Highlights

Robust design optimization of expensive stochastic simulators under lack-of-knowledge

Conradus van Mierlo, Augustin Persoons, Matthias G.R. Faes, David Moens

- In this work, a robustness under lack-of-knowlegde method for noisy function responses is proposed
- The proposed method is demonstrated on both analytical cases and numerical crashworthiness simulations
- The proposed method is shown to provide good results based on only a very limited number of model evaluations
- The extension is based on a GP with homocedastic noise variance that is assumed or calibrated during optimisation

Robust design optimization of expensive stochastic simulators under lack-of-knowledge

Conradus van Mierlo[a,*], Augustin Persoons[a], Matthias G.R. Faes[b], David Moens[a]

[a] *KU Leuven, Department of Mechanical Engineering, Jan De Nayerlaan 5 2860, Sint-Katelijne-Waver, Belgium*
[b] *TU Dortmund University, Chair for Reliability Engineering, Leonhard-Euler-Strasse 5, 44227 Dortmund, Germany*

Abstract

Robust design optimisation of stochastic black-box functions is a challenging task in engineering practice. Crashworthiness optimisation qualifies as such problem especially ~~in~~ regards ~~with~~ the high computational costs. Moreover, in early design phases there may be significant uncertainty about the ~~parameters used in the numerical models used to predict the systems response.~~ Therefore, this paper proposes an adaptive surrogate-based strategy for robust design optimisation of noise-contaminated models under lack-of-knowledge uncertainty. This approach is a significant extension to the Robustness under Lack-of-Knowledge method (RULOK) previously introduced by the authors, which was limited to noise-free models. In this work it is proposed to use a Gaussian Process as a regression model based on a noisy kernel. The learning process is adapted to account for noise variance either imposed and known or empirically learned as part of the learning process. The method is demonstrated on three analytical benchmarks and one engineering crashworthiness optimisation problem. In the case studies, multiple ways of determining the noise kernel are investigated: (1) based on a coefficient of variation, (2) calibration in the Gaussian Process model, (3) based on engineering judgement, including a study of the sensitivity of the result with respect to these parameters. The results highlight that the proposed method is able to efficiently identify a robust design point even with extremely limited or biased prior knowledge about the noise.

Keywords: robust optimisation, interval analysis, Gaussian Process modeling, efficient global optimisation, crashworthiness

1. Introduction

Robust design optimisation is a methodology that aims to create products and processes that are insensitive to variations from, e.g., applied loads, environmental conditions, manufacturing processes, and was pioneered by *Genichi Taguchi* who first applied his methodology on electrical circuits [1, 2]. This methodology has since been further developed and multiple definitions of robustness are found in literature. Two main classes of methods can be drawn from it: the first is aimed at minimizing the output variance, see, e.g., [3, 4, 5], while the second is aimed at optimizing of both the objective function and the variance associated with this optimum, see, e.g., [6, 7, 8]. Moreover, robust design methods differ in the conceptualisation of the source of variations that these designs are subjected to, which is best described by non-deterministic approaches. Typically, these non-deterministic modelling strategies are categorised as probabilistic and possibilistic

*Corresponding author
Email address: koen.vanmierlo[at]kuleuven.be (Conradus van Mierlo)

approaches [9]. Where probabilistic methods are best suited for aleatory uncertainties as they describe non-determinism via random variables defined by their joint probability distributions, possibilistic approaches are usually better suited to cover both aleatory and epistemic uncertainties, which can be modelled by techniques such as: interval [10], fuzzy sets [11], information gap methods [12], and imprecise probabilities [13, 14].

The authors of this work recently introduced the Robustness Under Lack-Of-Knowledge method (RULOK) [5]. This method is aimed at finding the design that causes the least amount of variation from a set of admissible design parameters $\mathbf{z} \in \mathcal{Z} \subseteq \mathbb{R}^{n_z}$ with $\mathcal{Z}$ the set of admissible designs and $n_z \in \mathbb{N}$. The design parameters represent quantities that are controlled by the analyst, sutch as e.g., plate thickness values, hole diameters. The uncontrolled parameters are modelled as purely epistemic interval parameters $\mathbf{x} \in \mathbf{x}^I \subseteq \mathbb{IR}^{n_x}$ with $n_x \in \mathbb{N}$ and $\mathbb{IR}$ the set of *real valued closed intervals*. They represents parameters affected by significant uncertainties, such as e.g., weld diameters, transmission parameters, material parameters. At the basis of the RULOK method an adaptively refined Gaussian Process (GP) is used to estimate the minimum interval width of the response for each of the designs. However, this approach is not well suited for non-linear noisy systems, as seen in e.g., crash analysis, since it assumes a deterministic behaviour of the underlying model. In these cases, the non-determinism about these systems should be considered in the Gaussian process to calibrate a meaningful surrogate. This remark is especially true when the meta-model is used for robust design optimisation and reliability based design optimisation (RBDO) [15, 16] as these methods require a meaningful surrogate to identify the correct optimum. Hence, in order to use industrial size multi-disciplinary numerical models such as those used in crash optimisation, see, e.g., [17] a more advanced meta-model is needed. One should note here that crashworthiness optimisation using these advanced numerical models has always been challenging, not only for meta-model assisted techniques. Crashworthiness simulations or other advanced non-linear finite element methods can be considered as *black-box* functions, as no closed-form formulation or gradient information is available. Moreover, it is well known for crash analysis that the deterministic simulations might experience *numerical inadequacies*, i.e., dynamic- and numerical instabilities that can cause a small (infinitesimal) change in the input to produce a major change in output [18, 19, 20]. In addition to these *numerical inadequacies*, in the specific case of explicit dynamic analysis there is a small but progressively increasing *numerical error* accumulation [21, 22, 23, 17] over the total duration of the simulation.The accumulated error term of both the *numerical inadequacies* and *numerical error* makes the deterministic simulation behave like a stochastic simulation model despite its deterministic nature. In other words, evaluation of the explicit numerical model returns different results for the same set of input parameters.

The previously introduced RULOK approach relies on an *interpolating* GP also known as Kriging [24, 25, 26] based on the assumption that the underlying systems behaviour is deterministic. However, due to the combination of the *numerical inadequacies* and *numerical errors* such systems exhibit noisy behaviour. The RULOK approach is not capable of representing the behaviour of a noisy system and induces significant over-fitting. The GP used in RULOK is adaptivly refined using a specific learning function, which identifies the next point to be evaluated by the expensive to evaluate *black-box* function. Note that the idea behind this adaptive strategy lies at the basis of Efficient Global Optimisation (EGO), as introduced by [27]. In this paper, an extension to the original RULOK method is proposed, which enables the method to work with both deterministic functions and non-deterministic functions. Therefore, this paper introduces the use of a GP with a noisy kernel, which is capable of truthfully representing stochastic function responses. The idea of using a GP with a noise kernel is not new and has gained an increasing in interest over the past decades, see, e.g., [28, 29] for an overview. In these works, the learning function used in

EGO is adapted to account for the noise contaminated responses. In this work the learning function introduced in [5] is slightly adapted. Especially the stopping criterion is changed to a more general formulation that accounts for the set or calibrated noise kernel of the GP. The paper is structured as follows: Section 2 describes the measure of robustness under lack-of-knowledge. In Section 3 the details about noisy GP's are provided, while Section 4 describes the new stopping criterion and provides an overview of the RULOK method. In Section 5, the method is tested on three noise contaminated analytical functions and in Section 6 an example about crashworthiness optimisation of a crashbox is given. Finally, in section 7 a discussion about the results is presented before conclusions are drawn in Section 8.

2. Robustness under lack-of-knowledge uncertainty

The uncertainty considered in this work is purely epistemic in nature and results from a lack-of-knowledge about the exact value of the parameter. The *real* value of the uncertain quantities, be it deterministic or variable, are modelled as an interval parameter [30]. Note that in this paper the following conventions are used: a vector is indicated as lower-case boldface characters $\mathbf{x}$, matrices are expressed as upper-case boldface characters $\mathbf{X}$ and interval parameters are indicated using apex I: x^I. Precisely, an interval is represented using the bounds of the interval:

$$\mathbf{x}^I = [\underline{\mathbf{x}}; \overline{\mathbf{x}}] = \{\mathbf{x} \in \mathbb{R}^{n_x} \mid \underline{\mathbf{x}} \leq \mathbf{x} \leq \overline{\mathbf{x}}\}, \tag{1}$$

where $\underline{\mathbf{x}}$ denotes the lower bound and $\overline{\mathbf{x}}$ denotes the upper bound. In addition, an interval can be represented by the centre point $\hat{\mathbf{x}} = \frac{\underline{\mathbf{x}}+\overline{\mathbf{x}}}{2}$ and radius $\Delta\mathbf{x} = \frac{\overline{\mathbf{x}}-\underline{\mathbf{x}}}{2}$ of the interval. In the multidimensional case a uncertain parameter vector $\mathbf{x}$ is represented as an interval vector $\mathbf{x}^I = [z_1^I, x_1^I, \ldots, x_{n_x}^I]$, with $x_i^I,\ i, \ldots, n_x$ the i^{th} parameter interval. An interval is considered *closed* when both the upper and lower bounds are a member of the interval with $\mathbb{IR}$ the domain of a real-valued intervals.

2.1. Propagation of interval valued uncertainty

In this work the model m is a continuous function on $\mathbb{R}$, which is parameterised by a parameter vector $\boldsymbol{\theta}$. The parameter vector is devided in two parts $\boldsymbol{\theta} = \{\mathbf{x}, \mathbf{z}\}$, with $\mathbf{x}$ the uncertain parameters and $\mathbf{z}$ the design parameters. The number of elements in the parameter vector are indicated by $n_{\boldsymbol{\theta}} = n_x + n_z$. By solving the model m the parameter vector $\boldsymbol{\theta}$ is transformed $\mathbb{R}^{n_{\boldsymbol{\theta}}} \mapsto \mathbb{R}$ to a scalar response quantity $y \in \mathbb{Y} \subset \mathbb{R}$, defined as:

$$m : y = m(\boldsymbol{\theta}), \tag{2}$$

with $\mathbb{Y}$ the set of admissible model parameters. The main goal of the interval analysis is to identify the extremes of the set of system responses $\tilde{y}$. Since finding the set $\tilde{y}$ is in general computationally intractable, the exact solution set is often approximated by a realisation set $\tilde{y}_s$ defined as [31]:

$$\tilde{y}_s = \left\{ y_j \mid y_j = m(\boldsymbol{\theta}_j); \mathbf{x}_j \in \mathbf{x}^I; j = 1, \ldots, n_q \right\}. \tag{3}$$

The set $\tilde{y}_s$ is typically constructed by performing n_q deterministic evaluations $y_j = m(\boldsymbol{\theta}_j)$ of the numerical model, with y_j the response of the j^{th} solution. For each of these n_q solutions, a sample is taken within the range of the interval $\mathbf{x}^I$. The main challenge herein is choosing the samples x_j such that $\tilde{y}_s$ is an accurate approximation of $\tilde{y}$. A first way to obtain such approximation is to

follow an optimisation approach. Here, the exact solution set $\tilde{y}$ is approximated by an accurate interval for the one dimensional case. The corresponding optimisation problem is defined as:

$$\begin{aligned} \underline{y} &= \min_{\mathbf{x}\in\mathbf{x}^I} m(\boldsymbol{\theta}), \\ \overline{y} &= \max_{\mathbf{x}\in\mathbf{x}^I} m(\boldsymbol{\theta}), \end{aligned} \tag{4}$$

where $y^I = [\underline{y}; \overline{y}]$ is the solution interval. For the higher dimensional case a conservative approximation is made about the hyper-cubic solution set in higher dimensions $\mathbf{y}^I = [y_1^I, y_2^I, \ldots, y_{n_y}^I]$, with $\tilde{y} \subseteq \mathbf{y}^I$. The function m represents the numerical model or any representative surrogate model, as seen in, e.g., [32, 33, 34, 35]. The global minimum or maximum is found following the anti-optimisation framework [36] based on global optimisation. The exact optimisation strategy to use here is highly problem dependent as the behaviour of the goal function with respect to the uncertain parameters is unpredictable in the case of strongly non-linear problems [37]. Note here that other approaches exist for interval analysis, see, e.g., [31] for a recent review.

