KU LEUVEN

ARENBERG DOCTORAL SCHOOL
Faculty of Engineering Technology

Design for crashworthiness under epistemic uncertainty

Conradus van Mierlo

Supervisors:
Prof. dr. ir. D. Moens
Prof. dr. M. Faes
(TU Dortmund)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Technology (PhD)

March 2023

Design for crashworthiness under epistemic uncertainty

Conradus VAN MIERLO

Examination committee:
Prof. dr. ir. P. Leroux , chair
Prof. dr. ir. D. Moens, supervisor
Prof. dr. M. Faes, supervisor
(TU Dortmund)
Prof. dr. ir. D. Vandepitte
Prof. dr. ir. J. Ivens
Dr. P. Lava (MatchID)
Prof. dr.-ing. habil. F. Duddeck
(Technische Universität München)
Prof. dr. A. Sofi
(Mediterranean University of Reggio Calabria)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Technology (PhD)

March 2023

Uitgegeven in eigen beheer, Conradus van Mierlo, Jan Pieter de Nayerlaan 5, B-2860 Sint-Katelijne-Waver, B-3001 Leuven (Belgium)

Preface

Since I can remember I've had a passion for technology. This passion probably started while growing up surrounded by all kinds of technical devices and machines. My grandparents had a barn filled with old mopeds and curiosities, various old cars, and an uncle with a milling machine. So, it came as no surprise that I wanted to learn how all of this worked. At Sint-Jansberg I started to learn the basic techniques of manufacturing and production. However, it wasn't until I started at Thomas More that I really got excited to learn more about the latest techniques in design and manufacturing, and I really got a taste of creating products that operate at the cutting edge of technology. During my master at KU Leuven, I was able to learn more about the concepts of creating something that can reliably perform at a high level even under various operating conditions. This concept interested me so much that it has become the cornerstone of this thesis and I hope you enjoy reading it.

I'm quite proud of myself for having completed this research, as I never imagined myself doing this during this whole journey. Of course, I could not have done it alone, and I would like to use this preface to express my gratitude to several key people that have been there along the road showing me the way, or who were just there to listen.

First, I would like to express my gratitude towards David Moens. David without you I would not have been able to conquer this mountain. Thank you for being a source of inspiration, subtly saying the right things at the right moment. Especially at those difficult moments where the road seemed to go down a canyon. Above all, thank you for the opportunity to start this research and the many valuable lessons learned along the way. I enjoyed the journey together, thank you for this amazing experience!

Secondly, I would like to thank Matthias Faes, Matthias you are an endless well of interesting ideas. I really enjoyed the discussions we had and sometimes a bit less your always critical comments. In addition to being an academic inspiration,

we have travelled together quite a bit and made many memories along the way, thank you for these memories and for being an inspiration! Thirdly, I would like to thank professor Fabian Duddeck for the nice collaboration, valuable discussions and sharing of your extensive expertise in crash analysis, and finally the fruitful exchanges between Munich and Leuven. In addition, I would also like to thank the other members of the examination committee: professor Alba Sofi, professor Dirk Vandepitte, professor Jan Ivens, dr. Pascal Lava. Thank you for the thorough evaluation of my manuscript, the suggested improvements have increased the clarity of this thesis significantly. In addition, I would like to express my gratitude towards the doctoral committee for their constructive feedback, discussions and concerns raised during the research period to ensure that the work could be accumulated in this thesis. Finally, I would like to express my gratitude to professor Paul Leroux for chairing the sessions.

Of course, I couldn't have had such a nice time without my colleagues at Campus the Nayer. I would like to thank Lars, Augustin, Maurice, Céderic, Kurt, Robin, Konstantinos, Rafael, Bouwe, Miriam, and Damien, for the many fun, bizarre, and just non-sense things that we have done during these years. Even though it wasn't always easy to keep in touch with each other, we found a way to enjoy a nice cold beer at the end of a week working from home.

Ook wil ik mijn dankbaarheid uiten voor mijn familie en vrienden die er altijd zijn geweest wanneer ik ze nodig had. De onvoorwaardelijke steun van mijn ouders om mijn eigen pad te kiezen is een belangrijk element geweest om tot dit resultaat te komen. Maar ook mijn vrienden die er steeds waren om voor momenten van afleiding te zorgen. Ik hoop ook dat ik de kans krijg om deze mijlpaal te vieren met degen die hier niet meer zijn. Bedankt voor de kleine en grote dingen bewust of onbewust, zonder jullie was het maar een saaie reis geweest. Uiteraard mag ik de belangrijkste persoon hierin niet vergeten Eveline, bedankt dat je er altijd bent voor me en dat we deze reis samen hebben gemaakt.

> "Life's just one great journey. It's a road we travel as we go from point A to point B. What makes that journey worthwhile is the people we choose to travel with, the people we hold close as we take steps into the darkness and blindly make our way through life. They're the people who matter."
>
> —dr. Seuss

Abstract

Driven by environmental challenges, current engineering systems are increasingly designed to be more efficient. This is enabled using new materials, material combinations, increased integration of functionality, and advanced manufacturing processes. Moreover, to remain competitive in the global economy a short time-to-market and high-quality standards are required. In the case of the automotive industry, this challenge is further increased by the introduction of new fuel sources, e.g., electric, hydrogen, hybrid solutions. Hence, the development and design of these systems is increasingly more challenging. A typical solution is to breakdown the complexity of the complete system in smaller sub-problems. This approach is commonly described as Systems Engineering, which allows multiple departments to develop a certain sub-system. However, following the Systems Engineering approach, the overview of the complete system performance is lost, as parts might not have been developed yet. Therefore, system requirements are difficult to assess during development.

An example of such system is a passenger vehicle where safety requirements are defined for the complete system, i.e., vehicle, and changes made to individual components within the system could potentially impact this crash performance. Furthermore, if the crashworthiness is deemed insufficient at the final stage of development, the complete system should be redesigned, which is a very costly operation. In addition to this challenge, these crashworthiness requirements have been changing over the last decades to reflect the increase in vulnerable road users, e.g., cyclists, pedestrians, motorcyclists. Ultimately, the goal of these safety requirements is to limit the potential injuries to all involved parties, not limited to the occupants of the vehicle. Decisions that are made in an early development phase have the largest effect on the systems performance. However, due to uncertainties at various system levels and loss of the overall system during development these opportunities remain undiscovered.

To overcome these limitations in systems engineering, the work presented in this thesis introduces a set of new computational concepts to integrate uncertainty

effectively and efficiently throughout the system levels. The uncertainty at these stages is mainly caused by a lack-of-knowledge about the actual component or system layout. This information will only be available at a later or final development stage. In literature, this type of uncertainty is described as epistemic uncertainty, which means that the amount of non-determinism could potentially be reduced. The only way to reduce the non-determinism is to perform experiments or measurements that increase the available information. However, this might be impossible, as the non-deterministic parameter is not directly measurable, or the associated measurement error is too large. Furthermore, experimental data might be very sensitive to environmental changes. In the case of automotive development, impact or crash tests are quite expensive, and the measurements made in these circumstances contain usually high measurement errors. Therefore, during development a simulation-based approach is preferred, instead of experimental developments. A second reason, for the simulation-based approach, is that each test would require the vehicle to already be manufactured in a final stage, which is not possible at an early development stage and would be very expensive.

The developments presented in this work are based on the notion of creating a robustness towards these uncertainties. In other words, although there are several uncertain parameters, or uncertain interactions with other components, the component under development should be insensitive to this. This is enabled by the introduction of a robustness-based design optimisation strategy. To enable this robustness-based approach, this work introduces the following advancements: (1) an interval field framework that allows to represent non-deterministic material data, which can be obtained from a set of measurements, (2) a modelling approach that allows to model uncertainty about adjacent components, and finally, (3) a robustness-based optimisation strategy that allows to efficiently evaluate the former types of uncertainty. Each of these developments are illustrated on several cases related to automotive industry with a specific focus on crashworthiness. Moreover, it is shown that the robustness-based optimisation strategy can identify an appropriate optimum under lack-of-knowledge uncertainty, using only a very limited set of evaluations of the underlying non-linear finite element model. The latter is even challenging for the most advanced optimisation algorithms as the finite element analysis in these applications often proves to behave as a stochastic black-box function. In other words, even though the physics are well-understood the numerical solution scheme experiences numerical inadequacies during the solution of the problem. As such, the developed methods presented in this work present novel solutions to assist in the design and development of complex engineering structures.

Beknopte samenvatting

Om de huidige ecologische uitdagingen aan te kunnen worden technische systemen ontworpen om steeds efficiënter te zijn. Deze ontwikkelingen worden mogelijk gemaakt door het gebruik van nieuwe materialen, materiaal combinaties, geïntegreerde functionaliteit, en geavanceerde productieprocessen. Bovendien om concurrerend te blijven in de globale economie is een korte tijd-tot-de-markt en hoge kwaliteitseisen nodig. Deze uitdaging wordt extra lastig voor de automobielindustrie door de introductie van nieuwe brandstoffen zoals: elektrisch, waterstof, en hybride oplossingen. Hierdoor wordt de ontwikkeling en het ontwerp van voertuigen een alsmaar complexere taak. Een typische oplossing is om het complexe systeem op te splitsen in kleinere sub-problemen. Deze methode wordt ook wel Systems Engineering genoemd. Het grote voordeel is dat deze methode toelaat dat verschillende afdelingen onafhankelijk onderdelen kunnen ontwikkelen. Het grote nadeel is echter dat het algemene overzicht van het systeem verloren gaat. Hierbij kunnen verschillende onderdelen nog niet ontwikkeld zijn. Hierdoor is het lastig om systeem vereisten te controleren tijdens ontwikkeling. Het voorbeeld dat in deze thesis gebruikt wordt van een complex systeem is een passagiers voertuig waarbij de veiligheidseisen gedefinieerd worden op het complete voertuig. De crashbestendigheid is een van deze veiligheidseisen welke beïnvloed wordt door kleine veranderingen aan individuele componenten. Daarom als de crashbestendigheid onvoldoende wordt geacht aan het einde van de ontwikkeling zal het hele systeem een herontwerp moeten ondergaan, een erg kostelijke operatie. Deze veiligheidseisen worden ook met enkele regelmaat herzien om het veranderende straatbeeld in rekening te brengen. Zo is er in de afgelopen decennia een toename in het aantal kwetsbare weggebruikers zoals: motorrijders, fietsers, en wandelaars. Het ultieme doel van deze veiligheidseisen is om in het geval van een ongeluk alle betrokken partijen zo min mogelijk schade oplopen. Hiervoor is het essentieel om in een vroeg stadium van ontwikkeling de juiste beslissingen te kunnen nemen. Echter door onzekerheden op de verschillende niveaus en gebrek aan overzicht van het systeem tijdens ontwikkeling worden deze opportuniteiten vaak gemist.

Om deze beperkingen van Systems Engineering te overkomen worden in dit werk een aantal computationele concepten geïntroduceerd waarmee onzekerheid op deze verschillende niveaus effectief en efficiënt in rekening gebracht kan worden. De onzekerheid op deze niveaus komt vooral van een gebrek-aan-kennis over de onderdelen en indeling die daadwerkelijk gebruikt gaan worden. Deze informatie wordt namelijk pas duidelijk op een later stadium tijdens de ontwikkeling. In de wetenschappelijke literatuur wordt dit type onzekerheid omschreven als epistemische onzekerheid. De hoeveelheid niet-determinisme van deze onzekerheid kan dus verminderd worden door het uitvoeren van bijkomende testen en metingen die essentiële informatie toevoegen. Dit is echter niet altijd mogelijk omdat bijvoorbeeld de parameter niet direct meetbaar is of dat de meeting niet nauwkeurig genoeg uitgevoerd kan worden. Daarbij kan experimentele data erg gevoelig zijn aan kleine variaties in de omstandigheden. In het geval van voertuig ontwikkeling zijn crash testen zeer kostelijk en bevatten de metingen vaak veel ruis. Daarom gaat de voorkeur uit naar een modelgebaseerde ontwikkeling in plaats van een experimentele ontwikkeling. Een bijkomend propleem is dat het voertuig pas na de ontwikkeling volledig beschikbaar is, en hiervoor prototypes bouwen erg duur is.

De ontwikkelingen in dit werk zijn gebaseerd op het idee om een robuustheid tegen deze onzekerheden in het ontwerp te bouwen. In andere woorden, ook al zijn er onzekere parameters en interacties, het beste ontwerp heeft het liefst een zo groot mogelijke ongevoeligheid hieraan. Deze ontwikkeling wordt mogelijk gemaakt door de introductie van een robuustheid gebaseerde ontwerp optimalisatie. Om dit mogelijk te maken zijn volgende vooruitgangen gemaakt: (1) een interval veld methode waarmee niet-deterministische materiaal gegevens kunnen voorgesteld worden, welke experimenteel bekomen kunnen zijn, (2) een modelleer techniek om onzekerheid van aanliggende onderdelen in rekening te brengen, en (3) een robuustheid gebaseerde ontwerp optimalisatie waarmee de vorige punten efficiënt in rekening gebracht kunnen worden. Al deze ontwikkelingen zijn gedemonstreerd op verschillende cases gerelateerd aan de automobielindustrie met een focus op crashbestendigheid. Daarbij is aangetoond dat de robuustheid gebaseerde ontwerp optimalisatie in staat was om een voldoende robuust optimum te identificeren onder deze gebrek-aan-kennis onzekerheid. Hiervoor zijn slechts enkele evaluaties van het onderliggende niet-lineaire numerieke model nodig. Dit is zelfs voor de meest geavanceerde optimalisatie technieken een grote opgave omdat het numerieke model zich gedraagt als een stochastische “black box” functie. Dit houdt in dat zelfs als de onderliggende fysica goed gekend is de numerieke oplossing een aantal fouten doorrekend naar het uiteindelijke resultaat. Hierdoor dragen de voorgelede ontwikkelingen in dit werk bij aan het ontwikkelen en ontwerpen van complexe technische systemen.

List of Abbreviations

ATD Anthropometric Test Device.

BIFD B-spline Interval Field Decomposition.

C.O.V. Coefficient of Variance.

CDF Cumulative Distribution Function.

CI Confidence Interval.

DE Differential Evolution.

DIC Digital Image Correlation.

DOE Design Of Experiments.

DOF Degrees of Freedom.

EGO Efficient Global Optimisation.

EUI Extra Unitary Interval.

FE Finite Element (method).

FOV Field Of View.

GAMBIT Generalized Acceleration Model for Brain Injury Threshold.

GP Gaussian Process.

HIC Head Injury Criterion.

HIP Head Injury Power criterion.

HS-DIC High speed Digital Image Correlation.

IDW Inverse Distance weighting.

IGA Isogeometric analysis.

LHS Latin Hyper-Cube sampling.

LIFD Local Interval Field Decomposition.

NCAP New Car Assessment Program.

NVH Noise, Vibration and Harshness.

PDF Probability density function.

RBDO Robustness Based Design Optimisation.

RULOK Robustness under lack-of-knowledge.

SLS Selective Laser Sintering.

TO Topology Optimisation.

UQ Uncertainty Quantification.

List of Symbols

This is a list of all symbols used in the manuscript, ordered thematically. Listed in the context where the symbol was first used.

Probabilistic model parameters

$\mathbf{K}$	Covariance matrix
$\mathcal{N}$	Normal distribution
σ^2	Standard deviation
σ	Variance
$Cov[\bullet, \bullet]$	Covariance
$E[\bullet]$	Expectance

General conventions

$\mathbb{N}$	Set of natural numbers
$\mathbb{R}$	Set of real numbers
$\mathbb{R}^d$	Set of d-dimensional real vectors
$\mathbb{R}^{d \times d}$	Set of d-dimensional real matrices
$\mathbf{A}$	Matrix
$\mathbf{a}$	Vector
$\mathcal{G}$	Surrogate model
$\mathcal{M}$	Numerical model
a	Scalar

f	Analytical function
m	Analytical model
n_x	Number of uncertain variables
n_y	Number of response quantities
n_z	Number of design variables
n_θ	Number of basis functions
n_θ	Number of input parameters
x	Uncertain parameters
y	Output quantity
z	Design parameters

Possibilistic model parameters

$\Delta\bullet$	Interval radius
$\hat{\bullet}$	Interval midpoint
α^I	Interval scalars
ψ	Basis functions
$\tilde{\mathbf{y}}$	Exact solution set
$\tilde{\mathbf{y}}_s$	Approximated solution set
$\overline{\bullet}$	Upper bound
$\underline{\bullet}$	Lower bound
$\bullet^I$	Interval parameter

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In face of the environmental challenges, industry has a major responsibility to limit the ecological impact of their products during their entire lifetime. Therefore, over the past decades, products were further optimised to enhance the efficiency. This continuous improvement is enabled by the introduction of new materials, manufacturing processes, and considering the end of life potential of discarded products [143]. As such, due to changing policies, the automotive industry has experienced a revolution over the last decades [24] with the introduction of alternative fuel sources, e.g., electricity, hydrogen, and bio-fuels, while simultaneously increasing the efficiency of vehicles with more aerodynamic and lightweight designs. As a result of this revolution, modern day vehicles are increasingly complex systems, where each component is optimised to be as light as possible to reduce fuel and material consumption.

In addition to these environmental challenges faced by the automotive industry, safety is another major concern, with an estimated 1.35 million casualties per year worldwide due to accidents that involve passenger vehicles [178]. A disproportional amount of these casualties are among vulnerable road users, e.g., cyclists, motorcyclists, and pedestrians. The amount of these vulnerable road users has been increasing worldwide over the past years [117]. Presumably, a part of this can be attributed to initiatives that motivate people to consider alternative modes of transport, which are more environmental friendly. Therefore, to limit casualties, policy makers and automotive manufacturers are enforcing ever stricter safety requirements. Typically, safety is increased in two ways: active safety measures, e.g., antilock braking system (ABS), electronic stability control (ESC), adaptive cruise control, line assist, and passive safety systems that are built into the vehicle, e.g., cripple zones, crashboxes, other

energy absorbing structures. In essence, the idea is to first avoid a collision and, when a collision occurs, provide sufficient safety to all involved parties.

The design of complex engineering systems is a common way to describe the process of developing and designing a passenger vehicle. Following this framework, the total design is divided over multiple specialised departments due to complexity of the overall product, and some parts may even be outsourced to external companies. However, at the end of the design phase, the product should uphold the system requirements, which were set at the start of the design process. In the case of vehicle development, various authorities will check the compliance of the vehicle to the latest standards ranging from, e.g., emissions, safety, and Noise, Vibration and Harshness (NVH). These requirements that are defined on the complete system should be translated to design or optimisation goals for each component, in a way that when the complete system is built the envisioned performance is obtained. However, a complex engineering system, i.e., passenger vehicle, is not easily sub-dividable especially as some quantities of interest can only be assessed on the overall system, such as, the crashworthiness. The latter is illustrated in Figure 1.1 where different components of the body-in-white, i.e., the chassis, are shown in different colours. The component highlighted in the detail above the vehicle is the crashbox, which is essential for the crash performance of the vehicle. However, crash performance is not only determined by the crashbox. The type of impact, the impact energy, and the weight of the vehicle play a crucial role as well.

The focus of this work resides in the development and application of uncertainty quantification (UQ) techniques. It is easy to imagine sources of uncertainty in the case of an impact or crash analysis, as shown in Figure 1.1, which shows a passenger vehicle on the left and on the right possible crash scenarios. In each of these accident scenarios, the preferred outcome is that no injuries are sustained by both the passengers of the vehicle, and other vulnerable road users. Therefore, a large number of test cases are standardised to guide vehicle development while ensuring road safety [64, 92, 175]. Most of these tests are performed by insurance companies with test results being published as NCAP ratings. It should be noted that these standards are regularly revised to reflect changes in global or local trends in transportation, e.g., an increase in cyclists, other small personal transportation methods, and new traffic situations. The latest large change is the use of deformable barriers [120, 197] used in frontal crash scenarios to more realistically represent a car-to-car collision.

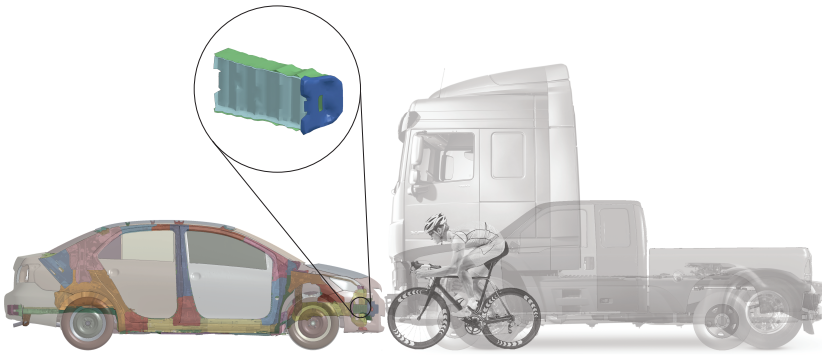


Figure 1.1: Illustration of the challenges in designing components for impact; on the left: a passenger vehicle body in white, on the right: three possible road users, e.g., cyclist, SUV and a semi-truck; on top: a detail of the crashbox (this figure is created with the models from [175])

1.1 Simulation driven crashworthiness optimisation

To work efficiently and minimise operational costs, the development of new products is usually steered based on computer simulations that assist in the design decisions that are made during development. In the case of impact or crash analysis, elaborated numerical codes are employed to provide detailed insights in the performance of each component, or the system as a whole. These numerical models are used to approximate a set of differential equations describing the time-dependent physical behaviour of the component during impact, e.g, crash. Without a doubt, the Finite Element method (FE) [15] is the most indispensable tool used in engineering practice, providing accurate predictions under given loading conditions. An additional method that has gained a lot of interest over the years, and is worth mentioning, is Isogeometric Analysis (IGA) [41], which can solve the set of differential equations without explicit discretisation of the domain. The advantage over the FE method is that changes made to the geometry can directly be analysed without re-discretisation

of the domain. The latter makes this technique interesting to be used in design optimisation [95, 146]. In this work, only the FE method will be used to perform numerical simulations. Moreover, the majority of numerical simulations discussed in this work are nonlinear simulations, meaning that the solution of the system depends on the history to reach that point. In general, there are three main types of non-linearities that are discussed in this work:

Material behaviour can only be approximated linearly in a limited region, e.g., before plastic deformation, damage and failure sets in. For example, metals are generally prone to plastic deformations when loaded beyond the yield point. On the contrary, elastomers, e.g. rubbers, behave highly non-linear under a load without plastic deformation. In every passenger vehicle, a combination of various plastics, rubbers, and metals is used.

Geometric effects often described as *large deformations* are non-linearities resulting from deformations, which are no longer correctly approximated by the small strain approximation, which is typically the case during crash analysis.

Boundary conditions The main contribution here is contact, which can mean contact with other components, contact of the component itself, or both. Especially in crash analysis, these contact conditions play an important role as multiple components are crushed or a single component is collapsing upon itself.

Furthermore, in the case of crash analysis, an explicit solution scheme is used where the solution at time t_n depends on the solution of the previous time t_{n-1} with n the time-step. The time step depends on the specific case, and typically ranges from 10^{-7}s to 10^{-9}s, and is usually automatically calculated by the numerical solver based on the Courant-Friedrichs-Lewy (CFL) condition [50]. Nevertheless a small error term is accumulated at each time-step of the solver contributing to the *numerical error* [14, 15, 59, 223]. In addition, crash analysis is well known to exhibit *numerical inadequacies*, i.e., dynamic- and numerical instabilities that can cause a small (infinitesimal) change in the input to produce a major change in the output [4, 59, 124, 148]. An example of a numerical inadequacy is shown in Figure 1.2, which shows the deformation of a structural member for repeated computations with the same input parameters [59]. The reason for these inadequacies and errors can be caused by a number of reasons, e.g., the hardware used to compute the simulation, the mesh, round-off errors. Therefore, this work stresses that due to the combined error of both *numerical inadequacies* and *numerical error* a deterministic crash simulation behaves like a stochastic simulation model, despite its deceiving deterministic nature. In other words, evaluation of the explicit numerical model returns different results for

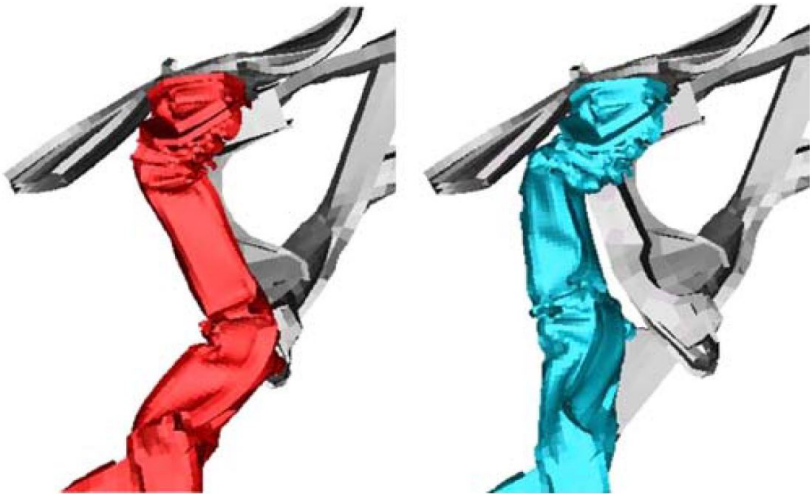


Figure 1.2: Example of a numerical inadequacy obtained by repeated computations of a structural member during high-speed frontal crash [27]

the same set of input parameters. Therefore, this work is focused on including this non-determinism in the design phase of impact critical structures.

