian  
 225 Markov Random Field [41]. However, some of these approaches require equispaced points,  
 226 and the others remain too expensive to allow substantial sampling of the  $\mathbf{q}$  random field.

227 As raised in [42], the complexity is particularly high when the domain is much larger  
 228 than the characteristic correlation length of the random field. Intuitively, when a value is  
 229 drawn, it gives information in its surrounding at a distance proportional to the correlation  
 230 length.

231 Based on this idea of local information, an inducing point strategy is proposed in the next  
 232 section for drawing realisations  $\mathbf{q}^{(i)}$  at an extremely high number of locations simultaneously.  
 233 This strategy is presented in Section 3.2.1, and a quantitative quality indicator is proposed  
 234 in Section 3.2.2 for use in an iterative strategy in Section 3.2.3.

### 235 3.2.1. Inducing points strategy

236 The central idea is to draw realisations  $\mathbf{q}^{(i)}$  of the Gaussian random field on few well-  
 237 spread points, from which a smooth functional realisation can be extended to the whole  
 238 input space. These well-chosen points will be referred to as inducing points.

239 In the following, for conciseness, coordinates  $(\mathbf{x}, \boldsymbol{\xi})$  in the coupled space are denoted  
 240 with  $\mathbf{z}$ . With a set of training designs  $\mathbf{z}_{train} = (\mathbf{x}_{train}, \boldsymbol{\xi}_{train})$  and a set of inducing points  
 241  $\mathbf{z}_{ind} = (\mathbf{x}_{ind}, \boldsymbol{\xi}_{ind})$ , we denote  $\mathbf{z}_{tot} = \mathbf{z}_{train} \cup \mathbf{z}_{ind}$ .

242 We assume that a GP has been constructed on  $\mathbf{z}_{train}$ , as plotted in Figure 1, on values  
 243  $\mathbf{q}_{train}$  and with an optimised kernel  $k_{train}$ . Its predictive mean and covariance are written

244  $\boldsymbol{\mu}_{train}$  and  $\Sigma_{train}$ . This surrogate model is the current estimation of  $\mathbf{q}$  over the whole coupled  
 245 space.

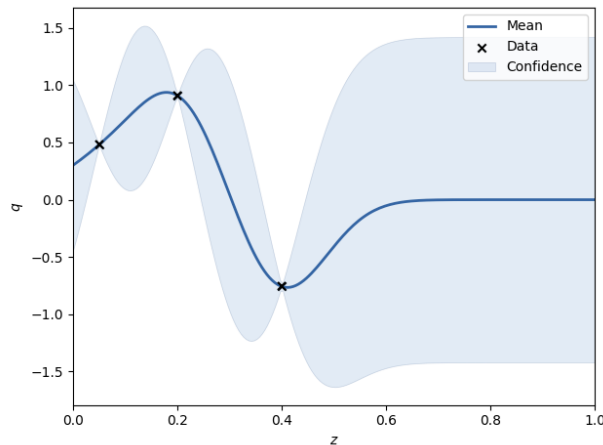


Figure 1: Gaussian process model built on the training data.

246 We aim to generate random realisations of this Gaussian field on an extensive set of points  
 247  $\mathbf{z}_{draw}$ . Direct Gaussian sampling would perform Cholesky decomposition on the matrix  
 248  $K_{draw} = \Sigma_{train}(\mathbf{z}_{draw})$  of size  $N_{draw}^2$ , which quickly becomes unfeasible.

We propose to only draw realisations on the inducing points:

$$\mathbf{q}_{ind}^{(i)} = \boldsymbol{\mu}_{train}(\mathbf{z}_{ind}) + L_{ind}\mathbf{Y}^{(i)}, \quad \mathbf{Y} \sim \mathcal{N}(\mathbf{0}, I_{N_{ind}}), \quad (17)$$

249 with  $L_{ind}$  the Cholesky decomposition of matrix  $K_{ind} = \Sigma_{train}(\mathbf{z}_{ind})$  of much smaller size  
 250  $N_{ind}^2$ . Figure 2 gives an example of such drawn values at the inducing points.

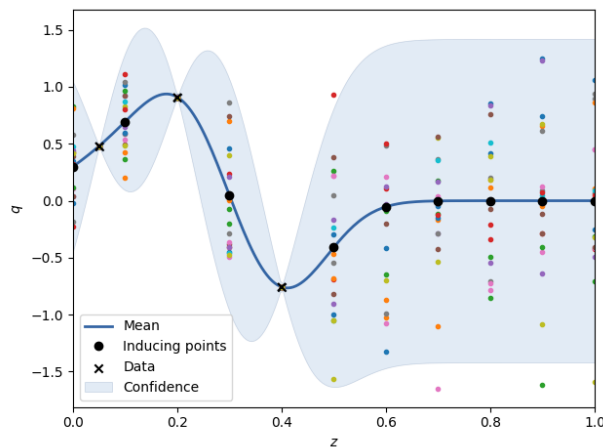


Figure 2: Realisations drawn at the inducing points.

Similarly to the locations  $\mathbf{z}_{tot}$ , we gather training values  $\mathbf{q}_{train}$  and drawn ones  $\mathbf{q}_{ind}^{(i)}$  in  $\mathbf{q}_{tot}^{(i)}$ . The realisations drawn at the inducing points are then extended to all locations  $\mathbf{z}_{draw}$

by constructing a new GP on the  $(\mathbf{z}_{tot}, \mathbf{q}_{tot}^{(i)})$  database. Note that the new GP kernel is not optimised as  $k_{train}$  is directly used. Final realisations estimates  $\mathbf{q}_{draw}^{(i)}$  are obtained as follows:

$$\mathbf{q}_{draw}^{(i)} = \boldsymbol{\mu}_{tot}^{(i)}(\mathbf{z}_{draw}) = \mathbf{k}_*^\top K_{tot}^{-1} \mathbf{q}_{tot}^{(i)} \quad (18)$$

251 where  $\mathbf{k}_* = k_{train}(\mathbf{z}_{tot}, \mathbf{z}_{draw})$  and  $K_{tot}^{-1}$  only need to be computed once as they do not depend  
 252 on  $i$ . These realisations are represented in Figure 3, and the empirical 2- $\sigma$  band of the  
 253 random field is computed on  $10^5$  samples. One can note that they fit well the real confidence  
 254 band of the underlying GP.

255 Intuitively, this approach corresponds to parametrising the random field as a sum of  
 256 kernels located at all training and inducing points. This is noticeable in the above Equation  
 257 (18). By denoting  $K_{tot}^{-1} \mathbf{q}_{tot}^{(i)}$  as a vector  $\boldsymbol{\beta}^{(i)}$  of coefficients, the proposed realisations  $\mathbf{q}_{draw}^{(i)}$   
 258 are linear combinations of the kernel function  $k_{train}$  located at each training and inducing point.  
 259 Coefficients  $\boldsymbol{\beta}^{(i)}$  manage the randomness and kernel functions the spatial characteristic. The  
 260 functional realisations write  $k_{train}(\mathbf{z}_{tot}, \cdot)^\top \boldsymbol{\beta}^{(i)}$ .

261 In practice, the proposed approach has a lot of similarities with using a Karhunen-Loève  
 262 Expansion (KLE) of a Gaussian random field with Nyström approximation of the basis  
 263 functions. Computational details are given in Appendix A.

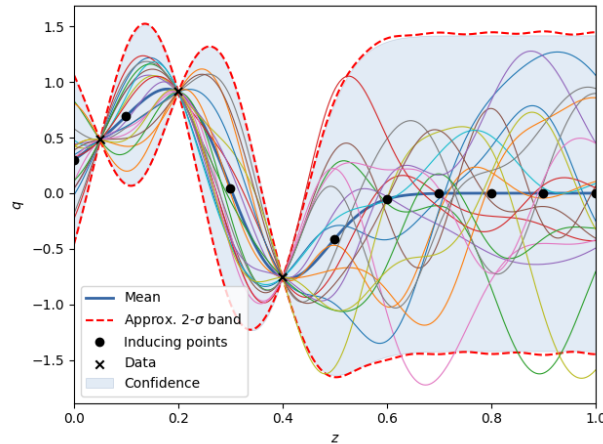


Figure 3: Functional realisations from draws in Figure 2. Approximated 2- $\sigma$  band of the stochastic process in dashed red.

### 264 3.2.2. Quality indicator

To attain a good accuracy for this random field approximation, a key element is the number of inducing points. We propose in this section an indicator to assess the approximation quality. Intuitively, if the density of inducing points is high enough with respect to the lengthscales, the predictive means  $\mathbf{q}_{draw}^{(i)}$  (Eq. (18)) will be able to represent all scales of the random process. In such cases, the associated predictive variance remains very low over the whole design space  $\mathcal{X}$ . The GP covariance reads:

$$\Sigma_{tot}(\mathbf{z}_{draw}) = K_{**} - \mathbf{k}_*^\top K_{tot}^{-1} \mathbf{k}_*, \quad (19)$$

with  $K_{**} = k_{train}(\mathbf{z}_{draw}, \mathbf{z}_{draw})$ . The predictive variance vector  $\sigma_{tot}^2(\mathbf{z}_{draw})$  is the diagonal of  $\Sigma_{tot}(\mathbf{z}_{draw})$ . We propose to base the quality indicator on the mean predictive variance over all points  $\mathbf{z}_{draw}$ . More precisely, as  $K_{tot}^{-1}$  is positive definite,  $\mathbf{k}_*^\top K_{tot}^{-1} \mathbf{k}_* > 0$  and thus  $\forall i, \sigma_{tot i}^2(\mathbf{z}_{draw}) \in [0, K_{**ii}]$ , we propose the following normalised indicator:

$$\alpha = 1 - \frac{1}{N_{draw}} \sum_{i=1}^{N_{draw}} \frac{\sigma_{tot i}^2(\mathbf{z}_{draw})}{K_{**ii}}, \quad (20)$$

265 where  $\alpha \in [0, 1]$  and  $\alpha \simeq 1$  denotes a low predictive variance, meaning a good approximation  
 266 quality.

Note that for most of the usual kernel functions, the diagonal of  $K_{**}$  is filled with  $\sigma^2$ , the variance hyperparameter of the chosen kernel. This is notably the case with the exponential, squared exponential and Matern kernels. In these cases, using  $\sigma^2 N_{draw} = \text{Tr}(K_{**})$  and linearity of the trace, the indicator reads:

$$\alpha = \frac{\text{Tr}(\mathbf{k}_*^\top K_{tot}^{-1} \mathbf{k}_*)}{\sigma^2 N_{draw}}. \quad (21)$$

One may recognise a Frobenius norm at the numerator of the proposed indicator:

$$\begin{aligned} \text{Tr}(\mathbf{k}_*^\top K_{tot}^{-1} \mathbf{k}_*) &= \text{Tr}(\mathbf{k}_*^\top (L_{tot} L_{tot}^\top)^{-1} \mathbf{k}_*), \\ &= \text{Tr}((L_{tot}^{-1} \mathbf{k}_*)^\top L_{tot}^{-1} \mathbf{k}_*), \\ &= \|L_{tot}^{-1} \mathbf{k}_*\|_F^2, \end{aligned} \quad (22)$$

267 where  $L_{tot}$  refers again to the Cholesky decomposition of  $K_{tot}$ .

268 The proposed indicator is pictured in Figure 4 with nine and three inducing points, re-  
 269 spectively. The approach gives an empirical  $2\text{-}\sigma$  band very close to the real one in Figure  
 270 4(a) associated with a value of the quality indicator of 0.995. On the contrary, the approxi-  
 271 mated random field in Figure 4(b) is quite different from the underlying GP, and the quality  
 272 indicator only has a value of 0.825.

For deeper understanding of this indicator, we study the behaviour of  $\alpha$  on an analytical example. With  $D$  the dimension of the input parameters  $\mathbf{z}$ , we study the following quantity of interest:

$$q(\mathbf{z}) = \sin(\mathbf{L}^\top \mathbf{z}) = \sin\left(\sum_{i=1}^D L_i z_i\right). \quad (23)$$

273 To vary the complexity of  $q$ , we regard parameters  $D$  and  $\mathbf{L}$  are random variables, which  
 274 yields a random function  $q_{D,\mathbf{L}}$ . Here, we take  $D \sim \mathcal{U}([1, 4])$  and  $\mathbf{L} \sim \mathcal{U}([3, 20])^D$ .