2.2. Robustness for interval analysis

Robustness under lack-of-knowledge uncertainty is defined in [5] as the ratio of input uncertainty to the output uncertainty, which can be regarded as an interval counterpart to robustness measures that minimize the variance of the performance. The robustness measure is illustrated for a case with one interval valued input parameter, of which the input and output uncertainty are represented respectively by the scalar interval radius Δx and the associated scalar output interval radius Δy, which is a function of the design parameter $\mathbf{z}$. Hence, the output radius should be obtained for multiple designs $\mathbf{z} \in \mathcal{Z}$. The robustness is defined as:

$$R(\mathbf{z}) = \frac{\Delta x}{\Delta y(\mathbf{z})} = \frac{\overline{x} - \underline{x}}{\overline{y}(\mathbf{z}) - \underline{y}(\mathbf{z})}. \tag{5}$$

Since the uncertainty Δx is independent of the design $\mathbf{z}$, finding the most robust design $\mathbf{z}^*$ is reformulated to the minimisation of the output uncertainty, defined by:

$$\mathbf{z}^* = \operatorname*{argmin}_{\mathbf{z}\in\mathcal{Z}} [\overline{y} - \underline{y}] = \operatorname*{argmin}_{\mathbf{z}\in\mathcal{Z}} [\max_{\mathbf{x}\in\mathbf{x}^I} m(\boldsymbol{\theta}) - \min_{\mathbf{x}\in\mathbf{x}^I} m(\boldsymbol{\theta})], \tag{6}$$

with $\max_{\mathbf{x}\in\mathbf{x}^I} m(\boldsymbol{\theta})$ the predicted upper-bound and $\min_{\mathbf{x}\in\mathbf{x}^I} m(\boldsymbol{\theta})$ the predicted lower-bound derived from the GP surrogate. Note there that Equation (6) can be evaluated for multiple outputs $\mathbf{y}$. In the specific case of a stochastic function the location of the upper- and lower-bound can only be estimated by the mean of the process. This point is illustrated in Figure 1 showing the robust design point indicated in green based on the mean upper- and lower bound in red and blue, respectfully. Obtaining these bounds from a stochastic function is not trivial, especially correct estimations of the variance might be challenging to obtain. Hence, in this work a GP is used to estimate the mean responses based on a limited number of evaluations. The variance of the process is then reflected by the noise variance of the GP.

3. Gaussian process model for noisy responses

This section provides a short theoretical summary of Gaussian Process (GP) models or Kriging [24][38], an introduction with examples is also available in [39]. A GP model is a stochastic meta-model that assumes $m(\boldsymbol{\theta})$ to be a realisation of a Gaussian process, which is defined as [40]:

$$\mathcal{G} = \beta^T \mathbf{f}(\alpha) + \sigma^2 F(\mathbf{x}, \Omega), \tag{7}$$

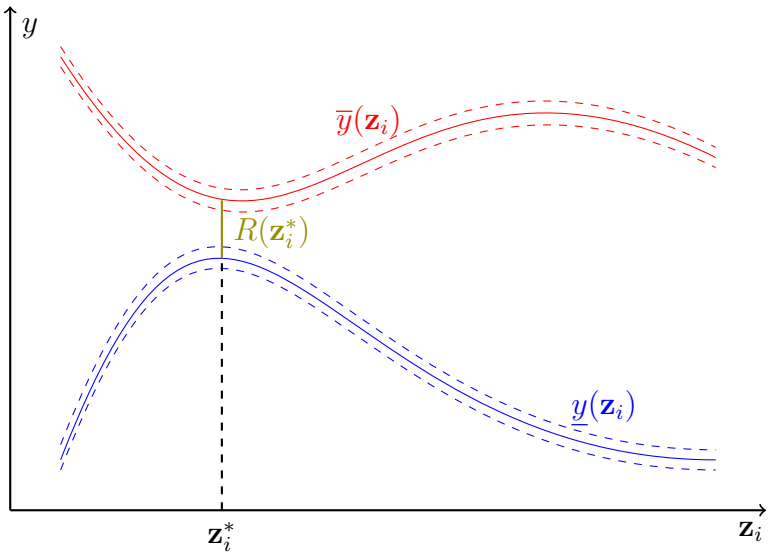

Figure 1: Illustration of the optimal robust design points $R(\mathbf{z}^*)$ (orange) for the with noise contaminated upper- and lower bounds $\overline{y}$ and $\underline{y}$ for a specific design parameter $\mathbf{z}_i$, adapted from [CITE self]

with the first term being a deterministic regression model with $\mathbf{f}(\alpha) = \{f_1(\alpha), \ldots, f_k(\alpha)\}$ a set of arbitrary basis functions, and β^T a vector of regression coefficients. The second term consists of a zero-mean, unit variance, stationary Gaussian process $\digamma(\mathbf{x}, \Omega)$ scaled with a constant variance of the Gaussian process σ^2. The underlying probability space of the Gaussian process is represented by Ω and the correlation between two points $\mathbf{r}$ and $\mathbf{r}'$ is defined by the covariance function $K(\mathbf{r}, \mathbf{r}', l_c)$, with l_c the characteristic length or other hyper-parameters. In general, one refers to the covariance matrix $\mathbf{K}$ where the covariance is determined for all points in a domain. The reader may refer to [41] for details about different covariance functions in Gaussian processes. In this paper two well-known covariance functions are used: The Gaussian kernel (also known as squared-exponential covariance function) and the Matérn $\frac{5}{2}$ kernel.

3.1. Noise Gaussian Process predictions

In the specific case where a GP is used to predict a stochastic function, a noise term can be defined. In general, the noise of the stochastic function can be defined in the following ways:

$$y = m(\boldsymbol{\theta}) + \zeta, \tag{8}$$

where the additive noise is assumed to follow a zero-mean Gaussian distribution:

$$\zeta = \mathcal{N}(0, \Sigma_{\text{gp}}), \tag{9}$$

with Σ_{gp} the covariance matrix of the noise term. Depending on the definition of Σ_{gp} different classes of noise are identified:

$$\Sigma_{\text{gp}} = \sigma_{\text{gp}}^2 \mathbf{I}, \tag{10}$$

with $\mathbf{I}$ the identity matrix, for the case of homogeneous (*homoscedastic*) noise. It is also possible that for each observed response an independent noise variance is observed, which is defend as:

$$\Sigma_{\text{gp}} = \text{Diag}(\sigma_{\text{gp}}^2), \tag{11}$$

for the case of independent heterogeneous (*heteroscedastic*) noise. In the most general case described as *general heteroscedastic* the noise matrix has the shape of a general covariance maxtrix $\Sigma_{\text{gp}} = \Sigma_{\text{gp}}$ where for each observation a different noise variance can be obtained and correlations of this noise are possible. The work presented in this paper is limited to homoscedastic noise. In other words, it is assumed that all observed responses have the same noise variance without any underlying correlations. Due to the noise variance the covariance matrix is defined as $\mathbf{C} = \sigma^2\mathbf{K} + \Sigma_{\text{gp}}$. The GP-model is then calibrated on an initial design of experiments $\mathbf{x}_{DOE}$ obtained from, i.e., Latin hyper-cube sampling and their observed results $\mathbf{y}_{DOE}$. Conditional on the observed data the mean and the variance of the Gaussian process can be estimated [40]:

$$\mu_{gp} = \mathbf{f}(\mathbf{x})^T\hat{\beta} + \mathbf{r}(\mathbf{x})^T\mathbf{C}^{-1}(\mathbf{y}_{DOE} - \mathbf{F}\hat{\beta}), \tag{12}$$

$$\sigma^2_{\mu} = \left(\sigma^2 - \mathbf{c}^T(\mathbf{x})\mathbf{C}^{-1}\mathbf{c}(\mathbf{x}) + \mathbf{u}_c^T(\mathbf{x})(\mathbf{F}^T\mathbf{C}^{-1}\mathbf{F})^{-1}\mathbf{u}_c(\mathbf{x})\right), \tag{13}$$

with $\mathbf{F}$ the matrix of the observed trend, $\mathbf{c}(\mathbf{x})$ a cross covariance vector between predicted points x and observed points and with:

$$\hat{\beta} = (\mathbf{F}^T\mathbf{C}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{C}^{-1}\mathbf{y}_{DOE}, \tag{14}$$

the general least-squares estimate of regression coefficients β and

$$\mathbf{u}_c(\mathbf{x}) = \mathbf{F}^T\mathbf{C}^{-1}\mathbf{c}(\mathbf{x}) - \mathbf{f}(\mathbf{x}). \tag{15}$$

Note that in the special case of homoscedastic noise $\mathbf{C} = \sigma^2\mathbf{K} + \sigma^2_{\text{gp}}\mathbf{I}$, with as total GP variance:

$$\sigma^2_{\text{total}} = \sigma^2_{\text{gp}} + \sigma^2. \tag{16}$$

Equations (12) and (13) are referred to as the mean and variance of the GP predictor, respectively. The parameters of the GP, e.g., β,σ^2,l_c, are optimised using maximum likelihood estimation, which maximizes the likelihood of observing the points in $\mathbf{y}_{\text{DOE}}$. In the case of unknown homoscastic noise an additional noise parameter σ^2_{gp} is added to the maximum likelihood estimation [39]. Note that, unlike the noise free case, the variance of the prediction at an experimental design point $x \in \mathbf{x}_{DOE}$ does not collapse to zero, and the GP predictor becomes a regression model as it is no longer interpolating through the observations.