1.2 The role of uncertainty quantification

The topic of Uncertainty Quantification (UQ) is very broad and can be applied in many ways in multiple fields of research. In general, uncertainty quantification is the practice of quantifying, assessing and reducing variations, defects, and failures, resulting from variations caused by manufacturing processes, environmental changes. Hereto, a number of tools have been developed that analysts can use to describe and work with these vaguely described or variable quantities. The most well-known tools are based on the axioms of probability theory, which describe uncertainty by means of probability distributions. Here, by assigning a probability to the occurrence of a certain event the analyst can make an informed decision about the probability that the design is sufficiently strong. Below, an example is given of how people make these decisions in their daily lives.

On a Sunday evening, you might ask yourself *'Will it rain tomorrow?, or could I take my bike to work?'*. A first solution is to follow the weather forecast during the eighth o'clock news, where the weather expectation is provided by

meteorological services. However, the models used here are quite general and usually cover a large area, e.g., province, country. Hence, there is quite some uncertainty in these predictions. Therefore, if you are not satisfied with the prediction, e.g., a 30% chance of rain in the morning, you might wait until the morning to check the forecast on your phone, with models that are usually more frequently updated, and predictions covering a shorter time frame. Therefore, a more accurate prediction of rainfall is made.

Although simplistic, the example above illustrates the essential ingredients used in UQ: input parameters, i.e., your location, a model, and the predicted rainfall as output. In engineering practice other more complex questions are being asked, e.g, how should a building be constructed to withstand an occasional earthquake? How thick should the construction of a windmill be designed to withstand a storm? The manufacturer specified the stiffness of the material to be between 200GPa and 210GPa, what is the effect hereof on our products? In an attempt to answer these questions, a more elaborate framework is needed to quantify the uncertainty. The following categories of non-determinism are often described in UQ [77]:

- **Aleatory uncertainty** which can be described as irreducible uncertainty. In other words, this is the seemingly inherent variability of a system, which appears to us as being random. Note that it depends on the scale at which one looks at the system whether this is actually inherent randomness. Modelling this type of uncertainty is best done following the axioms of probability theory, which describes the randomness by means of probabilities.

- **Epistemic uncertainty** can be described as reducible uncertainty, which means that it stems from a lack-of-knowledge, incomplete data, unknown correlations, limited sample sizes, or other vague or ambiguous sources of data. Therefore this uncertainty is reducible as one could in theory reduce the uncertainty by, e.g., collecting more data, doing more experiments, or get more accurate sensors. However, capturing this data can be extremely expensive and might not even be possible using the current state-of-the-art methods. Since this uncertainty is not random, it is mainly described by set-based approaches such as intervals, fuzzy numbers or convex sets.

- **Mixed uncertainties** refer to cases where epistemic and aleatory uncertainty are combined, e.g., one knows the distribution of an input variable but can only bound the mean and variance. This type of uncertainty is described as imprecise probabilities or by possibility theory [57], and can also be referred to as deep uncertainty, in the literature.

In the framework of Systems Engineering, the main cause for uncertainty is the lack-of-knowledge about component or sub-system developments at other departments, teams, or companies. Especially at an early design stage, where the impact of, e.g., new materials, layouts, or components is not yet well understood. Nevertheless, engineers and analysts must deal with this inherent uncertainty to design the product with the desired properties. Only after this initial design and production the actual performance can be assessed or more detailed investigations can be initiated that provide sufficient data to fully describe the naturally occurring variability. However, this is not trivial, especially when only a limited set of tests can be performed. The limitations are often the costs of these test campaigns, the time it takes to perform these tests, or it might even simply be impossible to measure the quantity of interest. Therefore the developments in this work rely on numerical models that are used to predict the performance beforehand, while accounting for the uncertainty at the current design stage.

1.3 Thesis objectives and outline

The objective of this thesis is to develop methods that allow for efficient and effective design and optimisation of impact critical components under epistemic uncertainty. In current engineering practice, designing these complex engineering systems is a challenging task. Although major automotive manufacturers have a legacy starting over a century ago, the introduction of new fuel sources, materials, and safety requirements has led to a revolution in the industry. Furthermore, in this global economy there is a general drive to shorten development times and introduce new produces faster. To enable this the industry is increasingly using advanced simulation methods, which reduce the number of physical tests that are conducted. However, the following main challenges are encountered during the design of impact critical systems:

- Data from tests is often scarce, difficult to interpret, and extremely costly to obtain. Moreover, most commonly used data acquisition techniques, e.g., load cells, accelerometers, influence the dynamics of the structure. Furthermore, crashworthiness testing of an individual component provides no guarantee towards meeting the system crashworthiness requirements. The main reason for this are the interactions between components.
- The material of the component is essential to the performance of system, especially as most energy is dissipated during plastic deformation. Therefore the material models must be accurate to correctly estimate the

amount of dissipated energy. Due to the non-linearity's this is trivial to model especially when new materials are being used.

- At an early design stage, other adjacent components might still be under development, at another department, or even outsourced to other companies. Nevertheless, in order to assess and guarantee overall system performance, these components have to be accounted for.
- Currently used state-of-the-art numerical solvers for non-linear transient dynamical analysis are known to experience numerical inadequacies and the models that are used are often quite sensitive, e.g., a small change in input parameter may result in a large difference in output

Therefore, the aim of this thesis is the development, implementation, and application of epistemic uncertainty quantification techniques to capture, propagate, and optimise the design of a complex system. As a benchmark case, a non-linear transient numerical crash model is used to demonstrate the proposed modelling techniques. Specifically three main contributions are made in this work:

1. A novel interval field technique is developed to describe epistemic uncertainty. This technique is then applied to describe non-linear material behaviour, which can be used in FE simulations.
2. A method is developed to account for epistemic uncertainty of adjacent components in an early design stage. The uncertainty of these adjacent components stems from a lack-of-knowledge about these components as they are under simultaneous development.
3. A robustness based optimisation method is introduced to optimise a component with parameters subjected to epistemic uncertainty. In the developed method numerical errors and inadequacies can be accounted for.

The work in this manuscript is summarised as follows:

- Chapter 2 gives an overview of the current state-of-the-art methods that are used in this thesis. The chapter starts with a detailed overview of epistemic uncertainty modelling techniques followed by a description of Gaussian process and robust optimisation techniques specifically used with epistemic uncertainty. This chapter provides the foundations for the following chapters.

- Chapter 3 describes an interval technique with scaled basis functions. As a result, the realisations of this interval field are kept within an envelope. Based on an optimisation approach, the interval field realisations are fitted to a set of experimentally obtained stress-strain curves. Epistemic uncertainty in material data is often overlooked in modern day crash analysis.

- Chapter 4 introduces a novel kind of interval field where the gradients at control points are controlled. Opposed to the previous method here is no need for optimisation while the control over the gradients is obtained. The method is described for two ways of obtaining the gradients, first from a set of measurements and second estimated form the data set.

- Chapter 5 is aimed at tackling a systems engineering problem, where the performance of a component changes between standalone performance to integrated performance. Therefore this work deals with assessing the performance of a single component like it is integrated in the complete system. For this the uncertainty of the adjacent components that might still be under development is taken into account. This work also demonstrates that these adjacent components have a significant influence on the performance of the component.

- Chapter 6 introduces a robustness based design optimisation approach capable of identifying a design that is robust towards sources of epistemic uncertainty, which have been described in the previous chapters. The optimisation method is based on an underlying Gaussian Process that locally captures the functional relations between input and output parameters. The method is presented on a set of analytical problems to demonstrate the applicability

- Chapter 7 builds upon the previously introduced global optimisation technique to work on stochastic functions, e.g., numerical impact simulations. In this chapter it is demonstrated that the proposed method is capable of identifying an appropriate design in case of crash analysis with interval valued uncertainty.

- Chapter 8 is dedicated to exploring the valorisation potential of the developments made in this research.

- Chapter 9 concludes the work presented in this thesis and provides an outlook of future opportunities based on the knowledge gained in this research project.

Chapter 2

Robust design of complex systems

This chapter provides an overview of the methods that are used in the developments made in the research project. In the first Sections 2.1 to 2.3 an introduction to the design of complex engineering systems is given. In the second part, a detailed description about epistemic uncertainty and robustness measures is provided in Sections 2.4 until 2.6. Finally Section 2.7 provides an overview of stochastic emulators used in this work.

2.1 Design of complex systems under uncertainty

In engineering practice, complex systems are developed and optimised on a daily basis, guided by the framework of *Systems Engineering* [201]. A central aspect in systems engineering is *System Thinking* [97, 128], where a system is considered a set of components that work together. By defining the relations between the components, one can view each component as an individual system as well. System thinking can be applied to many products in our daily lives, e.g., toothbrushes, computers, and passenger vehicles. In this thesis, the main example will be the development of a passenger vehicle. Although the basic principles of automobiles have already been established over a century ago, a revolution is happening within the modern-day automotive sector. This revolution is driven by environmental challenges that require sophisticated engineering solutions to create more efficient vehicles that are lighter, safer, and provide the necessary comfort to users. Examples of recent innovations are,

e.g., hybrid or electric vehicles, autonomous driving. In addition to the impact of these innovations on the use and infrastructure, they also have an impact on the body-in-white design of these vehicles, which is further amplified by changing road conditions, with an increased attention for vulnerable road users. Therefore, consumer agencies such as Euro-NCAP [64] or Global-NCAP [92] are increasing passive safety requirements.

To successfully design complex engineering systems, such as a passenger vehicle, the complex system is thought to be an assembly of multiple interconnected systems, e.g., the body-in-white, engine, transmission. This idea is illustrated in Figure 2.1, where the idea of systems engineering is applied to a passenger vehicle where the doors, wheels, and exhaust system are dismantled, among others. Note that each of these components can be further broken down, e.g., the door is composed out of exterior- and interior panels, a locking mechanism, a window mechanism, and hinges. All of these sub-components are again subjective to systems engineering, as they are composed out of a number of sensors, components, and actuators [128]. A particular well-known systems engineering method used in the development of complex systems is the V-model approach [102], which is shown in Figure 2.2. Note that other methods exist and a recent review can be found in [97]. The V-model approach is based on the idea of decomposing the system requirements into sub-system requirements, which are then further decomposed out of a number of components. By assembling verified components, sub-systems are built, which can then be verified and hereafter integrated into the complete system. Therefore, by allowing development of individual components, the V-model enables design decisions to be made more flexible and agile. Due to the breakdown of complexity, this method allows for faster and more straightforward development. However, there are pitfalls to this method, one of which is the loss of a general overview of the systems behaviour, i.e., each decision made has an effect on the overall system performance that is only assessed in hindsight. The crashworthiness is one of these system performances that is typically assessed in hindsight. As crash analysis requires to know the complete layout of the vehicle with all relevant components installed at the correct location with the associated weight. Hence, optimisation of the layout towards crash performance is challenging, as this would require all other departments to redesign their systems, changing the layout, influencing the crash behaviour. This design loop is usually avoided by using existing designs and knowledge of previous products to guide developments.

2.2 Component solution spaces

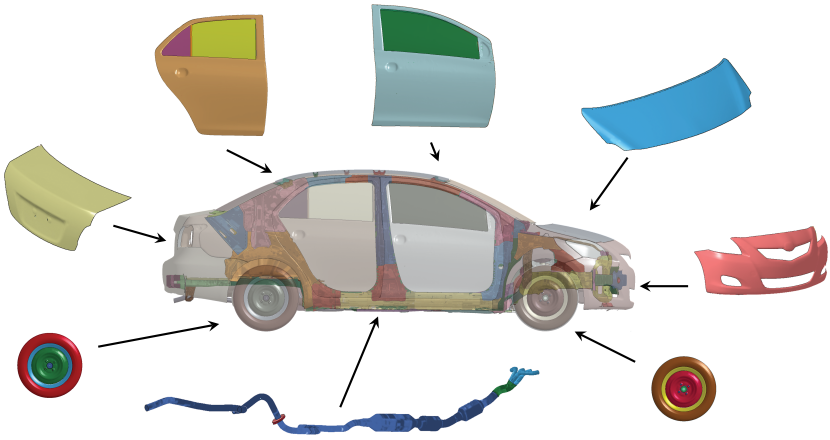


Figure 2.1: Illustration of a vehicle and some of the relevant parts seen from the outside, note that the body in white is also made out of multiple sheet-metal parts

A state-of-the-art method to break this iterative design loop is briefly described and illustrated in this section. The component solution space method was introduced [249] at BMW, as a means to break down the system requirements and provide each designer with clear design targets while guaranteeing overall system performance. The method starts from a high level energy based description of the systems behaviour where the kinetic energy is absorbed by plastic deformation of the parts. In addition, it is assumed that the parts deform in a progressive order, i.e., the first section fully deforms before the second starts deforming. A short example of the component solution spaces is provided below, as given in[249]. Figure 2.3 shows the case of a frontal impact case. The impact energy in this scenario is given by the kinetic energy of a vehicle of mass m moving at an initial speed of v_0. The total kinetic energy is then given as $\frac{1}{2}mv_0^2$. The deformation length of the first section is indicated as d_1, and for the second section as d_2, and the sections can deform up to the deformation limit d_{1c}, and d_{2c} respectively. The total deformation of the sections is given by $d_c = d_{1c} + d_{2c}$. Associated with this deformation, a force f_1 and f_2 for sections 1 and 2 is obtained. Therefore, the deformation energy in section one is given as $e_1 = f_1 d_{1c}$. Figure 2.3 illustrates these two sections and the typical frontal impact case.

The goal of the overall crash performance is to keep the deceleration below

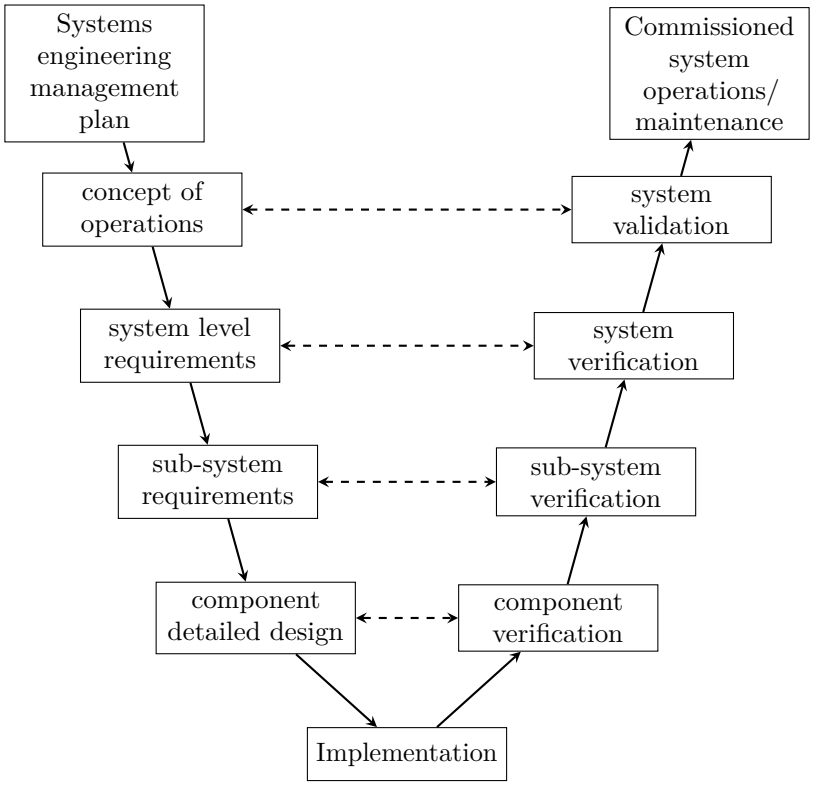


Figure 2.2: Schematic overview of the V-model, as adapted from [102]

a critical threshold a_c, while assuming that a solution exists $a_c > \frac{v_0^2}{2d_c}$. In other words, the total deformation length should be sufficient to provide a deceleration from the initial speed v_0 until 0 without exceeding the critical threshold. Additionally, a progressive order of deformation is assumed, which means that section 1 fully deforms before section 2, i.e., $f_1 < f_2$. Taking these two constraints into account the performance of the system z is given by:

$$z(f_2) = (f_2/m - a_c)/a_c \tag{2.1}$$

while the following constraints should be respected:

$$\frac{1}{2}mv_0^2 < f_1 d_{1c} + f_2 d_{2c} \tag{2.2}$$

$$f_1 < f_2 \tag{2.3}$$

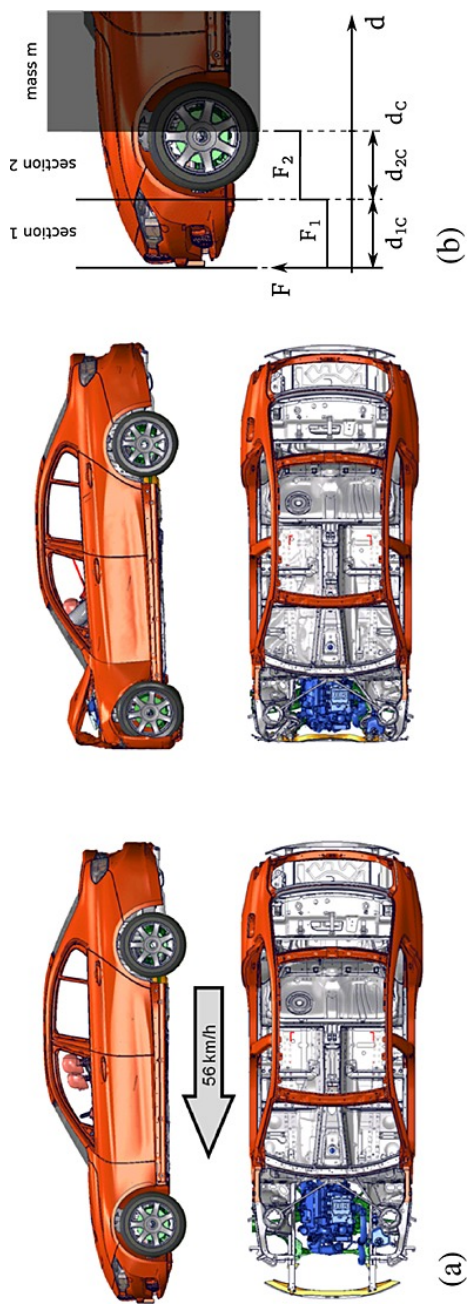


Figure 2.3: Typical NCAP [92] frontal full overlap crash test, illustrating the two sections used in the component solution space example [249]; while F_1 and F_2 refer to the force of section 1 and 2, respectively

Note that for this example the performance only depends on f_2, as f_1 is determined by the constraints. The obtained solution space is shown in Figure 2.4, where a lower and upper bound of the force f^l and f^u are given for sections 1 and 2. The optimum is found at $z(\frac{mv_0^2}{2d_c})$, which means a constant force over both sections until reaching the deformation limit $d_c = d_{1c} + d_{2c}$. Note that this optimum is located right at the design limits and is therefore not very robust, e.g., a small variation results in a violation of the constraints. Furthermore, as components 1 and 2 should preferably be developed independently one has to ensure that independent decisions still guarantee system performance. This is shown with the green area, which contains the system behaviour allowing for concurrent development of components. In essence, the component solution

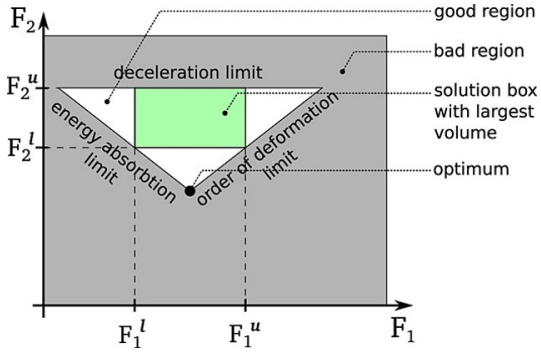


Figure 2.4: Illustration of the component solution spaces for a frontal crash example [249]

space method identifies the maximum volume hypercube, which would not violate the imposed constraints. The method has been shown to be capable of reducing costs and time during development [250] after having calibrated the solution spaces to more realistic models for frontal impact, see, e.g., [82]. Moreover, the method was further improved by making it possible to couple design decisions [46, 48], where the solution space is no longer restricted to be box-shaped. The interested reader is referred to [45] for an overview of the latest methods for creating component solution spaces. This framework already introduces performance intervals to the designers that should ensure that the component behaviour falls within these intervals. The main challenge is now to correctly assess the components performance and the associated uncertainty about this performance. Therefore the developments in this work can be used to assess this uncertainty and ensure that there is no constraint validation.

2.3 Verification and validation of complex systems

Following the approaches of systems engineering and the V-model, as shown in Figure 2.2, each step in the implementation phase should be verified against the requirements of that specific level. Seemingly straightforward, this is an extremely delicate task especially for dynamic systems where the system requirements at the highest level should be validated. These good practices and implementation guidelines are well described the ASME V&V standards [5] in which the essential question is:

"how trustworthy is the numerical model with respect to the intended use of the model".

The general workflow to validate and verify this workflow is shown in Figure 2.5. Note that this flowchart starts at a certain reality of interest, which might be at a system or component level, e.g., the plastic behaviour of high-strength steels in coupon tests, or full vehicle impact tests. Note that this depends on the domain, and that depending on these domains other standards and guidelines are provided, see, e.g., [6] for computational fluid dynamics. Furthermore, V&V highlights the central role of UQ, as this is required for both simulation and experimental activities. In this work the main discussion is about UQ in numerical simulations, while it should be pointed out that experimental work involves a significant amount of uncertainty, which stems from, e.g., noise in measurements, variations in the experimental setup, or influences of the measurement devices on the measurement, and even external influences might contaminate the measurements. The quantitative comparison illustrates that both the simulations and the experiments should be compared, including the spread in results caused by the uncertainties.