275 For each sampled function  $q_{D^{(k)}, \mathbf{L}^{(k)}}$ , a GP is built on 10 training data  $q_{D^{(k)}, \mathbf{L}^{(k)}}(\mathbf{z}_{train})$ .  
 276 A set of points  $\mathbf{z}_{draw}$  is drawn uniformly within the input space. The predictive covariance  
 277  $K$  on these points is given by Equation (13). A set of well spread inducing points of size  
 278  $N_{ind} \sim \mathcal{U}([10, 60])$  is also drawn (*e.g.* with a maximin Latin Hypercube Sampling (LHS)

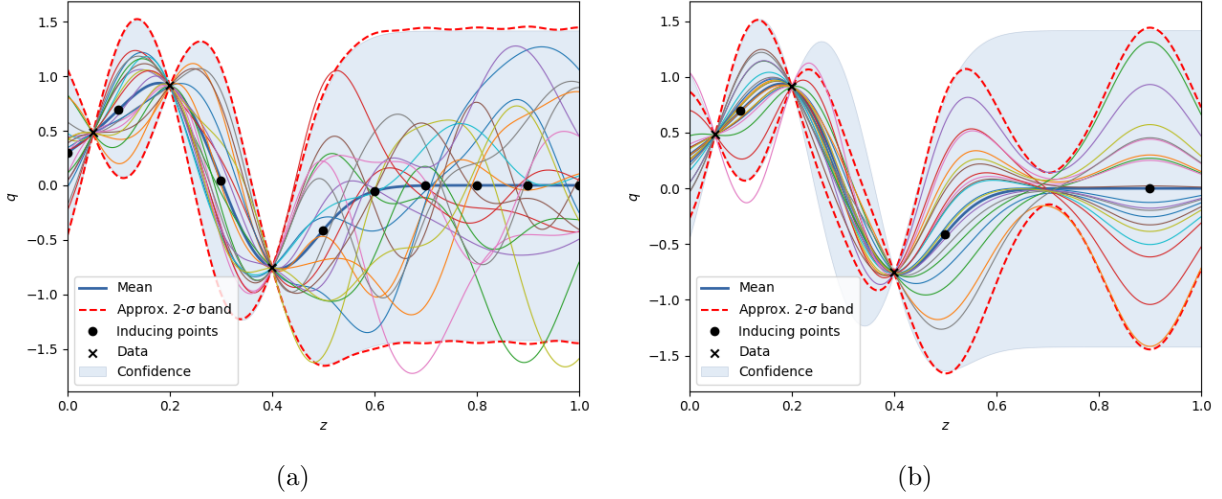


Figure 4: Approximated random field with (a) 9 inducing points,  $\alpha = 0.995$  and (b) 3 inducing points,  $\alpha = 0.825$ .

279 strategy) and  $10^5$  realisations  $q_{D^{(k)}, \mathbf{L}^{(k)}}^{(i)}(\mathbf{z}_{draw})$  are computed at  $10^3$  locations  $\mathbf{z}_{draw}$ . The  
 280 empirical covariance between all these realisations gives an approximation  $\tilde{K}$  of  $K$  and the  
 281 indicator  $\alpha$  from Equation (20) is computed.

282 All this process is repeated 300 times, and we plot in Figure 5 the evolution of the  
 283 normalised Frobenius distance between  $K$  and  $\tilde{K}$  with respect to  $\alpha$ .

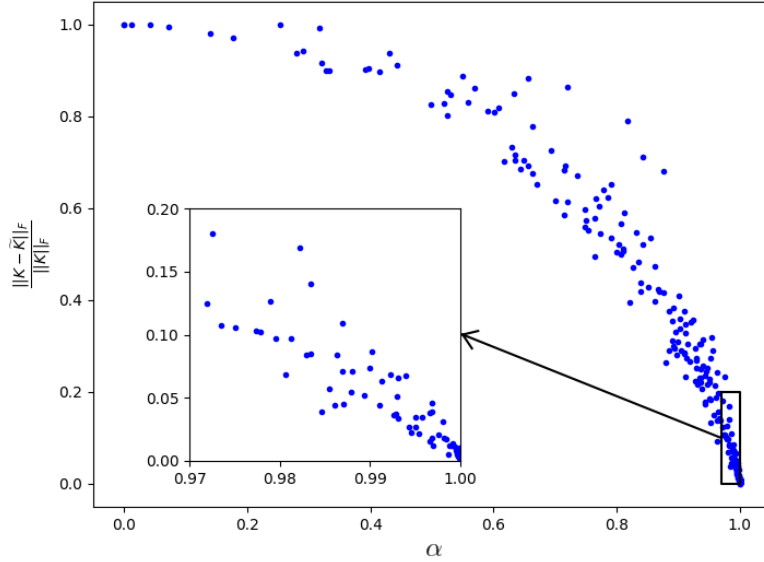


Figure 5: Accuracy of the empirical covariance  $\tilde{K}$  approximated with inducing points with respect to the proposed indicator.

284 Figure 5 shows that a strong correlation exists between the proposed indicator  $\alpha$  and  
 285 the accuracy of the empirical covariance. A value of the indicator very close to 1 seems to  
 286 ensure an accurate representation of the random field. Notably, the zoomed window permits  
 287 to see that when  $\alpha$  is greater than 0.99 or 0.995, the worst recorded empirical covariances  
 288 only yield 5 to 10% relative error.

289 Note that several appropriate norms for covariance matrices exist. Namely, the loga-  
 290 rithmic, Cholesky-based, and Riemannian distances between covariance matrices allow to  
 291 preserve positive definiteness and have many desirable properties [43]. However, because  
 292 we draw very badly conditioned empirical covariances, very small negative eigenvalues may  
 293 arise, that prohibit the use of such distances. For its simplicity and interpretability, the  
 294 Frobenius distance is the most reliable choice.

### 295 3.2.3. Adaptive strategy

296 To take advantage of the proposed quality indicator, we propose an iterative algorithm  
 297 ensuring that  $\alpha$  is sufficiently close to 1 before returning the approximated realisations.  
 298 The number of inducing points is hence increased as long as a prescribed value for  $\alpha$  has  
 299 not been reached. At each iteration, these points are spread in the input space using the  
 300 classical maximin criterion on a Latin Hypercube Sampling (LHS). The algorithm is given  
 301 in Algorithm 1.

---

#### Algorithm 1 Algorithm for choosing the number of inducing points

---

```

1: Choose threshold  $\alpha_* \in [0, 1]$ 
2: Choose initial number of inducing points  $N_{init}$  and added per iteration  $N_{add}$ 
3: Build a GP on the training data
4:  $N_{ind} = N_{init}$ 
5: while True do
6:   Construct a DoE of  $N_{ind}$  points using a maximin LHS
7:   Compute  $\alpha$  from Equation (20)
8:   if  $\alpha \geq \alpha_*$  then
9:     Draw realisations with Equation (18) using the DoE of inducing points
10:    Return realisations
11:   else
12:      $N_{ind} += N_{add}$ 
13:   end if
14: end while

```

---

302 In the following, the threshold  $\alpha_*$  will take values of 0.99 or 0.995, basing on Figure 5.  
 303 The simple approach proposed here is very intuitive and reveals very cost-efficient.

## 304 4. Application to Optimisation Under Uncertainty

305 In this section, we give a general technique, heavily inspired from the Surrogate-Assisted  
 306 Bounding-Box approach introduced in [44, 31], with the aim of making use of the non-

307 parametric approximations presented in the previous section in an Optimisation Under Un-  
 308 certainty (OUU) context.

309 A general formulation of the proposed approach is presented in Section 4.1, and its  
 310 application with non-parametric measure approximation is made explicit in Section 4.2.

#### 311 4.1. General formulation

312 The proposed approach integrates the concept of tunable accuracy within the OUU pro-  
 313 cess. This concept has notably been presented in [17] and exploited in [30, 44, 31], with the  
 314 aim of tuning the accuracy of each computation of  $\boldsymbol{\rho}$  on the fly in order to focus computa-  
 315 tional power on the most promising designs.

316 In practice, robustness and reliability measures are computed using some evaluations  
 317 of  $\mathbf{q}$ , and their approximation error can be estimated. We propose to treat these approxi-  
 318 mated measures as random vectors, that can be compared under the  $\epsilon$ -relaxed probabilistic  
 319 constrained Pareto dominance paradigm, proposed in Section 2.3, in order to discard all  
 320 dominated approximations

321 Following the developments from [31], we further propose to rank all remaining measures  
 322 with the  $POP_{min}$  metric, which is an approximation of the Pareto Optimal Probability  
 323 (POP) with desirable properties. Using the notations from Section 3, the metric writes as  
 324 follows:

$$POP_{min}(\mathbf{P}(\mathbf{x}_k)) = \min_{\substack{j \in \llbracket 1, N_x \rrbracket \\ j \neq k}} \left( \mathbb{P}_{\mathbf{P}}[\mathbf{P}(\mathbf{x}_j) \not\prec_c \mathbf{P}(\mathbf{x}_k)] \right). \quad (24)$$

325 The accuracy is then adaptively tuned by improving these approximations only on the  
 326 most promising individuals. We refer to this approach as Measure Approximation with Tun-  
 327 able Accuracy (MATA), as different levels of refinement coexist, depending on the estimated  
 328 performances.

329 In addition, we construct a surrogate model on these measure approximations in the  
 330 design space. This model is referred to in the following as the Surrogate-Assisting (SA)  
 331 model. It directly yields predictions of the robustness and reliability measures at any design  
 332  $\mathbf{x}$ , so as to bypass  $\boldsymbol{\rho}$  estimations when the SA model is sufficiently accurate. In essence, it  
 333 permits to further lower the required number of function evaluations when the measures  $\boldsymbol{\rho}$   
 334 are easier to model in  $\mathcal{X}$  than  $\mathbf{q}$  in the coupled  $\mathcal{X} \times \Xi$  space.

335 These two methods (SA and MATA) couple very efficiently to lower the cost of an  
 336 uncertainty-based optimisation process. Figure 6 gives a flowchart of the global strategy,  
 337 called SAMATA. The specific notations are presented in the following sections 4.1.1 and  
 338 4.1.2.

##### 339 4.1.1. Surrogate-Assisting strategy (SA)

340 Surrogate-Assisting (SA) strategies are common practice for accelerating the optimisation  
 341 process. In the proposed approach, a surrogate model is constructed and updated throughout  
 342 the optimisation process directly on the robustness and reliability measures  $\boldsymbol{\rho}$  in the design  
 343 space. For each visited design  $\mathbf{x}_i$ , an aleatoric approximation  $\mathbf{P}(\mathbf{x}_i)$  is computed. The SA



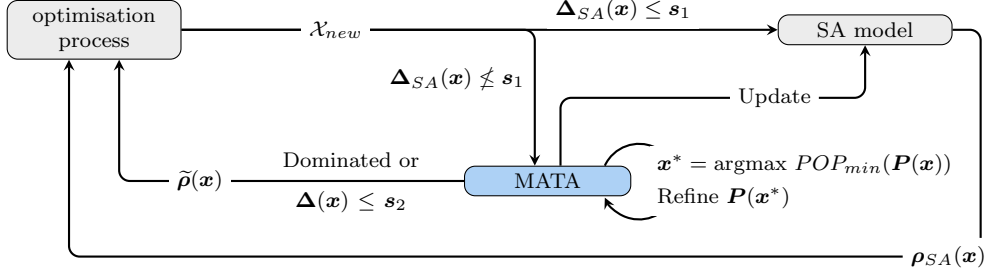


Figure 6: SAMATA strategy flowchart.

344 model then builds on the training set  $\{(\mathbf{x}_i, \mathbf{P}(\mathbf{x}_i))\}_i$ . When enough information is gathered  
 345 at a new design  $\mathbf{x}_{new}$ , the SA model bypasses the computation of  $\mathbf{P}(\mathbf{x}_{new})$  and returns the  
 346 current prediction  $\boldsymbol{\rho}_{SA}(\mathbf{x}_{new})$ . One can expect this approach to be extensively exploited at  
 347 the end of the optimisation process, when a lot of training points are available and most new  
 348 designs are gathered in the optimal area.

349 The SA model takes as training data aleatoric approximations of  $\boldsymbol{\rho}$ , and must yield an  
 350 accuracy metric at any location. We propose to write the accuracy of the SA model at a new  
 351 design  $\mathbf{x}_{new}$  as a vector  $\boldsymbol{\Delta}_{SA}(\mathbf{x}_{new})$ . The SA prediction  $\boldsymbol{\rho}_{SA}(\mathbf{x}_{new})$  is then directly returned  
 352 to the optimisation process whenever  $\boldsymbol{\Delta}_{SA}(\mathbf{x}_{new}) \leq \mathbf{s}_1$ , where  $\mathbf{s}_1$  is a user-defined threshold  
 353 chosen beforehand..