3.2. Predicting interval bounds with a Gaussian Process model

Based on an initial set of evaluations, the GP is calibrated and the model responses can be obtained based on the easy to evaluate GP. To this end, μ_{gp} is considered to be the best GP-estimate and σ^2_{gp} is the variance over this estimate. For the specific application of estimating an output interval based on the GP-model the main interest lies in estimation of the maximum and the minimum response over the complete range of uncertainty. Therefore, the bounds of the response are estimated by:

$$\overline{y}_{gp}(\mathbf{z}) = \overline{\mu}_{gp}(\mathbf{z}) = \max_{\mathbf{x}\in\mathbf{x}^I} \mu_{gp}(\boldsymbol{\theta}), \tag{17}$$

$$\underline{y}_{gp}(\mathbf{z}) = \underline{\mu}_{gp}(\mathbf{z}) = \min_{\mathbf{x}\in\mathbf{x}^I} \mu_{gp}(\boldsymbol{\theta}). \tag{18}$$

A similar approach can be taken to identify the maximum and minimum of the confidence bounds:

$$\overline{\delta}_{\mu+\sigma_\mu}(\mathbf{z}) = \max_{\mathbf{x}\in\mathbf{x}^I} (\mu_{gp}(\boldsymbol{\theta}) + c\sigma_\mu(\boldsymbol{\theta})), \tag{19}$$

$$\underline{\delta}_{\mu+\sigma_\mu}(\mathbf{z}) = \min_{\mathbf{x}\in\mathbf{x}^I} (\mu_{gp}(\boldsymbol{\theta}) + c\sigma_\mu(\boldsymbol{\theta})), \tag{20}$$

$$\overline{\delta}_{\mu-\sigma_\mu}(\mathbf{z}) = \max_{\mathbf{x}\in\mathbf{x}^I} (\mu_{gp}(\boldsymbol{\theta}) - c\sigma_\mu(\boldsymbol{\theta})), \tag{21}$$

$$\underline{\delta}_{\mu-\sigma_\mu}(\mathbf{z}) = \min_{\mathbf{x}\in\mathbf{x}^I} (\mu_{gp}(\boldsymbol{\theta}) - c\sigma_\mu(\boldsymbol{\theta})), \tag{22}$$

with $c\sigma$ confidence bounds. The bounds of the response are estimated for each design point $\mathbf{z}$, based on Equations (17-22). Note that although the GP is cheap to evaluate finding the minimum and maximum response as in Equations (17) until (22) is non-trivial as this is a non-convex problem. However, successful strategies have been proposed to efficiently optimise such problems e.g., using branch and bound algorithms as proposed in [27]. In this work, the continuous problem is discretised over a fine grid with a fixed number of points, which was also done in the previous work of the authors. The complex problem of identifying the maximum and minimum in a continuous setting reduces to identifying the highest value in a set of candidates in a grid-shaped design. Note that this only works efficiently with a low number of parameters, as the computational burden increases exponentially $\mathcal{O}(n^d)$ with dimension d for a full grid.

4. Adaptive refinement of the noisy Gaussian process model

In this section the learning function introduced in [5] is described, with the new stopping criterion. The aim of the learning function is to identify points that improve the GP estimate of the robust design point. In this regard a balance should be found between, exploration (low prediction confidence) and exploitation (identified areas of possible optimum). The main goal of the optimisation procedure is to identify the most robust design point in $\mathbf{z} \in \mathcal{Z}$, such that this design provides a minimum variation in the output interval for all $\mathbf{x} \in \mathbf{x}^I$. This is enabled by adapting the maximum improvement [32] to work directly on the minimum interval width:

$$MI_z(\mathbf{z}) = \frac{\min_{\mathbf{z}\in\mathbf{z}^I} \left(\overline{y}_{gp}(\mathbf{z}) - \underline{y}_{gp}(\mathbf{z})\right) - \left(\overline{\delta}_{\mu-\sigma}(\mathbf{z}) - \underline{\delta}_{\mu+\sigma}(\mathbf{z})\right)}{\min_{\mathbf{z}\in\mathbf{z}^I} \left(\overline{y}_{gp}(\mathbf{z}) - \underline{y}_{gp}(\mathbf{z})\right)}, \tag{23}$$

with $\overline{\delta}_{\mu-\sigma}(\mathbf{z}) - \underline{\delta}_{\mu+\sigma}(\mathbf{z})$ the predicted minimum interval width $2\Delta\delta(\mathbf{z})$ with a confidence of $c\sigma$ about this bound, and $\min_{\mathbf{z}\in\mathbf{z}^I} \left(\overline{y}_{gp}(\mathbf{z}) - \underline{y}_{gp}(\mathbf{z})\right)$ the current best estimate of the robust design point z^{opt}. The learning function in Equation (23) is illustrated in Figure 2 were the GP predicted upper- and lower-bound are shown in the top graph. Here a design point z^* at $\min 2\Delta\delta$ is possible more robust then then the current optimum z^{opt} at $\min \Delta y_{\text{gp}}$. This is also shown in the graph below where $MI_z(z^*) > MI_z(z^{opt})$, illustrating that it is likely to improve the estimated robustness at $\min \Delta\delta$. Note here that by reaching a $MI_z(\mathbf{z}) \leq 0$ the two intervals are equal. Hence, one can state that it is not expected with, e.g, 95% confidence for $c = 1.96$, that there is a smaller bound of Δy within the current range of design parameters $\mathbf{z} \in \mathcal{Z}$.

4.1. Maximum improvement of the predicted bounds

The learning function in Equation (23) finds a promising design point $\mathbf{z} \in \mathcal{Z}$, based on the estimates of the GP-model. However, to improve the estimated interval width for each design

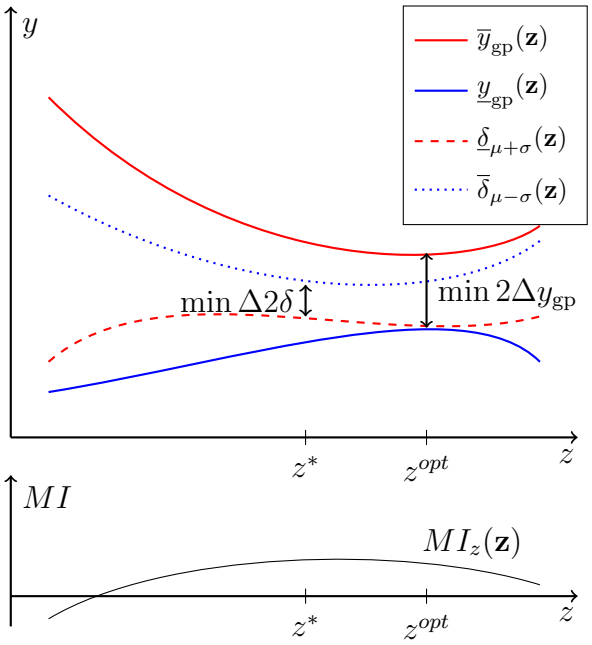

Figure 2: Illustration of the predicted mean bound $\Delta y_{\text{gp}}(\mathbf{z}) = \overline{y}_{\text{gp}}(\mathbf{z}) - \underline{y}_{\text{gp}}(\mathbf{z})$ and the minimum bound based on the confidence interval $\Delta\delta(\mathbf{z}) = \underline{\delta}_{\mu+\sigma}(\mathbf{z}) - \overline{\delta}_{\mu-\sigma}(\mathbf{z})$, adapted from [5]

a second learning function is used. This second function can be seen as an estimation of the relevance of candidates with respect to their coordinates in the $\boldsymbol{\theta}$ uncertain dimensions. The maximum improvement [32] is adapted to obtain the best estimate of the upper and lower bound for each design point. The maximum improvement of the lower bound of the interval is given as:

$$MI_{\min}(\boldsymbol{\theta}) = \min_{\mathbf{x}\in\mathbf{x}^I} \left[\mu_{gp}(\boldsymbol{\theta}) + c\sigma_{\mu}(\boldsymbol{\theta})\right] - \mu_{gp}(\boldsymbol{\theta}), \tag{24}$$

and the maximum improvement of the upper bound is given as:

$$MI_{\max}(\boldsymbol{\theta}) = \mu_{gp}(\boldsymbol{\theta}) - \max_{\mathbf{x}\in\mathbf{x}^I} \left[\mu_{gp}(\boldsymbol{\theta}) - c\sigma_{\mu}(\boldsymbol{\theta})\right], \tag{25}$$

both of which are not normalized as seen in Equation (23). This to guarantee a possible improvement even if the global minimum and maximum are identified. This learning function is illustrated in Figure 3, were the maximum improvement is given for a candidate point $\mathbf{x}^* \in \mathbf{x}^I$. The improvement of the minimum bound $MI_{\min}(\mathbf{z}, \mathbf{x}^*)$ at x^* is unlikely (negative value) while it seems likely to improve the upper limit $MI_{\max}(\mathbf{z}, \mathbf{x}^*)$. However, only one candidate point can be chosen to improve the estimation of the bounds. Therefore, for each evaluated point the highest improvement value is used, which can either improve the lower bound or the upper bound:

$$MI_x = \max(MI_{\min}, MI_{\max}). \tag{26}$$

This means that for the illustration in Figure 3 only the value of MI_{max} is saved for the point $\mathbf{x}^*$.

Finally, the candidate point that performs best over the sum of the two improvement functions Equation (23) and Equation (26) is selected. Hence, the next candidate point $\boldsymbol{\theta}_{\text{candidate}}$ is obtained by:

$$\boldsymbol{\theta}_{\text{candidate}} = \underset{\mathbf{z}\in\mathcal{Z}\ \mathbf{x}\in\mathbf{x}^I}{\operatorname{argmax}} \left[MI_z(\mathbf{z}) + MI_x(\boldsymbol{\theta})\right]. \tag{27}$$

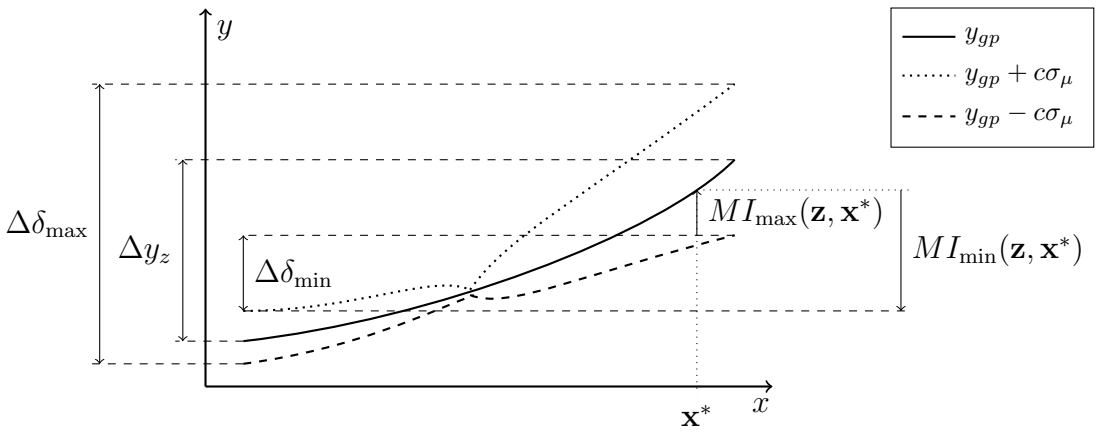

Figure 3: Illustration of the learning function for a candidate point x^*, showing the MI of the lower and upper bound; here the improvement of the lower bound is negative [5]

4.2. Stopping criterion for adaptive refinement of noisy responses

The role of a stopping criterion is to indicate when the algorithm reached a desired level of convergence. In this work, the stopping criterion is defined on the improvement of the robustness MI_z, which means that based on the current GP-model it is unlikely to identify a point that is more robust than the current best estimate $\min_{z\in\mathbf{z}^I}\left(\overline{y}_{gp}(z) - \underline{y}_{gp}(z)\right)$. However, this estimate of the interval width is affected by the noise variance of the GP, illustrated in Figure 4. The dashed blue and red lines indicate the Gaussian noise about the mean bounds at -5 and 20, which corresponds with the bounds of function $f_a(z_1 = 0)$, as shown in the case studies. The full lines are the prediction given by the GP model where the total variance is the sum of the GP noise and variance $\sigma^2_{\text{gp}} + \sigma^2$. In accordance with the learning function in Equation (23) the smallest interval width with 95% confidence is illustrated by $\min\Delta\delta$, which can never be larger than the interval width based on the noise $\Delta\sigma_n$. Hence, to account for the homoscedastic noise the stopping criteria is defined as:

$$MI_s = \frac{2c\sigma_{\text{gp}}}{\min_{z\in\mathbf{z}^I}\left(\overline{y}_{gp}(z) - \underline{y}_{gp}(z)\right)} + \epsilon, \tag{28}$$

with σ_{gp} the noise variance of the GP, which is either known or unknown. Note that in the case of unknown noise variance the stopping criterion changes over each iteration of the adaptive scheme. The adaptive refinement is complete when the possible improvement is smaller then the maximal improvement given the noise of the GP:

$$MI_z \leq MI_s. \tag{29}$$

By the end of the adaptive refinement one can state that according the the current GP there is with 95% confidence no point R within the domain smaller than $R(1+\epsilon)$. To prevent premature termination of the algorithm the method is only stopped when the criterion is satisfied by two consecutive iterations. Note that the stopping criterion in Equation (29) would not work in the case of *heteroscedastic* noise. Hence, the remainder of this paper focuses on homogeneous or *homoscedastic* noise.