2.3.1 Validation of numerical crashworthiness simulations

Validation of a crashworthiness simulation is a tremendous challenge, which requires multiple challenges to be overcome. The main reason is the complexity of the problem; with a typical vehicle consisting of thousands of components made out of numerous materials, each of these components have a certain variability due to the manufacturing processes. Furthermore, testing of these vehicles is extremely expensive, with highly advanced setups. Nevertheless, there will be differences between the experimental setup and the numerical model of this setup, which are challenging to differentiate. Finally, it should be noted that in real life accidents one rarely or never impacts a rigid wall, e.g, accidents happen between cars, with a guardrail, or involve vulnerable road users. Despite this complexity, validation of predictions based on numerical

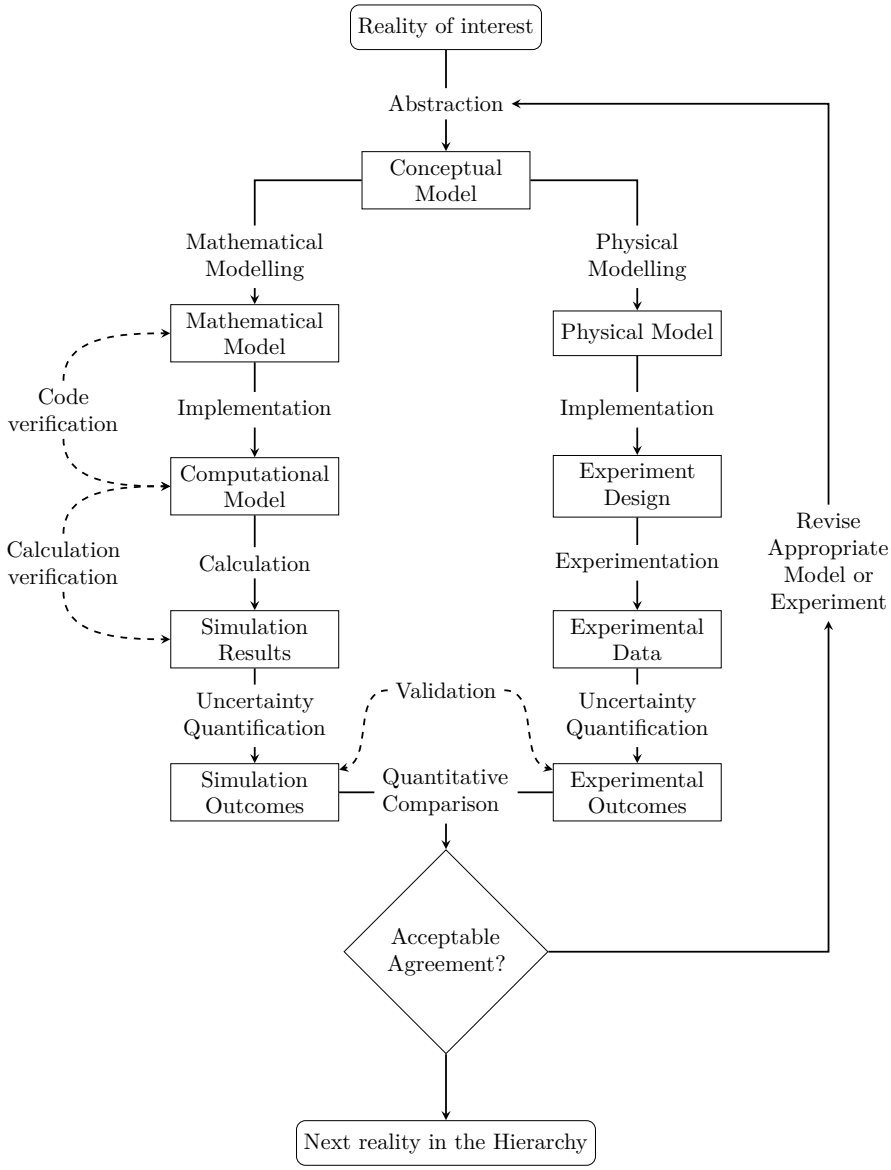


Figure 2.5: Verification and Validation activities as seen within the V&V standard [5] covering all steps within the development phase of complex systems

models is a basic requirement in any engineering context, and therefore has to

be addressed.

The crash-related simulations in this work have always been about protecting a precious good. Therefore, during a car crash, the safety of the occupants is the main concern, while at the same time considering the safety of vulnerable road users. Details of these requirements and appropriate testing procedures are defined by various authorities such as the United Nations Economic Commission for Europe (UN-ECE) [60], or based on consumer tests, e.g., the New Car Assessment Programmes (NCAP) like Euro NCAP [64] or Global NCAP [92]. These standards change quite regularly, as they should reflect a set of realistic scenarios that correspond to the change in road use, e.g., increasingly number of cyclists, the use of alternative energy sources, and novel infrastructure.

Injury criteria The main goal of crashworthiness optimisation is to keep people safe. Hence, the requirements are mostly related to bio-mechanical injury criteria, which relate multi-directional accelerations, velocities, deformations, forces, and moments to the sustained injuries during a crash scenario. In experimental procedures, these requirements are typically assessed using Anthropometric Test Devices (ATD's), commonly referred to as "*dummies*". Examples of these criteria are shown below starting with the well-known and typically used Head Injury Criterion (HIC) [226], empirically defined as:

$$\mathrm{HIC} = \left\{ (t_2 - t_1) \left[\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} a(t) dt \right]^{2.5} \right\}_{max}, \tag{2.4}$$

with $a(t)$ the resultant translational head acceleration measured at the heads center of gravity and $t_2 - t_1$ a time interval over which HIC is typically set at $t_2 - t_1 = 15$ ms as recommended by the International Organisation for Standardisation (ISO). The downside of this measure is that it does not account for rotational accelerations, impact force and has no directional dependency. Therefore, the Generalized Acceleration Model for Brain Injury Threshold (GAMBIT) [173] was introduced, which includes both translational and rotational accelerations. The criterion is defined as:

$$\mathrm{GAMBIT} = \left[\left(\frac{a(t)}{a_c} \right)^n + \left(\frac{\alpha(t)}{\alpha_c} \right)^m \right]^{\frac{1}{s}}, \tag{2.5}$$

with $a(t)$ and $\alpha(t)$ the instantaneous values of the translational and rotational acceleration, a_c and α_c are the critical values for the translational and rotational acceleration and n, m, s are empirical constants used to fit the data. A critical parameter that is missing in GAMBIT is the duration of the impact. To mitigate this, the author later proposed an extension of the HIC criterion, called the Head Injury Power criterion (HIP)[174]. This extension relates the probability

of an injury to the rate of change in kinetic energy, which is calculated for all six degrees of freedom of a head, formally defined as:

$$\text{HIP} = ma_x \int a_x dt + ma_y \int a_y dt + ma_z \int a_z dt + I_{xx}\alpha_x \int \alpha_x dt + I_{yy}\alpha_y \int \alpha_y dt + I_{zz}\alpha_z \int \alpha_z dt, \quad (2.6)$$

with m the mass, a_i and α_i the translational and rotational accelerations along their respective axis and I_{ii} the moments of inertia about the same axis, respectively. A more comprehensive overview of the head injury criteria can be found in [56], while a comparison of different predictors with real data is made in [126].

The experimental results are inherently uncertain as these ATDs should represent *'the average user'*, which typically contains high variations. Therefore, several studies are conducted on the design and validity of ATDs [87, 144] and also the loading conditions in which valid results are obtained [245]. An illustration of the use of ATDs in a numerical model is given in Figure 2.6. However, based on the details in this section and the experimental challenges in the use and development of ATDs, it should be clear that these measurements are almost impossible to directly link to the performance of a single component in the front structure of the vehicle. Furthermore, the costs of performing these tests and the time it would delay development are the main reasons that automotive industry is working towards simulation based development, and certification of vehicles.

2.3.2 Verification of crashworthiness simulations

For completeness this section briefly describes the idea behind verification steps that should be taken as illustrated in Figure 2.5. In practice, most effort is spend on code verification which should be performed by the user of the finite-element software package. The main goal is to verify that the installation and software code perform as intended. This verification is performed by comparing a number of test cases to analytical solutions. Usually the software developer provides these test cases to the user, which must execute these and compare the results. In addition, a number of manufactured solutions exist for more complex cases such as crash analysis. Where a well described case is simulated and solutions are compared to an agreed result. In this work LS-DYNA is used for the numerical simulations and the relevant verification steps can be found in the LS-DYNA documentation, which is included by the software.

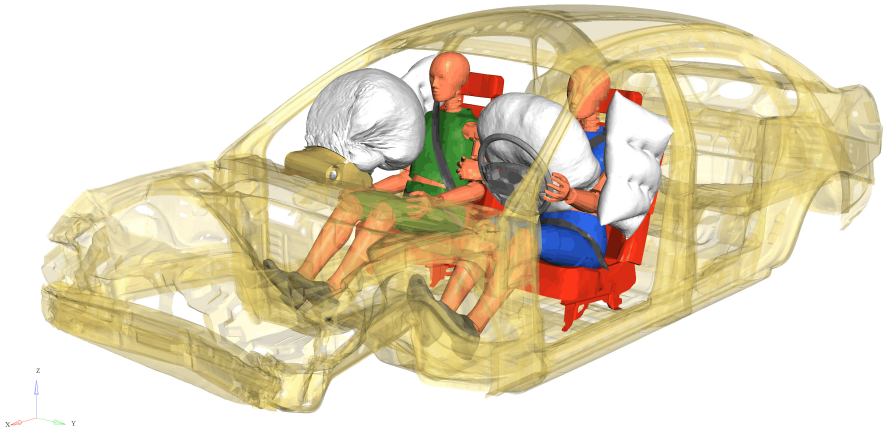


Figure 2.6: Illustration of a crash simulation with inflated airbag systems and ATD's placed in the vehicle [93]

2.4 Model definitions

In this section, the mathematical conventions and symbols are introduced that are used throughout this thesis. By convention, a vector is indicated as lower-case boldface character $\mathbf{x}$ and matrices are expressed as upper-case boldface characters $\mathbf{X}$. A design is characterised by a set of design parameters $\mathbf{z} \in \mathcal{Z} \subseteq \mathbb{R}^{n_z}$, with $\mathcal{Z}$ the set of admissible design parameters consisting of $n_z \in \mathbb{N}$ independent parameters. These parameters are *controllable* parameters, e.g., plate thickness, hole diameters, which are set or controlled by the analyst.

Opposed to the *controllable* parameters a set of *uncontrollable* uncertain parameters $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^{n_x}$, with $\mathcal{X}$ the set of admissible parameters and $n_x \in \mathbb{N}$ the number of independent uncertain variables. The type of uncertainty about these parameters is indicated by subsequent superscripts or should be clear from the context. The collection of design and uncertain parameters is indicated by $\boldsymbol{\theta} = \{z_1, \ldots, z_{n_z}, x_1, \ldots, x_{n_x}\}$ and referred to as the input parameter vector, with $n_\theta = n_z + n_x$ the total number of input parameters.

The model $m : \mathbb{R}^{n_\theta} \mapsto \mathbb{R}^{n_y}$ maps the input vector to a set of n_y system responses $\mathbf{y} \subseteq \mathbb{R}^{n_y}$. In this work, a distinction is made between different types of models, where analytical models are indicated as m, black-box models as $\mathcal{G}$, and numerical models as $\mathcal{M}$.

Furthermore, the spatial coordinates are explicitly indicated by $r \in \Omega$ with $\Omega \in \mathbb{R}^{n_d}$ the spatial domain of $n_d \in [1, 2, 3]$ physical dimensions. In the specific

case where the finite element method is used, this spatial domain Ω is often discretised to $n_k \in \mathbb{N}$ elements.

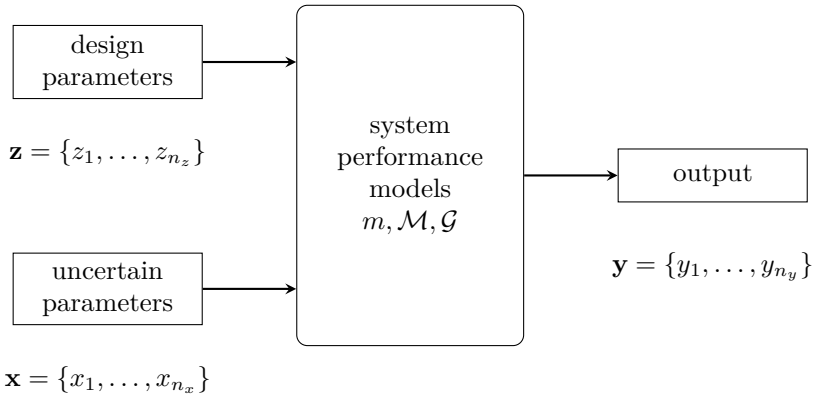


Figure 2.7: Diagram of a general model setup used in engineering

2.5 Interval techniques

The first application of interval arithmetic is traced back to Archimedes, who bounded the value of π to lie within the interval [223/71; 22/7]. Modern day interval use of interval techniques is based on Moore's interval arithmetic [163], who was one of the first to apply interval calculus to real problems. Independent contributions were presented by Warmus [230] and Sunaga [216].

An *interval* or *interval scalar* is a convex subset of the domain of real numbers $\mathbb{R}$. An interval-valued parameter is indicated using apex $I : x^I$. The interval is closed when both the upper and lower bounds are a member of the interval. The domain of closed real valued intervals is denoted as $\mathbb{IR}$. Intervals are explicitly defined as:

$$x^I = [\underline{x}, \overline{x}] = \{x \in \mathbb{R} | \underline{x} \leq x \leq \overline{x}\} \tag{2.7}$$

The centre or midpoint of an interval is defined as:

$$\hat{x} = \frac{\underline{x} + \overline{x}}{2}, \tag{2.8}$$

and the corresponding interval radius as:

$$\Delta x = \frac{\overline{x} - \underline{x}}{2}. \tag{2.9}$$

An interval vector is defined as a vector in which each element is an interval:

$$\mathbf{x}^I = \begin{Bmatrix} x_1^I \\ x_2^I \\ \vdots \\ x_n^I \end{Bmatrix} = \{\mathbf{x} \in \mathbb{R}^n | x_i \in x_i^I\}, \tag{2.10}$$

with $\mathbf{x}^I \in \mathbb{IR}^n$, the domain of closed real-valued interval vectors of size n. Analogously, the interval matrices are defined on $\mathbb{IR}^{n \times l}$. Note that all elements of the interval vector and the matrices are independent. Therefore, an n-dimensional interval vector describes an n-dimensional hypercube in n-dimensional space. The vertices of this hypercube are determined by the lower and upper bounds of the scalar interval entries in the interval vector or matrix [159].

Let m be a continuous function on $\mathbb{R}$. The function is then evaluated in the interval sense as:

$$m(\mathbf{x}^I) = \left[\inf_{\mathbf{x} \in \mathbf{x}^I} m(\mathbf{x}); \sup_{\mathbf{x} \in \mathbf{x}^I} m(\mathbf{x}) \right], \tag{2.11}$$

with inf the infimum and sup the supremum, which can be replaced by the minimum and maximum operators, if the interval is *closed*. Moreover, by assuming that the model m is monotonically increasing with input $\mathbf{x}$, the search for the infimum and supremum can be replaced by:

$$m(\mathbf{x}^I) = \left[m(\underline{\mathbf{x}}); m(\overline{\mathbf{x}}) \right], \tag{2.12}$$

while for a monotonically decreasing function this becomes:

$$m(\mathbf{x}^I) = \left[m(\overline{\mathbf{x}}); m(\underline{\mathbf{x}}) \right]. \tag{2.13}$$

These equations form the basis of the vertex method [101]. It is important to note that the vertex method assumes the function to behave monotonically towards an input $\mathbf{x}$, which is rather strict and not true in general.

For completeness, it should be noted here that an alternative approach known as "Info-gap theory" can be found in literature. It is especially useful in the domain

of decision theory, when an actuator should make a choice among multiple alternatives under deep-uncertainty [21, 22]. Applications of this method can be found in [23, 127, 130]. The formulation of info-gap theory is very general and is complementary to interval techniques.

2.5.1 Interval analysis

Let us consider a deterministic numerical model $\mathcal{M}$ that is used to approximate $\mathbf{y} \in \mathbb{R}^{n_y}$, the solution of a (set of) differential equations, through a set of (usually) real-valued function operators $\mathbf{g} = \{g_i | i = 1, \ldots, n_y\}$:

$$\mathcal{M}(\mathbf{x}) : y_i = g_i, \quad g_i : \mathbb{R}^{n_x} \mapsto \mathbb{R}, \quad i = 1, \ldots, n_y, \tag{2.14}$$

with $\mathbf{x} \in \mathcal{F} \subset \mathbb{R}^{n_x}$ the vector of the model parameters and $\mathcal{F}$ the sub-domain of feasible parameters. The main objective of the interval FE is to identify the solution set $\tilde{\mathbf{y}}$ containing the extreme realisations of $\mathbf{y}^I$, by propagation of $\mathbf{x}^I$, defined as:

$$\tilde{\mathbf{y}} = \{\mathbf{y} | \mathbf{y} = \mathcal{M}(\mathbf{x}), \mathbf{x} \in \mathbf{x}^I\}, \tag{2.15}$$

which reads as: $\tilde{\mathbf{y}}$ is the set containing all output vectors $\mathbf{y}$, obtained by performing a deterministic numerical procedure to all vectors $\mathbf{x}$, contained in $\mathbf{x}^I$. In general, this set $\tilde{\mathbf{y}}$ spans a non-convex manifold in $\mathbb{R}^{n_y}$, since the numerical model $\mathcal{M}$ provides a possibly nonlinear coupling between at least a subset of y_i. Figure 2.8 illustrates the hyper-cubic approximation $\mathbf{y}^I$ of $\tilde{\mathbf{y}}$ as a Cartesian product of two arbitrary output vectors y_i and y_j. The hatched area illustrates solutions that are not physical, whilst they are included in the solution set $\mathbf{y}^I$. Therefore, neglecting the underlying dependency can result in over-conservative output estimations.

A closed form solution to the problem of interval FE can only be obtained when there is an explicit analytical solution for $\mathbf{y} = f(\mathbf{x})$. Nevertheless, even in such a case, the dependency problem must be tackled. Note that a solution to this problem using numerical solutions is NP-hard [66, 159]. Therefore, the exact solution set $\tilde{\mathbf{y}}$ is usually approximated by the construction of a uncertain realisation set $\tilde{\mathbf{y}}_s$, which is obtained by propagating n_q deterministic realisations of the interval valued variable $\mathbf{x}^I$:

$$\tilde{\mathbf{y}}_s = \{\mathbf{y}_{sj} | \mathbf{y}_{sj} = \mathcal{M}(\mathbf{x}_j), \mathbf{x}_j \in \mathbf{x}^I; j = 1, \ldots, n_q\}, \tag{2.16}$$

with $\mathbf{y}_{sj} \in \mathbb{R}^{n_y}$ a vector containing the n_y model responses of the j^{th} deterministic model solution.

In practice, the response quantities that constitute $\mathbf{y}_{sj}$ depend on the considered model $\mathcal{M}$. Furthermore, the n_q realisations should represent the solution set

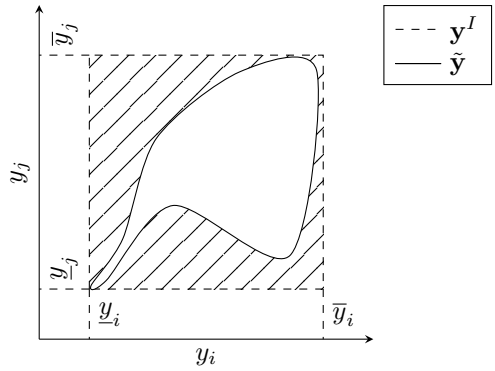


Figure 2.8: Illustration of the cross-section on two arbitrary output quantities y_i and y_j of the output vector $\mathbf{y}$ showing the hyper-cubic approximation $\mathbf{y}^I$ of the uncertain solution set $\tilde{\mathbf{y}}$, adapted from [71]

$\tilde{\mathbf{y}}$ as close as possible. Therefore, the focus of recent works is put on finding numerical procedures to find the smallest conservative convex approximation of $\tilde{\mathbf{y}}$. Until now, two main strategies have been explored in this regard: *Interval arithmetic strategies* and *Global optimisation solutions*, which are discussed in more details below. Also, the extension of these interval techniques to fuzzy analysis is briefly discussed hereafter.

Interval arithmetic strategies

A distinct or naive way of solving the interval FE is by direct application of interval arithmetic techniques as described in Moore's interval calculus [163, 164, 165] to the problem in equation (2.15). However, before elaborating further on this method, one should first introduce a number of arithmetic operations augmenting the standard arithmetic with definitions for the addition, subtraction, multiplication, and division of intervals:

$$x^I + y^I = \left[\underline{x} + \underline{y}; \overline{x} + \overline{y}\right] \tag{2.17}$$

$$x^I - y^I = \left[\underline{x} - \overline{y}; \overline{x} - \underline{y}\right] \tag{2.18}$$

$$x^I . y^I = \left[\min\left(\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\right); \max\left(\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\right)\right] \tag{2.19}$$

$$x^I / y^I = \begin{cases} x^I \left[\frac{1}{\overline{y}}; \frac{1}{\underline{y}}\right] & \text{if } 0 \notin y^I \\ \text{undefined} & \text{if } 0 \in y^I \end{cases} \tag{2.20}$$

Note that the addition and subtraction operations directly translate towards interval vectors and matrices, provided that the dimensions of the operands are compatible. One of the main shortcomings of interval arithmetic is the *dependency problem*. Consider the following example $x^I = [a, b] = [1, 2]$ is subtracted from itself in the following function $f(x) = x - x$ the result $[a - b; b - a] = [-1; 1]$, is not as expected the interval $[0; 0]$. The issue here is that interval arithmetic is unable to recognise that the interval variable x appears multiple times in $f = x - x$ [169]. Thus, this function is treated as the subtraction of two independent interval variables. In a more general setting, the *dependency problem* is expressed as:

$$x^I(y^I + z^I) \leq (x^I y^I) + (x^I z^I) \tag{2.21}$$

The *dependency problem* prohibits the direct use of interval arithmetic to solve the interval FE analysis. Therefore, multiple studies describe possible ways to mitigate the dependency problem for the specific use of interval FE, see, e.g., [63, 169, 170, 210].

Affine interval arithmetic

The use of *affine interval arithmetic* is an alternative way of keeping track of dependencies, with a more versatile extension to interval arithmetic, as introduced by Comba and Stol [40] and further applied by Manson [139]. The principal idea is to represent interval parameters by their affine form, which allows tracking their dependency in operands and sub-formulae, which helps to reduce the dependency problem. The affine form of an interval parameter is defined as follows:

$$\langle x^I \rangle = x_0 + \sum_{i=1}^{n_x} x_i \hat{\epsilon}_i^I + x_e \hat{\epsilon}_e^I, \tag{2.22}$$

with $\hat{\epsilon}_i^I \in [-1, 1]$ unknown symbolic real independent interval parameters, which allow tracking of the dependency through addition, subtraction and scalar multiplication, and $\hat{\epsilon}_e^I$ is an error term introduced to account for possible non-linear dependencies [53]. Muscolino and Sofi [172] extended the idea of a symbolic interval variable by Manson and defined an extra unitary interval EUI,

which is defined such that the following properties hold:

$$\hat{\epsilon}_i^I - \hat{\epsilon}_i^I = 0, \tag{2.23}$$

$$\hat{\epsilon}_i^I \times \hat{\epsilon}_i^I \equiv (\hat{\epsilon}_i^I)^2 = [0, 1], \tag{2.24}$$

$$\hat{\epsilon}_i^I \times \hat{\epsilon}_j^I = [-1, 1] \quad i \neq j, \tag{2.25}$$

$$x_i \hat{\epsilon}_i^I \pm y_i \hat{\epsilon}_i^I = (x_i \pm y_i)\hat{\epsilon}_i^I, \tag{2.26}$$

$$x_i \hat{\epsilon}_i^I \times y_i \hat{\epsilon}_i^I \equiv x_i y_i (\hat{\epsilon}_i^I)^2 = x_i y_i [0, 1], \tag{2.27}$$

with x_i and y_i finite numbers associated to the i^{th} EUI,$\hat{\epsilon}_i^I$. An interval is converted into its affine form by:

$$\langle x^I \rangle = \frac{1}{2}(\underline{x} + \overline{x}) + \frac{1}{2}(\overline{x} - \underline{x})\hat{\epsilon}_x^I, \tag{2.28}$$

with $\hat{\epsilon}_x^I$ symbolising the EUI corresponding to the interval variable x^I. By associating an EUI to each interval variable, the dependency trough computations is taken into account. Moreover, the conservatism introduced when assembling the global stiffness matrix $\mathbf{K}^I$ is alleviated as the interval radius of the stiffness $\mathbf{\Delta K}$ can be written as a superposition of the contribution of each interval parameter; see [210] for the proof. The applicability of the so-called *improved interval analysis via extra unitary interval* has been demonstrated in the context of interval perturbation [172], interval arithmetic computations of truss structures [171], Timoshenko beams and Euler-Bernoulli beams subjected to spatial non-determinism [203, 208], or the computation of natural frequencies of structures containing interval-valued nondeterminism [204].

Although recent developments have shown that Interval Arithmetical techniques are capable of approximating the bounds of the interval-valued responses of a numerical model within reasonable computational cost, their general application is still limited [66]. The main reason for this is that these techniques are *intrusive*, requiring dedicated FE solvers to handle interval valued uncertainty. This hinders the use of well-known and robust commercially available numerical codes, which is a major backdraw for the general, i.e., industrial, application of these methods.