354 Similarly to [31], we propose to carry out the whole optimisation process several times  
 355 with decreasing values of  $\mathbf{s}_1$ . All training data  $\{(\mathbf{x}_i, \mathbf{P}(\mathbf{x}_i))\}_i$  are transferred from one iteration  
 356 to another. This procedure allows to get several intermediate results, that compromise  
 357 between computational cost and accuracy. With high  $\mathbf{s}_1$  value, the number of evaluations of  
 358  $\mathbf{q}$  is kept quite low as the SA model prediction  $\boldsymbol{\rho}_{SA}$  is quickly exploited. On the contrary,  
 359 with a small  $\mathbf{s}_1$  value, many training data are required before employing the SA model, thus  
 360 ensuring a better accuracy of  $\boldsymbol{\rho}_{SA}$ .

361 The structure of this approach is depicted in Algorithm 2. The computation and refine-  
 362 ment of the probabilistic approximations  $\mathbf{P}(\mathbf{x})$  (lines 4 and 12 of the algorithm) are described  
 363 in the following section 4.1.2.

#### 364 4.1.2. Measure Approximation with Tunable Accuracy (MATA)

Throughout the optimisation process, we denote  $\mathcal{X}_c$  the set of designs at which an ap-  
 proximation  $\mathbf{P}$  has been computed. At each iteration, among the new designs  $\mathcal{X}_{new}$  given  
 by the optimisation algorithm, all designs which computation is not bypassed by the SA  
 strategy are added to  $\mathcal{X}_c$  and first approximations  $\mathbf{P}$  and  $\boldsymbol{\Delta}$  are computed. Within this  
 set, we denote by  $\mathcal{X}_r$  the set of designs that must be further refined. These designs satisfy  
 two criteria: (i) they belong to the current Pareto front  $\mathcal{P}(\mathbf{P}(\mathcal{X}_c))$  and (ii) the user-defined  
 accuracy threshold  $\mathbf{s}_2$  is not reached. Formally:

$$\mathcal{X}_r = \{\mathbf{x} \in \mathcal{X}_c \mid \boldsymbol{\Delta}(\mathbf{x}) \not\leq \mathbf{s}_2 \ \& \ \mathbf{P}(\mathbf{x}) \in \mathcal{P}(\mathbf{P}(\mathcal{X}_c))\}, \quad (25)$$

365 where  $\boldsymbol{\Delta}(\mathbf{x})$  is a chosen measure of the variability of  $\mathbf{P}(\mathbf{x})$ , here its standard deviation, and  
 366  $\mathcal{P}$  corresponds to the non-dominated measures under the  $\epsilon$ -relaxed probabilistic constrained

---

**Algorithm 2** Algorithm of the Surrogate-Assisting strategy

---

```
1: Set initial thresholds  $\mathbf{s}_1$ 
2: Initialise the training set  $\mathcal{T} = \{(\mathbf{x}_{train_i}, \mathbf{P}(\mathbf{x}_{train_i}))\}_i$  empty
3: for  $k$  from 0 to  $N_{threshold}$  do
4:   Refine approximations in  $\mathcal{T}$  ▷ Alg. 3
5:   Launch optimiser
6:   while optimisation running do
7:     Update the SA model on  $\mathcal{T}$ 
8:     Get new designs  $\mathcal{X}_{new}$  to visit
9:      $\mathcal{X}_{SA} = \{\mathbf{x} \in \mathcal{X}_{new} \mid \Delta_{SA}(\mathbf{x}) \leq \mathbf{s}_1\}$ 
10:     $\mathcal{X}_c = \mathcal{X}_{new} \setminus \mathcal{X}_{SA}$ 
11:     $\forall \mathbf{x} \in \mathcal{X}_{SA}$ , return  $\rho_{SA}(\mathbf{x})$  to the optimiser
12:     $\forall \mathbf{x} \in \mathcal{X}_c$ , compute  $\mathbf{P}(\mathbf{x})$  and return  $\tilde{\rho}(\mathbf{x})$  to the optimiser ▷ Alg. 3
13:     $\forall \mathbf{x} \in \mathcal{X}_c$ , add  $(\mathbf{x}, \mathbf{P}(\mathbf{x}))$  to  $\mathcal{T}$ 
14:  end while
15:  Take next  $\mathbf{s}_1$  value
16: end for
```

---

367 *Pareto dominance* paradigm.

Given this set of designs to refine, we propose to use the POP ranking (Eq. 24) to select the most promising design within  $\mathcal{X}_r$ . The variability of the associated approximation  $\mathbf{P}(\mathbf{x})$  is then reduced, at the cost of some evaluations of  $\mathbf{q}$ . We denote  $\mathbf{x}^*$  the design maximising the  $POP_{min}$  metric.

$$\mathbf{x}^* = \operatorname{argmax}_{\mathbf{x} \in \mathcal{X}_r} POP_{min}(\mathbf{P}(\mathbf{x})). \quad (26)$$

368 The MATA refinement technique, for a given set of designs  $\mathcal{X}_c$  and a fixed threshold  $\mathbf{s}_2$ ,  
369 is detailed in Algorithm 3.

---

**Algorithm 3** Algorithm of the Measure Approximation with Tunable Accuracy technique

---

```
1: Read threshold  $\mathbf{s}_2$  and design set  $\mathcal{X}_c$ 
2: Compute first approximations  $\mathbf{P}(\mathbf{x})$  or retrieve the ones that already exist
3: Extract  $\mathcal{P}(\mathbf{P}(\mathcal{X}_c))$ 
4: Compute  $\mathcal{X}_r$  with Equation (25)
5: while  $\mathcal{X}_r$  is non-empty do
6:   Find  $\mathbf{x}^*$  with Equation (26)
7:   Refine approximation  $\mathbf{P}(\mathbf{x}^*)$  with some evaluations of  $\mathbf{q}$ 
8:   Update  $\mathcal{P}(\mathbf{P}(\mathcal{X}_c))$  and  $\mathcal{X}_r$ 
9: end while
10: return  $\mathbf{P}(\mathbf{x})$  for all  $\mathbf{x}$  in  $\mathcal{X}_c$ 
```

---

370 Similarly to the SA strategy, we propose to iterate over decreasing values of  $\mathbf{s}_2$  in order  
371 to obtain intermediate approximations of the Pareto front.

372 *4.2. Application to non-parametric measure approximations*

373 The above SAMATA strategy is presented in its most general form. In practice, specific  
 374 choices must be made on the computation of approximations  $\mathbf{P}(\mathbf{x})$ , metrics  $\tilde{\boldsymbol{\rho}}(\mathbf{x})$  and  $\boldsymbol{\Delta}(\mathbf{x})$ ,  
 375 and SA predictions  $\boldsymbol{\rho}_{SA}(\mathbf{x})$  and  $\boldsymbol{\Delta}_{SA}(\mathbf{x})$ .

376 The inducing point strategy presented in Section 3.2.1 permits to generate realisations  
 377 of  $\mathbf{q}$  on an extensive set of points  $\mathbf{z}_{draw}$ , to empirically estimate realisations of the measures  
 378  $\mathbf{P}$ , *e.g.* using Equation (16). In short, the obtained empirical distribution of  $\mathbf{P}$  comes from  
 379 the Gaussian assumption of the GP surrogate model built on  $\mathbf{q}$ , composed with the specific  
 380 statistical estimator.

381 The mean of this approximated distribution is returned to the optimisation process as  
 382  $\tilde{\boldsymbol{\rho}}$ , and three times the standard deviation is used as accuracy metric  $\boldsymbol{\Delta}$ . Concerning the  
 383 refinements, we use the criterion proposed in [31], corresponding to a linear combination of  
 384 measure-specific criteria.

Finally, the SA model is constructed in a specific way to best exploit all the available  
 information. SA-based measure approximations should take the form of set of samples and  
 be jointly drawn with the realisations obtained in (16). To this extent, we propose to define  
 the SA-based realisations as random draws of a GP surrogate model conditioned on the  
 computed joint realisations  $\mathbf{P}_c^{(i)}$ . In practice, on any set of design  $\mathbf{x}_{SA}$ , SA-based measure  
 realisations  $\mathbf{P}_{SA}^{(i)}(\mathbf{x}_{SA})$  are computed as such:

$$\mathbf{P}_{SA}^{(i)}(\mathbf{x}_{SA}) = \boldsymbol{\mu}_{SA}^{(i)}(\mathbf{x}_{SA}) + L_{SA}\mathbf{Y}^{(i)}, \quad \mathbf{Y} \sim \mathcal{N}(\mathbf{0}, I), \quad (27)$$

where the predictive mean writes:

$$\boldsymbol{\mu}_{SA}^{(i)}(\mathbf{x}_{SA}) = \mathbf{k}_*^\top K_c^{-1} \mathbf{P}_c^{(i)}, \quad (28)$$

385 and  $L_{SA}$  results from a Cholesky decomposition performed on the predictive covariance  
 386  $\Sigma_{GP}(\mathbf{x}_{SA})$  defined in Equation (13). In the above,  $\mathbf{P}_c^{(i)}$  refers to realisations drawn from  
 387 Equation (16), on the set designs  $\mathbf{x}_c$ .  $K_c$  corresponds to the autocovariance matrix  $k(\mathbf{x}_c, \mathbf{x}_c)$   
 388 and  $\mathbf{k}_*$  to the covariance  $k(\mathbf{x}_c, \mathbf{x}_{SA})$ . Note that hyperparameters are only optimised once for  
 389 numerical efficiency.

390 This SA model construction is illustrated on an example set of samples in Figure 7, where  
 391 realisations of  $\mathbf{P}$  are jointly drawn at ten locations in ascending order. Two bimodal sets  
 392 of samples are notably drawn at  $x = 0.15$  and  $x = 0.45$  in order to emphasise the need for  
 393 non-parametric reconstruction. Some SA-based functional realisations  $\mathbf{P}_{SA}^{(i)}$  from Equation  
 394 (27) are represented. The two bi-modal densities are naturally well captured using the  
 395 proposed approach. All computed  $\mathbf{P}_c^{(i)}$  and SA-based  $\mathbf{P}_{SA}^{(i)}$  realisations are jointly redrawn  
 396 at each computed measure refinement, in order to use the joint dominance probability from  
 397 Equation (12).

398 These formulations can be readily incorporated within the SAMATA strategy, presented  
 399 in Section 4.1 and pictured in Figure 6. The specific algorithm is given in Algorithm 4.

400 Finally, specific numerical choices such as the optimisation algorithm and the threshold  
 401 values  $\mathbf{s}_1$  and  $\mathbf{s}_2$  will be made case by case.

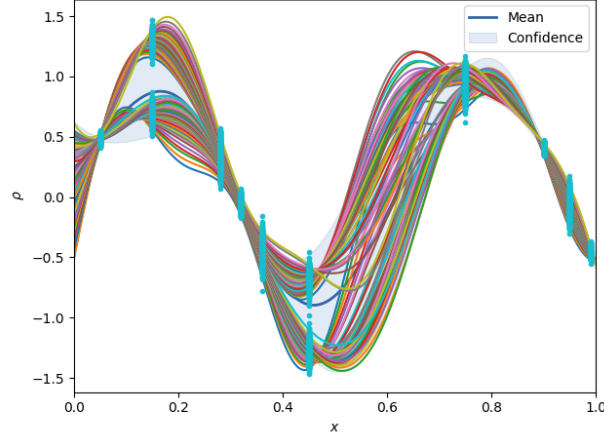


Figure 7: Realisations of the SA model. Training data in light blue and underlying GP model in dark blue, in the background.