4.3. Overview of the method

In Figure 5, a flowchart of the method is provided. The flowchart describes in detail the steps needed to perform the optimisation as proposed in this paper. The method starts at the

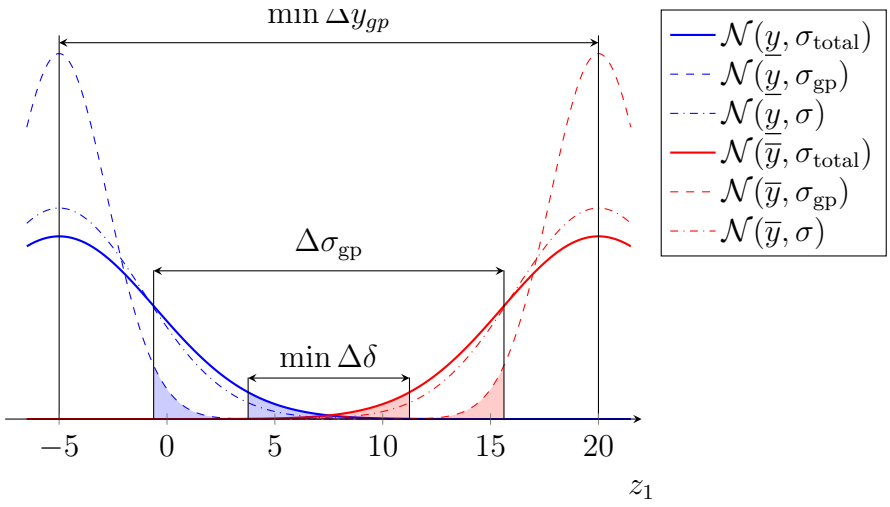

Figure 4: Illustration of the stopping criteria for a GP with noise; The illustration shows that the $\Delta\delta$ canFor both the upper and lower bound three normal distributions are drawn $\mathcal{N}(\overline{y}, \sigma_n)$,$\mathcal{N}(\overline{y}, \sigma_{gp})$, $\mathcal{N}(\overline{y}, \sigma_{\text{total}})$ indicated by the dashed, dash-dotted and full lines, respectively

initialisation where all parameters are selected by the user, i.e., using a set GP variance or calibrate for the GP variance, the correlation function that is used, size of the initial design of experiments, value for ϵ. After this initialisation is made, the initial design of experiments is evaluated by the model m and the GP is calibrated. Hereafter, the GP is adaptively refined to identify a new potential robust designs point based on the learning function in Section 4. For each newly identified point the model is evaluated $m(\boldsymbol{\theta}_{\text{candidate}})$ and the results are added to the Design of Experiments. This loop continues until the stopping criterion Equation (29) is met for two consecutive times. Finally, after finishing the optimisation, it is considered good practice to validate the results of the GP.

5. Analytical test functions with noise

To study the basic properties of the proposed method a set of analytical test functions is used, which are identical to the analytical functions used in [5]. However, in this work a random noise term is added. The analytical test functions with noise are defined as:

$$f_a(z_1, x_1) = z_1^2 x_1 - x_1^2 + \zeta_{\text{fn}}, \tag{30}$$
$$f_b(z_1, x_1) = x_1 z_1 - \sin(z_1) x_1^2 + z_1^2 + \zeta_{\text{fn}}, \tag{31}$$
$$f_c(z_1, x_1) = \cos(4\pi z_1) - \sin(z_1 x_1) + x_1 + \zeta_{\text{fn}}, \tag{32}$$

with $z_1 \in [-5, 5]$ the design parameter, $x_1^I = [-5, 5]$ the uncertain parameter and ζ_{fn} represents a random component. It is assumed that the random errors are i.i.d. random errors with $\mathbb{E}[(\zeta_{\text{fn}})] = 0$ and $\mathbb{V}[(\zeta_{\text{fn}})] = \sigma_{\text{fn}}^2$, thus σ_{fn}^2 represents the imposed homoscedastic noise variance independent of z_1 and x_1. Figure 6 illustrates the effect of the added noise to the functions f_a, f_b and f_c, which is illustrated by the red and blue areas around the mean upper- and lower-bound indicated by full red and blue lines. The proposed method is tested on these case under both known- and unknown-homogeneous noise.

5.1. Analytical functions with known homogeneous noise

In this case homogeneous variance of the GP σ_{gp}^2 is assumed a priori, which is independent of the noise term put on the analytical functions σ_{fn}^2. In the cases below the effect of different noise

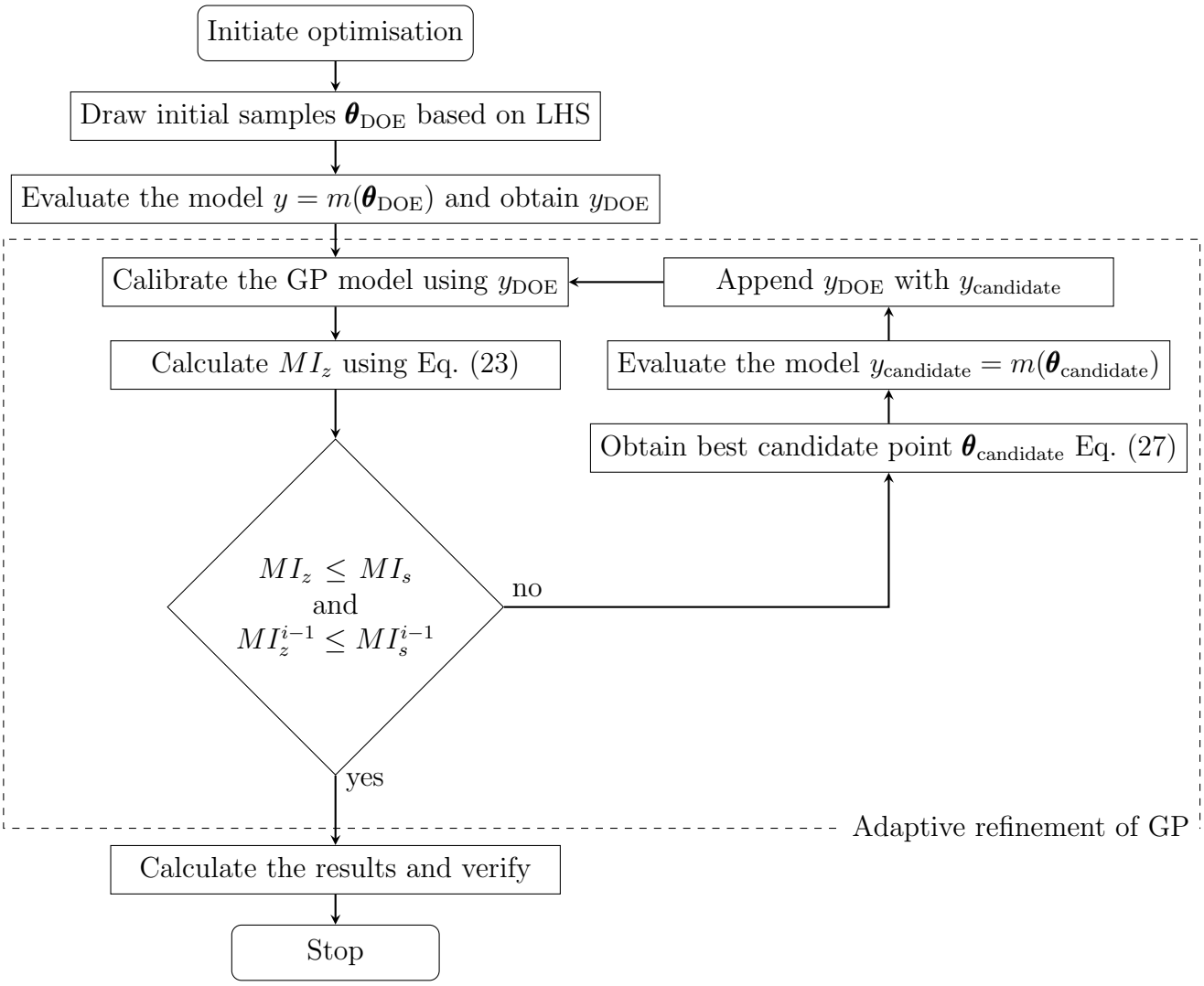

Figure 5: Flowchart of the robustness under lack-of-knowledge method for noisy functions

terms on both the analytical function σ^2_{fn} and GP σ^2_{gp} is demonstrated. The proposed approach is stochastic in nature since it depends on the noise-affected realizations of the system. Therefore, each of the cases shown in this section are repeated ten times, and the mean and envelope off all runs are shown. The first case illustrated in Figure 7 shows the effect of increasing the imposed noise variance σ^2_{fn} for function f_a Equation (30). The figure on the left shows an increasing error for an increased imposed noise variance σ^2_{fn}, indicated with the mean relative error in a blue line and the blue area showing the minimal and maximal error that was obtained for ten runs. The figure on the right shows in a similar way the total number of function evaluations, which includes the initial 20 design of experiment evaluations.

For the second case, the variance of the GP σ^2_{gp} is set at different values while the imposed noise variance is kept at $\sigma^2_{\text{fn}} = 10$, again for function f_a Equation (30). The results are shown in Figure 8, which is identical in setup to the previous case. It is clear that with an increase of GP variance σ^2_{GP} the number of function evaluations n_{total} increases, while the relative error decreases slightly. In addition, the results indicate that when using an GP with almost no variance, i.e., interpolating GP, on a function with noise the obtained results are subjected to higher errors, if convergence is even possible. In the opposite case, where the GP is set with a high variance, one