Global optimisation approach

The main goal of the global optimisation approach is to actively search the input space bounded by $\mathbf{x}^I$ for the smallest hyper-cubic approximation $\mathbf{y}^I$ of

$\tilde{\mathbf{y}}$. Here, based on independent optimisations $\overline{y}_i$ and $\underline{y}_i$ are identified for each output quantity y_i of the solution interval $\mathbf{y}^I$. This optimisation problem is explicitly defined as:

$$\underline{y}_i = \min_{\mathbf{x}\in\mathbf{x}^I} g_i(\mathbf{x}) \quad i = 1, \ldots, n_y \tag{2.29}$$

$$\overline{y}_i = \max_{\mathbf{x}\in\mathbf{x}^I} g_i(\mathbf{x}) \quad i = 1, \ldots, n_y \tag{2.30}$$

where $y_i^I = [\underline{y}_i; \overline{y}_i]$ is the i^{th} output quantity, and g_i the function operator for y_i that provides the functional relation between the input and output parameters, which can be given by an analytical expression, numerical model, or appropriate surrogate models. The advantage of this approach is that the dependency of the non-deterministic parameters is implicitly taken into account by sampling the input space. Moreover, the global optimisation approach is *non-intrusive* decoupling the uncertainty in the parameters and evaluating the function operators. This allows the use of robust and high-performance commercially available codes and packages, which facilitates the application of interval analysis to large-scale problems, as shown in [85, 96, 99].

One major disadvantage is that conservatism towards the interval parameters is not guaranteed, unless the exact bounds of the, in general non-convex, goal function are identified. Also, the computational effort and the convergence of the optimisation strategy are highly problem-dependent. Although Moens and Hanss argue in [159] that the goal function often exhibits a smooth behaviour towards the uncertain parameters, this statement can not be generalised. Furthermore, in the case of crash analysis, the non-linearity of the goal function is further emphasised by the uncertain parameters. However, due to the availability of advanced *black-box* optimisation strategies, global optimisation is considered the standard approach to solve interval problems. Examples of optimisation algorithms that are applied in this context are directional search [187, 188, 189], linear programming [62], genetic algorithms [25, 34, 148, 162], efficient global optimisation [52, 118, 148], or the recently introduced Bayesian global optimisation method [43]. In addition to these *black-box* optimisation algorithms, one could also use *adjoint state* methods, where local gradient information about the goal function is obtained in the adjoint state [80, 90]. The main disadvantage of these adjoint methods is that they are *intrusive* as one would need to adapt the underlying model to obtain gradient information.

In engineering practice, the computational cost of finding an accurate solution is often very high, as the underlying models are already expensive to evaluate. This is especially true in crash analysis, where a single full vehicle model generally needs upto several days to compute. Furthermore, the number of evaluations required by the optimisation algorithm increases exponentially with

the number of uncertain parameters [66], which is described as the *curse of dimensionality*. Therefore, even models with reasonable computation time become intractable at some point. Hence, the use of surrogate models to replace the computationally expensive models is highly recommended to achieve accurate results within a reasonable time. These surrogate models represent the output domain of a deterministic model as a continuous function, calibrated on a set of observations. The main challenge is to find an appropriate surrogate model that can be calibrated using the lowest amount of observations. This is achieved by replacing the function operator g_i in Equation (2.29) by a surrogate model. Surrogate models found in the literature with application to interval propagation include: Gaussian process [52, 123], Artificial Neural networks [28, 31, 68], Interval Predictor Models [42, 74, 195], Polynomial Response Surfaces [205, 209], Chebyshev-based series expansions [136, 236], Taylor Series Expansions [88], or a dimension-wise approach [227].

On the other hand, in the case that the deterministic model response is monotonic with respect to the uncertain parameters, the vertex method, introduced by Dong and Shah [55] provides the exact result for interval problems. The method is a special case of the global optimisation approach, as it provides a first-order response surface model approximation of the model responses, which is obtained only by propagating the vertices of the hyper-cubic uncertain input set $\mathbf{x}^I$. Hence, the method only needs 2^{n_x} evaluations for n_x uncertain interval parameters. The method is widely applied in interval analysis, see, e.g., [1, 37, 148, 152, 177, 186, 187, 231, 238]. However, it should be noted that, in general, one cannot assume that the response is monotonic.

Fuzzy arithmetic

The concept of fuzzy numbers is based on *fuzzy sets*, as introduced by Zadeh [242], as an approach to represent vague linguistic information. Fuzzy sets represent an extension to the interval concept, where a membership function describes the degree to which a variable belongs to a set. This is modelled by the membership function $\mu_{\tilde{x}}(\mathbf{x})$ describing the membership of each element $x \in \mathcal{X}$ with the fundamental set $\mathcal{X} = \mathbb{R}^{n_x}$ to the fuzzy set $\tilde{x}$. The fuzzy set $\tilde{x}$ is then described by:

$$\tilde{x} = \{(x, \mu_{\tilde{x}}(x)) | x \in \mathcal{X}; \mu_{\tilde{x}}(x) \in [0;\ 1]\}, \tag{2.31}$$

with $\mu_{\tilde{x}}(x)) = 1$ meaning that x is certainly belonging to the fuzzy set $\tilde{x}$, and $\mu_{\tilde{x}}(x)) = 0$ meaning that x is definitely not a member of $\tilde{x}$. In the case that $0 < \mu_{\tilde{x}}(x)) < 1$ the membership is uncertain. Multiple membership functions can be applied, while the most commonly used are the triangular and Gaussian membership functions, depicted in Figure 2.9. When multiple parameters are modelled as fuzzy variables, a joint fuzzy membership function is defined for n_x

fuzzy input parameters as:

$$\mu_{\tilde{\mathbf{x}}}(x_1, \ldots, x_{n_x}) = \min(\mu_{\tilde{x}_1}, \ldots, \mu_{\tilde{x}_{n_x}}). \tag{2.32}$$

The resulting fuzzy membership function $\mu_{\tilde{y}}(y)$ of the output $\tilde{y}$ is identified using Zadeh's extension principle [241, 243] based on the results of a model $m()$, which can be a numerical model. The extension principle is defined as:

$$\mu_{\tilde{\mathbf{y}}}(y) = \begin{cases} \sup_y(\mu_{\tilde{\mathbf{x}}}(x_1, \ldots, x_n)) & \text{if} \quad \exists y = m(x_1, \ldots, x_n) \\ 0 & \text{otherwise} \end{cases}. \tag{2.33}$$

To obtain a solution one usually has to rely on a computationally expensive multidimensional optimisation scheme, which is not that efficient. Hence, over the last decades, a number of alternative approaches have been proposed in literature. The most comprehensive review about these methods can be found in the book by Hanss [99]. Here only the α-cut method will be briefly discussed, as this provides an extension of the interval arithmetic approach to solve numerical models with fuzzy variables. In the α-cut method the membership function $\mu_{\tilde{x}_1}(x_1)$ is subdivided into n_α equally spaced intervals. An interval for each membership level μ_{α_i} is then obtained as follows:

$$x_{i,\alpha}^I \{x_i \in \mathcal{X}_i | \mu_{\tilde{x}_i}(x_i) \geq \alpha\} \tag{2.34}$$

with the discrete values of the $n_\alpha + 1$ intervals equal to:

$$\mu_{\alpha_j} = \frac{j}{n_\alpha}, \quad j = 0, \ldots, n_\alpha. \tag{2.35}$$

As such the α-cut method contains all elements x_i that belong to $\tilde{x}_i$ at least to a degree α. By performing commonly applied interval arithmetic techniques at each α-cut the output membership function is constructed. It was shown in [159] that the output membership function is an intersection of the output intervals at a certain α-cut. Figure 2.9 illustrates three α-cuts that are made in the two membership functions with the blue dot at $\mu_{\tilde{x}_i}(x_i) = 1$ in both figure (a) and (b) a deterministic model evaluation and the red and orange line are two intervals at $\mu_{\tilde{x}_i}(x_i) = 0.5$ and $\mu_{\tilde{x}_i}(x_i) = 0$.

2.5.2 Interval field analysis

Interval analysis as described above is by definition unable to take dependency into account, which results in two extremes when spatial uncertainty, e.g., the stiffness of a material, needs to be modelled. In the first extreme, the stiffness

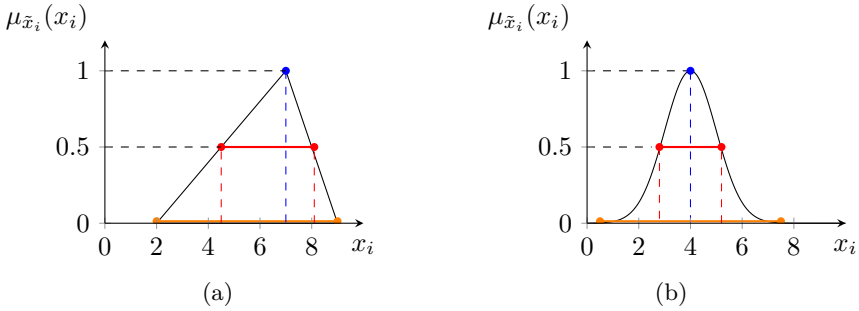


Figure 2.9: Illustration a triangular membership function (a) and a Gaussian membership function (b); the blue dot indicates a deterministic analysis and the red and orange lines the *alpha*-cuts at 0 and 0.5 with the corresponding intervals

of the material is fully coupled, e.g., the component has a high or low stiffness without any spatial variation, which would lead to a serious under-estimation of the spatial complexity of the uncertainty. In the second extreme, all elements in the model have an independent stiffness value, which neglects any kind of dependence through the spatial domain Ω. Here, the computational time would become intractable for industrial size models, which could easily contain more than 10^6 DOF. Furthermore, by neglecting the dependence between two adjacent locations in Ω, discontinuous and possible nonphysical realisations are implicitly included in the mathematical description. In an attempt to represent spatial non-determinism in a more truthful manner *explicit interval fields* were introduced in [158], which can be seen as a possibilistic counterpart to random fields [224]. It should be mentioned that other techniques have also been proposed to model spatial or coupled intervals, based on a local averaging technique [234, 235], or by defining the inter-dependence [73].

Explicit interval fields

The definition of an explicit interval field is given in Equation (2.36), where the field consists of a superposition of $n_b \in \mathbb{N}$ base functions $\psi_i(\mathbf{r}) : \Omega \mapsto \mathbb{R}$ defined over the geometrical domain $\Omega \subset \mathbb{R}^d$, where d is defined as the physical dimension of the problem. These base functions describe the spatial nature of the uncertain parameter x, distributed along the coordinate $\mathbf{r} \in \Omega$. An interval field scales these basis functions $\psi(\mathbf{r})$ with independent interval scalars $\alpha_i^I \in \mathbb{IR}$,

formally defined as [158]:

$$\mathbf{x}^I(\mathbf{r}) = \hat{\mathbf{x}}(\mathbf{r}) + \sum_{i=1}^{n_b} \psi_i(\mathbf{r})\alpha_i^I, \tag{2.36}$$

with $\hat{\mathbf{x}}(\mathbf{r}) \in \mathbb{R}$ the midpoint function of the interval field. When Ω is discretised into n_e finite elements $\Omega_i^{n_e} \subseteq \Omega$, these base functions $\psi_i(\mathbf{r})$ interpolate the independent interval scalars α_i^I to dependent intervals for each $\Omega_i^e,\ i = 1, \ldots, n_e$ by projecting them onto a non-orthogonal vector space [65]. Furthermore, the dimension of the input space is reduced if $n_b << n_e$, which reduces the computational cost of propagating the interval uncertainty towards bounds on the response quantity of interest.

The main goal of the interval field analysis is to identify the set of system responses $\tilde{\mathbf{y}}$ that bounds the possible range of responses $\mathbf{y}$ given the interval field $\mathbf{x}^I(\mathbf{r})$. Since finding the exact set is generally computationally intractable, the exact solution set $\tilde{\mathbf{y}}$ is usually approximated by a realisation set $\tilde{\mathbf{y}}_s$ defined as:

$$\tilde{\mathbf{y}}_s = \left\{ \mathbf{y}_j | \mathbf{y}_j = m_i(\mathbf{x}_j(\mathbf{r})); \mathbf{x}_j(\mathbf{r}) \in x^I(\mathbf{r}); j = 1, \ldots, n_q \right\}. \tag{2.37}$$

The set $\tilde{\mathbf{y}}_s$ is typically constructed using n_q deterministic solutions $\mathbf{y}_j = \mathcal{M}(\mathbf{x}_j)$. For each of these n_q solutions, the interval field realisations $\mathbf{x}_j(\mathbf{r})$ are generated by drawing a realisation from the interval scalars constituting the interval field. The main challenge here is to choose $\mathbf{x}_j(\mathbf{r})$ such that $\tilde{\mathbf{y}}_s$ is a conservative approximation of $\tilde{\mathbf{y}}$.

In the specific case of crash analysis, the functional relationship between $\mathbf{x}$ and $\mathbf{y}$, as given by the numerical model $\mathcal{M}$, is strongly non-convex. Therefore, the analyst has to resort to global optimisation schemes to solve Equation (2.29). A particularly well-known non-gradient-based algorithm is Differential Evolution (DE) [214], which was successfully applied to crash analysis in [32, 148] .

Definition of the basis functions

The basis functions ψ_i of the interval field described in Equation (2.36) can be defined in various ways and should reflect the knowledge or physics of the underlying problem. To meet the various requirements multiple definitions have been proposed. In addition, basis functions can also be defined on the basis of direct [111] or indirect [67, 71] measurement data. Realisations of the interval field as defined in Equation (2.38) are obtained through discretisation of the basis functions. In the following sections, an overview is given about different basis functions, based on literature reviews [66, 70].

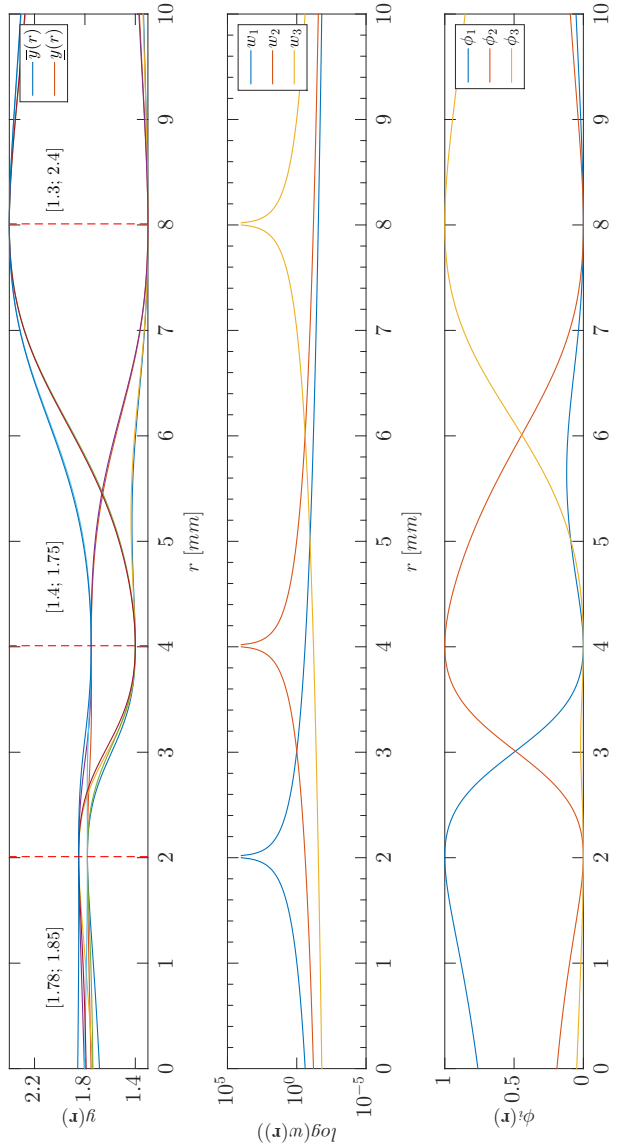


Figure 2.10: Illustration of an interval field using IDW basis functions; on the top the realisations are shown with the control points $\mathbf{r} = \{2,\ 4,\ 8\}$ indicated by a red line, the middle figure shows the corresponding weight functions plotted on a logarithmic scale, and the lower figure shows the basis functions, taken from [108]

Inverse Distance Weighting interpolation basis functions are defined through the intuitive Inverse Distance Weighing (IDW) framework, which was introduced in [65]. In this case, the basis functions $\psi_i(\mathbf{r})$ use the IDW framework to model the spatial dependence of the interval scalars α^I proportional to the inverse distance from predefined locations, denoted as control points $\mathbf{r}_i$. In practice, the interval field is discretised over $\mathbf{r}_{n_k}$ with $n_k \in \mathbb{N}$ for instance the element centre points, Gauss integration points, or nodal locations of the FE model under consideration. Following the IDW framework, the basis functions are based on a set of normalised weight functions $w_i(\mathbf{r}) \in \Omega$, defined as:

$$\psi_i(\mathbf{r}) = \frac{w_i(\mathbf{r})}{\sum_{j=1}^{n_b} w_j(\mathbf{r})}, \tag{2.38}$$

with $i = 1, \ldots, n_b$. The weight functions w_i are inversely proportional to a distance measure $d(\cdot)$. This distance is measured to all other coordinates in the domain. The weight function w_i is denoted as:

$$w_i(\mathbf{r}) = \frac{1}{[d(\mathbf{r}_i, \mathbf{r})]^p}, \tag{2.39}$$

with the power $p \in \mathbb{R}^+$ as a non-negative parameter that the analyst can set to influence the decay rate from the control point $\mathbf{r}_i$. Note that for a power $p < 1$ no derivative of the basis function exists at the control points, while in the case $p > 2$ the basis functions flatten and higher gradients at the transitions are obtained. Empirical evidence suggests that, in general $p = 2$ is a good starting point [71], if no further information about the spatial nature is available. The distance measure $d(\cdot)$ is measured in Euclidean space, which in the standard IDW framework, is defined as:

$$d(\mathbf{r}_i, \mathbf{r}) = \|\mathbf{r}_i - \mathbf{r}\|_2 , \tag{2.40}$$

with $\|\cdot\|_2$ denoting the L_2 norm. To construct the interval field based on IDW, the normalised weight functions are multiplied by independent interval scalars α_i^I and summed to the midpoint of the interval field. This is explicitly denoted as follows:

$$\mathbf{x}^I(\mathbf{r}) = \hat{\mathbf{x}}(\mathbf{r}) + \sum_{i=1}^{n_b} \frac{\alpha_i^I w_i(\mathbf{r})}{\sum_{j=1}^{n_b} w_j(\mathbf{r})}. \tag{2.41}$$

An example of the resulting interval field is shown in Figure 2.10, where the vertex realisations are shown on the top figure. Also the figure shows the weight functions as defined in Equation (2.39) and the basis functions following Equation (2.38). Note in this figure that the weight functions have *"non-vanishing"* weights, retaining a non-negligible weight throughout the domain. The latter is one of the main reasons hindering the application of IDW to large

finite element models. Furthermore, this hinders the application of IDW based interval fields to inhomogeneous phenomena. Hence, the author proposed an extension to IDW in [153], which can be found in Chapter 4.

Local interval field decomposition was introduced by Imholz et al. [110, 112] to limit the spatial complexity of interval field realisations by placing an upper limit on the gradients. This method also starts by denoting an interval field as the sum of the mean field value μ_x and the deviation from that point:

$$\mathbf{x}^I(\mathbf{r}) = \mu_x^I + \mathbf{s}_x^I(\mathbf{r}) \tag{2.42}$$

with $\mu_x = \frac{1}{\Omega}\int_\Omega x(\mathbf{r})d\Omega$. The deviation of the interval field is bounded by a maximum absolute deviation $s_{\mathrm{x,max}}$. The dependency within the field is modelled based on a maximum difference between two points, which will cause the field to vary within reasonable limits. This property is taken into account for a continuous field by the first spatial derivative. As the uncertainty is assumed to be homogeneous with respect to the spatial domain this limit is given by a single value $\frac{\partial x}{\partial \mathbf{r}}|_{max}$, which bounds the first derivative by:

$$-\frac{\partial x}{\partial \mathbf{r}}|_{max} \leq \frac{\partial x_j(\mathbf{r})}{\partial \mathbf{r}} \leq \frac{\partial x}{\partial \mathbf{r}}|_{max} \tag{2.43}$$

The basis functions ψ_i for this method need to be identically shaped, piece-wise second order polynomial functions, defined for each element or node within the domain. This is shown in Figure 2.11 where the radial basis functions are shown at each on the nodes N within the domain y. Note that the influence of these basis functions is determined by an influence radius R. This is a major advantage of the method compared with IDW where an influence of the basis functions remains within the domain. The main draw-back of this method with respect to the application to industrial size models is that each node requires a basis function. For a detailed comparison with the IDW technique, the reader is referred to [108].

Local explicit interval fields for non-homogeneous uncertainty The concept introduced in [33] is aimed at modelling a localized increase in non-determinism, which is referred to as non-stationary or non-homogeneous uncertainty. In practice this for example observed in cast products, were typical regions are likly to contain porosity's. The interval field is created with a limited influence, whcih is contained within a subdomain $\mathcal{K}_i \subseteq \Omega$. Furthermore, by limiting the domain in which the interval field is defined, the computational burden to calculate the realisations is significantly reduced, for which the author also

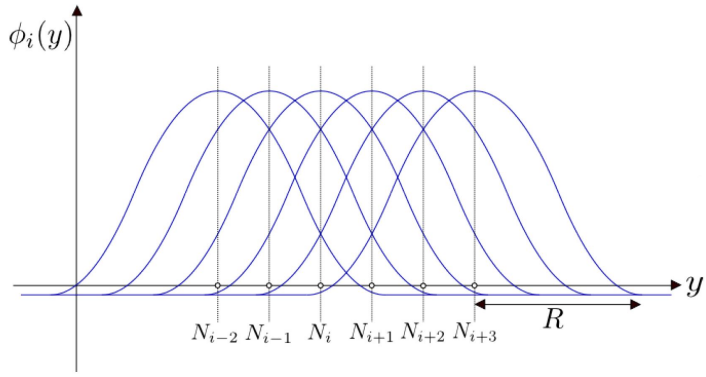


Figure 2.11: Illustration of the radial basis functions Φ_i of the LIFD method; with R the radius of influence, and N_i the FE node numbers within the domain y, taken from [110]

proposed solutions in [69]. The main requirement is that the defined basis functions satisfy the following constraints:

$$\frac{\partial w_i}{\partial d(\mathbf{r}_i, \mathbf{r})} \leq 0, \tag{2.44}$$

$$\frac{\partial w_i}{\partial d(\mathbf{r}_i, \mathbf{r})} = 0, \quad \textit{if } d(\mathbf{r}_i, \mathbf{r}) = \mathbf{B}_i, \tag{2.45}$$

$$\begin{cases} w_i(\mathbf{r}) \geq 0, & \textit{if } d(\mathbf{r}, \mathbf{r}_i) \neq \mathbf{B}_i, \\ w_i(\mathbf{r}) = 0, & \textit{if } d(\mathbf{r}, \mathbf{r}_i) = \mathbf{B}_i, \end{cases} \tag{2.46}$$

with $\mathbf{B}_i$ the width of the local support zone $\mathcal{K} \subseteq \Omega$ around a control point $\mathbf{r}_i$. An example of these basis functions is provided in Figure 2.12. The local domain and the points within this domain are indicated by $\mathcal{K}$, which is also done for the three weight functions within the sub-domain $w^{\mathcal{K}}_{1,2,3}$. Note that the basis function ψ_i is zero throughout the domain except within the sub-domain with radius B.