---

**Algorithm 4** SAMATA algorithm with joint measure approximations

---

- 1: Launch  $N_{init}$  evaluations  $\mathbf{q}(\mathbf{x}_i, \boldsymbol{\xi}_i)$
  - 2: Generate a first coupled space surrogate model  $\hat{\mathbf{q}}$
  - 3: Loop over values of  $\mathbf{s}_1$  and  $\mathbf{s}_2$
  - 4: **while** optimisation running **do**
  - 5:   Get new designs  $\mathcal{X}_{new}$
  - 6:   Initialise  $\mathcal{X}_c$  empty
  - 7:   Compute  $\mathbf{P}_{SA}$  for all  $\mathbf{x} \in \mathcal{X}_{new}$  from (27)
  - 8:   Compute  $\mathcal{X}_{SA}^0 = \{\mathbf{x} \in \mathcal{X}_{new} \mid \Delta_{SA}(\mathbf{x}) \leq \mathbf{s}_1\}$  and  $\mathcal{X}_c^0 = \mathcal{X}_{new} \setminus \mathcal{X}_{SA}^0$
  - 9:   Deduce  $\mathcal{X}_r^0$  from  $\mathcal{X}_c^0$  using (25)
  - 10:    $k = 0$
  - 11:   **while**  $\mathcal{X}_r^k \neq \emptyset$  **do**
  - 12:     Compute  $\forall \mathbf{x} \in \mathcal{X}_{new}, POP_{min}(\mathbf{x})$  w.r.t. all designs using (24) and (12)
  - 13:     Find  $\mathbf{x}^*$  from  $\mathcal{X}_r^k$  using (26)
  - 14:     **if**  $\mathbf{x}^* \notin \mathcal{X}_c$  **then**
  - 15:       Add  $\mathbf{x}^*$  in  $\mathcal{X}_c$
  - 16:     **end if**
  - 17:     Find  $\boldsymbol{\xi}^*$  that maximises the refinement criterion from [31]
  - 18:     Launch evaluation of  $\mathbf{q}(\mathbf{x}^*, \boldsymbol{\xi}^*)$
  - 19:     Update the coupled space surrogate model  $\hat{\mathbf{q}}$
  - 20:     Update all approximations  $\mathbf{P}_c^{(i)}(\mathbf{x})$  for  $\mathbf{x} \in \mathcal{X}_c$  using empirical estimators like (16)
  - 21:     Update  $\mathbf{P}_{SA}$  for all  $\mathbf{x} \in \mathcal{X}_{new} \setminus \mathcal{X}_c$
  - 22:     Compute  $\mathcal{X}_c^{k+1}$ ,  $\mathcal{X}_{SA}^{k+1}$  and  $\mathcal{X}_r^{k+1}$
  - 23:      $k = k + 1$
  - 24:   **end while**
  - 25:   Return  $\mathbb{E}[\mathbf{P}(\mathbf{x})]$  to the optimiser for all  $\mathbf{x} \in \mathcal{X}_{new}$
  - 26: **end while**
  - 27: Return all measure approximations  $\mathbf{P}(\mathbf{x})$  and their associated  $POP_{min}$  score
-

402 **5. Analytic comparison**

403 In this section, the proposed SAMATA approach depicted in Alg. 4 is quantitatively  
 404 compared to SABBa from [31] with Coupled-Space surrogate models, which was shown very  
 405 cost efficient. As a reference, we also compare these techniques to an A Priori Metamodel  
 406 strategy, where a Coupled-Space surrogate model is constructed in the coupled  $\mathcal{X} \times \Xi$  space  
 407 on a given number of data points and is readily employed as is for measure estimation and  
 408 optimisation.

The three approaches are compared in terms of the expected modified Hausdorff distance between the approximated Pareto optimal designs and the true ones. Such a metric was proposed in [31] in the context of Bounding-Box approach. Here, by construction, joint realisations  $\mathbf{P}^{(i)}(\mathbf{x})$  of the aleatory measure approximations can be drawn simultaneously on a set of designs  $\mathbf{x}$  using (16) and (27). This readily permits to employ the aforementioned metric, which formula is recalled here:

$$Q_{\mathcal{B}} = \mathbb{E}_{\tilde{\mathcal{X}}_{\mathcal{P}}} [d'_H(\mathcal{X}_{\mathcal{P}}, \tilde{\mathcal{X}}_{\mathcal{P}})], \quad (29)$$

409 where  $d'_H$  is the modified Hausdorff distance from [45],  $\mathcal{X}_{\mathcal{P}}$  is the Pareto front pre-image,  
 410 and  $\tilde{\mathcal{X}}_{\mathcal{P}}$  is an aleatory variable corresponding to the pre-image of the currently approximated  
 411 aleatory Pareto front. This expected value is empirically estimated, exploiting the capability  
 412 of drawing joint realisations of  $\mathbf{P}$ . Note that in a normalised input space, a value  $Q_{\mathcal{B}} = 1$   
 413 corresponds to a bad score while  $Q_{\mathcal{B}} = 10^{-2}$  reveals a very accurate approximation of the  
 414 Pareto front.

415 The NSGA-II optimisation algorithm is used in the following test-cases. The number of  
 416 individuals per generation is set to 32. Both the optimisation and refinement processes are  
 417 sequential, and the coupled-space surrogate  $\hat{q}$  is initialised with a Latin Hypercube Sampling  
 418 (LHS) of 30 points.

419 Each technique is run ten times to get the mean and standard deviation of the convergence  
 420 curves.

421 *5.1. Test-case 1: Unconstrained Taguchi optimisation*

This problem is a bi-objective robust optimisation proposed in [30]. There are two design variables  $\mathbf{x}$  and one uncertain parameter  $\xi$ . The problem reads:

$$\begin{aligned} \text{minimise: } \quad & \boldsymbol{\rho}_{\mathbf{f}}(\mathbf{x}) = \begin{pmatrix} \mu(\mathbf{x}) \\ \sigma^2(\mathbf{x}) \end{pmatrix} \\ \text{where: } \quad & \mu(\mathbf{x}) = \mathbb{E}_{\xi}[q(\mathbf{x}, \xi)] \\ & \sigma^2(\mathbf{x}) = \mathbb{V}_{\xi}[q(\mathbf{x}, \xi)] \\ \text{with: } \quad & q(\mathbf{x}, \xi) = \xi - x_1\xi^5 + \cos(2\pi x_2\xi) + 5 \\ & \xi \sim \mathcal{U}([0, 1]) \\ \text{by changing: } \quad & (x_1, x_2) \in [1, 2]^2 \end{aligned} \quad (30)$$

422 The Pareto front associated with this problem is discontinuous, and the optimal set in  
 423 the design space consists in segment and a point (see Fig. 8).

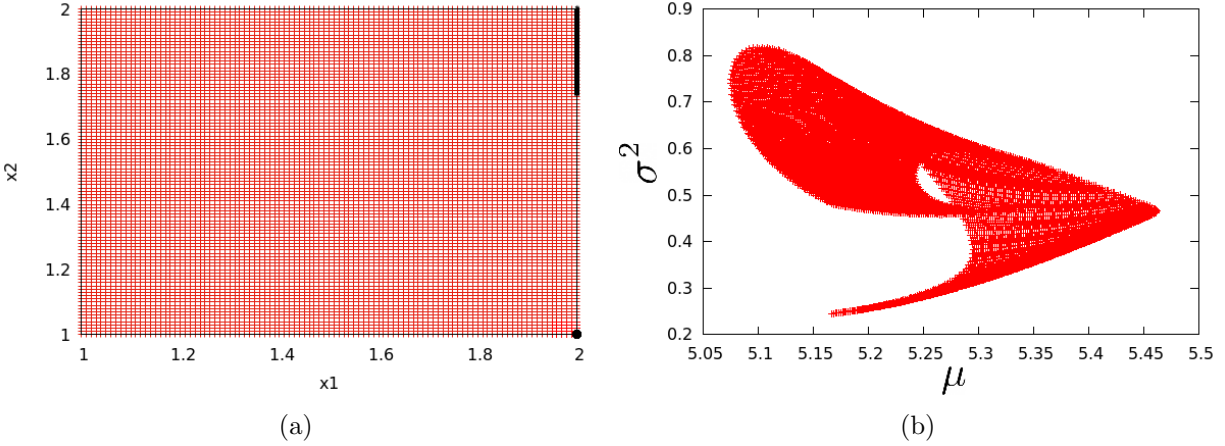


Figure 8: Test-case 1: a) Discretisation of the design space in red and Pareto optimal sets in black. b) Image of the discretised points in the objective  $(\mu, \sigma^2)$  space.

424 The performance of the three compared strategies is depicted in Figure 9, showing the  
 425 mean convergence over ten runs and the Bollinger Band corresponding to the associated  
 426 standard deviation. Both the Bounding-Box and the proposed non-parametric SAMATA  
 427 approaches show a steeper convergence than the A Priori Metamodel technique. In addition,  
 428 the non-parametric approach reveals significantly more parsimonious than the other two  
 429 techniques, reaching a  $Q_B$  metric of  $10^{-2}$  around 60 function evaluations, corresponding to  
 430 a gain of nearly one order of magnitude over the reference methods.

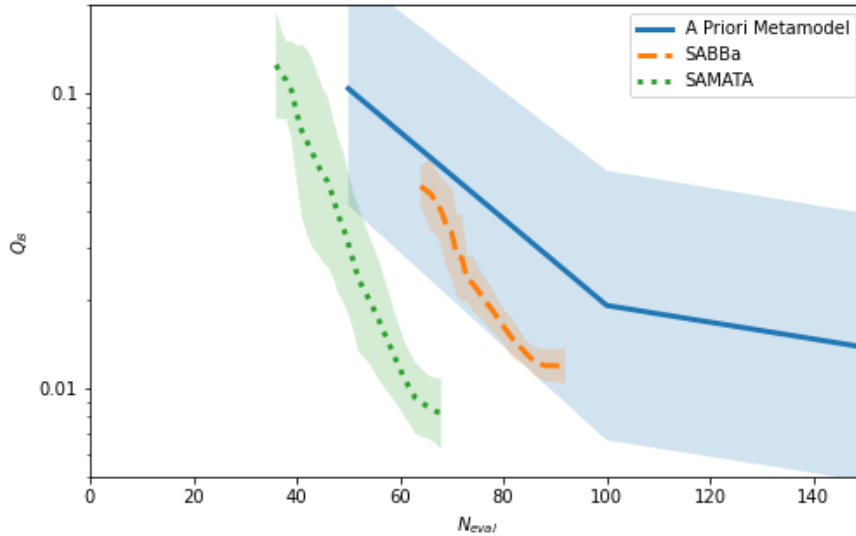


Figure 9: Test-case 1:  $Q_B$  metric for the A Priori Metamodel (solid blue), SABBa (dashed orange) and the proposed non-parametric SAMATA (dotted green) approaches.

431 *5.2. Test-case 2: Quantile-constrained mean optimisation*

This second test-case, proposed in [31], assesses the performance of the proposed non-parametric approach on a mono-objective optimisation problem with a reliability based constraint, in a higher dimensional context. The objective is a robustness measure derived from a simplified Rosenbrock function and the constraint a quantile formulation of the Six-Hump Camel function. It features four design variables and three uncertain parameters, and is stated as follows:

$$\begin{aligned}
&\text{minimise: } \boldsymbol{\rho}_f(\mathbf{x}) = \mu(\mathbf{x}) \\
&\text{satisfying: } \boldsymbol{\rho}_g(\mathbf{x}) = q^{95\%}(\mathbf{x}) \leq 1 \\
&\text{where: } \mu(\mathbf{x}) = \mathbb{E}_{\boldsymbol{\xi}}[q_1(\mathbf{x}, \boldsymbol{\xi})] \\
&\quad q^{95\%}(\mathbf{x}) = q_{\boldsymbol{\xi}}^{95\%}[q_2(\mathbf{x}, \boldsymbol{\xi})] \\
&\text{with: } q_1(\mathbf{x}, \boldsymbol{\xi}) = \sum_{i=1}^3 \left[ (1 - x_i) + 3 \left( 1 + \frac{\arctan(5(\xi_i - 0.5))}{2} \right) (x_{i+1} - x_i^2)^2 \right] \\
&\quad q_2(\mathbf{x}, \boldsymbol{\xi}) = \left( 4 - 2.1x_1^2 + \frac{x_1^4}{3} \right) x_1^2 + x_1x_2 + (-4 + 4x_2^2)x_2^2 \\
&\quad\quad + \frac{\cos(2\pi\xi_1) - \sin(\frac{\pi}{2}\xi_1) - \xi_1 - (\cos(2\pi \cdot 0.05) - \sin(\frac{\pi}{2} \cdot 0.05) - 0.05)}{5} \\
&\quad \boldsymbol{\xi} \sim \mathcal{U}([0, 1]^3) \\
&\text{by changing: } \mathbf{x} \in [-0.2, 1.2]^4 \tag{31}
\end{aligned}$$

432 The optimum design is  $\mathbf{x}_* \approx (0.7033, 0.7035, 0.6212, 0.3859)$  with  $q^{95\%}(\mathbf{x}_*) = 1$  and  
433  $\mu(\mathbf{x}_*) \approx 0.4981$ .