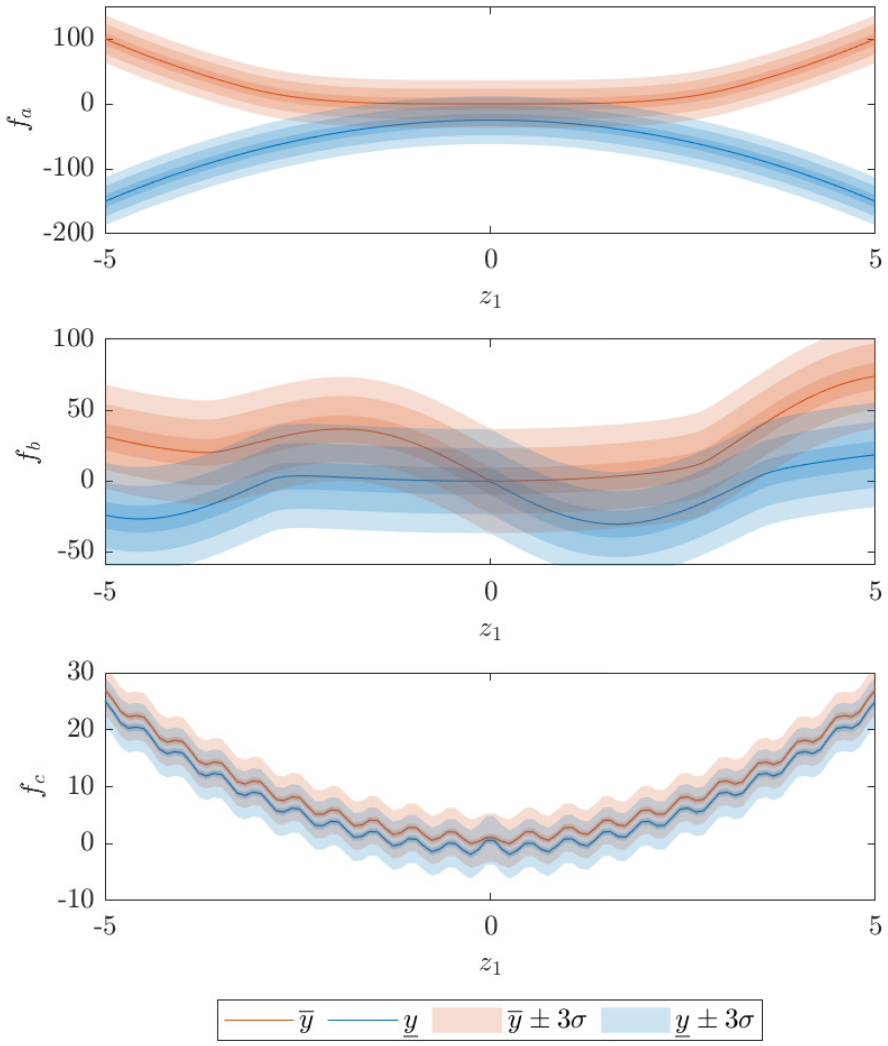

Figure 6: Illustration of the effect of noise on function f_a,f_b,f_c; The red and blue lines indicate the upper- and lower-bound of the functions, while the red and blue areas represent the effect of noise on the upper- and lower-bound illustrated by the 3σ CI for three noise variances $\sigma^2_{\text{fn}} = 10, 60, 150$ for f_a,f_b and $\sigma^2_{\text{fn}} = 0.01, 0.05, 2$ for f_a

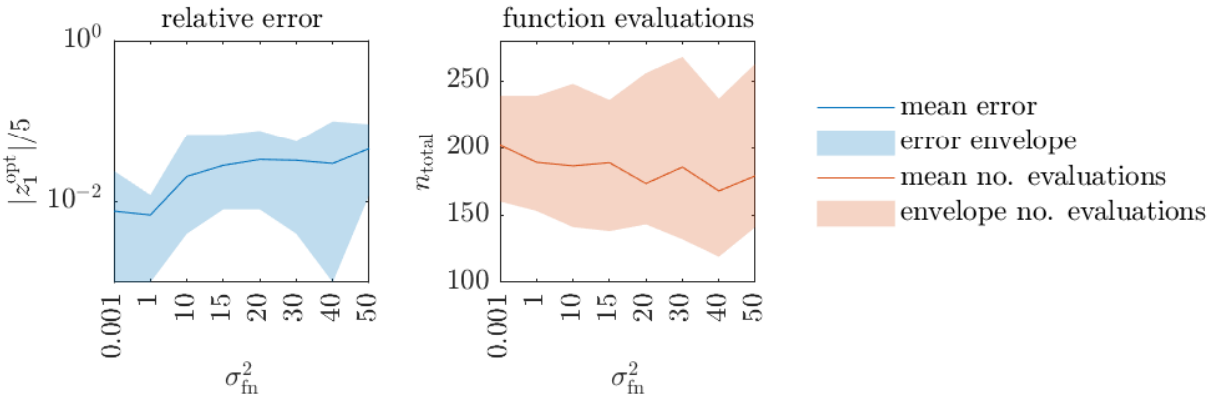

Figure 7: The mean and envelope of ten runs for function f_a Equation (30) with an increased imposed noise variance σ^2_{fn}

needs a large number of evaluations to reach the desired accuracy.

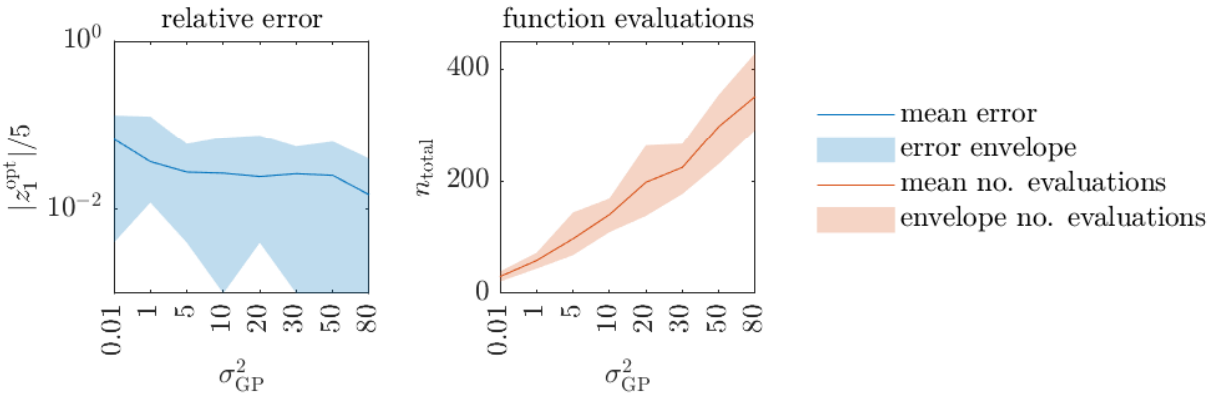

Figure 8: The mean and envelope of ten runs for function f_a Equation (30) with a increased set GP variance σ^2_{GP}

For the third analytical case, function f_b Equation (31) is used following a similar approach. The results of this case are shown in Figure 9 where the noise imposed on the function is increased and the variance of the GP model is kept at $\sigma^2_{\text{fn}} = 10$. These results are a bit different then expected from the previous results, as in this case, the number of function evaluations decreases with an increase of imposed noise variance σ^2_{fn}. This decreasing trend has not been observed in the previous case in Figure 7. The main reason can be found in the underlying function. Where f_a has a smooth transition to a global minimum, f_b experiences two local minima and a global minimum at $z_1 = 0$, which can also been seen in Figure 6. One possible interpretation of that result is that the high noise levels mask the local minima of function f_b.

Finally, Figure 10 shows the results of function f_c Equation (32) where in a similar way the imposed noise variance σ^2_{fn} is increased. It is already clear from the results on the left that the error term is very large, indicating that the obtained results are not satisfactory. Note that f_c in Figure 7 shows the complexity with multiple local minima covered by noise with very small difference between the lower and upper bound. In this case the function posses to much of a challenge when noisy responses are considered.

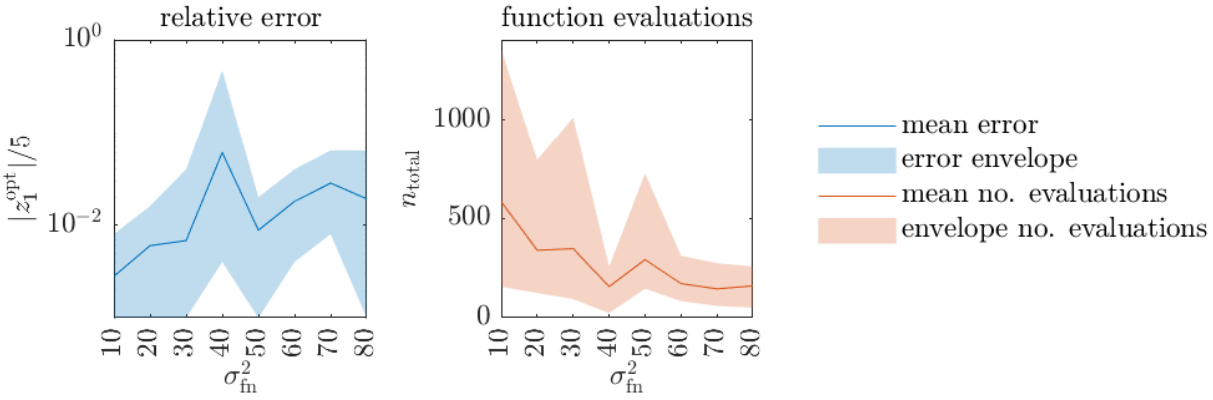

Figure 9: The mean and envelope of ten runs for function f_b Equation (31) with a increased imposed noise variance σ^2_{fn}

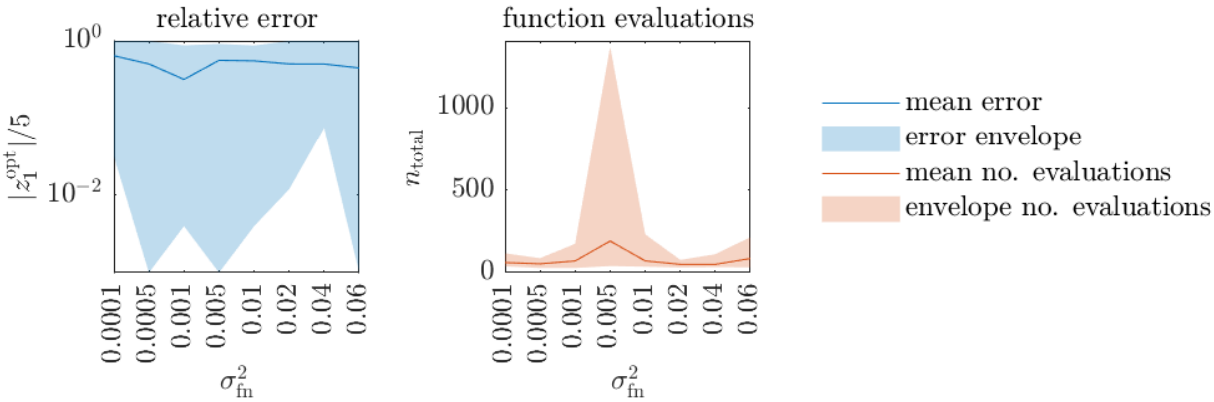

Figure 10: The mean and envelope of ten runs for function f_c Equation (32) with increasing imposed noise variance σ^2_{fn}

5.2. Analytical functions with unknown homogeneous noise

This section focuses on cases with unknown noise variance σ^2_{gp}. For such cases the possibility of learning/ estimating the noise parameter from the observations is investigated. The calibration of the noise variance is part of the GP calibration using a maximum-likelihood approach and without any further changes in the presented method. It is expected that the variance can only be estimated correctly based of a sufficently large number of observations. Hence, in the first case shown in Figure 11 investigates this effect by increasing the points in $\mathbf{y}_{\text{DOE}}$ for function f_a. The results indicate that for a very low number of initial samples the results are not always satisfactory, which can be seen by the high error, and the difference between the calibrated and imposed noise. In these cases the optimisation strategy converges to fast, as with the limited number of evaluations no correct estimates of the imposed noise variance σ^2_{fn} are made. This behavior changes when 16 or more initial samples are used with a reduction in the variance of the algorithm output and error suggesting a correct convergence. Furthermore, it can be noticed that a high number of initial samples not directly results in a high amount of total function evaluations. Nevertheless, it should be noted that there will be a penalty when larger amounts of initial samples are being used as the DOE will not place all points at optimal locations, increasing the amount of function evaluations that do not contribute to the final goal of the optimisation.