B-spline based interval field decomposition Recently introduced in [105] the method uses B-spline basis functions of degree k built upon a non-decreasing knot vector $\Xi = \{t_1, t_2, \ldots, t_{n_b+k+1}\}$, with t_1 the i^{th} knot entry and n_b the total number of B-spline basis functions. The k^{th} degree univariate B-spline basis function is recursively obtained through the Cox-de Boor formula [182]

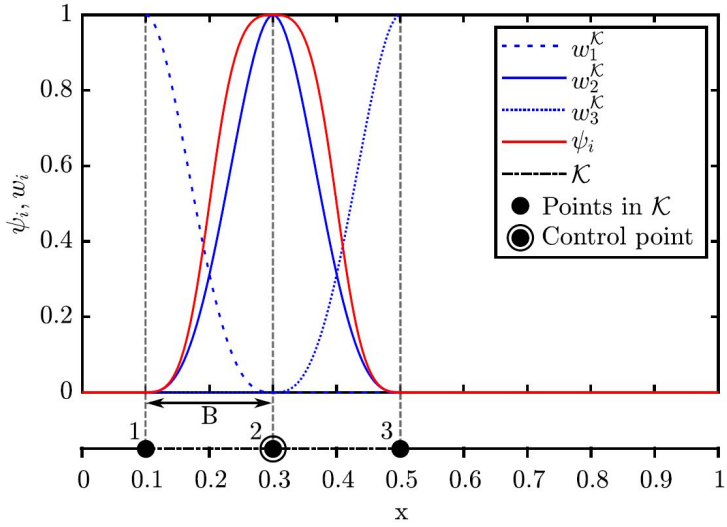


Figure 2.12: Illustration of the local explicit interval fields method basis functions within the sub-domain $\mathcal{K}$, taken from [33]

for $k = 0$:

$$\mathbf{B}_{i,0}(r) = \begin{cases} 1, & t_i \leq r < t_{i+1} \\ 0, & \text{otherwise} \end{cases}, \tag{2.47}$$

and for $k > 0$,

$$\mathbf{B}_{i,k}(r) = \frac{r - t_i}{t_{i+k} - t_i}\mathbf{B}_{i,k-1}(r) + \frac{t_{i+k+1} - r}{t_{i+k+1} - t_{i+1}}\mathbf{B}_{i+1,k-1}(r). \tag{2.48}$$

Inclusion of these basis functions in the explicit interval field formulation, as defined in Equation (2.7), the expression is slightly adapted:

$$\mathbf{x}^I(\mathbf{r}) = \hat{\mathbf{x}}^I + \Delta\mathbf{x}^I \sum_{i=1}^{n_b} \mathbf{B}_{i,k}(\mathbf{r})\alpha_i^I, t_{k+1} \leq x \leq t_{n_b}, \tag{2.49}$$

with $\mathbf{B}_{i,k}, i = 1, \ldots, n_b$ the k^{th} degree B-spline basis functions defined on Ξ, and α_i^I unitary interval scalars that are referred to as *interval field coordinates* in this method. The advantage of this method is that it is very flexible with easy extensions to non-homogeneous interval fields, and higher-dimensional interval fields. However, this requires that the analysis is able to provide detailed properties of the field.

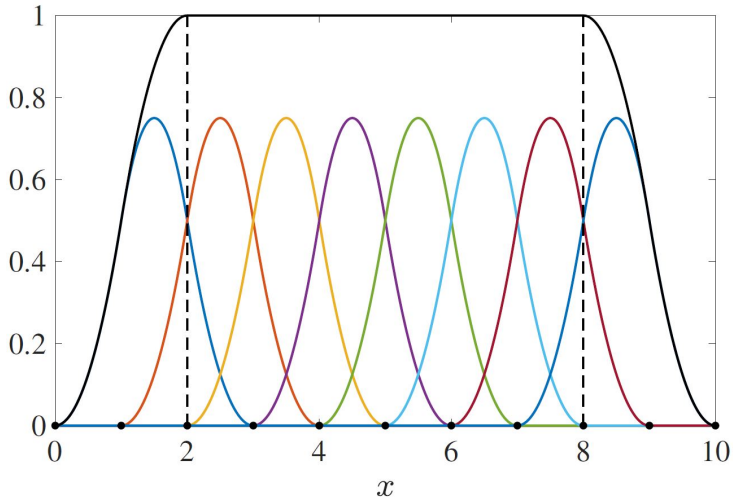


Figure 2.13: Illustration of the quadratic B-spline basis functions over the domain $x \in [0;\ 10]$ [105]

Affine arithmetical interval fields

Extending the concepts of affine arithmetic to model dependency with the aim of modelling spatial uncertainty, the dimensionless interval field $\mathbf{A}^I(\mathbf{r})$ with unit range was introduced in [203, 204, 207, 208] with applications described in [206, 211]. The parameter interval field $\mathbf{x}^I(\mathbf{r})$ is obtained as:

$$\mathbf{x}^I(\mathbf{r}) = \hat{\mathbf{x}}_0 \left(1 + \mathbf{A}^I(\mathbf{r})\right), \tag{2.50}$$

With $\hat{\mathbf{x}}_0 \in \mathbb{R}$ a midpoint taken constant over the domain $\mathbf{r}$. The amplitude of the interval field can be related to the interval field radius $\Delta\mathbf{x}(\mathbf{r})$ via:

$$\Delta\mathbf{x}(\mathbf{r}) = \frac{\overline{x}(\mathbf{r}) - \underline{x}(\mathbf{r})}{2} \equiv \hat{\mathbf{x}}_0 \Delta\mathbf{A}(\mathbf{r}), \tag{2.51}$$

with $\mathbf{A}^I$ a dimensionless interval function having zero midpoint and a radius $\Delta\mathbf{A}(\mathbf{r}) < 1$. Furthermore, by defining a deterministic, symmetric, non-negative, bounded spacial dependency function $\Gamma_A(\mathbf{r}_i, \mathbf{r}_j)$ on the dimensionless interval field $\mathbf{A}^I$:

$$\Gamma_A(\mathbf{r}_i, \mathbf{r}_j) = \text{mid}\left[\mathbf{A}^I(\mathbf{r}_i), \mathbf{A}^I(\mathbf{r}_j)\right] \equiv \frac{\text{mid}\left[\mathbf{A}^I(\mathbf{r}_i), \mathbf{A}^I(\mathbf{r}_j)\right]}{(\hat{\mathbf{x}}_0)^2} - 1, \tag{2.52}$$

with $\text{mid}(\cdot)$ an operator that returns the midpoint of the interval. This function only works when the EUI's are used as defined in Equations (2.23)-(2.27) to omit

the dependency problem. Arguably, $\Gamma_A(\mathbf{r}_i, \mathbf{r}_j)$ can be regarded as a possibilistic counterpart for the stochastic auto-correlation function [207], which is used to describe auto-covariance in a random field [224]. Therefore, following the random field approach, the interval field is expressed as a Karhunen-Loève-like decomposition of $\Gamma_A(\mathbf{r}_i, \mathbf{r}_j)$. Hence, the dimensionless interval function $\mathbf{A}^I(\mathbf{r})$ is expanded as a series of n_a deterministic functions and extra unitary intervals:

$$\mathbf{A}^I(\mathbf{r}) = \sum_{i=1}^{n_a} \sqrt{\lambda_i}\psi_i(\mathbf{r})\hat{e}_i^I \tag{2.53}$$

with λ_i and ψ_i the eigenvalues and eigenvectors obtained from the following eigenvalue problem:

$$\int_\Omega \Gamma_A(\mathbf{r}_i, \mathbf{r}_j)\psi_i(\mathbf{r}_i)d\mathbf{r} = \lambda_i\psi_i(\mathbf{r}) \tag{2.54}$$

which is known in the context of random fields as the homogeneous Fredholm integral of the second kind [224]. The parametric interval field following this approach is defined as:

$$\mathbf{x}^I(\mathbf{r}) = \hat{\mathbf{x}}_0 \left[1 + \sum_{i=1}^{n_a} \sqrt{\lambda_i}\psi_i(\mathbf{r})\hat{e}_i^I\right] \tag{2.55}$$

It should be noted that by replacing the independent identically distributed (i.i.d.) random variables in the truncated KL series expansion with interval variables the convergence properties of the expansion, as described in [212], are no longer guaranteed [66]

2.5.3 Overview and considerations using interval fields

This section provides guidelines and considerations for practical application of the previously discussed interval field techniques. First, an overview is given about the various interval field techniques and in the second part application examples are given. The main challenge in the selection of an interval field technique is that it depends on the available data and the problem at hand. To provide guidance in choosing between the various options the following points should be considered:

1. **User input** is the data that should be provided by the user or analyst. Depending on this data different field definitions might be preferred. The IDW technique requires intervals at the control points, which can be measured or based on engineering judgement. Other techniques require a global or sub-domain interval.

2. **Dependence structure** determines how two variables at different locations within the domain are related. For IDW these are independent at the control points and depend on their mutual distance within the domain. Other techniques use a support radius, or characteristic length parameter that determines the relation.

3. **Gradient control** means that the analyst can directly influence the gradients of the realisations. The reason for this is that high gradients might be non-physical, or information about the gradients within the domain at a specific location is available. Some techniques like LIFD allow the user to bound the gradients of the realisations.

4. **High dimensions** refers to the possibilities to use the technique with multiple input intervals, or in domains larger than the three dimensional physical domain. This mostly depends on the way that the fields are represented and discretised. For example, IDW can be used in three-dimensions but discretisation over, e.g., 10^6 nodes for large industrially sized models quickly becomes troublesome.

5. **Non-homogeneous** fields refer to changing the nature of the uncertainty of the field properties within the domain. This means that, for example, the characteristic length changes within the domain or the influence radius of the knot vector as shown with the BIFD method.

Finally, based in these points Table 2.1 provides a detailed overview of the practical considerations of the different interval field methods, highlighting the data that should be provided and where the technique is best applied.

Application of interval field techniques for process variation

In this case study an interval field is used to model the influence of process variance caused by additive manufacturing [73]. In this work, a rocker for the Formula Electric Belgium team was topology optimised and produced in Poly-Amide 12 (PA-12) using Selective Laser Sintering (SLS). It is well-known that the mechanical performance and dimensional stability are heavily impacted by a number of process-related variables; the interested reader can find more information about this in [30, 94, 180, 232]. Here it was found that the main contribution to thickness variation is the inter-layer cooling time. In other words, the time between sintering of a layer and the next layer, which determines the local cooling rate. The main cause of varying inter-layer cooling time is when multiple parts are built within the same building plate. The larger the area that the laser should process the longer it takes before the next layer is applied. Therefore, as there is some knowledge about the other parts on the building

field technique	**user input**	**dependence structure**	**gradient control**	**high dimensions**	**non-homogeneous**
Inverse distance weighing IDW	intervals at control-points r_i, parameter p	based on distance	possible at the control points, see Chapter 4	additional control points or dimensions highly increase input dimension	not directly possible, technique proposed in Chapter 3
Local Interval Field Decomposition LIFD	global field midpoint and radius, maximum gradient $\frac{\partial x}{\partial \mathbf{r}}\|_{\max}$, influence radius r	based on the influence radius of the radial basis functions	bounds the gradients homogeneous within the domain	computational burden increases with the number of nodes	not directly possible
local explicit interval field decomposition	intervals at the location of the local domain, size of domain	based on distance within local domain	not directly possible	specifically aimed at local effects in high dimensions	possible using multiple domains
B-spline Interval Field Decomposition BIFD	global field midpoint and radius, degree of the basis k	based on the influence knot span of the k^{th} degree B-spline basis	not directly possible	very efficient in moderate dimensions, i.e, less than 3	possible using multi patch stitching
Affine arithmetical interval fields	global midpoint and radius, dependency function, truncation terms n_a	based on the dependency function, characterised by characteristic length l_c	not directly possible	difficult	not directly possible

Table 2.1: overview table of interval field techniques

plate an estimation about the thickness variation can be made and modeled with an interval field. In this study the IDW based interval field is used as this allows the incorporation of this knowledge at certain points in the domain, i.e., control points. The resulting interval field is shown in Figure 2.14 where the upper and lower bound of the thickness are shown. The domain of this interval field Ω is described by two coordinates r_1 and r_2.The stack direction of the component during production was along r_1, which is the direction that is mainly influenced by the inter-layer cooling time. These interval field realisations were obtained using the in-house developed code FIRST.

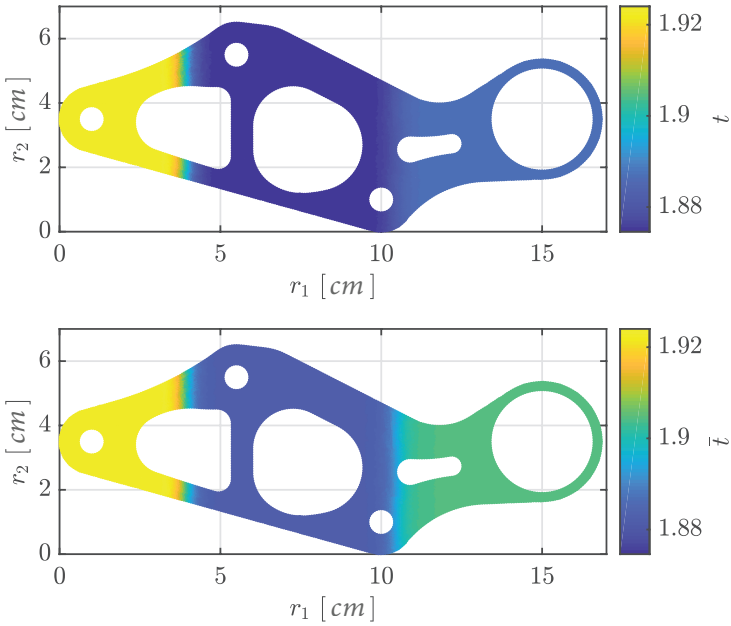


Figure 2.14: Illustration of interval field realisations on a rocker used by Formula Electric Belgium; This part is topology optimised and produced in PA-12 using (SLS), taken with permission from [73]

Application of interval field techniques for material variation

To highlight the intuitive framework that IDW based interval fields offer, a case with material variation is considered. Here a complex door hinge is obtained using Topology Optimisation TO. It is assumed that the Young's

modulus of the material spatially varies throughout the final product. Based on expert knowledge, the Young's modulus could be estimated at two locations of the product. However, as there is only limited information at hand these measurements are bounded by an upper and lower limit. The first point $x^I_{r_1} = [65;\ 72]$GPa and the second point $x^I_{r_2} = [67;\ 75]$GPa. The location of these control points is shown in Figure 2.15 indicated as yellow points, with r_1 at the bottom and r_2 at the top of the hinge.

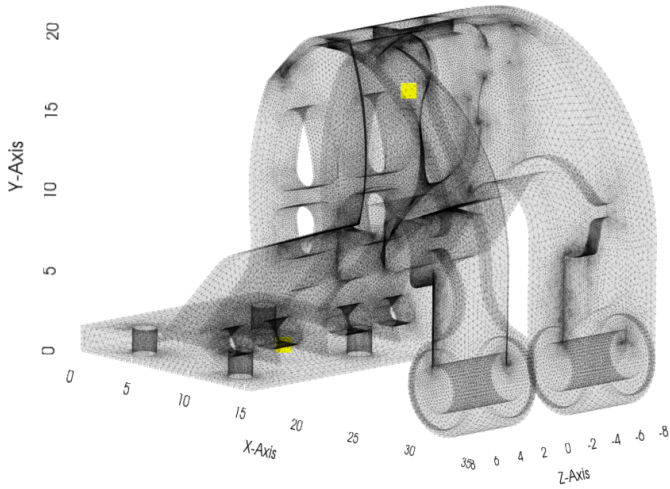


Figure 2.15: illustration of the control point placement in a medium size numerical model of a TO hinge; the two contol points are indicated by the yellow dots; created by the graphical user interface of FIRST

After placing the control points, the basis functions can be determined by setting the parameter p from Equation (2.39) by the user. Typical, a starting point for this parameter is $p = 2$ especially if there is no further information available [71]. Hereafter the field must be discretised, which can be to the nodes, Gauss integration points, or element locations depending on the problem and the possibilities at hand. Figure 2.16 illustrates one of the vertex realisations of the interval field.

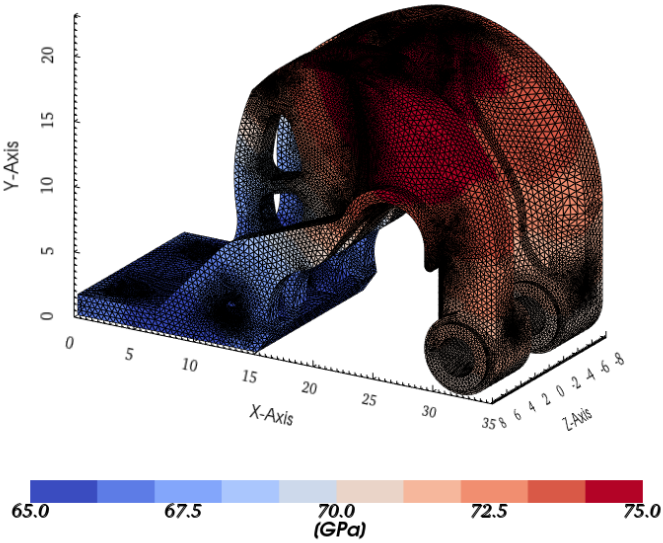


Figure 2.16: illustration of a single interval field realisation on Young's modulus, on a numerical model of a topology optimised door hinge; created by the graphical user interface of FIRST

Application of interval field techniques for structural dynamics

The final application example is the use of an interval field to model the uncertain stiffness of a composite solar panel used in a lunar lander [39]. In this study, the transient dynamic response during the landing procedure was modelled and the difference between interval- and random fields are investigated. The model that was used is shown in Figure 2.17 where the point of interest is shown at the tip of the composite solar panel. This study highlighted that there are quite some parallels between the concepts of random fields and interval fields, with as main conclusion that the correlation length of a random field and the maximum gradients of the LIFD method have a very similar effect on the global parameter behaviour and therefore produce comparable realisations. This is not so surprising as the correlation length determines the relative change in random variables and, in some sense limiting the gradients has a similar effect. However, due to the infinite support of the normal Gaussian random variables the random field results do not provide crisp bounds, which is the case for interval fields.

Other applications of interval field techniques are found in the fields of

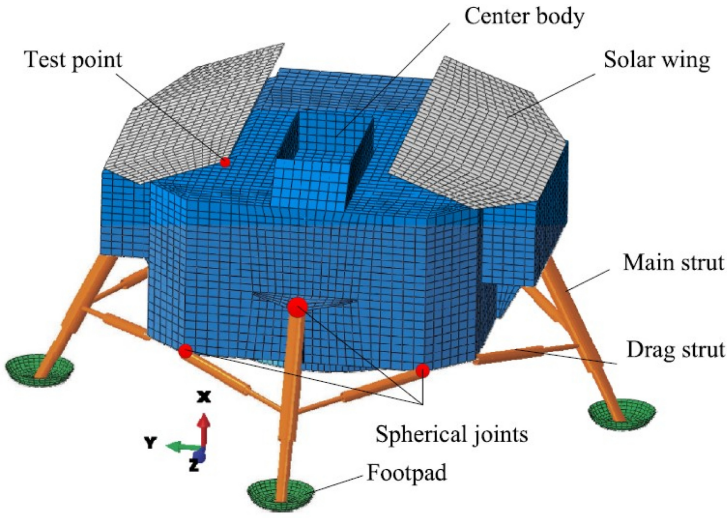


Figure 2.17: Numerical model of a lunar lander used to investigate the transient dynamical behaviour during landing [38]

geomechanics [83], where the effect of uncertain soil properties on slope stability is investigated, hence, improving the safety of constructions built in coastal or mountain areas. Another application can be found in electromagnetic problems [229], where it is found that the electromagnetic performance is very sensitive to uncertainty in the geometry, material, and environment. The final example is the application of interval analysis to assess the vibro-acoustic performance of a passenger vehicle [237].

2.6 Robustness based design optimisation

Robust design optimisation is a primary requirement to ensure that a system fulfils its intended purpose over a long period of time. To reach this goal, one often has to consider significant uncertainty in the problem specification. For example in an early design stage, one might not even know how adjacent components look like and what their specifications are. The objectives of general optimisation techniques, e.g., to minimise weight or maximise output, are cumbersome, as it is well known that the optimal performance of a system is often very sensitive to small changes in input parameters [59]. Hence, due to this sensitivity, either the optimal performance is only occasionally reached or the variation causes a constraint violation. Robust design optimisation considers

the output variation and is aimed at finding designs that are insensitive to input variation. The idea of products and processes that are insensitive to variations, i.e., in manufacturing, was pioneered by *Genichi Taguchi* who first applied his methodology on electrical circuits [218, 219], and is often quoted [20, 113] characterising a robust design as:

> *"Not just strong. Flexible! Idiot Proof! Simple! Efficient! A product/process that produces consistent, high level performance despite being subjected to a wide range of changing client and manufacturing conditions"* –Genichi Taguchi

In the decades after the ideas by Genichi Taguchi two viewpoints about robustness have emerged. The first viewpoint considers the performance of a system under extreme and exceptional conditions. The second viewpoint considers the performance of the system under normally varying conditions, e.g., manufacturing tolerances. This second viewpoint is based on the quote of Genichi Taguchi, which is also the viewpoint taken in this work. Note that the latter viewpoint is closely related to sensitivity analysis. However, there is a subtle difference between the two. Sensitivity analysis describes the dependency of the output of a system on the input parameters and aims to identify the most influential input parameter. In general, sensitivity analysis is a local measure often related to gradient information, which can be obtained by various techniques. Robustness, on the other hand, is a global measure of the degree of system variability due to input variations. Figure 2.18 illustrates the main ideas of a robust design, where the design parameter z_i has a normally distributed variation $\mathcal{N}(z_i, 0.2)$ illustrated for $z_i = 2$ and $z_i = 6$. It is clear that the output variation is smaller for $z_i = 2$ with an extremely low constraint violation. The constraint in this case is that the output should be lower than $y_i < 36$. Figure 2.18 also illustrates the role of sensitivity analysis for $z_i = 6$ where the red line shows the gradient. Note that this gradient is not consistent over the total variation of the design parameter. To differentiate between these robust design strategies, four classes of robustness measures are defined that are conceptually different [166]:

1. **Functional expectancy and dispersion** metrics are based on the evaluation of expectancy and the variation around this value. As an example, *Genichi Taguchi* proposed a signal-to-noise ratio to quantify robustness based on the idea of loss of quality [218, 219]. Mostly focused on aleatory descriptions of uncertainty these methods are often too difficult to calculate analytically for industrial cases. Therefore, often approximations are made using surrogate modelling techniques.

2. **Sensitivity** based robustness measures are the most straightforward robustness measures with a well-understood framework to relate the change in independent variables to dependent variables. These metrics are usually based on the calculation of quotients or partial derivatives in the case of multiple input parameters.

3. **Size of feasible domain** measures are aimed at finding the largest domain in which the functional output is considered within tolerance. In other words, a design that is very sensitive to variations while the output variation remains within certain bounds can still be considered a robust design. These measures usually include the calculation of a distance, area, volume, or polyhedron in one or more dimensions.

4. **Probability of functional compliance** metrics evaluate the probability that one or more functions fulfil their requirements under stochastic variation, which requires detailed knowledge about the probability density functions of the input parameters and design limits.

Note here that the four categories above are mainly considered with aleatory uncertainties, while often parallels can be made in the case of epistemic uncertainties. In this work, these parallels are explored, and below a more comprehensive overview is given about robust design optimisation techniques under epistemic uncertainty.