434 The convergence curves are depicted in Figure 10. Similarly to Figure 9, the A Priori  
435 Metamodel is the less efficient strategy, and the proposed non-parametric SAMATA approach  
436 is the most parsimonious. The standard deviation over the ten runs are greater in this test-  
437 case, but the proposed method permits to reach a metric of  $10^{-1}$  for less than 70 function  
438 evaluations, against 110 for the Bounding-Box approach and more than 200 with the A Priori  
439 Metamodel.

440 *5.3. Benefit of sampling-based approximations*

441 The so-called projected processes techniques [32, 33, 34, 35] permit to solve Optimisation  
442 Under Uncertainty (OUU) problems in a Bayesian Optimisation setting (also called Efficient  
443 Global Optimisation, EGO). They however rely on independent Gaussian approximations of  
444 the robustness and reliability measures. Such Gaussian approximations are cheap to estimate  
445 and particularly adapted for expectation-like measures, but might fail to accurately capture  
446 more complex or correlated distributions. For deeper insights on this method, see [33].

447 We run ten optimisations of the first test-case using the proposed approach, but replacing  
448 the jointly drawn realisations of  $\mathbf{P}$  with samples from a Gaussian distribution which mean  
449 and variance correspond to the empirical moments from the initial samples. In this manner,

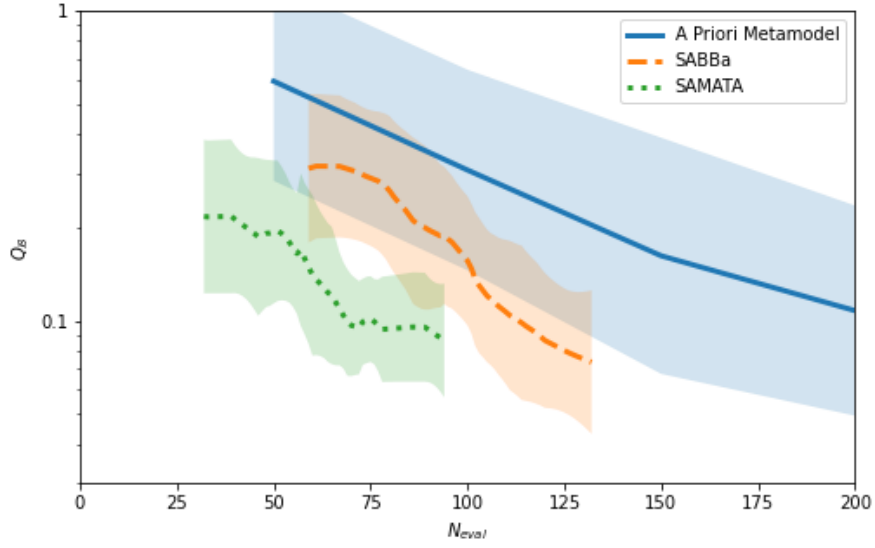


Figure 10: Test-case 2:  $Q_B$  metric for the A Priori Metamodel (solid blue), SABBa (dashed orange) and the proposed non-parametric SAMATA (dotted green) approaches.

450 we aim to investigate solely the impact of using a different density estimation, with either  
 451 the proposed sample-based strategy or a simple moment-based Gaussian fit.

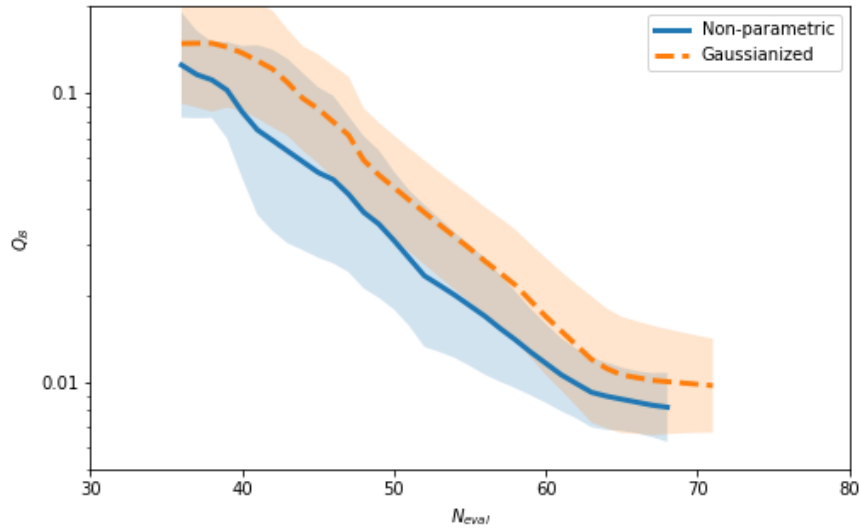


Figure 11: Convergence curves of the non-parametric SAMATA strategy (solid blue) against its Gaussianised equivalent (dashed orange). Variability represented with the translucent  $\pm\sigma$  band.

452 Figure 11 shows the two convergence curves, with their Bollinger bands. Although the  
 453 convergence seems to plateau around the same metric value, the Gaussianised approach



454 shows slightly inferior accuracy compared to the original non-parametric strategy along the  
 455 whole study. Intuitively, we argue that accurate distribution shapes and dependencies are  
 456 lost when Gaussian density is forced, which may deteriorate the intermediate approximations  
 457 of the Pareto optimal set.

## 458 6. ORC application

459 We perform here a robust shape optimisation of a typical converging-diverging turbine  
 460 nozzle for ORC applications. It was extensively studied in the context of deterministic  
 461 [46, 47, 48] and multi-point optimisation [49]. Some recent works tackled such design under  
 462 epistemic uncertainties [50] due to turbulence modelling, and investigated the robustness of  
 463 the blade design under aleatoric uncertainties in the operating conditions, thermodynamic  
 464 model parameters and geometric tolerances [51]. We propose here to minimise the mean of the  
 465 performance function while the mean mass-flow rate is constrained to lie within a prescribed  
 466 range centered on the nominal value.

467 The Biere is a reference two-dimensional benchmark geometry to test the design of devices  
 468 operating with the siloxane fluid MDM (Octamethyltrisiloxane,  $C_8H_{24}O_2Si_3$ ). This blade  
 469 profile is meant to obtain a convergent-divergent cascade passage which serves to accelerate  
 470 the fluid up to supersonic speed. Compressibility effects play a crucial role and a typical  
 471 fish-tail shock pattern is generated downstream the trailing edge. Strong shocks may cause  
 472 significant losses, design of the trailing edge region is thus critical for the turbine efficiency.  
 473 The Mach field is depicted in Figure 12 with the baseline blade profile.

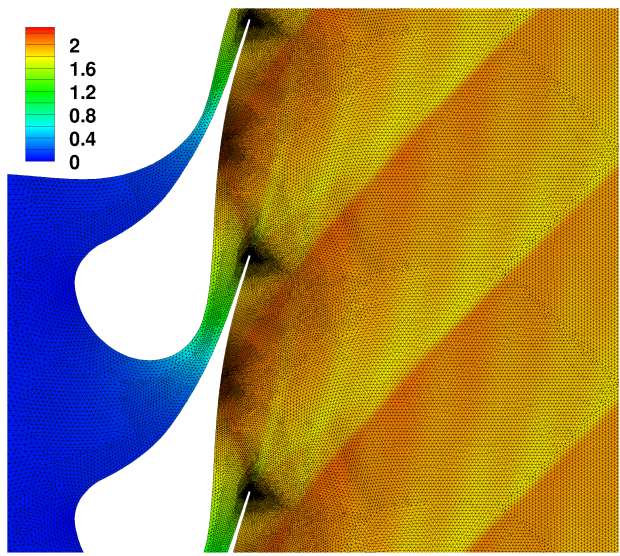


Figure 12: Mach contours at nominal conditions for the baseline profile and computational grid of 36k cells.

### 474 6.1. Problem setting

475 The objective function  $\Delta P$  is defined as the standard deviation of the azimuthal distri-  
 476 bution of static pressure half an axial chord downstream of the blade Trailing Edge (TE).

477 Minimising  $\Delta P$  should yield a significant reduction of the shock strength. As mentioned  
 478 above, a constraint is imposed on the mass-flow rate per unit span  $\dot{m}$ , normalised with  
 479 respect to the nominal value.

480 Uncertainties on the operating conditions are considered both at the inlet and outlet of the  
 481 turbine, through changes of boundary conditions. Following [52], we model the operational  
 482 variability as independent and uniform uncertain variables, gathered in  $\boldsymbol{\xi} = [P_{in}^t, T_{in}^t, P_{out}^s]$ .  
 483 The associated ranges are reported in Table 1, alongside Nominal and Mean conditions.  
 484 Uncertainties on the parameters of the thermodynamical model are neglected since previous  
 485 studies [53] provided evidence of their limited impact.

$\boldsymbol{\xi}$	$P_{in}^t$ [Pa]	$T_{in}^t$ [K]	$P_{out}^s$ [Pa]
Nominal ( $\boldsymbol{\xi}_0$ )	$8 \times 10^5$	545.15	$1.072 \times 10^5$
Mean ( $\boldsymbol{\xi}_\mu$ )	$8 \times 10^5$	545.15	$1.5 \times 10^5$
Random	$\mathcal{U}[7.95 \times 10^5, 8.05 \times 10^5]$	$\mathcal{U}[541.15, 549.15]$	$\mathcal{U}[1 \times 10^5, 2 \times 10^5]$

Table 1: Operating Conditions: Nominal, Mean and Random.

The optimisation problem follows a multi-objective Taguchi formulation:

$$\begin{aligned}
 &\text{minimise: } \boldsymbol{\rho}_f(\mathbf{x}) = \begin{pmatrix} \mu_1(\mathbf{x}) \\ \sigma^2(\mathbf{x}) \end{pmatrix} \\
 &\text{satisfying: } \boldsymbol{\rho}_g(\mathbf{x}) = \mu_2(\mathbf{x}) \in [0.98\dot{m}_0, 1.02\dot{m}_0] \\
 &\text{where: } \mu_1(\mathbf{x}) = \mathbb{E}_\xi[\Delta P(\mathbf{x}, \boldsymbol{\xi})] \\
 &\quad \sigma^2(\mathbf{x}) = \mathbb{V}_\xi[\Delta P(\mathbf{x}, \boldsymbol{\xi})] \\
 &\quad \mu_2(\mathbf{x}) = \mathbb{E}_\xi[\dot{m}(\mathbf{x}, \boldsymbol{\xi})] \\
 &\quad \dot{m}_0 = 344.843 \text{ kg.s}^{-1} \\
 &\text{with: } \boldsymbol{\xi} \sim \mathcal{U}(\Xi) \\
 &\text{by changing: } \mathbf{x} \in \mathcal{X} \tag{32}
 \end{aligned}$$

486 where  $\Xi = [7.95, 8.05] \times [541.15, 549.15] \times [1, 2]$ . The design space  $\mathcal{X}$  allows to modify  
 487 the shape of the blade as much as possible while ensuring convergence of the numerical  
 488 simulation. In practice, it consists of 9 of the 30 Control Points (CP) of a B-spline curve  
 489 of degree 3, that can be displaced in the direction normal to the baseline geometry. Such  
 490 parametrisation was proposed in [54] and is illustrated in Figure 13.

491 A mono-objective version of (32) is also treated for comparison purposes. Results will also  
 492 be compared with two deterministic optima obtained at the Nominal and Mean operating  
 493 conditions from Table 1.

494 Simulation is performed using the the Non-Ideal Compressible-Fluid Dynamics solver  
 495 included in the SU2 [55, 56, 57, 58] suite on unstructured grids, generated using an in-house  
 496 meshing tool. A dedicated mesh deformation tool based on Radial Basis Functions (RBF)  
 497 allows high flexibility and robustness while maintaining the grid connectivity throughout the  
 498 optimisation steps.

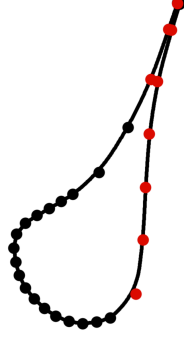


Figure 13: B-splines parametrisation. Fixed CP in black, moving CP in red.

499 *6.2. Mono-objective results*

500 Because nominal conditions are not centered within the uncertain range, we conduct the  
 501 reference deterministic optimisations both at Nominal and Mean conditions. Similarly to  
 502 [59], a state-of-the-art Bayesian Optimisation strategy is employed, with an Expected Im-  
 503 provement (EI) criterion multiplied by the Probability of Feasibility (PF) [60]. Convergence  
 504 is reached after roughly **100** function evaluations.