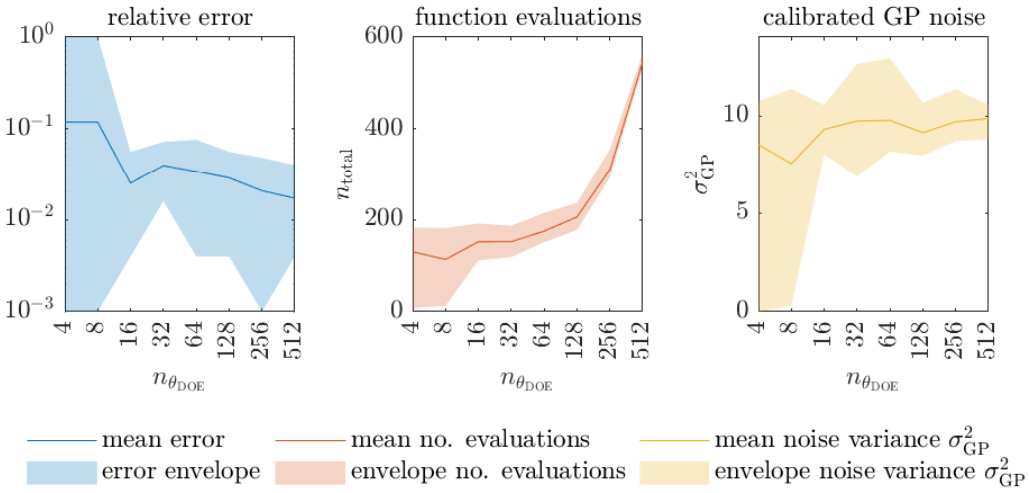

Figure 11: Results of function f_a Equation (30) with on the left the relative error, in the middle the total number of function evaluations including the initial evaluations, and on the right the calibrated noise of the GP in a full yellow line.

In addition to the effect of the initial samples, the stopping criterion depends on both the GP variance and the error parameter ϵ, which controls when to stop the adaptive refinement. Therefore, based on the previous results this effect is checked using 20 initial samples, while varying ϵ. The results are shown in Figure 12, with again the same structure as before. The decrease of epsilon and associated decrease of the allowed error are shown in the top left figure. Here, it is seen that the effect on the precision of changing ϵ is relative low. However, it can be seen that for very low values the number of evaluations starts increasing. It is also noted that for very low values of ϵ convergence becomes unlikely even with a very high number of evaluations. However, to prevent premature stopping the value of ϵ should be kept as low as possible.

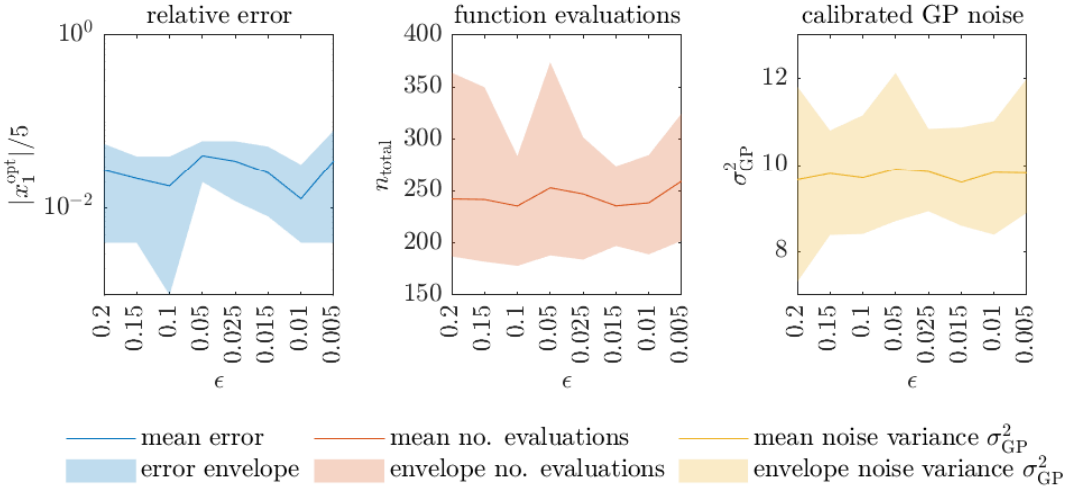

Figure 12: Results of Function f_a Equation (30) for eight cases with increasing values for Epsilon, with on the left the relative error, in the middle the total number of function evaluations including the initial evaluations, and on the right the calibrated noise of the GP in a full yellow line.

In accordance with the previous cases of function f_a the noise variance σ^2_{fn} imposed on f_b is

increased. However, this time the GP will take this increase of noise into account as it calibrates for the noise. Figure 13 shows the results for an increased imposed noise variance σ^2_{fn}. Note that these results were obtained for $\epsilon = 0.15$ and 20 initial samples for each run. The figure on the right shows the imposed noise variance σ^2_{fn} and the mean calibrated GP noise as a full line, with the area indicating the calibrated GP variances for all ten runs. It is clear from Figure 13 that the method is capable of tracking these high imposed variances without an increase of function evaluations.

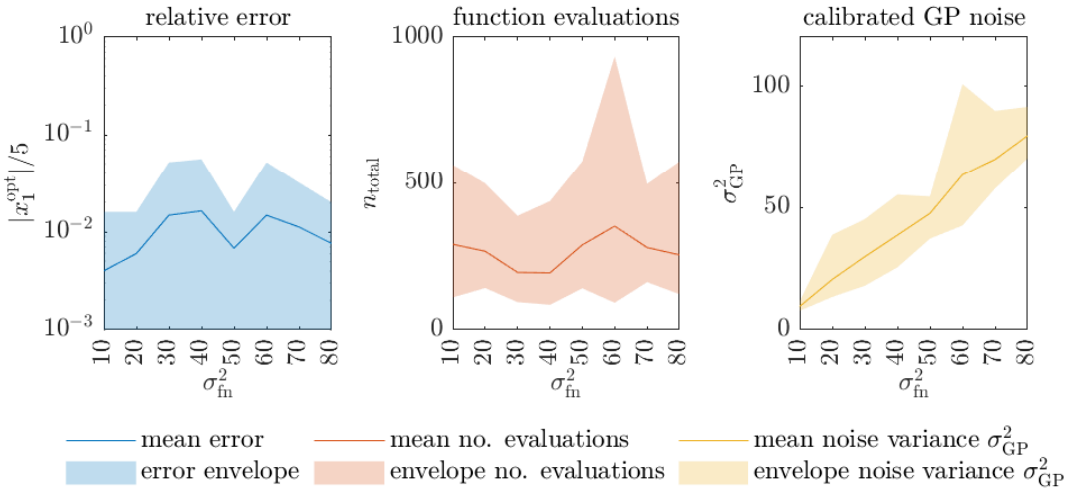

Figure 13: Results of function f_b Equation (31) for eight cases with increasing imposed noise variance σ^2_{fn}, with on the left the relative error, in the middle the total number of function evaluations including the initial evaluations, and on the right the calibrated noise of the GP in a full yellow line.

5.3. Conclusions based on the analytical functions

In this section two distinct ways of using the RULOK method for noisy functions are shown: first with a noise variance given a priori and second with unknown noise calibrated as part of the GP maximum likelihood estimation. These methods are both capable of providing satisfactory results for the analytical functions defined in this section. However, general conclusions are not easily made based on the obtained results as performance of the method heavily depends on the underlying problem. Conclusions that can be made are: (1) the number of evaluations is higher when considering a noisy function response; (2) the method is tolerant to over- and under-estimation of the actual noise variance; (3) convergence is not guaranteed in complex cases with many local minima. Furthermore, when the noise variance is estimated by the GP in the calibration step the size of the initial design of experiments should be sufficiently large. Although calibration of the noise variance is possible, better results, with less evaluations were obtained by a priori estimated noise.

6. Application to robust crashworthiness optimisation

In this section the proposed RULOK method for noisy function responses is demonstrated on a frontal crash example. Here, the output of a numerical impact simulation, as shown in Figure 14, is regarded as a noisy function response. This crashbox is a typical component that can be found in the front structure of a vehicle. The main objective of a crashbox is to dissipate a certain

Table 1: Significant parameters and their ranges as used in the numerical simulations of the crashbox

Material model properties used for the component					
initial speed	v_0	15 m/s	mass	m	600 kg
thickness plate 1	T_1	[2; 3] mm	thickness plate 2	T_2	[2; 4] mm
spotweld diameter	T_{sw}	[1; 3] mm	density of steel	ρ	7.89 kg/m^3
Young's modulus	GPA	200	Poisson ratio	ν	0.3

amount of energy during frontal impact, and to prevent further structural damage at low speed impact events. The numerical model to represent the crashbox is taken from the publicly available Toyota Yaris model, downloaded from [42], and consists out of three sheet metal parts that are held together by a number of spotwelds. The specific part numbers (PID's) are 2000137, 2000121, 2000142 and part 2000486, of which the latter is used to model the spotwelds. The setup of the numerical model, as shown in Figure 14, illustrates these parts as also two rigid surfaces, the red surface is fixated at the back of the component and the blue surface is impacting the crashbox with a prescribed kinetic energy, as shown by the arrow. The kinetic energy of the blue surface is scaled to $67,5$kJ with a mass of 600kg and an impacting speed of 15m/s, as there are two crashboxes in a full vehicle model. Other parameters that are used in this analysis can be found in Table 1. Note that T_1 refers to the thickness of the green plate in Figure 14, which has PID 2000121, and T_2 to the blue plate in the back with PID 2000142.

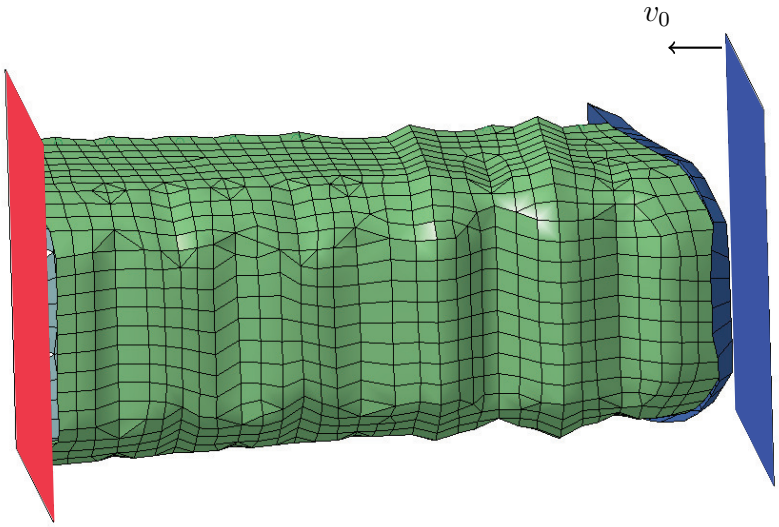

Figure 14: Finite Element Model of the crashbox with a rigid plane attached to the nodes in the back (red) and impacting plane right (blue); adapted from the Toyota Yaris model [42]

Optimisation of components for the front structure of a vehicle is quite challenging as there are multiple objectives from different development teams that should be met. For the structural requirements the mean force during impact is often regarded as a quantity of interest. Figure 15 shows a typical force-deformation curve for the crashbox with the dashed line indicating the mean force. In this case the objective is to identify the design that results in the smallest variation of the mean force for a given uncertainty. The uncertainty in the two cases below stems from a lack-of-knowledge about the weld diameter, and the thickness of the back-plate, which are modelled by an interval as described in Table 1.