2.6.1 Functional expectancy and dispersion

This section covers the more classical view on robust design as proposed by *Taguchi* with ideas such as quality loss and the mean square deviation [220]. Here no information about the requirements is regarded as the focus lies in minimisation of the dispersion, i.e., variance, or expected mean value, or a combination of both. Hence, the probability density function (PDF) of the input variables should be known a priori. Using an appropriate design of experiments or expert knowledge, these PDF's can be obtained and the mean value of the output can be calculated as:

$$E(y) = \int f(x)p(x)dx \tag{2.56}$$

with $f(x)$ the functional relation and $p(x)$ the PDF of x. The variance about this estimate is then determined by:

$$\sigma(y) = \int (f(x) - E(y))^2 p(x)dx, \tag{2.57}$$

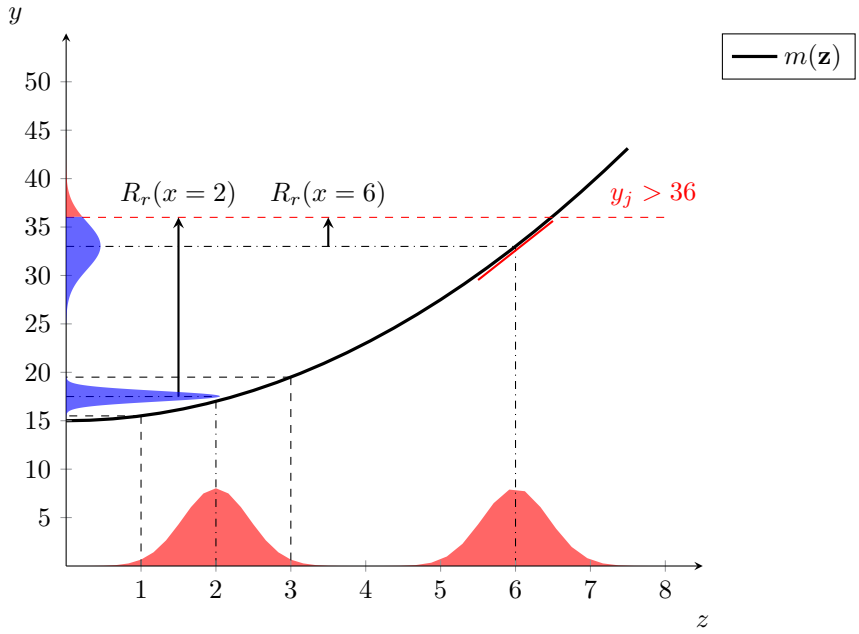


Figure 2.18: Illustration of multiple robustness measures; the performance function is shown in black with two possible inputs $z = 2$ or $z = 6$ the varaition of these parameters is shown in red, the resulting output variance is shown in blue

which is referred to as the standard deviation. Although seemingly straightforward, the calculation of the expected value and variance of the output under multiple uncertain inputs is usually untraceable analytically [26], which has led to the use of advanced surrogate modelling strategies.

2.6.2 Sensitivity based robustness

Robustness measures that are based on the sensitivity of dependent variables to independent variables are well understood and often used. In this class of metrics, one usually uses a finite quotient for one variable or partial derivatives in the case where there are numerous variables. When the underlying function $m()$ is known analytically one can derive an analytical expression for the derivatives. However, this is not true in general, where the underlying function is a black-box model without any information about the gradients. In this case, the gradients must be based on samples, which is only valid when there are small

differences between input parameters. Figure 2.18 illustrates this point at $z_i = 6$ where the derivative is shown tangentially to the performance function. The gradient obtained is only locally valid, not over the complete domain of possible input values. Hence, these methods should only be used with relatively small input variations and smooth functional relations, as otherwise problems and misjudgments might arise.

Interval sensitivities

In the specific case of interval-valued input variables, the concept of interval sensitivity was introduced [161] where sensitivity is measured over the entire range of the interval and a global measure is obtained. For interval sensitivities the relation between input and output of the model m can be written in terms of the interval radius Δx and center $\hat{x}$ as:

$$\Delta y = h(\hat{x}, \Delta x) = \frac{1}{2}\left(\sup_{x \in x^I} m(x) - \inf_{x \in x^I} m(x)\right) \tag{2.58}$$

By considering the interval center $\hat{x}$ as the fixed nominal case, the interval sensitivity of the output interval y^I with respect to the input interval x^I is defined as:

$$\delta_{x^I}^{y^I} = \frac{\partial(\Delta y)}{\partial(\Delta x)} \tag{2.59}$$

with $\delta_{x^I}^{y^I}$ the interval sensitivity relating a change in absolute input interval width to a change in output interval width. This derivative exists, if Δy at $\hat{x}$ is a C^1-continuous function of Δx, for parameter values within the interval. For further details and calculation procedure, the interested reader is referred to [161].

Interval sensitivity index

Recently, an interval sensitivity index C was proposed in [35]. This index is a more general approach, where the sensitivities of one or more parameters can be compared. In a one-dimensional case the index is based on the difference between the interval valued upper and lower bound and the actual upper and lower bound within the domain. Figure 2.19 illustrates the areas that are defined for this case with A_u between the upper-bound and $\overline{y}_i$, A_{process} between $\overline{y}_i$ and $\underline{y}$, and A_l between the upper-bound and $\underline{y}$. The total area A_{total} is given by the sum: $A_{\text{total}} = A_u + A_{\text{process}} + A_l$. Based on these measures the index is defined as [35]:

$$C_i = \frac{A_u(i) + A_l(i)}{A_{\text{total}}(i)} \tag{2.60}$$

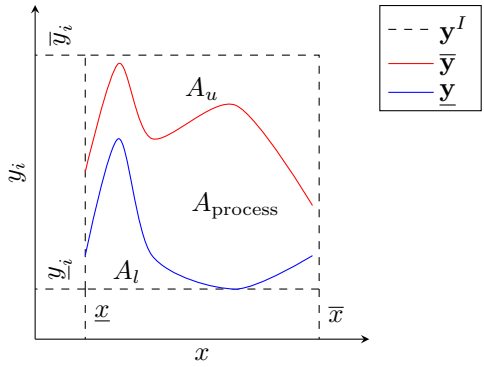


Figure 2.19: Illustration of the interval input output relation to highlight the area's used in the interval sensitivity index

Before the index is calculated, the interval radius of the parameters is normalized. Therefore, C_i will always be within $0 \leq C_i \leq 1$. In the case that $C_i = 0$, interval parameter x_i has no influence on the result, when $C_i = 1$ parameter x_i is the only influential parameter [35]. The main difference with the interval sensitivities is that the interval index is not limited to the bounds, as it quantifies what happens within the interval as well.

2.6.3 Size of feasible design space metrics

A feasible design space is obtained by quantifying the variance allowed for the input parameters that keeps the design within predefined functional limits. Robustness measures in this class usually involve the calculation of a high-dimensional volume or manifold that satisfies the design specifications. Note that the feasible design space is determined by relating the dependent variable to the independent variable. These metrics are usually independent of the input variation, as designs are preferred with the highest amount of permissible variation. One approach to achieve this is to search for the design that respects the largest distance away from the constraints, this can be done using the robustness radius, defined as:

$$R_r = \min \left(|\min_{x \in \mathcal{X}} m(x) - E(m(x))|; |\max_{x \in \mathcal{X}} m(x) - E(m(x))| \right), \tag{2.61}$$

with $E(m(x))$ the expected mean of the response and the minimum and maximum of the model are limited by the imposed constraints. Therefore, R_r describes the minimum distance between the constraints and the nominal

input value. The robustness radius is also shown in Figure 2.18, which makes clear that the robustness of $R_r(x = 2) > R_r(x = 6)$. More measures are provided in the overview of [166]. However, the remainder of this section will discuss the robustness measures introduced for fuzzy variables.

Robustness measures for fuzzy variables

A robustness measure for fuzzy numbers was introduced in [20, 246], which is related to the *"fuzziness"* of the input variables. This fuzziness is measured using *Shannon's* entropy [49, 248] applied as:

$$H(\tilde{\mathbf{x}}) = -k \int_{x_i \in \tilde{\mathbf{x}}} \cdots \int \left[\mu_{\tilde{x}}(x_i) \ln(\mu_{\tilde{x}}(x_i)) + (1 - \mu_{\tilde{x}}(x_i)) \ln(1 - \mu_{\tilde{x}}(x_i))\right] dx_i \tag{2.62}$$

with k a positive constant [49]. This measure can be interpreted as the "steepness" of the membership function $\mu_{\tilde{x}}(x)$. Application to a crisp set yields $H = 0$, whereas the most uncertain set with all elements evaluated by a membership function $\mu_{\tilde{x}} = 0.5$ except for the mean value yields the maximum value for H. Moreover, there is a dependency between the support of the fuzzy number and the entropy measure, which increases with a larger support. Except in the case of an interval variable $\mu_{\tilde{x}} = 1 \forall x \in \mathcal{X}$, which yields $H = 0$. Therefore, intervals are excluded from this robustness measure.

The robustness $R(\cdot)$ if now measured by the ratio of input entropy and response entropy, were parameter k in Equation (2.62) is omitted. In the case of a fuzzy input vector $\tilde{\mathbf{x}}$ and associated multi-variate fuzzy responses $\tilde{y_j}$, the structural robustness is defined as:

$$R(\tilde{\mathbf{x}}, \tilde{y_j}) = \frac{H(\tilde{\mathbf{x}})}{H(\tilde{y_j})}. \tag{2.63}$$

Note that the computation of $H(\cdot)$ is associated with the numerical computation of a high-dimensional integral. Hence, the authors proposed separate computations of the elements in $\tilde{\mathbf{x}}$, which poses a minor influence as long as the elements in $\tilde{\mathbf{x}}$ only weakly interact. Therefore, a numerically more efficient robustness measure is defined as:

$$R(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}_j) = \frac{\sum_{i=1}^{n_x} H(\tilde{x}_i)}{\sum_{j=1}^{n_y} u_j H(\tilde{y}_j)}, \tag{2.64}$$

with u_j a weight factor for certain output events and n_x,n_y the number of input and output variables in $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$.

Design optimisation based on this robustness measure was proposed by evaluating several alternative designs [20]. First, the uncertain structural

responses are obtained from an uncertainty analysis, which yields a set of output responses based on a sampling strategy. The set of output responses is then divided into permissible and non-permissible responses, which can be used to inversely quantify sets of design alternatives. Now the robustness measure, among others, is used to identify the most robust design. Note that in the case of *black-box* functions, the results depend on the underlying samples that are taken, as do the design variants. Although applied to a heavy non-linear crash example, one has to assume that the samples of the domain are sufficient to guarantee no violations occur within. The authors listed several shortcomings in a subsequent work [246] where an extension is proposed by calculating the robustness measure in Equation (2.63) at multiple α-cuts, with an application of the method for pipe corrosion [221].

Robustness in the framework of convex set theory

Measures for robustness to use in the framework of convex sets were proposed in [22], where convex sets are described in the framework of info-gap theory, which has analogies with convex sets [8]. This robustness definition is actually identical to the robustness radius, as described in Equation (2.61). In addition, info-gap robustness adds a definition for opportuneness, where the analysis defines an output that would be nice to realize. This is also defined identical to the robustness radius only now the value is maximized instead of minimized.

Robustness under interval uncertainty

In the case of interval valued uncertainty due to a lack-of-knowledge the robustness is usually defined as the resulting output interval width, which would correspond with the output variance at a certain confidence level. In general, robustness is defined as:

$$R(f(\mathbf{x}, \mathbf{z}^*)) = \max_{\mathbf{x} \in \mathbf{x}^I} f(\mathbf{x}, \mathbf{z}^*) - \min_{\mathbf{x} \in \mathbf{x}^I} f(\mathbf{x}, \mathbf{z}^*), \tag{2.65}$$

with $R(\cdot)$ the robustness measured for a specific design variable z^*. This type of robustness measure is seen in [18, 131, 154], which is mostly concerning interval-valued variables. More details and cases based on this robustness measure can be found in Chapter 6.

2.6.4 Probability of functional compliance

In this class of robustness, measures assess the probability that one or more system requirements are met under stochastic variation. Hence, these methods require a detailed description of the probabilistic uncertainty in the input variables, the system performance limits, and the model should accurately describe the underlying physical phenomena. This class of methods is limited to probabilistic methods, as one needs to obtain a probability of exceeding the requirements. In most cases found in literature, input variables are assumed to follow a normal distribution, an assumption mainly based on the ease of use and supported by the central limit theorem. In this case, the probability of exceeding the upper performance limit (UPL) and lower performance limit LPL is calculated as:

$$Pr_{ij}[\mathrm{LPL}_j \leq f_j(x_i) \leq \mathrm{UPL}_j], \tag{2.66}$$

where the probability is measured that the output f_j depending on variable x_i remains within tolerance. In the case of multiple requirements, the likelihood is measured that all these requirements are met.

To this point, depending on the robustness measure that is deployed, multiple techniques exist to evaluate the robustness. Examples using analytical expressions are found in the literature [21], while most industrial applications are more demanding. Therefore, the most common strategies found in the literature are: based on sets of samples [20], using surrogate model approaches [154] or general means of optimisation [130].

2.7 Stochastic process emulators

A meta-model, emulator, or surrogate model refers to a mathematical function that generally has no underlying physical constraints and is capable of emulating a set of observations. In the case of crash analysis, these observations can be related to the velocity of certain parts [17], or other quantities of interest. One of the main advantages is that these mathematical functions are easy to evaluate at a fairly low computational cost. In other words, one can evaluate surrogate models a lot faster than the numerical model. This point is also illustrated in Figure 2.20 where the black lines illustrate the path that is the easiest to evaluate. The dashed lines illustrate that the numerical model should still be evaluated to produce a set of observations. The surrogate model $\mathcal{G}$ is then trained based on these observations. Note that the model $\mathcal{G}$ should be as simple as possible to allow quick approximations. The non-determinism can be regarded in two ways at this point: (1) the underlying physics are stochastic in nature and the model captures this variability; (2) the model can approximate

the observations until a certain measurement error. The two are structurally the same and can be analysed in identical ways by a stochastic process [89], while the context is different. In this work the focus lies on Gaussian processes, which are a specific type of stochastic surrogate models. The interested reader is referred to [11] for a recent introduction and review on the use of stochastic modelling approaches.

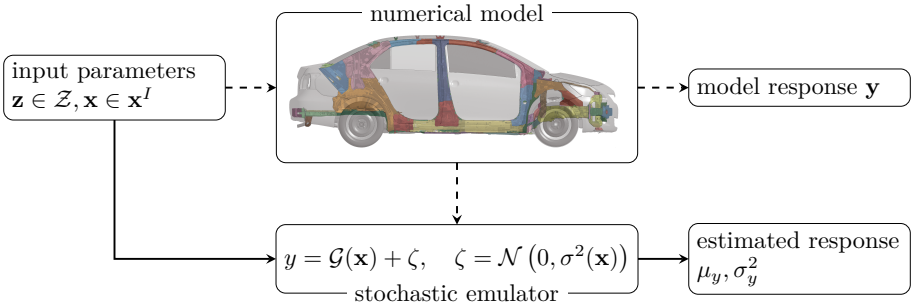


Figure 2.20: Illustration of a stochastic emulator replacing the expensive to evaluate numerical model

To illustrate the idea of a stochastic process emulator and notation hereof, the example as stated in [198] is followed. The goal of the example is to demonstrate a *draw* or *realisation* of a quadratic random function defined in Equation (2.67). In this example $\mathcal{X}$ is denoted as the input space for the unknown output $g(x)$. A draw or realisation of the stochastic process is denoted by $g(\cdot)$, and is accomplished by generating realisations b_0, b_1 and b_2 of random variables B_0, B_1 and B_2. Here, the realisation of a random variable is the mapping of a set of outcomes Ω to a real number $\mathbb{R}$.

To illustrate the process of creating draws $g(x) : x \in \mathcal{X} = [-2;\ 2]$ a quadratic random function is defined:

$$g(x) = b_0 + b_1 x + b_2 x^2, \tag{2.67}$$

where the realisations b_0, b_1 and b_2 are mutually independent normally distributed random variables $B_i \sim \mathcal{N}(0, \sigma_i^2)$ for $i = 1, 2, 3$. It can be shown that this normally or Gaussian distributed process realised over $x \in [-2;\ 2]$ has a zero mean:

$$E[g(x)] = E[b_0 + b_1 x + b_2 x^2], \tag{2.68}$$

$$= E[b_0] + E[b_1]x + E[b_2]x^2, \tag{2.69}$$

$$= 0 + 0x + 0x^2 = 0. \tag{2.70}$$

Note that Equation (2.68) is not applicable to a single realisation, as this property is only achieved for a very large number of realisations. Likewise, the variance of the process is given by:

$$Var[g(x)] = E[(b_0 + b_1 x + b_2 x^2)(b_0 + b_1 x + b_2 x^2)], \tag{2.71}$$

$$= \sigma_0^2 + \sigma_1^2 x^2 + \sigma_2^2 x^4. \tag{2.72}$$

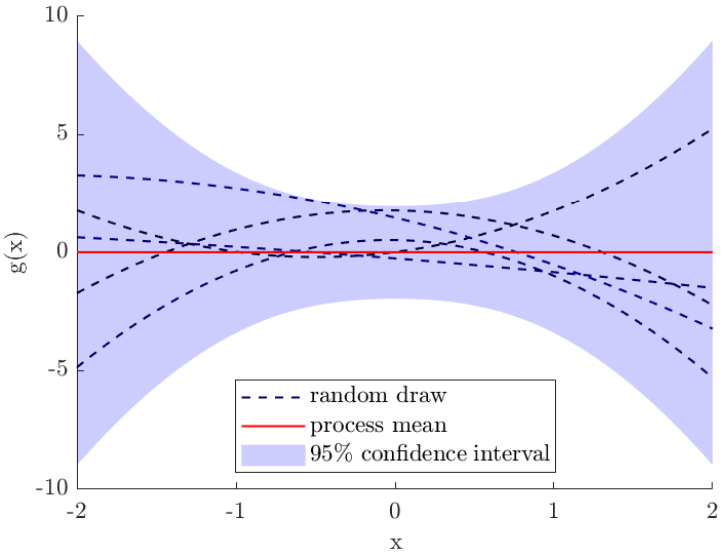


Figure 2.21: Illustrations of random draws of the function in Equation (2.67)

Based on the process variance, Confidence Intervals (CI) are calculated as:

$$\text{CI} = c\sigma, \tag{2.73}$$

with c a constant that indicates the number of standard deviations, e.g., $c = 1,96$ indicates a 95% confidence interval. This means that 95% of the realisations fall within the CI. Figure 2.21 shows a number of realisations in dashed lines, the process mean as a red line, and the 95% CI as the blue area in the background. It is clear from this figure, by the increased blue area, that the variance is larger at $x = 2$ compared to $x = 0$.

Note that there is a relation between $x = 2$ and $x = 0$ as they are both related to the same process. This relation is known as covariance, which is a measure of the joint variability of two random variables. The covariance is calculated for

$x_1, x_2 \in [-2;\ 2]$ as:

$$Cov[g(x_1), g(x_2)] = E[(b_0 + b_1 x_1 + b_2 x_1^2)(b_0 + b_1 x_2 + b_2 x_2^2)], \tag{2.74}$$

$$= \sigma_0^2 + \sigma_1^2 x_1 x_2 + \sigma_2^2 x_1^2 x_2^2. \tag{2.75}$$

Based on this example, the reader should have an idea about random process functions. Note that the quadratic function defined in this section has a number of limitations, e.g., $g(x)$ only produces quadratic realisations. In practice the main challenge is to find a functional relation and covariance structure that correctly predicts observations. Therefore the remainder of this text is devoted to a more general and flexible formulation of a random process, which is made conditional on a set of observations.

2.7.1 Gaussian process models

Multiple approaches have been proposed to enhance the capabilities of random process to represent functional relations. The simplest approach is to allow the mean of the stochastic process to depend on x in the form of a regression equation while assuming that the residual variation follows a stationary stochastic process. This GP has the form:

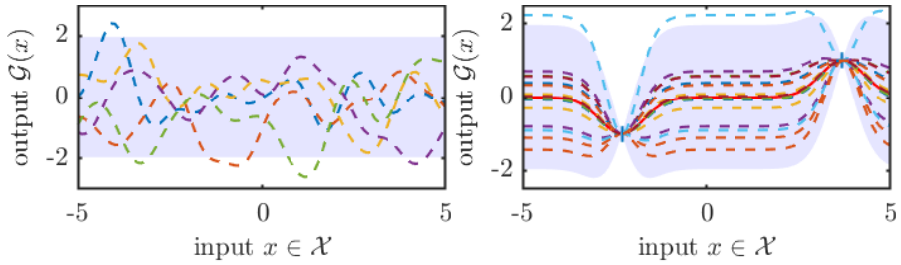
$$\mathcal{G}(x) = \boldsymbol{\beta}^T f(x) + \sigma^2 F(x, \Lambda), \tag{2.76}$$

with the first term being a deterministic regression model with $f(x)$ an arbitrary basis function, and $\boldsymbol{\beta}^T$ a vector of regression coefficients. The second term, is a stochastic process that maps the probability space $(\Lambda, \mathcal{F}, P)$ to a real value, with sample space Λ, a σ-algebra $\mathcal{F}$, and probability measure P. This zero-mean, unit variance stochastic process is scaled with a constant variance of the Gaussian process σ^2. The correlation between two points x and x' is defined by the covariance function $K(x, x', \mathbf{p})$, with $\mathbf{p}$ presenting parameters of the covariance function, generally referred to as hyper-parameters. The correlation length (see later) is one of those typical hyper-parameters.

For a GP, $F(\cdot)$ is considered to be stationary, which means that $(F(x_1), \ldots, F(x_{n_x}))$ and $(F(x_1 + h_1), \ldots, F(x_{n_x} + h_{n_x}))$ for $\mathbf{h} \in \mathbb{R}^{n_x}$ have the same distribution for any $x_1, \ldots, x_{n_x}$. The covariance matrix must satisfy the following conditions:

$$Cov[g(x_1), g(x_2)] = K(x_1 - x_2), \tag{2.77}$$

Where $K(\cdot)$ is the covariance- or "autocovariance" function of the process. Equation (2.77) means that the covariance function of a stationary Gaussian Process only depends on the distance $x_1 - x_2$. Furthermore, the constant



(a) Prior (b) Posterior

Figure 2.22: Trajectories sampled from the prior distribution (a) and trajectories sampled from the posterior distribution (b) where two observations have been added. For both illustrations the correlation length is $l_c = 0.6$. The blue area indicates the 95% confidence area, and for the right figure the mean is given in red.

variance of a stationary stochastic process F can be expressed in terms of the covariance function $Var[g(x)] = Cov[g(x), g(x)] = K(0)$. The correlation is then given as:

$$Cor[g(x_1), g(x_2)] = K(x_1 - x_2)/K(0). \tag{2.78}$$

Equation (2.77) means that points with the same inter-point distance between x_1 and x_2, and with the same orientation will have the same covariance. Further details about the properties of correlation functions fall outside the scope of this work and can be found in [198]. However, there exists a useful way of creating correlation functions given a collection of known correlation functions. Suppose that there are two valid correlation functions $K_1(\cdot)$ and $K_2(\cdot)$ then their product $K_1(\cdot) \times K_2(\cdot)$ is also a valid correlation function, as is any convex combination. Correlation functions that are the product of one-dimensional marginal correlation functions are called separable correlation functions.

Gaussian correlation function

This correlation function is based on the normal Gaussian density and is one of the most commonly used correlation functions. The separable Gaussian correlation function is given by:

$$K(x_1, x_2, l_c) = \exp\left(-\frac{1}{2}\left(\frac{||x_1 - x_2||_2}{l_c}\right)^2\right), \tag{2.79}$$

with $||\cdot||_2$ the Euclidean distance and l_c the correlation length, which is a hyper-parameter controlling the rate of correlation.

Exponential correlation functions

The Gaussian correlation function can be seen as a special case of the power exponential correlation function. The function:

$$K(x_1, x_2, l_c) = \exp\left(-\frac{1}{2}\left(\frac{||x_1 - x_2||_2}{l_c}\right)^p\right), \tag{2.80}$$

is said to be a power exponential correlation function for $l_c \geq 0$ and $0 < p \leq 2$.

Whittle- Martérn family of correlation functions

In the thesis of Bertil Martérn [141, 225] a correlation function was introduced that was deduced to be the most general representation of a correlation function. The general form of the Whittle-Martérn correlation function is given as [76]:

$$K(x_1, x_2, l_c, \nu) = \frac{2^{\nu-1}}{\Gamma(\nu)}\left(\sqrt{2\nu}\frac{||x_1 - x_2||_2}{l_c}\right)^{\nu} \mathcal{K}_v\left(\sqrt{2\nu}\frac{||x_1 - x_2||_2}{l_c}\right), \tag{2.81}$$

with $\nu \geq 1/2$ the "smoothness" parameter, Γ Euler's Gamma function, and $\mathcal{K}_v$ is the modified Bessel function of the second kind. It can be shown that for different values of the "smoothness" parameter ν the single and squared exponential correlation functions are obtained [76]. The most commonly used Martérn correlation functions are the analytical solutions derived for $\nu = 3/2$ and $\nu = 5/2$. The analytical expression of the Martérn 3/2 correlation function is defined as:

$$K(x_1, x_2, l_c, \nu = 3/2) = \left[1 + \sqrt{3}\frac{||x_1 - x_2||_2}{l_c}\right]\exp\left(-\sqrt{3}\frac{||x_1 - x_2||_2}{l_c}\right). \tag{2.82}$$

The Martérn 5/2 correlation function is defined as:

$$K(x_1, x_2, l_c, \nu = 5/2) = \left[1 + \sqrt{5}\frac{||x_1 - x_2||_2}{l_c} + \frac{5}{3}\left(\frac{||x_1 - x_2||_2}{l_c}\right)^2\right] \exp\left(-\sqrt{5}\frac{||x_1 - x_2||_2}{l_c}\right). \tag{2.83}$$

An interesting property of this correlation function is that it is $\lceil \nu - 1 \rceil$ times differentiable, with $\lceil \cdot \rceil$ the ceiling function for when $\nu = 1/2$ [132, 190]. A comment often made about the exponential correlation function is that it is infinite differentiable, which is not common for functional relations in engineering practice. Therefore, this function might possibly be too smooth.