505 The robust optimisation problem is solved using the proposed non-parametric SAMATA  
 506 strategy. The NSGA-II optimisation procedure is again exploited, and constraints are han-  
 507 dled through penalisation. The initial DoE consists of 30 evaluations drawn with a maximin  
 508 Latin Hypercube Sampling (LHS) in the coupled space. Each population contains 32 indi-  
 509 viduals, and measure refinements are performed sequentially. We iteratively take threshold  
 510 values of 10%, 5%, 3%, 2% and 1%.

511 Convergence values are gathered in Table 2, with a final optimum centered on  $\rho_f = 8087$ ,  
 512 with a standard deviation of the approximation error of 44.26. This optimum satisfies the  
 513 1% accuracy imposed by the normalised thresholds and only required **145** simulation calls,  
 514 which is not significantly higher than the deterministic optimisation. The proposed ap-  
 515 proach permits to reduce by around one order of magnitude the number of required function  
 516 evaluations with respect to [59] that employed a more classical nested approach.

$N_{eval}$	$\rho_f$ [Pa]		$\rho_g$ [ $kg.s^{-1}$ ]	
	Mean	Std. dev.	Mean	Std. dev.
76	$8.328 \times 10^3$	$6.355 \times 10^2$	$3.464 \times 10^2$	$1.065 \times 10^{-1}$
91	$8.854 \times 10^3$	$3.986 \times 10^2$	$3.514 \times 10^2$	$3.149 \times 10^{-2}$
106	$8.869 \times 10^3$	$1.124 \times 10^2$	$3.387 \times 10^2$	$1.534 \times 10^{-2}$
118	$8.288 \times 10^3$	$1.068 \times 10^2$	$3.500 \times 10^2$	$1.144 \times 10^{-2}$
145	$8.087 \times 10^3$	$4.426 \times 10^1$	$3.505 \times 10^2$	$1.812 \times 10^{-2}$

Table 2: Optimal outputs associated to the different threshold values.

517 The point-wise ( $\xi_0$  and  $\xi_\mu$ ) and mean values of the objective function are computed at  
 518 the Nominal  $\mathbf{x}_0^*$ , Mean  $\mathbf{x}_\mu^*$  and Robust  $\mathbf{x}_{OUU}^*$  operating conditions. Results are gathered in

519 Table 3 and the full distributions of  $\Delta P$  over  $\boldsymbol{\xi}$  at these given designs are plotted in Figure  
 520 14. These distributions are obtained with an extensive Monte Carlo on a GP built over  $\Xi$   
 521 at a given  $\boldsymbol{x}$ , with 50 training points.

Blade profile	$\Delta P(\boldsymbol{x}, \boldsymbol{\xi}_0)$ [Pa]	$\Delta P(\boldsymbol{x}, \boldsymbol{\xi}_\mu)$ [Pa]	$\mathbb{E}_\xi[\Delta P(\boldsymbol{x}, \boldsymbol{\xi})]$ [Pa]
Nominal optimum $\boldsymbol{x} = \boldsymbol{x}_0^*$	$1.958 \times 10^3$	$1.526 \times 10^4$	$1.624 \times 10^4$
Mean optimum $\boldsymbol{x} = \boldsymbol{x}_\mu^*$	$1.079 \times 10^4$	$1.773 \times 10^3$	$8.472 \times 10^3$
Robust optimum $\boldsymbol{x} = \boldsymbol{x}_{OUU}^*$	$1.210 \times 10^4$	$3.082 \times 10^3$	$8.084 \times 10^3$

Table 3: Nominal values and statistics for different blade profiles.

522 As expected, the Nominal optimum yields very high variability in Fig. 14(a) but very  
 523 low nominal output. The Mean and Robust optima are quite similar, Figs 14(b) and 14(c).  
 524 However, the Robust optimal blade manages to reach a lower value for the expectation of  
 525  $\Delta P$ , with a slightly modified distribution shape.

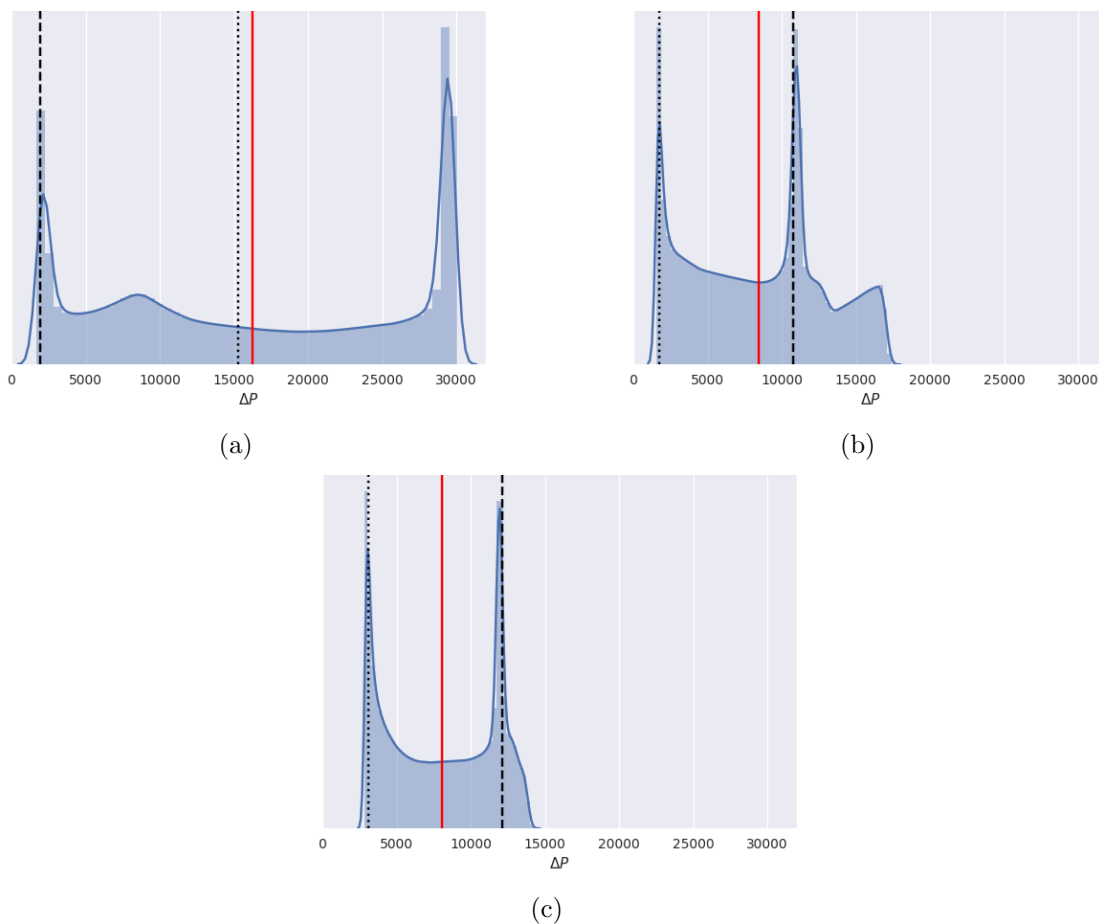


Figure 14: Distribution of  $\Delta P$  at (a)  $\boldsymbol{x}_0^*$ , (b)  $\boldsymbol{x}_\mu^*$  and (c)  $\boldsymbol{x}_{OUU}^*$ .  $\Delta P(\boldsymbol{x}, \boldsymbol{\xi}_0)$  represented as a dashed black line,  $\Delta P(\boldsymbol{x}, \boldsymbol{\xi}_\mu)$  as a dotted black line and  $\mathbb{E}_\xi[\Delta P(\boldsymbol{x}, \boldsymbol{\xi})]$  as a plain red line.

526 *6.3. Bi-objective results*

527 We then tackle the complete Taguchi formulation (32) of the Optimisation Under Un-  
 528 certainty of the ORC turbine blade. This formulation aims to find the Pareto front between  
 529 good expected performance and low variability. The proposed non-parametric SAMATA al-  
 530 gorithm is applied, with again the NSGA-II optimisation method and normalised thresholds  
 531 of 10%, 5%, 2% and 1%.

532 The Pareto fronts associated with these threshold values are depicted in Figure 15, and  
 533 reveal great improvements in the Pareto front accuracy throughout the iterations. The total  
 534 computational cost for attaining the chosen thresholds is of **72**, **122**, **276** and **419** simulation  
 535 calls respectively. Each cloud of samples associated with a given design has its transparency  
 536 driven by its POP value (1: fully visible, 0: transparent). The mono-objective optimum  
 537 found in the preceding section is represented as a red dot and lies on the Pareto front, as  
 expected.

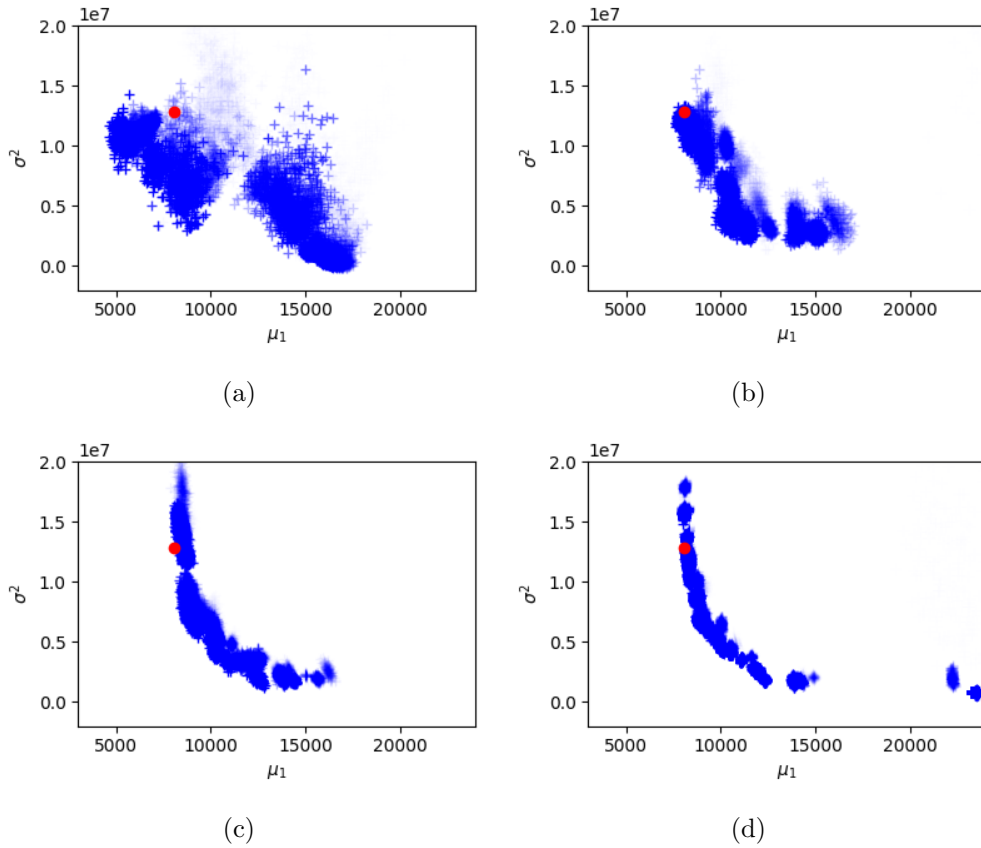


Figure 15: Pareto front of the bi-objective OUU problem with normalised threshold values of (a) 10% (**72** evaluations), (b) 5% (**122** evaluations), (c) 2% (**276** evaluations) and (d) 1% (**419** evaluations). Transparency driven by POP value and mono-objective optimum from previous section as a red dot.

538

## 539 7. Conclusions

540 This paper describes an efficient strategy for solving multi-objective and constrained  
541 Optimisation Under Uncertainty problems. It consists in coupling two primary techniques:  
542 i) a Measure Approximation with Tunable Accuracy (MATA) approach that only refines  
543 promising individuals up to a given threshold, ii) a Surrogate-Assisting (SA) strategy that  
544 constructs a surrogate model of the objective and constraint functions in the design space  
545 to bypass further evaluations when it is sufficiently accurate.

546 To apply this framework in practice, we propose non-parametric approximations of the  
547 error distributions using a sampling-based technique. The robustness and reliability measures  
548 for a given design are represented as a set of realisations that may exhibit complex shapes  
549 and dependencies. These realisations usually reveal more representative and can be jointly  
550 drawn between several designs. Issues associated with drawing such joint realisations are  
551 highlighted, and an adaptive algorithm is proposed to keep the associated cost manageable.  
552 This strategy show similarities with Nyström approximation for Karhunen-Loève Expansion  
553 (KLE) but allows for intuitive control of the overall accuracy and adaptive sampling. A  
554 Surrogate-Assisting (SA) model that takes advantage of these joint realisations is introduced  
555 to apply the SAMATA strategy. Quantitative comparisons with state-of-the-art approaches  
556 on several algebraic test cases reveal a drastic increase in the overall parsimony.