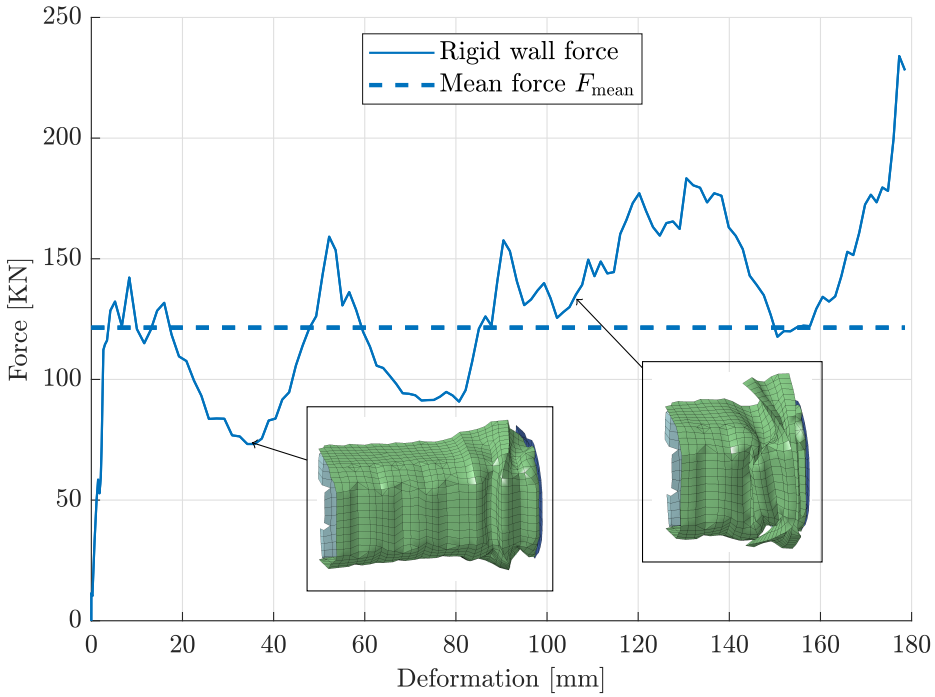

Figure 15: Typical force-displacement curve, obtained from the numerical simulation, with the mean-force F_{mean} as a dashed line; in addition, two deformed states of the crashbox are provided

6.1. Crashbox with uncertain spotweld diameter

In this section the robust optimisation is performed with as design parameter T_1 and uncertain parameter T_{sw} with the ranges as described in 1, while $T_2 = 1.8$ mm is fixed. As a reference, the existing RULOK method using an interpolating GP without noise is used to identify the robust design point. However, the original RULOK method failed to converge and was interrupted after 1500 model evaluations. The results obtained by these 1500 evaluations are plotted as the blue dots in Figure 16. When applying the RULOK method for noisy functions convergence was reached after 30 model evaluations including 20 initial evaluations. The results of this are also shown in Figure 16 with the upper- and lower bound as predicted by the GP including the 95% CI about these estimations, based on a set GP variance of $\sigma^2_{\text{GP}} = 5$ kN. The robust design point for this case was determined to be $T_1 = 2,39$ mm, which is shown by the green line. Bases on the 1500 points evaluated by the original method this optimum is clearly in the correct region.

In the previous example the variance of the GP was set at an arbitrary value with $\sigma^2_{\text{gp}} = 5$ kN, which would correspond to a coefficient of variation (COV) of about 0.045 on average within the domain. Hence, to illustrate the applicability of the method in an industrial setting the variance of the GP is determined by the mean response of the 20 initial evaluations multiplied by an assumed COV. In Figure 17 the results are shown for different assumed COV's and a mean response of the 20 initial evaluations of 110 kN. The top figure shows the robust design point for each of the cases with the blue line indicating the mean of the ten evaluations and the blue area the envelope. On the bottom figure the number of evaluations is shown in red, with the line indicating the mean number of evaluations and the area covering all obtained results. The obtained results indicate that the method is not very sensitive to the assumed GP noise variance, and that even with COV = 1 correct results are obtained. However, for both very low and high COV's the number of evaluations start increasing and sometimes wrong optima are identified, while the mean predicted

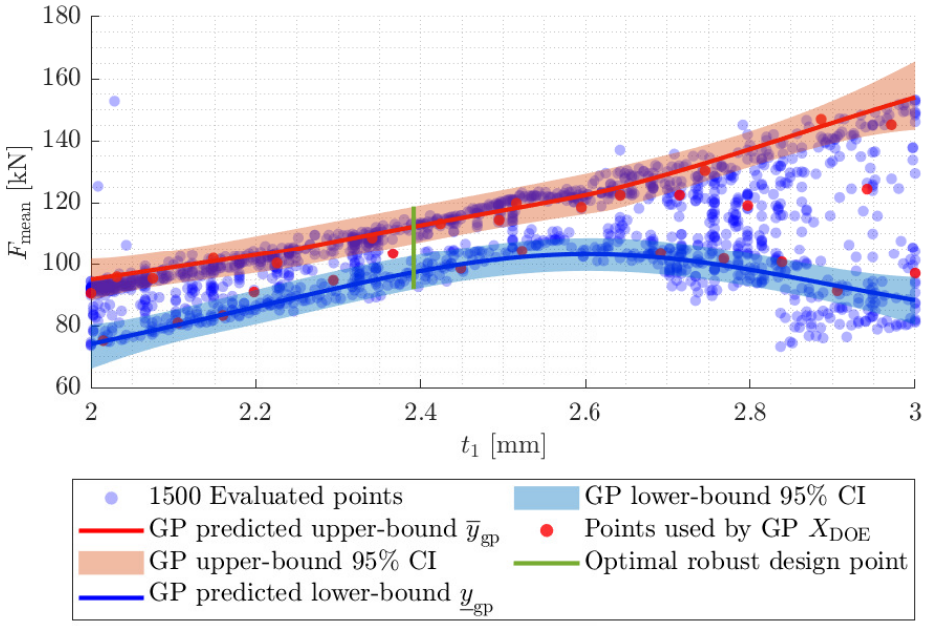

Figure 16: Evaluations by the ROLUK method without noise kernel, in blue vs. the evaluations, in red, and prediction of the method with a noise kernel

optimum is always in the correct range.

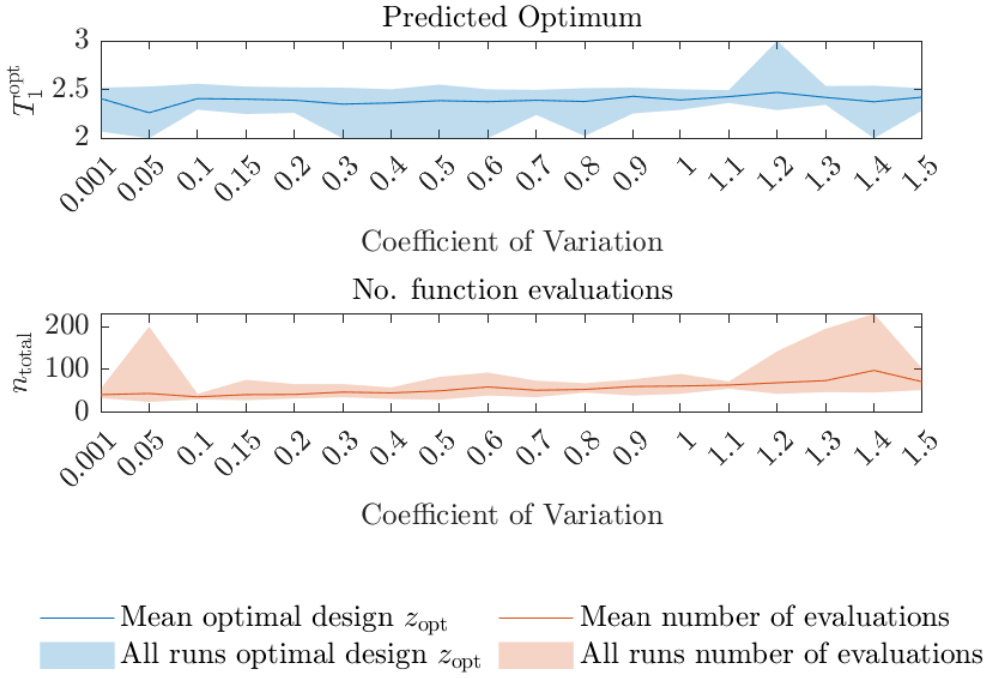

Figure 17: Obtained results for the crashbox with the GP noise depending on the COV of the initial 20 evaluations

6.2. Crashbox with uncertain spotwelds and plate thickness

To demonstrate the proposed method on a case with multiple uncertain parameters both the spotweld diameter T_{sw} and the thickness of the back-plate T_2 are regarded uncertain, following the intervals as listed in Table 1. By introducing an additional uncertain parameter the location of the robust design point has changed as well. Therefore, a reference is created based on 1000 Latin-Hyper-Cube (LHS) samples, before initiating the optimisation by ROLUK. The obtained results are shown in Figure 18 with the LHS sampels in blue, the GP predicted upper- and lower-bound in red and blue lines, respectively. The evaluations used to calibrate the GP are shown in red and the predicted optimum is highlighted in green. It is clear that the the predicted optimum has shifted towards the lower bound of T_1. The results shown here are obtained for a set variance of $\sigma^2_{\text{gp}} = 5$ kN and convergence was reached after only 63 iterations. The obtained optimum $T_1 = 2.03$ mm is indicated in green, seems correct based on the LHS samples.

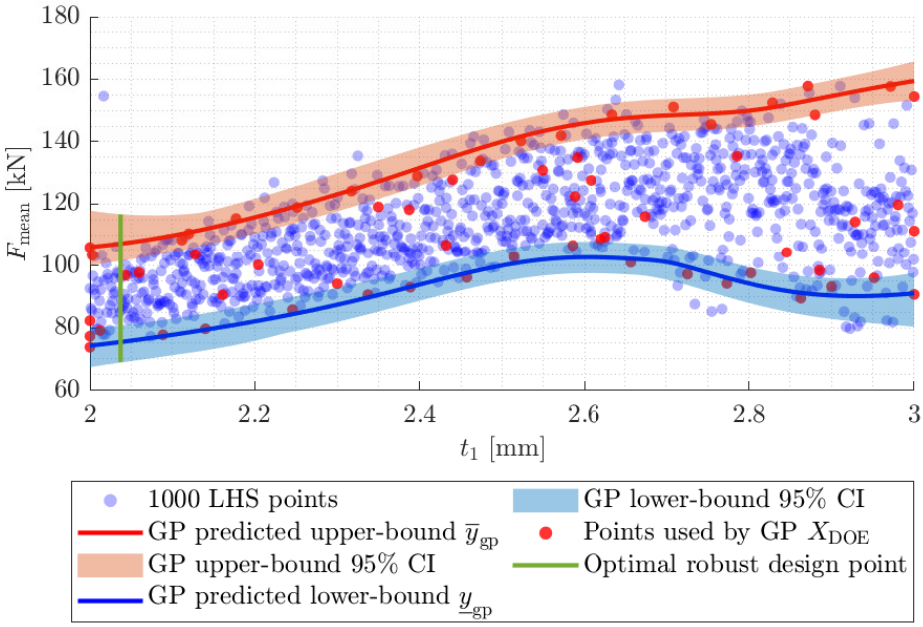

Figure 18: 1000 LHS samples of the numerical model vs. the GP prediction of the upper- and lower bound based on only 62 function evaluations

7. Discussion

The obtained results are very promising, especially those for the crashbox case, which demonstrate the added value of this method for the use in non-linear explicit numerical codes. However, as demonstrated on the analytical functions the results are not always satisfactory as seen for f_c, where often local minima were obtained for. It should be mentioned here that the analytical function f_c presents an extremely difficult problem, which as seen in [5] where a Genetic Algorithm (GA) needed 2760857 function evaluations to find the robust design point of function f_c. By imposing i.i.d. random noise on this already complex function the complexity rises further, which poses a real challenge for most commonly used optimisation strategies. The main added value of this method is shown in Figure 16 and Figure 18 where the method arguably shows some kind of *ignorance* towards bifurcations or *numerical inadequacies*. Furthermore, in crash analysis finding

the exact optimum is extremely challenging and proving that one found the global optimum is even more so. Therefore, the obtained optimum is certainly not *optimal* in the mathematical sense. However, based on the very limited information about the obtained highly non-linear response that is available, a good estimate is made towards the location of the robust design point, which is already a large improvement and provides guidance for further developments. Note here that in previous works of the authors optimisation of similar crash cases took about 160 as a minimum, until more then 3000 function evaluations for a single design [19, 43]. Reducing this to only about 50 evaluations for a range of designs is a huge improvement in terms of efficiency.