The interested reader is refereed to [3] for more details regarding the different covariance functions in Gaussian processes. In this thesis, two well-known covariance functions are used: the Gaussian kernel, as defined in Equation (2.80), and the Matérn $\frac{5}{2}$ kernel, as defined in Equation (2.83).

2.7.2 Predictions made by noise free Gaussian processes

The main aim of using a GP is to make a prediction $\hat{\mathcal{M}}(x)$ of the underlying model output $\mathcal{M}(x)$, at a new unexplored point $x \in \mathcal{X}$, based on noise-free observations of the process in $\mathbf{y}_{\text{DOE}} = \{y^1_{\text{DOE}} = \mathcal{M}(x^1_{\text{DOE}}), \ldots, y^{n_y}_{\text{DOE}} = \mathcal{M}(x^{n_n}_{\text{DOE}}) \}$. This type of GP is also known as the best linear unbiased predictor [58, 198] or universal Kriging within the geo-statistics community. Here, the predictions are based on the Gaussian assumption, which states that the prediction at an unobserved point $y = \hat{\mathcal{M}}(x)$ follows a conditional Gaussian distribution based on the observed responses y_{DOE} [58]:

$$\begin{Bmatrix} \hat{\mathcal{M}}(x) \\ \mathbf{y}_{\text{DOE}} \end{Bmatrix} \sim \mathcal{N}_{n_{\text{DOE}}+1} \left(\begin{Bmatrix} f^T(x)\boldsymbol{\beta} \\ \mathbf{F}\boldsymbol{\beta} \end{Bmatrix}, \sigma^2 \begin{Bmatrix} 1 & \mathbf{k}^T(x) \\ \mathbf{k}(x) & \mathbf{K} \end{Bmatrix} \right), \tag{2.84}$$

with the covariance matrix $\mathbf{K}$ the covariance for all points in the domain, $\mathbf{F}$ the matrix of the observed trend, and $\mathbf{k}(x)$ a vector of cross-correlations between predicted point x and observed points $\mathbf{x}_{\text{DOE}}$ and f, σ^2, $\boldsymbol{\beta}$ correspond to Equation (2.76). Conditional on the observed data, the mean and variance of the Gaussian process can be estimated by [198]:

$$\mu_{gp}(x) = f^T(x)\hat{\boldsymbol{\beta}} + \mathbf{k}^T(x)\mathbf{K}^{-1}(\mathbf{y}_{DOE} - \mathbf{F}\hat{\boldsymbol{\beta}}), \tag{2.85}$$

$$\sigma^2_{gp}(x) = \sigma^2 \left(1 - \mathbf{k}^T(x) + \mathbf{u}^T(x)(\mathbf{F}^T\mathbf{K}^{-1}\mathbf{F})^{-1}\mathbf{u}(x)\right), \tag{2.86}$$

and with

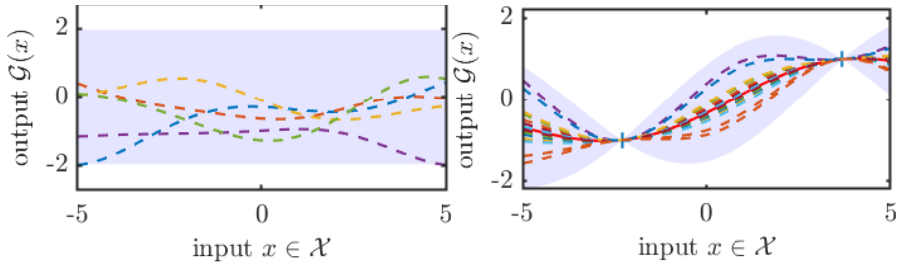
$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^T\mathbf{K}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{K}^{-1}\mathbf{y}_{DOE} \tag{2.87}$$

the general least-squares estimate of $\boldsymbol{\beta}$ and

$$\mathbf{u}(x) = \mathbf{F}^T\mathbf{K}^{-1}\mathbf{k}(x) - f(x). \tag{2.88}$$

Equations (2.85) and (2.86) are referred to as the mean and variance of the GP predictor, respectively. The GP described in this section is an interpolating

GP, which means that the prediction of the variance at an experimental point $\mathbf{x} \in \mathbf{x}_{DOE}$ tends to zero. This is also shown in Figure 2.23 where the prior 2.23a and posterior distributions 2.23b of the GP are shown. These figures show the mean, in red, and 95% CI as a blue area, including a number of realisations of the GP in dashed lines.



(a) Prior (b) Posterior

Figure 2.23: Trajectories sampled from the prior distribution (a) and trajectories sampled from the posterior distribution (b) . Similar as in Figure 2.22. However the correlation length parameter is changed to $l_c = 3$. The blue area indicates the 95% confidence area, and for the right figure the mean is given in red.

2.7.3 Predictions made by a noisy Gaussian process

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

$$\mathcal{M}(\mathbf{x}) = m(\mathbf{x}) + \zeta, \tag{2.89}$$

where the model output $m(\mathbf{x})$ is contaminated with an additional noise term ζ. It is convenient to assume that this noise follows a zero-mean Gaussian distribution:

$$\zeta = \mathcal{N}(0, \Sigma_{\text{gp}}), \tag{2.90}$$

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

$$\Sigma_{\text{gp}} = \sigma_{\text{gp}}^2 \mathbf{I}, \tag{2.91}$$

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:

$$\Sigma_{\text{gp}} = \text{diag}(\sigma_{\text{gp}}^2), \tag{2.92}$$

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 matrix Σ_{gp} where for each observation a different noise variance can be obtained and correlations of this noise are possible. The work presented here is limited to homoscedastic noise. In other words, it is assumed that all observed responses have the same noise variance without any underlying correlations. For this specific case of homoscedastic noise, the total GP variance is given as:

$$\sigma^2_{\text{total}} = \sigma^2_{\text{gp}} + \sigma^2, \tag{2.93}$$

with σ^2 from Equation (2.76), and:

$$\tau = \frac{\sigma_{\text{gp}^2}}{\sigma^2_{\text{total}}}, \tag{2.94}$$

which will allow us to estimate the noise variance of the GP-model. Here the GP is again calibrated on an initial design of experiments $\mathbf{x}_{DOE}$ obtained from, e.g., Latin hyper-cube sampling and their observed results $\mathbf{y}_{DOE}$. The posterior Gaussian distribution of the unobserved point $\hat{\mathcal{M}}(x)$ is then given by:

$$\begin{Bmatrix} \hat{\mathcal{M}}(x) \\ \mathbf{y}_{\text{DOE}} \end{Bmatrix} \sim \mathcal{N}_{n_{\text{DOE}}+1} \left(\begin{Bmatrix} f^T(x)\boldsymbol{\beta} \\ \mathbf{F}\boldsymbol{\beta} \end{Bmatrix}, \sigma^2_{\text{total}} \begin{Bmatrix} (1-\tau)\ \tilde{\mathbf{k}}^T(x) \\ \tilde{\mathbf{k}}(x) \quad \sigma^2\tilde{\mathbf{K}} \end{Bmatrix} \right). \tag{2.95}$$

For the noisy GP the cross-covariance vector between the predicted point x and observed points $\mathbf{x}_{\text{DOE}}$ including the noise is given as:

$$\tilde{\mathbf{k}} = (1-\tau)\mathbf{k} \tag{2.96}$$

and the covariance matrix is given as:

$$\tilde{\mathbf{K}} = (1-\tau)\mathbf{K} + \tau\mathbf{I} \tag{2.97}$$

The mean and variance of a universal Gaussian process with homoscedastic noise conditional on the observations are estimated by [198]:

$$\mu_{gp}(\mathbf{x}) = \mathbf{f}^T(\mathbf{x})\hat{\boldsymbol{\beta}} + \tilde{\mathbf{k}}^T(\mathbf{x})\tilde{\mathbf{K}}^{-1}(\mathbf{y}_{DOE} - \mathbf{F}\hat{\boldsymbol{\beta}}), \tag{2.98}$$

$$\sigma^2_{gp}(\mathbf{x}) = \sigma^2_{\text{total}} \left(1 - \tilde{\mathbf{K}}^T(\mathbf{x})\tilde{\mathbf{K}}^{-1}\tilde{\mathbf{k}}(\mathbf{x}) + \mathbf{u}_c^T(\mathbf{x})(\mathbf{F}^T\mathbf{K}^{-1}\mathbf{F})^{-1}\mathbf{u}_c(\mathbf{x})\right), \tag{2.99}$$

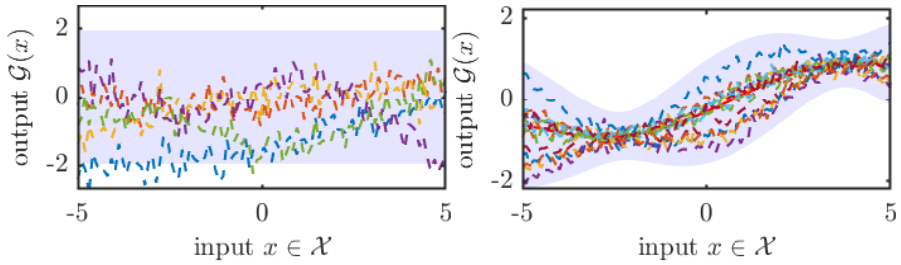
with $\mathbf{F}$ the matrix of the observed trend, with:

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^T\tilde{\mathbf{K}}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\tilde{\mathbf{K}}^{-1}\mathbf{y}_{DOE}, \tag{2.100}$$

the general least-squares estimate of regression coefficients $\boldsymbol{\beta}$ and

$$\mathbf{u}_c(\mathbf{x}) = \mathbf{F}^T\tilde{\mathbf{K}}^{-1}\tilde{\mathbf{k}}(\mathbf{x}) - \mathbf{f}(\mathbf{x}). \tag{2.101}$$

Equations (2.98) and (2.99) are referred to as the mean and variance of the GP predictor, respectively. The parameters of the GP, e.g., $\boldsymbol{\beta}$,σ^2, as well as the hyper-parameters of the covariance kernel, e.g., l_c, are optimised using maximum likelihood estimation, which is further explained in the next section. In the case of unknown homoscedastic noise, an additional noise parameter σ^2_{gp} is added to the maximum likelihood estimation [190]. Note that, unlike the noise-free case, the variance of the prediction at an experimental design point $\mathbf{x} \in \mathbf{x}_{DOE}$ does not collapse to zero, and the GP predictor becomes a regression model since it is no longer interpolating through the observations. This effect is also shown in Figure 2.24 with the GP in Figure 2.24a and the posterior in Figure 2.24b. In these figures again the mean is in red, the 95% CI in blue, and a number of realisations are in the dashed lines. Note that unlike Figure 2.23 the observations are not crisp at the points $\mathbf{x}_{\text{DOE}} = \{-2; 3\}$ as there remains an, e.g., observation or process, variance σ_{gp} at the sampled points.



(a) Prior (b) Posterior

Figure 2.24: Trajectories sampled from a noisy GP with the prior distribution (a) and posterior distribution (b). Similar as in Figure 2.22, with a correlation length of $l_c = 3$. The blue area indicates the 95% confidence area, and for the right figure the mean is given in red.

2.7.4 Maximum-likelihood estimation of hyper-parameters

The hyper-parameters, e.g. l_c, used to build a GP model are usually unknown and should be determined based on the information at hand. It is common practice to use an optimisation strategy to estimate these parameters. In the literature, multiple dedicated strategies are described based on the information that is known and the parameters that should be optimised. Currently there are two main hyperparameter estimation methods used in the context of GP modelling, maximum-likelihood estimation and cross validation [10], which can be found in more comprehensive works [58, 190, 198]. In this work only maximum-likelihood estimation is considered.

The idea behind maximum-likelihood estimation is to find the set of hyper-parameters $\mathcal{H} = \{\boldsymbol{\beta}, \sigma^2, l_c\}$ such that the likelihood of the observations in $\mathbf{y}_{\text{DOE}}$ are maximised. The marginal likelihood is based on the evidence $p(\mathbf{y}_{\text{DOE}}|\mathcal{H})$, given by:

$$p(\mathbf{y}_{\text{DOE}}|\mathcal{C}) = \int p(\mathbf{y}_{\text{DOE}}|\mathcal{H})p(\mathcal{H}|\mathcal{C})d\mathcal{H}, \tag{2.102}$$

with $p(\mathbf{y}_{\text{DOE}}|\mathcal{H})$ the likelihood of observing $\mathbf{y}_{\text{DOE}}$ given the parameters $\mathcal{H}$ and $p(\mathcal{H}|\mathcal{C})$ the prior parameter distribution, which is characterised by $\mathcal{C}$. Based on the fact that the GP-model uses a Gaussian prior, an analytical formula can be derived for multiple combinations of hyper-parameters that must be estimated. The interested reader is referred to [190, 198] for more details and proofs.

For noise-free GPs the parameters that should be estimated are: $\boldsymbol{\beta}$, σ^2 and l_c. Note that Equation (2.87) represents the least squares estimate $\hat{\boldsymbol{\beta}}$. In addition, $\hat{\sigma}^2$ is estimated as:

$$\hat{\sigma}^2 = \sigma^2(l_c) = \frac{1}{n_{\text{DOE}}} \left(\mathbf{y}_{\text{DOE}} - \mathbf{F}\boldsymbol{\beta}\right)^T \mathbf{K}^{-1} \left(\mathbf{y}_{\text{DOE}} - \mathbf{F}\boldsymbol{\beta}\right)^T, \tag{2.103}$$

with $\boldsymbol{\beta}$ and $\mathbf{K}$ depending on l_c. The latter parameter l_c in turn is found by means of the maximum likelihood estimation, which for noise-free GP is defined as:

$$\hat{l}_c = \underset{l_c \in \mathcal{D}_{\mathbf{l_c}}}{\text{argmin}} \frac{1}{2} \left[\log\det(\mathbf{K}) + n_{\text{DOE}} \log(2\pi\hat{\sigma}^2) + n_{\text{DOE}}\right], \tag{2.104}$$

with $\mathcal{D}_{\mathbf{l_c}}$ a set bounding the possible values for l_c.

In the case of unknown homoscedastic noise variance, the hyper-parameters are obtained in a similar fashion, where now the covariance matrix $\mathbf{K}$ is replaced by $\tilde{\mathbf{K}}$ as defined in Equation (2.97). Furthermore, the noise ratio τ should also be estimated with the maximum-likelihood estimation for this case defined as:

$$\hat{\mathbf{l}}_c, \hat{\tau} = \underset{\mathbf{l_c} \in \mathcal{D}_{\mathbf{l_c}}, \tau \in [0;1]}{\text{argmin}} \frac{1}{2} \left[\log\det(\tilde{\mathbf{K}}) + n_{\text{DOE}} \log(2\pi\hat{\sigma}^2_{\text{total}}) + n_{\text{DOE}}\right], \tag{2.105}$$

with the parameter τ bounded within $[0;\ 1]$, and $\hat{\sigma}^2_{\text{total}}$ the estimated total variance, which is estimated as stated in Equation (2.103) where $\mathbf{K}$ is replaced by $\tilde{\mathbf{K}}$. The set of optimal parameters is then identified by finding the minimum of Equation (2.104) or Equation (2.105), which can be accomplished using various optimisation techniques, e.g., Interior point methods [233], or Differential Evolution [185]. The choice of an optimisation algorithm in this context depends on the complexity of the underlying problem and the associated computational costs.

2.7.5 Optimisation using a Gaussian Process

As highlighted in Figure 2.20, a GP can replace the complex and difficult to evaluate numerical model. The GP is analytical and is therefore relatively easy to evaluate. Therefore, in this work a GP is used as a meta-model in optimisation. It should be clear that this section does not cover a general discussion regarding optimisation approaches, as this falls outside the scope of this work. The interested reader is referred to [104] for a more comprehensive overview and practical considerations. The focus of this section is on global optimisation approaches applicable to the non-linear nature of crash analysis. Furthermore, as the numerical model of the crash case behaves as a stochastic black-box function only a limited set of optimisation strategies can be used. Generally speaking, there are two approaches for this type of problem *meta-modelling* strategies, or *evolutionary algorithms*. The former uses a meta-model as a surrogate to speed up the search algorithm in the optimisation process, e.g. using a GP. The latter uses a large population, i.e., samples, which slowly evolve towards a global minimum, e.g., Differential Evolution (DE) [185]. In this work, an efficient global optimisation approach is proposed, capable of identifying the most robust design with a minimum number of evaluations. As this approach is based on the idea of EGO the rest of this section will introduce EGO as introduced in [118].

Efficient global optimisation

Introduced in [118] Efficient Global Optimisation EGO is a strategy to optimise black-box functions by updating a surrogate model in the background. In essence, the idea is to adaptively update a surrogate model by minimising the discrepancy with a physical model $f(\cdot)$ and the surrogate $g(\cdot)$. Hence, the surrogate provides a functional relation between output and input variables. In general, the surrogate is a Gaussian Process GP, as described in Chapter 2.7. Hence, the method starts with an initial DOE to obtain an initially trained GP, which is then enriched by a new point at each iteration. These new points are identified by the expected improvement:

$$E(I(x)) = E(f_{\min} - g(x)), \tag{2.106}$$

where $f_{\min} = \min(\mathbf{y}_{\text{DOE}})$ is the current minimum of the evaluated DOE points and $g(x)$ the surrogate prediction. By using a GP as surrogate model g, the outputs are normally distributed $\mathcal{N}(\mu, \sigma)$. Therefore, the expected improvement is expressed as:

$$E(I(x)) = (f_{\min} - \mu(x))\Phi\left(\frac{f_{\min} - \mu(x)}{\sigma(x)}\right) + \sigma(x)\phi\left(\frac{f_{\min} - \mu(x)}{\sigma(x)}\right) \tag{2.107}$$

with $\phi(\cdot)$ the normal PDF and $\Phi(\cdot)$ the cumulative distribution function (CDF) of $\mathcal{N}(0,1)$. An illustration of EGO is given in Figure 2.25 where the GP posterior is shown in the left top, with underneath the expected improvement. For the left hand figure the expected improvement is the highest at $x = 0$, indicated by the blue dot. Therefore, in the following iteration $f(x = 0)$ is evaluated and the GP is updated with this information. The expected improvement on the right side shows now that the next point that promises to be the largest improvement is $x = -5$. Note that the absolute value of the expected improvement is lower in the right figure, which will decrease further as the minimum is approached. Variations of expected improvement were introduced [115, 134, 168] for different needs, with a specific application on interval variables in [52].

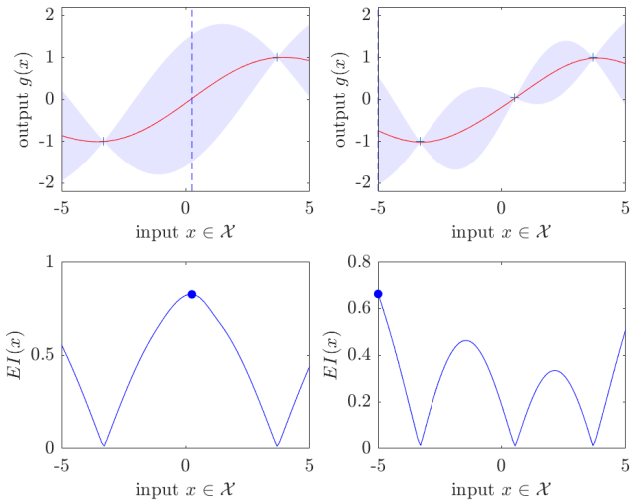


Figure 2.25: Illustration of Efficient Global Optimisation with on top the GP posterior distribution and on the bottom the expected improvement; the blue dot indicates the point with maximum expected improvement indicated by the dashed blue line in the top figures

Chapter 3

Inhomogeneous interval fields based on scaled inverse distance weighting interpolation

This chapter was previously published as:
C. van Mierlo et al. “Inhomogeneous interval fields based on scaled inverse distance weighting interpolation”. In: *Computer Methods in Applied Mechanics and Engineering* 373 (2021). Publisher: Elsevier, p. 113542

3.1 Abstract

This chapter introduces a novel method to model non-deterministic quantities based on experimental measurement data. The focus of this work is on quantities that vary over a continuous domain, e.g., material properties, time-dependent strain rate effects, or stress-strain curves. These quantities are modelled by means of the recently introduced concept of interval fields. An interval field defines intervals that are defined throughout the continuous domain and have dependence in this domain by expanding them over a set of basis functions, describing the spatial nature of the non-determinism of the modelled quantities. One of the more intuitive concepts of defining basis functions in an interval field

is through inverse distance weighting interpolation (IDW), which starts from known intervals at specific control points within the domain. For each of these control points, a corresponding basis function is defined, the relative weight of which is decreasing inversely with the distance. Through this definition, all intervals have non-vanishing basis functions throughout the model domain. This makes the application of standard IDW extremely challenging when the interval uncertainty varies inhomogeneously over the domain, i.e., when local effects are present in the model.

Therefore, in this work standard IDW is adapted by changing the distance measure. More specifically, the weight of intervals is increased locally, while diminishing the weight in other regions. For this purpose, a function is introduced that maps the domain to a higher dimension feature space, in which the distances that determine the weight are measured. This mapping function is based on either the size of the intervals at the control points or experimental data, which both yield additional control resulting in increased agreement with experimental data.

This chapter demonstrates that this method outperforms standard IDW in controllability, while limiting the number of control points. This is illustrated in three case studies: a first case concerning modelling local non-determinism; a second case where a mix of global and local effects is modelled; and the third case, where the interval field is based on experimental stress strain curves. In all these cases, multiple configurations demonstrate the effects of the parameters, and how the new technique is applied. The proposed technique outperforms standard IDW in all three case studies, with an increased coefficient of determination, R^2, between 22% and 56%, in comparison to standard IDW.

3.2 Introduction

In practice, engineers are faced with the task of designing functional components, that should be capable of performing under challenging situations. The performance of these components is mainly assessed by means of numerical approximations of sets of differential equations describing the physical behaviour of the component. However, in many cases the parameters that govern these equations are only known vaguely, with limited information about how these values are distributed, or what their exact value is. The main reasons for this are: the corresponding quantities are inherently variable, e.g. wind loads, or there is incomplete knowledge about the quantity, e.g. direct measurement is challenging, or a combination of both [84]. In order to account for these vague or non-deterministic quantities in engineering practice, large safety factors are

used, which might introduce a large degree of conservatism to ensure reliable performance. To account for these non-deterministic quantities, numerous techniques have been introduced during the last decades. Typically these techniques are categorised as probabilistic [213] and possibilistic approaches such as: intervals [68], fuzzy sets [100], and imprecise probabilities [19]. Where probabilistic methods describe non-determinism as the likelihood that parameters assume a value via a joint probability density function, possibilistic methods as, i.e. interval methods, consider non-deterministic quantities to be bounded.