557 Finally, the proposed approach is applied to an engineering problem dealing with the ge-  
558 ometric optimisation of the blade profile in an Organic Rankine Cycle turbine. Convergence  
559 is attained for a cost of 100 to 150 evaluations in a mono-objective setting, which is one or-  
560 der of magnitude less than the state-of-the-art approach. The Pareto front of a bi-objective  
561 formulation is accurately recovered at the cost of 100 to 400 evaluations.

562 Several perspectives can be drawn. Regarding overall parsimony, cost improvement could  
563 be achieved by improving the Uncertainty Propagation (UP) techniques associated with spe-  
564 cific statistical measures. For example, extreme quantile estimation and refinement should  
565 be performed with appropriate techniques such as [61] or [62]. Such approaches can be  
566 readily incorporated into the proposed framework. Similarly, we only performed sequential  
567 refinement in this work, while techniques such as Kriging believer can be exploited to return  
568 a set of new locations. Multi-point refinement is critical to take advantage of nowadays' high  
569 parallelisation capacities of computational clusters. Finally, the proposed non-parametric  
570 error estimation relies on coupled-space surrogate models that may become unfeasible in  
571 high dimensional uncertain space or when the simulator is intrinsically stochastic. Spe-  
572 cific developments for using SAMATA in this context would effectively widen its range of  
573 application.

## 574 Appendix A. Link between inducing points and Nyström method for KLE

575 In practice, the proposed approach is equivalent to constructing a Karhunen-Loève Ex-  
576 pansion (KLE) of a Gaussian random field, when computing the basis functions with Nyström  
577 approximation. Computational details are given hereafter to show the equivalence.

To this extent, note that we can write the eigendecomposition (also called spectral decomposition) of the autocovariance matrix  $K_{tot}$  from Equation (18), which is Symmetric Positive Definite (SPD) by definition:

$$K_{tot} = U_{tot}\Lambda_{tot}U_{tot}^\top, \quad (\text{A.1})$$

where  $\Lambda_{tot}$  is the diagonal matrix containing the eigenvalues associated with the eigenvectors stored in the columns of  $U_{tot}$ . To make this decomposition unique, the eigenvalues are sorted in decreasing order,  $\Lambda_{tot_{11}} \geq \Lambda_{tot_{22}} \geq \dots \geq 0$ , and the eigenvectors are orthonormal, so that  $U_{tot}^\top = U_{tot}^{-1}$ , which gives:

$$U_{tot_{.i}}^\top U_{tot_{.j}} = \sum_k U_{tot_{ki}} U_{tot_{kj}} = \delta_{ij}, \quad (\text{A.2})$$

578 with  $\delta_{ij}$  the Kronecker delta.

The inverse of  $K_{tot}$  can thus be written:

$$K_{tot}^{-1} = U_{tot}\Lambda_{tot}^{-1}U_{tot}^\top, \quad (\text{A.3})$$

and the square root  $K_{tot}^{-1/2}$  of this matrix is:

$$K_{tot}^{-1/2} = U_{tot}\Lambda_{tot}^{-1/2}, \quad (\text{A.4})$$

579 which verifies  $K_{tot}^{-1} = K_{tot}^{-1/2}K_{tot}^{-1/2\top}$ .

*KLE and Nyström approximation.* A realisation of a random field described with a truncated Karhunen-Loève Expansion reads as follows:

$$\mathbf{q}_{KLE}^{(i)}(\mathbf{z}) = \sum_{j=1}^{N_{KLE}} \sqrt{\lambda_j} \phi_j(\mathbf{z}) \theta_j^{(i)}, \quad (\text{A.5})$$

where the eigenvalues  $\lambda_j$  and eigenvectors  $\phi_j$  are solution of the following homogeneous Fredholm integral equation of the second kind:

$$\int k(\mathbf{z}, \mathbf{z}') \phi_j(\mathbf{z}) d\mathbf{z} = \lambda_j \phi_j(\mathbf{z}') \quad (\text{A.6})$$

with  $\int \phi_i(\mathbf{z}) \phi_j(\mathbf{z}) d\mathbf{z} = \delta_{ij}$ .

As presented in [63], the Nyström approximation consists in computing an empirical estimation of the above integral. This Monte Carlo approximation based on points  $\mathbf{x}_k$  gives:

$$\frac{1}{N_{MC}} \sum_{k=1}^{N_{MC}} k(\mathbf{z}_k, \mathbf{z}') \phi_j(\mathbf{z}_k) \approx \lambda_j \phi_j(\mathbf{z}') \quad (\text{A.7})$$

with  $\frac{1}{N_{MC}} \sum_{k=1}^{N_{MC}} \phi_i(\mathbf{z}_k) \phi_j(\mathbf{z}_k) \approx \delta_{ij}$ .

Replacing  $\mathbf{z}'$  with all  $\mathbf{z}_{tot_k}$ , and setting  $N_{MC} = N_{tot}$  so that the Monte Carlo estimation is performed on locations  $\mathbf{z}_{tot}$ , Equation (A.7) yields:

$$\begin{aligned} \frac{1}{N_{tot}} K_{tot} \Phi &\approx \Lambda \Phi \\ \text{with } \frac{1}{N_{tot}} \Phi_{\cdot i}^\top \Phi_{\cdot j} &\approx \delta_{ij}. \end{aligned} \quad (\text{A.8})$$

Matching (A.8) against (A.1) and (A.2) permits to derive:

$$\begin{aligned} \Phi &\approx \sqrt{N_{tot}} U_{tot}, \\ \Lambda &\approx \frac{\Lambda_{tot}}{N_{tot}}, \end{aligned}$$

which gives the Nyström approximation of the  $j$ -th eigenfunction from Eq. (A.7):

$$\phi_j(\mathbf{z}) \approx \frac{\sqrt{N_{tot}}}{\Lambda_{totjj}} \sum_{k=1}^{N_{tot}} k(\mathbf{z}, \mathbf{z}_{tot_k}) U_{tot_{kj}}. \quad (\text{A.9})$$

Using these approximations, KLE-based realisations write:

$$\begin{aligned} \mathbf{q}_{KLE}^{(i)}(\mathbf{z}) &= \sum_{j=1}^{N_{tot}} \left( \sqrt{\frac{\Lambda_{totjj}}{N_{tot}}} \frac{\sqrt{N_{tot}}}{\Lambda_{totjj}} \sum_{k=1}^{N_{tot}} (k(\mathbf{z}, \mathbf{z}_{tot_k}) U_{tot_{kj}}) \theta_{tot_j}^{(i)} \right) \\ &= \sum_{j=1}^{N_{tot}} \left( \frac{\theta_{tot_j}^{(i)}}{\sqrt{\Lambda_{totjj}}} \sum_{k=1}^{N_{tot}} (k(\mathbf{z}, \mathbf{z}_{tot_k}) U_{tot_{kj}}) \right). \end{aligned} \quad (\text{A.10})$$

580 *Inducing points.* We proposed to draw realisations using Equation (18). To simplify the  
581 following, we assume here that  $N_{train} = 0$  and  $\mathbf{q}_{tot} = \mathbf{q}_{ind}$ . This implies that matrix  $L_{ind}$   
582 corresponds to the square root matrix  $K_{tot}^{1/2}$ .

By definition, in this context,  $\mathbf{q}_{tot}^{(i)} = K_{tot}^{-1/2} \boldsymbol{\theta}_{tot}^{(i)}$ , with  $\boldsymbol{\theta}_{tot}^{(i)} \sim \mathcal{N}(\mathbf{0}, I)$ . A realisation at location  $\mathbf{z}$  then writes:

$$\begin{aligned} \mathbf{q}^{(i)}(\mathbf{z}) &= k(\mathbf{z}, \mathbf{z}_{tot}) K_{tot}^{-1} \mathbf{q}_{tot}^{(i)} \\ &= k(\mathbf{z}, \mathbf{z}_{tot}) K_{tot}^{-1/2} \boldsymbol{\theta}_{tot}^{(i)} \\ &= k(\mathbf{z}, \mathbf{z}_{tot}) U_{tot} \Lambda_{tot}^{-1/2} \boldsymbol{\theta}_{tot}^{(i)} \\ &= \sum_{j=1}^{N_{tot}} \left( \frac{\theta_{tot_j}^{(i)}}{\sqrt{\Lambda_{totjj}}} \sum_{k=1}^{N_{tot}} (k(\mathbf{z}, \mathbf{z}_{tot_k}) U_{tot_{kj}}) \right), \end{aligned} \quad (\text{A.11})$$

583 which corresponds exactly to Equation (A.10).

584 Note that in the case  $\mathbf{q}_{tot} \neq \mathbf{q}_{ind}$ , which permits to condition the realisations on some  
585 training data, the same equivalence can be shown by comparing (18) to a conditional  
586 Karhunen-Loève model [64].



587 **References**

- 588 [1] V. Keshavarzzadeh, R. G. Ghanem, D. A. Tortorelli, Shape optimization under uncer-  
589 tainty for rotor blades of horizontal axis wind turbines, *Computer Methods in Applied*  
590 *Mechanics and Engineering* 354 (2019) 271–306.
- 591 [2] P. Seshadri, P. Constantine, G. Iaccarino, G. Parks, A density-matching approach for  
592 optimization under uncertainty, *Computer Methods in Applied Mechanics and Engi-*  
593 *neering* 305 (2016) 562–578.
- 594 [3] J. Zhang, M. Xiao, P. Li, L. Gao, Quantile-based topology optimization under uncer-  
595 tainty using kriging metamodel, *Computer Methods in Applied Mechanics and Engi-*  
596 *neering* 393 (2022) 114690.
- 597 [4] V. Keshavarzzadeh, R. M. Kirby, A. Narayan, Stress-based topology optimization un-  
598 der uncertainty via simulation-based gaussian process, *Computer Methods in Applied*  
599 *Mechanics and Engineering* 365 (2020) 112992.
- 600 [5] X. Liu, T. Li, Z. Zhou, L. Hu, An efficient multi-objective reliability-based design op-  
601 timization method for structure based on probability and interval hybrid model, *Com-*  
602 *puter Methods in Applied Mechanics and Engineering* 392 (2022) 114682.
- 603 [6] C. Comis Da Ronco, R. Ponza, E. Benini, Aerodynamic shape optimization of air-  
604 craft components using an advanced multi-objective evolutionary approach, *Computer*  
605 *Methods in Applied Mechanics and Engineering* 285 (2015) 255–290.
- 606 [7] E. Zitzler, D. Kalyanmoy, L. Thiele, Comparison of multiobjective evolutionary algo-  
607 rithms: Empirical results, *Evolutionary computation* 8 (2) (2000) 173–195.
- 608 [8] H. Ishibuchi, N. Tsukamoto, Y. Nojima, Evolutionary many-objective optimization:  
609 A short review, in: *2008 IEEE Congress on Evolutionary Computation (IEEE World*  
610 *Congress on Computational Intelligence)*, IEEE, 2008, pp. 2419–2426.
- 611 [9] X.-S. Yang, *Engineering optimization: an introduction with metaheuristic applications*,  
612 John Wiley & Sons, 2010.
- 613 [10] Y. Jin, J. Branke, et al., Evolutionary optimization in uncertain environments-a survey,  
614 *IEEE Transactions on evolutionary computation* 9 (3) (2005) 303–317.
- 615 [11] C.-K. Goh, K. C. Tan, Evolutionary multi-objective optimization in uncertain environ-  
616 ments, *Issues and Algorithms, Studies in Computational Intelligence* 186 (2009) 5–18.
- 617 [12] K. C. Tan, C. K. Goh, Handling uncertainties in evolutionary multi-objective opti-  
618 mization, in: *IEEE World Congress on Computational Intelligence*, Springer, 2008, pp.  
619 262–292.