The results obtained by calibration of the GP noise variance showed that the method could be used for unknown homoscedatic noise variance. However, this comes at the cost of increased function evaluations, starting with a larger initial set of samples. It should be mentioned that the authors attempted to calibrate the GP noise variance for the crashbox example. However, after a large number of function evaluations convergence was deemed unlikely. The problem here is that the signal is contaminated with a combination of *numerical errors* and *numerical inadequacies*, which are challenging to differentiate using only a limited number of evaluations. However, it was demonstrated on the analytical function that the calibrated GP noise variance can be tracked well for different noise variances. Arguably the *numerical inadequacies* do not follow the Gaussian noise assumptions, which is followed in the analytical cases.

The results in this work are based on the GP model as implemented in UQlab [44] for all case studies. However, using the stopping criterion proposed in this paper the method is applicable to all implementations of Gaussian Processes. This was not the case before as multiple implementations always use a small GP noise variance, called *nugget*, for numerical stability [45], which can be taken into account as Gaussian noise. Finally, it should be noted that just as the original RULOK method a structured grid is used. Hence, the computational cost to evaluate all points on this grid increases exponential in d-dimensions $\mathcal{O}(n^{-d})$ for a full grid. Therefore, in high dimensional cases this becomes a bottleneck without sacrificing the resolution of the grid, and one should consider the possible dependency of the solution to the discretisation of the grid.

8. Conclusion

In this paper an extension to the robustness under lack-of-knowledge method is proposed, focusing on function responses that are contaminated by i.i.d. Gaussian noise. A learning function with a new stopping criterion is proposed capable of taking *homoscedastic* noise into account. The applicability of the method is demonstrated on a set of analytical cases. Furthermore, the proposed method is demonstrated on a highly non-linear crashworthiness case, which arguably contains a certain amount of Gaussian noise on the response. The results of this case show that the proposed method is capable to identify a robust design point, with fewer model evaluations than what would be expected from a general optimisation algorithm.

Acknowledgements

The authors gratefully acknowledge the support of the Research Foundation Flanders (FWO) under grant 1SA3919N (C. van Mierlo) and GOC2218N (A. Persoons). In addition, we acknowledge the European Union's Horizon 2020 Research and Innovation program GRAYDIENT under grant agreement n° 955393.

References

[1] G. Taguchi, Quality engineering (taguchi methods) for the development of electronic circuit technology, IEEE transactions on reliability 44 (2) (1995) 225–229.

[2] G. Taguchi, Performance analysis design, International journal of production research 16 (6) (1978) 521–530.

[3] M. Beer, M. Liebscher, Designing robust structures–a nonlinear simulation based approach, Computers & Structures 86 (10) (2008) 1102–1122.

[4] A. T. Beck, W. J. Gomes, F. A. Bazán, On the robustness of structural risk optimization with respect to epistemic uncertainties, International Journal for Uncertainty Quantification 2 (1) (2012).

[5] C. van Mierlo, A. Persoons, M. G. Faes, D. Moens, Robust design optimisation under lack-of-knowledge uncertainty, Computers & Structures 275 (2023) 106910. doi:https://doi.org/10.1016/j.compstruc.2022.106910.
URL https://www.sciencedirect.com/science/article/pii/S0045794922001705

[6] C. Zang, M. Friswell, J. Mottershead, A review of robust optimal design and its application in dynamics, Computers & structures 83 (4-5) (2005) 315–326.

[7] N. Hu, B. Duan, An efficient robust optimization method with random and interval uncertainties, Structural and Multidisciplinary Optimization 58 (1) (2018) 229–243.

[8] I. P. Mitseas, I. Kougioumtzoglou, M. Beer, E. Patelli, J. Mottershead, Robust design optimization of structural systems under evolutionary stochastic seismic excitation, 2014, pp. 215–224. doi:10.1061/9780784413609.022.

[9] G. Stefanou, The stochastic finite element method: Past, present and future, Computer Methods in Applied Mechanics and Engineering 198 (9-12) (2009) 1031–1051. doi:10.1016/j.cma.2008.11.007.

[10] M. Faes, M. Broggi, E. Patelli, Y. Govers, J. Mottershead, M. Beer, D. Moens, A multivariate interval approach for inverse uncertainty quantification with limited experimental data, Mechanical Systems and Signal Processing 118 (2019) 534 – 548.

[11] M. Hanss, Applied fuzzy arithmetic, Springer, 2005.

[12] Y. Ben-Haim, Info-gap decision theory: decisions under severe uncertainty, Elsevier, 2006.

[13] M. G. Faes, M. Daub, S. Marelli, E. Patelli, M. Beer, Engineering analysis with probability boxes: a review on computational methods, Structural Safety 93 (2021) 102092.

[14] M. Beer, S. Ferson, V. Kreinovich, Imprecise probabilities in engineering analyses, Mechanical systems and signal processing 37 (1-2) (2013) 4–29.

[15] R. Teixeira, M. Nogal, A. O'Connor, Adaptive approaches in metamodel-based reliability analysis: A review, Structural Safety 89 (2021) 102019.

[16] C. Dang, P. Wei, M. G. Faes, M. A. Valdebenito, M. Beer, Parallel adaptive bayesian quadrature for rare event estimation, Reliability Engineering & System Safety (2022) 108621.

[17] F. Duddeck, Multidisciplinary optimization of car bodies, Structural and Multidisciplinary Optimization 35 (4) (2008) 375–389.

[18] J. P. Kleijnen, Kriging metamodeling in simulation: A review, European journal of operational research 192 (3) (2009) 707–716.

[19] C. van Mierlo, L. Burmberger, M. Daub, F. Duddeck, M. G. Faes, D. Moens, Interval methods for lack-of-knowledge uncertainty in crash analysis, Mechanical Systems and Signal Processing 168 (2022) 108574.

[20] N. Andricevic, F. Duddeck, S. Hiermaier, A novel approach for the assessment of robustness of vehicle structures under crash, International Journal of Crashworthiness 21 (2) (2016) 89–103.

[21] K.-J. Bathe, Finite element procedures, Klaus-Jurgen Bathe, 2006.

[22] K. Bathe, E. Wilson, Stability and accuracy analysis of direct integration methods, Earthquake Engineering & Structural Dynamics 1 (3) (1972) 283–291.

[23] V. V. Toropov, U. Schramm, A. Sahai, R. D. Jones, T. Zeguer, Design optimization and stochastic analysis based on the moving least squares method, 6th World Congresses of Structural and Multidisciplinary Optimization (2005).

[24] D. G. Krige, A statistical approach to some basic mine valuation problems on the witwatersrand, Journal of the Southern African Institute of Mining and Metallurgy 52 (6) (1951) 119–139.

[25] J. Sacks, W. J. Welch, T. J. Mitchell, H. P. Wynn, Design and analysis of computer experiments, Statistical science 4 (4) (1989) 409–423.

[26] T. W. Simpson, T. M. Mauery, J. J. Korte, F. Mistree, Kriging models for global approximation in simulation-based multidisciplinary design optimization, AIAA journal 39 (12) (2001) 2233–2241.

[27] D. R. Jones, M. Schonlau, W. J. Welch, Efficient global optimization of expensive black-box functions, Journal of Global optimization 13 (4) (1998) 455–492.

[28] V. Picheny, D. Ginsbourger, Noisy kriging-based optimization methods: a unified implementation within the diceoptim package, Computational Statistics & Data Analysis 71 (2014) 1035–1053.

[29] J. P. Kleijnen, Regression and kriging metamodels with their experimental designs in simulation: a review, European Journal of Operational Research 256 (1) (2017) 1–16.

[30] R. E. Moore, Interval analysis, Vol. 4, Prentice-Hall Englewood Cliffs, 1966.

[31] M. Faes, D. Moens, Recent trends in the modeling and quantification of non-probabilistic uncertainty, Archives of Computational Methods in Engineering 27 (2019) 633–671.

[32] M. De Munck, D. Moens, W. Desmet, D. Vandepitte, An efficient response surface based optimisation method for non-deterministic harmonic and transient dynamic analysis, Computer Modeling in Engineering & Sciences 47 (2) (2009) 119–166.

[33] M. Faes, J. Sadeghi, M. Broggi, M. De Angelis, E. Patelli, M. Beer, D. Moens, On the robust estimation of small failure probabilities for strong nonlinear models, ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg 5 (4) (2019).

[34] L. Bogaerts, M. Faes, D. Moens, A fast inverse approach for the quantification of set-theoretical uncertainty, in: 2019 IEEE Symposium Series on Computational Intelligence (SSCI), IEEE, 2019, pp. 768–775.

[35] L. G. Crespo, D. P. Giesy, S. P. Kenny, Interval predictor models with a formal characterization of uncertainty and reliability, in: 53rd IEEE conference on decision and control, IEEE, 2014, pp. 5991–5996.

[36] Z. Qiu, I. Elishakoff, Antioptimization of structures with large uncertain-but-non-random parameters via interval analysis, Computer methods in applied mechanics and engineering 152 (3-4) (1998) 361–372.

[37] D. Moens, M. Hanss, Non-probabilistic finite element analysis for parametric uncertainty treatment in applied mechanics: Recent advances, Finite Elements in Analysis & Design 47 (1) (2011) 4–16.

[38] C. Lataniotis, D. Wicaksono, S. Marelli, B. Sudret, UQLab user manual – Kriging (Gaussian process modeling), Tech. rep., Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich,Switzerland, report# UQLab-V1.4-105 (2021).

[39] C. E. Rasmussen, Gaussian processes in machine learning, in: Summer school on machine learning, Springer, 2003, pp. 63–71.

[40] T. J. Santner, B. J. Williams, W. I. Notz, B. J. Williams, The design and analysis of computer experiments, Vol. 1, Springer, 2003.

[41] M. A. Alvarez, L. Rosasco, N. D. Lawrence, Kernels for vector-valued functions: A review, arXiv preprint arXiv:1106.6251 (2011).

[42] National highway traffic safety administration (nhtsa), crash simulation vehicle models.
URL `https://www.nhtsa.gov/crash-simulation-vehicle-models`

[43] R. Barzanooni, C. van Mierlo, M. Pabst, C. Boegle, M. Faes, D. Moens, F. Duddeck, Evaluation of uncertain boundary conditions for analysis of structural components with respect to crashworthiness, in: Online proceedings ISMA2022-USD2022, 2022.

[44] S. Marelli, B. Sudret, UQLab: A Framework for Uncertainty Quantification in Matlab, pp. 2554–2563. arXiv:https://ascelibrary.org/doi/pdf/10.1061/9780784413609.257, doi:10.1061/9780784413609.257.
URL `https://ascelibrary.org/doi/abs/10.1061/9780784413609.257`

[45] Matlab and statistics and machine learning toolbox release 2020b.
URL `https://nl.mathworks.com/help/stats/fitrgp.html`