- 620 [13] C. P. Ho, P. Parpas, Multilevel optimization methods: Convergence and problem struc-  
621 ture (2016).
- 622 [14] M. El-Beltagy, A. Keane, Optimisation for multilevel problems: A comparison of various  
623 algorithms, in: Adaptive computing in design and manufacture, Springer, 1998, pp.  
624 111–120.
- 625 [15] A. March, K. Willcox, Provably convergent multifidelity optimization algorithm not  
626 requiring high-fidelity derivatives, AIAA journal 50 (5) (2012) 1079–1089.
- 627 [16] M. K. Zahir, Z. Gao, Variable-fidelity optimization with design space reduction, Chinese  
628 Journal of Aeronautics 26 (4) (2013) 841–849.
- 629 [17] V. Picheny, D. Ginsbourger, Y. Richet, Noisy expected improvement and on-line com-  
630 putation time allocation for the optimization of simulators with tunable fidelity (2010).
- 631 [18] J. Teich, Pareto-front exploration with uncertain objectives, in: International Confer-  
632 ence on Evolutionary Multi-Criterion Optimization, Springer, 2001, pp. 314–328.
- 633 [19] E. J. Hughes, Evolutionary multi-objective ranking with uncertainty and noise, in: In-  
634 ternational Conference on Evolutionary Multi-Criterion Optimization, Springer, 2001,  
635 pp. 329–343.
- 636 [20] P. Limbourg, D. E. S. Aponte, An optimization algorithm for imprecise multi-objective  
637 problem functions, in: 2005 IEEE Congress on Evolutionary Computation, Vol. 1, IEEE,  
638 2005, pp. 459–466.
- 639 [21] H. Eskandari, C. D. Geiger, R. Bird, Handling uncertainty in evolutionary multiobjec-  
640 tive optimization: Spga, in: 2007 IEEE Congress on Evolutionary Computation, IEEE,  
641 2007, pp. 4130–4137.
- 642 [22] D.-w. Gong, N.-n. Qin, X.-y. Sun, Evolutionary algorithms for multi-objective optimiza-  
643 tion problems with interval parameters, in: 2010 IEEE Fifth International Conference  
644 on Bio-Inspired Computing: Theories and Applications (BIC-TA), IEEE, 2010, pp.  
645 411–420.
- 646 [23] G. L. Soares, F. G. Guimarães, C. A. Maia, J. A. Vasconcelos, L. Jaulin, Interval  
647 robust multi-objective evolutionary algorithm, in: 2009 IEEE Congress on Evolutionary  
648 Computation, IEEE, 2009, pp. 1637–1643.
- 649 [24] M. Li, R. Silva, F. Guimarães, D. Lowther, A new robust dominance criterion for  
650 multiobjective optimization, IEEE Transactions on Magnetics 51 (3) (2015) 1–4.
- 651 [25] F. Khosravi, M. Müller, M. Glaß, J. Teich, Uncertainty-aware reliability analysis and  
652 optimization, in: 2015 Design, Automation & Test in Europe Conference & Exhibition  
653 (DATE), IEEE, 2015, pp. 97–102.

- 654 [26] F. Khosravi, M. Borst, J. Teich, Probabilistic dominance in robust multi-objective op-  
655 timization, in: 2018 IEEE Congress on Evolutionary Computation (CEC), IEEE, 2018,  
656 pp. 1–6.
- 657 [27] F. Khosravi, A. Raß, J. Teich, Efficient computation of probabilistic dominance in robust  
658 multi-objective optimization, arXiv preprint arXiv:1910.08413 (2019).
- 659 [28] F. Khosravi, A. Rass, J. Teich, Efficient computation of probabilistic dominance  
660 in multi-objective optimization, *ACM Trans. Evol. Learn. Optim.* 1 (4) (oct 2021).  
661 doi:10.1145/3469801.  
662 URL <https://doi.org/10.1145/3469801>
- 663 [29] M. Mlakar, T. Tušar, B. Filipič, Comparing solutions under uncertainty in multiobjec-  
664 tive optimization, *Mathematical Problems in Engineering* 2014 (2014).
- 665 [30] F. Fusi, P. M. Congedo, An adaptive strategy on the error of the objective functions for  
666 uncertainty-based derivative-free optimization, *Journal of Computational Physics* 309  
667 (2016) 241–266.
- 668 [31] M. Rivier, P. Congedo, Surrogate-assisted bounding-box approach applied to con-  
669 strained multi-objective optimisation under uncertainty, *Reliability Engineering System*  
670 *Safety* 217 (2022) 108039. doi:<https://doi.org/10.1016/j.ress.2021.108039>.  
671 URL <https://www.sciencedirect.com/science/article/pii/S0951832021005445>
- 672 [32] S. Da Veiga, F. Delbos, Robust optimization for expensive simulators with surrogate  
673 models: Application to well placement for oil recovery, *Safety, Reliability, Risk and*  
674 *Life-Cycle Performance of Structures and Infrastructures* 11 (2013) 3321–3328.
- 675 [33] V. Baudoui, Optimisation robuste multiobjectifs par modèles de substitution, Ph.D.  
676 thesis, Toulouse, ISAE (2012).
- 677 [34] J. Janusevskis, R. Le Riche, Simultaneous kriging-based estimation and optimization of  
678 mean response, *Journal of Global Optimization* 55 (2) (2013) 313–336.
- 679 [35] B. J. Williams, T. J. Santner, W. I. Notz, Sequential design of computer experiments  
680 to minimize integrated response functions, *Statistica Sinica* (2000) 1133–1152.
- 681 [36] R. F. Coelho, Probabilistic dominance in multiobjective reliability-based optimization:  
682 Theory and implementation, *IEEE Transactions on evolutionary computation* 19 (2)  
683 (2014) 214–224.
- 684 [37] GPy, GPy: A gaussian process framework in python,  
685 <http://github.com/SheffieldML/GPy> (since 2012).
- 686 [38] C. E. Rasmussen, C. K. I. Williams, Gaussian processes for machine learning., *Adaptive*  
687 *computation and machine learning*, MIT Press, 2006.

- 688 [39] C. E. Powell, et al., Generating realisations of stationary gaussian random fields by  
689 circulant embedding, *matrix* 2 (2) (2014) 1.
- 690 [40] M. Pourahmadi, Covariance estimation: The glm and regularization perspectives, *Sta-*  
691 *tistical Science* (2011) 369–387.
- 692 [41] C. Lantuéjoul, N. Desassis, Simulation of a gaussian random vector: a propagative  
693 version of the gibbs sampler, in: *The 9th international geostatistics congress, 2012*, pp.  
694 174–181.
- 695 [42] A. M. Panunzio, R. Cottureau, G. Puel, Large scale random fields generation using  
696 localized karhunen–loève expansion, *Advanced Modeling and Simulation in Engineering*  
697 *Sciences* 5 (1) (2018) 20.
- 698 [43] M. Herdin, N. Czink, H. Ozcelik, E. Bonek, Correlation matrix distance, a meaningful  
699 measure for evaluation of non-stationary mimo channels, in: *2005 IEEE 61st Vehicular*  
700 *Technology Conference*, Vol. 1, IEEE, 2005, pp. 136–140.
- 701 [44] M. Rivier, P. M. Congedo, Surrogate-assisted bounding-box approach for optimization  
702 problems with tunable objectives fidelity, *Journal of Global Optimization* 75 (4) (2019)  
703 1079–1109.
- 704 [45] M. P. Dubuisson, A. K. Jain, A modified hausdorff distance for object matching, in:  
705 *Proceedings of 12th International Conference on Pattern Recognition*, Vol. 1, 1994, pp.  
706 566–568 vol.1. doi:10.1109/ICPR.1994.576361.
- 707 [46] M. Pini, G. Persico, D. Pasquale, S. Rebay, Adjoint method for shape optimization in  
708 real-gas flow applications, *ASME Journal of Engineering for Gas Turbines and Power*  
709 137 (3) (2015).
- 710 [47] S. Vitale, T. A. Albring, M. Pini, N. R. Gauger, P. Colonna, Fully turbulent discrete  
711 adjoint solver for non-ideal compressible flow applications, *Journal of the Global Power*  
712 *and Propulsion Society* 1 (2017) Z1FVOI.
- 713 [48] G. Persico, P. Rodriguez-Fernandez, A. Romei, High-fidelity shape-optimization of non-  
714 conventional turbomachinery by surrogate evolutionary strategies, *Journal of Turboma-*  
715 *chinery* 141 (8) (2019) 081010.
- 716 [49] M. Pini, G. Persico, V. Dossena, Robust adjoint-based shape optimization of supersonic  
717 turbomachinery cascades, in: *ASME Turbo Expo 2014: Turbine Technical Conference*  
718 *and Exposition*, American Society of Mechanical Engineers, 2014, pp. V02BT39A043–  
719 V02BT39A043.
- 720 [50] N. Razaaly, G. Gori, G. Iaccarino, P. Congedo, Optimization of an orc supersonic nozzle  
721 under epistemic uncertainties due to turbulence models, in: *GPPS 2019*, 2019.

- 722 [51] N. Razaaly, G. Persico, P. M. Congedo, Impact of geometric, operational, and model  
723 uncertainties on the non-ideal flow through a supersonic orc turbine cascade, *Energy*  
724 169 (2019) 213–227.
- 725 [52] M. Pini, G. Persico, V. Dossena, Robust adjoint-based shape optimization of supersonic  
726 turbomachinery cascades, in: *ASME Turbo Expo 2014: Turbine Technical Conference*  
727 *and Exposition*, American Society of Mechanical Engineers, 2014, pp. V02BT39A043–  
728 V02BT39A043.
- 729 [53] P. Congedo, G. Geraci, R. Abgrall, V. Pediroda, L. Parussini, Tsi metamodels-based  
730 multi-objective robust optimization, *Engineering Computations (Swansea, Wales)* 30 (8)  
731 (2013) 1032–1053.
- 732 [54] G. Farin, *Curves and Surfaces for CAGD: A Practical Guide*, 5th Edition, Morgan  
733 Kaufmann Publishers Inc., San Francisco, CA, USA, 2002.
- 734 [55] F. Palacios, M. F. Colonno, A. C. Aranake, A. Campos, S. R. Copeland, T. D.  
735 Economon, A. K. Lonkar, T. W. Lukaczyk, T. W. R. Taylor, J. J. Alonso, Stanford  
736 University Unstructured (SU2): An open-source integrated computational environment  
737 for multi-physics simulation and design, in: *51st AIAA Aerospace Sciences Meeting and*  
738 *Exhibit*, 2013.
- 739 [56] T. D. Economon, D. Mudigere, G. Bansal, A. Heinecke, F. Palacios, J. Park,  
740 M. Smelyanskiy, J. J. Alonso, P. Dubey, Performance optimizations for scalable im-  
741 plicit {RANS} calculations with {SU2}, *Computers & Fluids* 129 (2016) 146 – 158.  
742 doi:<http://dx.doi.org/10.1016/j.compfluid.2016.02.003>.  
743 URL <http://www.sciencedirect.com/science/article/pii/S0045793016300214>
- 744 [57] M. Pini, S. Vitale, , P. Colonna, G. Gori, A. Guardone, T. Economon, J. Alonso, F. Pala-  
745 cios, Su2: the open-source software for non-ideal compressible flows, in: *NICFD 2016:*  
746 *1st International Seminar on Non-Ideal Compressible-Fluid Dynamics for Propulsion &*  
747 *Power*, Varenna, Italy, 2016.
- 748 [58] S. Vitale, G. Gori, M. Pini, A. Guardone, T. D. Economon, F. Palacios, J. J. Alonso,  
749 P. Colonna, Extension of the SU2 open source CFD code to the simulation of turbulent  
750 flows of fluids modelled with complex thermophysical laws, no. *AIAA Paper 2015-2760*,  
751 2015.
- 752 [59] N. Razaaly, Rare event estimation and robust optimization methods with applications  
753 to orc turbine cascade, Ph.D. thesis, Université Paris-Saclay (2019).
- 754 [60] J. R. Gardner, M. J. Kusner, Z. E. Xu, K. Q. Weinberger, J. P. Cunningham, Bayesian  
755 optimization with inequality constraints., in: *ICML*, 2014, pp. 937–945.
- 756 [61] B. Echard, N. Gayton, M. Lemaire, Ak-mcs: an active learning reliability method  
757 combining kriging and monte carlo simulation, *Structural Safety* 33 (2) (2011) 145–154.

- 758 [62] N. Razaaly, P. M. Congedo, Novel algorithm using active metamodel learning and im-  
759 portance sampling: application to multiple failure regions of low probability, *Journal of*  
760 *Computational Physics* 368 (2018) 92–114.
- 761 [63] C. K. Williams, M. Seeger, Using the nyström method to speed up kernel machines, in:  
762 *Advances in neural information processing systems*, 2001, pp. 682–688.
- 763 [64] R. Tipireddy, D. A. Barajas-Solano, A. M. Tartakovsky, Conditional karhunen-lo\eve  
764 expansion for uncertainty quantification and active learning in partial differential equa-  
765 tion models, arXiv preprint arXiv:1904.08069 (2019).