A recently introduced interval method to distribute properties spatially is the framework of interval fields, which can be regarded as a possibilistic counterpart to random fields [224] for quantities that are spatial or time dependent [158]. Following this framework of interval fields, locally defined intervals are expanded through the model domain based on a set of a priori defined basis functions. Multiple definitions of basis functions can be found in literature, which are based on inverse distance weighting [71], affine arithmetic [204, 208, 211], radial basis functions [112], a spatial averaging method [234], or set-theoretical approaches [116, 176]. The main focus of this work is on basis functions defined using inverse distance weighting or, Shepard's method, as described in [200]. In this article, Inverse Distance Weighting interpolation (IDW) is proposed as a method to interpolate two-dimensional irregularly spaced data, that could be applied in the fields of, e.g. meteorology, geology and urban planning. All these disciplines have in common that information is collected at specific locations, e.g. weather stations, and needs to be distributed towards a larger area, e.g. a city, or state. IDW is applied to construct the basis functions for defining interval fields in [51], where it is assumed that information about the non-deterministic quantity is obtained at specific locations. This local information about the non-determinism is reflected by the intervals at these points, referred to as control points. This local definition of intervals makes IDW an intuitive technique that is capable of representing non-determinism in different cases, as described in [71].

However, when considering inhomogeneous interval fields (i.e. interval fields that are constructed from intervals with large size variations at the control points), IDW suffers from several issues, which make it incapable of representing the observed non-determinism [150]. This is problematic when real-life cases such as local damage, gradual increasing variation of model parameters (e.g., time sensitive parameters such as creep parameters) or stress-strain measurements are to be modelled using the interval field framework.

One of the shortcomings of IDW is that it neglects the *direction* to other points, as reported in [200]. The effect of not taking the direction into account is that the basis functions are non-vanishing in the continuous domain. Thus,

even when other points are crossed in a straight line, the interval defined at the first control point still has an influence. This is counter-intuitive, as one would expect that after crossing other control points the effect should be negligible. This problem is inherent to the definition of the basis functions that are inversely proportional to a distance measure, computed to all other points within the model domain. Thus, when a locally increased amount of non-determinism is modelled, a larger interval is placed at the corresponding control point, which affects the entire interval field. The *effect* that is noticed depends on the information that is placed at the control points of the interval field, and can manifest itself in unpredictable manners. Usually this occurs in locations further away from the control points, as in these locations the weights of different control points all have an influence.

To overcome these shortcomings, this work introduces an adaptation of standard IDW, by measuring the distance between elements in the geometrical domain in a feature space. This is accomplished by an explicit mapping function that transforms the model domain into a higher-dimensional feature space, which changes the distance between control points. The key idea here is to use a set of experimental data to define this mapping function. In addition, when high dimensional data about the quantities of interest is available throughout the domain the spatial nature of the interval field can be matched to these measurements. For this work the focus lies on one-dimensional data, as measurements of two-dimensional data remain challenging even with the use of advanced techniques, such as Digital Image Correlation (DIC) [9], which can be used in conjunction with the virtual fields method [184], as applied in [75]. Therefore, the method is demonstrated on three case studies, including a one-dimensional example based on a real set of stress-strain curves. The structure of the paper is as follows: Section 2 provides the reader with the necessary background in interval fields. In Section 3, a detailed description of the methodology is given, while in Section 4 the optimisation procedure is described for obtaining a data-based interval field. The performance of the method and examples of the obtained results are provided in Section 5, with final remarks and conclusions given in Section 6.

3.3 Interval Field analysis

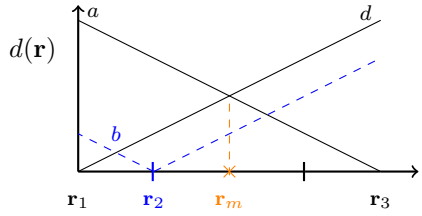
Omitted to avoid redundancy, see Chapter 2 Section 2.5.2

3.4 Scaled basis functions

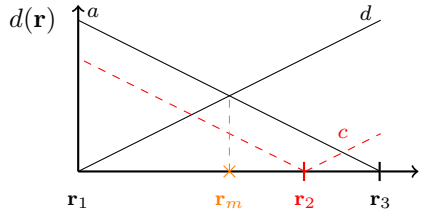
The classical definition of Inverse Distance Weighting (IDW) interpolation as introduced in Section 2.5.2 has a major shortcoming, which is attributed to the non-vanishing weights used to create the basis functions in Equation (2.38). The weights, defined in Equation (2.39) depend on a power p and a distance measure $d(\cdot)$ measured from the control points to all points within the geometrical domain Ω. Therefore, at each point within the domain at an identical distance from the control point has an equal weight, which causes undesired 'artefacts' between the control points. These artefacts depend on the location of the control point and the size of the interpolated values, which may cause the realisation of the interval field to *inflate* or *deflate* in particular situations. When the realisations of the interval field experience these undesired artefacts, only limited adjustments can be made by the analyst in order to mitigate these effects in the standard definition of IDW. These effects are best illustrated when the weights are calculated at a query point $\mathbf{r}_m$ (which is not a control point in the interval field) for two configurations of co-linear points. This is illustrated in Figure 3.1 where in the two configurations 3.1a and 3.1b the control point $\mathbf{r}_2$ is placed in two different locations: between control point $\mathbf{r}_1$ and query point $\mathbf{r}_m$, and between the query point $\mathbf{r}_m$ and $\mathbf{r}_3$, respectively. It would be expected that the location of this control point influences the weight of the other control points at the query location $\mathbf{r}_m$. However, as the distance between $\mathbf{r}_1$ and $\mathbf{r}_3$ to $\mathbf{r}_m$ is identical the weight of these points is equal at the query point, as illustrated by the weights w_a and w_d in illustration 3.1c. Hence, this is counter-intuitive as it would be expected that placing the control point between $\mathbf{r}_m$ and $\mathbf{r}_3$ would lower the influence of $\mathbf{r}_3$ at the query point. This is described as neglecting the *direction* by Shepard in his original paper [200], as it is clear that the point $\mathbf{r}_2$ is crossed on the path from $\mathbf{r}_m$ to $\mathbf{r}_3$ in configuration b.

In addition, if one assumes that the interval at $\mathbf{r}_3$ is significantly larger compared to the intervals at $\mathbf{r}_1$ and $\mathbf{r}_2$ and when IDW interpolation is used as described in Equation (2.38) at $\mathbf{r}_m$, then the interpolated value is larger than the values at $\mathbf{r}_1$ and $\mathbf{r}_2$ in both configurations. This effect or *inflation* can be counter-intuitive and undesired, especially in the configuration in Figure 3.1b where a control point $\mathbf{r}_2$ is placed between the control points $\mathbf{r}_m$ and $\mathbf{r}_3$. Moreover, the example in Figure 3.1 only uses three control points in each configuration, which would be very low for a case with industrial relevance. Nevertheless, it is clear that the problems associated with these non-vanishing weights become more pronounced and harder to mitigate with each control point that is added, as each control point has a non-vanishing weight throughout the domain.

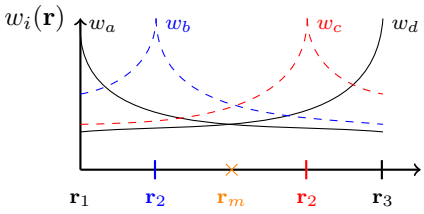
To overcome the shortcomings associated with these non-vanishing weight functions in standard IDW, a new framework for constructing interval fields



(a) Distance measures with control point $\mathbf{r}_2$ placed between $\mathbf{r}_1$ and $\mathbf{r}_m$



(b) Distance measures with control point $\mathbf{r}_2$ placed between $\mathbf{r}_m$ and $\mathbf{r}_3$



(c) Weight functions based on the location of the control point $\mathbf{r}_2$ in red and blue

Figure 3.1: Effect of changing the location of the control point on the weight functions

based on IDW interpolation is introduced in this section. The key idea is that the weight functions are *scaled* with respect to an envelope that represents experimental data or expert knowledge. This *scaling* is performed to alter the influence of certain control points within the domain. Specifically, this scaling is accomplished by transforming the physical domain Ω to a higher dimensional feature space $\Omega \subset \mathcal{F}$. This transformation is performed by means of an explicit scaling function $\mathcal{S} : \mathbb{R}^d \mapsto \mathbb{R}^{d+1}$, which maps Ω to the higher dimensional feature space $\mathcal{F}$. The idea behind computing the distance measures in the feature space $\mathcal{F}$ is illustrated in Figure 3.2, which corresponds to configuration b in Figure 3.1b. In this example, the assumption is made again that the

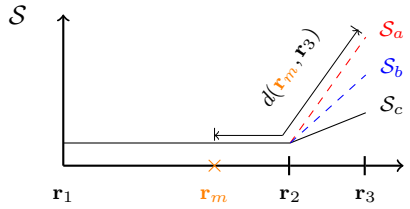
first control points $\mathbf{r}_1$ and $\mathbf{r}_2$ have interval values that are significantly smaller compared to this of the last control point $\mathbf{r}_3$, which leads to the *inflation* of the realisations. In order to reduce the influence of this large interval value at $\mathbf{r}_3$, a scaling in the feature space is performed $\mathcal{S}_{a,b,c} = \mathcal{S}(\mathbf{r}_3)$, illustrated in Figure 3.2a. By changing the location of the control point in the feature space the distance measure of the query point to the control point is changed $d(\mathbf{r}_m, \mathbf{r}_3)$, which is shown in Figure 3.2b where the distance measure increases from c to a, in correspondence with the scaling in Figure 3.2a. By increasing the distance between the query point $\mathbf{r}_m$ and the control point $\mathbf{r}_3$ the weight of the control point decreases at the query point and therefore mitigating the effect of *inflation* of the basis functions without changing the number or location of the control points.

3.4.1 Scaling function based on interval size

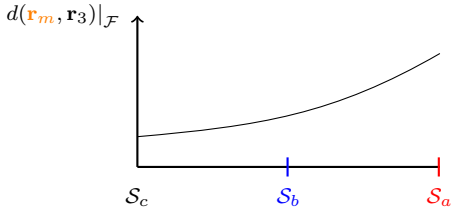
In this section, a more general case is considered where the scaling, as discussed in the introduction of this section, is performed only at the control points, which represent expert knowledge or local measurements. Especially in cases where the interval size is varying considerably within the domain to model localised phenomena, this can reduce the *inflation* or *deflation* of the interval field due to the large intervals. Therefore, the scaling is only performed on the interval size, which is determined through the interval radius Δx. This corresponds to the previously discussed case where very small intervals are placed at $\mathbf{r}_1$ and $\mathbf{r}_2$ and a large interval is placed at $\mathbf{r}_3$. The scaling in Figure 3.2a illustrates the results of Equation (3.1) where intervals with different sizes are placed at an increased distance in the feature space $\mathcal{F}$. Therefore the weight of larger intervals that differ in size from their neighbouring points is reduced, as illustrated in Figure 3.2c. The mapping function associated with this is explicitly defined as:

$$\mathcal{S}_i = \left(\frac{\Delta\alpha_i}{\max(\Delta\alpha)} \right)^t, \tag{3.1}$$

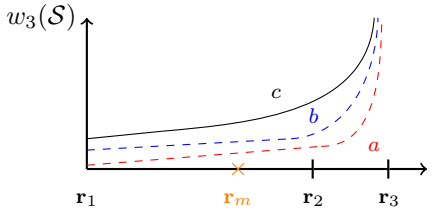
with the parameter $t \in \mathbb{R}_0^+$ being a scaling factor. The index i denotes that this mapping is only made at discrete control points of the interval field. Therefore, before the weight functions are created the distances are measured within the feature space as indicated by Figure 3.2a. From this illustration it is also clear that the path from $\mathbf{r}_2$ to $\mathcal{S}_{a,b,c}$ can be formed in different manners by means of interpolation (e.g. linear-, polynomial-, spline-interpolation, ...). Nevertheless without losing generality, this chapter specifically uses only linear interpolation between scaled control points $\mathcal{S}(\mathbf{r}_i)$ to obtain the scaled value at point $\mathcal{S}(\mathbf{r}_j)$ in



(a) Different magnitudes of scaling control point $\mathbf{r}_3$, indicated by a,b and c



(b) The effect of the different magnitudes of scaling on the distance measure



(c) Weight function of $\mathbf{r}_m$ dependent on the scaling that is used

Figure 3.2: The effect of increasing the magnitude of scaling, labelled c,b and a, of control point $\mathbf{r}_3$ on the distance measure and the corresponding effect on the weight functions

the feature space, which is explicitly denoted as:

$$\mathcal{S}(\mathbf{r}_j) = \mathbf{S}_i \frac{\mathbf{r}_j - \mathbf{r}_{i+1}}{\mathbf{r}_i - \mathbf{r}_{i+1}} + \mathbf{S}_{i+1} \frac{\mathbf{r}_j - \mathbf{r}_i}{\mathbf{r}_{i+1} - \mathbf{r}_i} \quad \forall \mathbf{r}_j \in [\mathbf{r}_i \ \mathbf{r}_{i+1}] \tag{3.2}$$

which is based on first order Lagrange polynomials. Note that these linear interpolation functions will make the basis functions approximately linear functions between control points. Obviously, also other interpolation schemes can be used to replace Equation (3.2).

3.4.2 Distance measures

In order to obtain the weight functions, the Euclidean distance between the control point $\mathbf{r}_i$ and other non-control points $\mathbf{r}_j$ is measured in the feature space $\mathcal{F}$ along the created path. This can be interpreted as measuring the distance over a topological surface described by the feature space $\mathcal{F}$, indicated by the super-script $\mathcal{S}$. The total distance is computed by summing the Euclidean distances between all elements from the control point $\mathbf{r}_i$ to the j^{th} element $\mathbf{r}_j$, denoted as:

$$d^{\mathcal{S}}(\mathbf{r}_i, \mathbf{r}_j) = \sum_{v=i}^{j-1} \left\| (\mathbf{r}_v, \mathcal{S}_v) - (\mathbf{r}_{v+1}, \mathcal{S}_{v+1}) \right\|_2, \quad \text{for } i < j \tag{3.3}$$

These measurements are directly used to determine the weights based on Equation (2.39):

$$w_i^{\mathcal{S}}(\mathbf{r}) = \frac{1}{[d^{\mathcal{S}}(\mathbf{r}_i, \mathbf{r})]^p}. \tag{3.4}$$

Here, the index i selects the appropriate weights for the corresponding i^{th} interval. The final interval field is obtained by implementing these weights in Equation (2.38), which yields following expression:

$$\mathbf{x}^I(\mathbf{r}) = \hat{\mathbf{x}} + \sum_{i=1}^{n_b} \frac{\alpha_i^I w_i^{\mathcal{S}}(\mathbf{r})}{\sum_{j=1}^{n_b} w_j^{\mathcal{S}}(\mathbf{r})}. \tag{3.5}$$

3.5 Data based basis functions

In the previous section it is assumed that information about the non-deterministic quantities is available at discrete points. This section deals with the case where measurements about the spatial nature are available throughout the domain. Specifically, the applicability of the mapping function to the experimental data is investigated in order to obtain a better overall fit with the available experimental data. This is achieved in two steps where first an envelope of experimental data is constructed that is used to define the mapping function, and second, an optimisation is performed to maximise the agreement between the interval field and the experimental data.

3.5.1 Scaling function based on measurement data

In order to define the basis functions to build an interval field that represents a set of experimental data, an envelope $\mathcal{A}^I(\mathbf{r})$ is defined which encapsulates all

experimental data. The envelope $\mathcal{A}^I(\mathbf{r}) : \Omega \mapsto \mathbb{IR}$ is defined as a function of the model domain, and when sampled, the interval bounding the measurements range at this point is returned. This allows for direct measurement of the interval centre point and radius at a location within $\mathbf{r}$, denoted as:

$$\hat{\mathcal{A}}(\mathbf{r}) = \frac{\underline{\mathcal{A}}(\mathbf{r}) + \overline{\mathcal{A}}(\mathbf{r})}{2}, \tag{3.6}$$

$$\Delta\mathcal{A}(\mathbf{r}) = \frac{\overline{\mathcal{A}}(\mathbf{r}) - \underline{\mathcal{A}}(\mathbf{r})}{2}. \tag{3.7}$$

It should be noted that when only $\mathcal{A}^I(\mathbf{r})$ is considered, the spatial dependence that is present in the data is omitted. This effect is especially pronounced when considering non-monotonic spatial functions that are represented by an envelope. Additionally, by decomposing the envelope $\mathcal{A}^I(\mathbf{r})$ into a midpoint and a radius, not only the spatial dependence is decoupled, but the inhomogeneities are decomposed as well. Where in general, both the midpoint and the radius can vary throughout the domain. Therefore, the scaling function in Equation (3.8) is aimed at maximising the difference between the measurements in the data set by combining both effects, and define the scaling on the sum of absolute value of the midpoint with the radius. This scaling based on the envelope of experimental data is explicitly denoted as:

$$\mathcal{S}(\mathbf{r}) = \left(\frac{|\hat{\mathcal{A}}(\mathbf{r})| + \Delta\mathcal{A}(\mathbf{r})}{\max_{\mathbf{r}}(|\hat{\mathcal{A}}(\mathbf{r})| + \Delta\mathcal{A}(\mathbf{r}))} \right)^q, \tag{3.8}$$

with $|\cdot|$ indicating the absolute value, $q \in \mathbb{R}_0^+$ a parameter to scale the influence, similar to t in Equation (3.1), and the max(.) is taken with respect to the vector $\mathbf{r}$ to normalise the scaling.

3.5.2 Optimisation of the basis functions

Scaling the basis functions will not directly provide the optimal agreement between realisations and the interval field. A good agreement is accomplished by optimising the realisations of the interval field towards the data set in a least squares sense, minimising the residual between experimental results and the interval field. The residual is calculated by computing the Euclidean distance $\|\cdot\|_2$ between the radius of the envelope of experimental data $\Delta\mathcal{A}(\mathbf{r})$ and the radius of the interval field $\sum_{i=1}^{n_b} \psi_i(\mathbf{r}_j, p, q)\Delta\alpha_j$ at a selected set of locations $\mathbf{r}_j$, explicitly denoted as:

$$\delta_j(p, q) = \left\| \Delta\mathcal{A}(\mathbf{r}_j) - \sum_{i=1}^{n_b} \psi_i(\mathbf{r}_j, p, q)\Delta\alpha_j \right\|_2. \tag{3.9}$$

Based on this distance metric, an optimisation can be performed based on these residuals by summing them for all elements n in the vector $\mathbf{r}$. Optimisation is performed with respect to the parameters p and q, from equation (3.4) and (3.8). During optimisation these parameters are constrained to ensure that $p > 0$ and $q \geq 0$, which can for instance be accomplished by means of Sequential Quadratic Programming (SQP) [91]. The optimisation problem is explicitly denoted as:

$$\begin{aligned} \min_{p,q} \quad & \sum_{j=1}^{n} \delta_n(\psi(p,q)) \\ \text{s.t.} \quad & p > 0 \\ & w \geq 0 \end{aligned} \tag{3.10}$$

The results shown in this chapter are obtained by initiating both parameters as $p = q = 1$. At this point, all information is provided in order to obtain scaled basis functions, which are used in the case studies in the following section. All the steps needed to obtain an interval field from scaled basis functions are illustrated in Figure 3.3.

This workflow starts from the analyst, placing the control points at the desired locations, and the calculation of the interval midpoint and radius of the envelope. The optimisation loop indicated here is kept general, as multiple constrained optimisation algorithms can be employed. At the end of the optimisation process the optimal parameters p^* and q^* are obtained, and are used to calculate the final interval field using Equation (3.5). Additionally, it should be noted that this optimisation step is also applicable when standard IDW is used, by tuning the p value for optimal correspondence with the data. This will be illustrated in section 3.6.

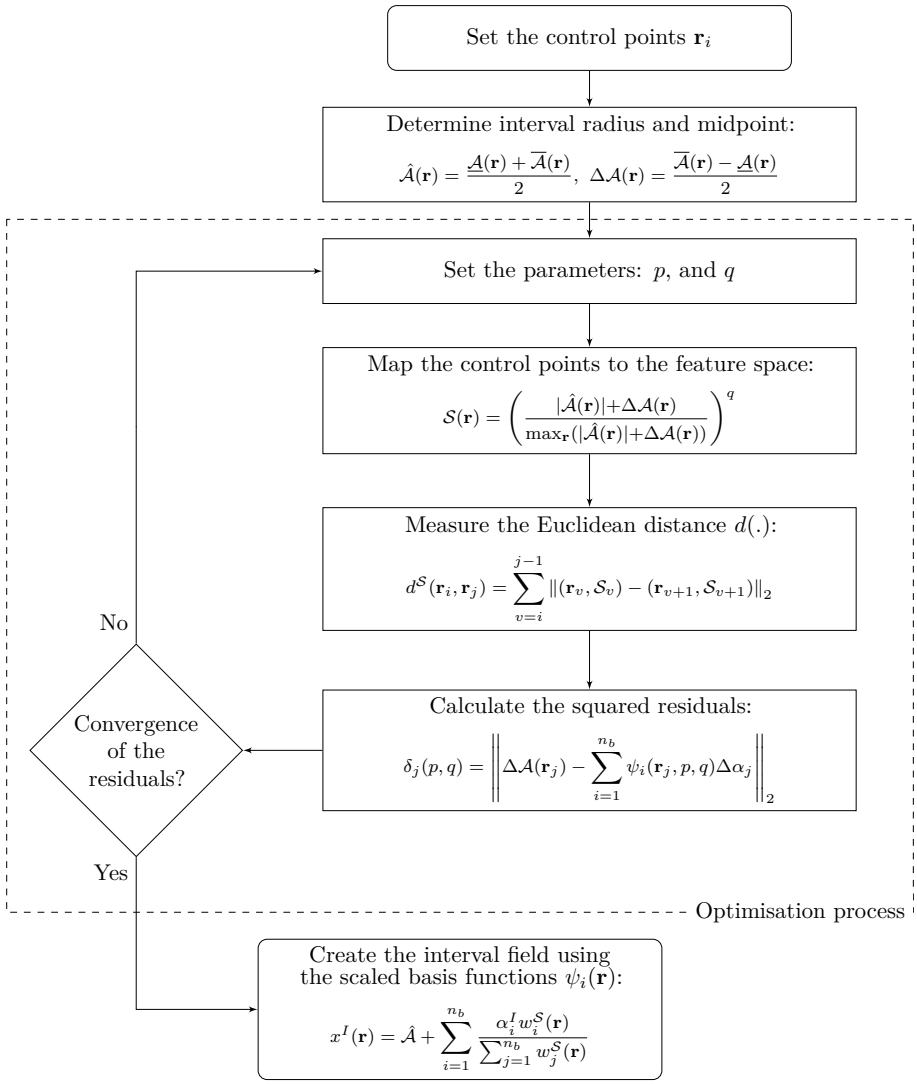


Figure 3.3: Workflow to optimise the interval field with respect to experimental data

3.6 Case studies

In order to demonstrate the proposed approach, three cases studies are defined:

Case 1: locally varying intervals

Case 2: varying both the midpoint and radius

Case 3: interval field model for stress-strain measurements

In the remainder of this section, results are mainly discussed using figures with a specific consistent lay-out, where the weights, basis functions and the upper and lower bound of the interval field are illustrated in rows, respectively. The columns represent different configurations, where additional control points are placed, or different factors are used in equations (3.8) and (2.38), which are determined through optimisation when indicated by an $*$. All parameters that are used to obtain the results are included on the last figure, which shows the realisations of the interval field. Note, that in all figures the weight functions are plotted with a logarithmic y-scale that is truncated at a value of 500, this in order to aid in the comparison of results since theoretically the weights $w \to \infty$ when $r \to r_i$. In addition, to provide quantitative comparison of the results an R^2 value is provided, calculated between the interval field and the experimental data set, denoted as:

$$R^2 = 1 - \frac{\sum_{j=1}^{n} \left(\Delta\mathcal{A}(\mathbf{r}_j) - \sum_{i=1}^{n_b} \psi_i(\mathbf{r}_j)\Delta\alpha_j\right)^2}{\sum_{j=1}^{n}(\Delta\mathcal{A}(\mathbf{r}_j) - \frac{1}{n}\sum_{i}^{n} \Delta\mathcal{A}_i)^2}, \tag{3.11}$$

where the average radius of the envelope is calculated using all data points n. To calculate the R^2 value the explained variation of the interval field is divided by the total variation. Therefore, an R^2 value of 1 tells the analyst that all the variation is explained by the interval field, i.e. perfect agreement is obtained.

3.6.1 Case 1: Locally varying intervals

This case represents local effects that could be caused by, e.g., porosity or cracks, which are modelled by a locally increased amount of uncertainty. This can be represented by an interval field, which describes the spatial distribution of a model parameter that represents this behaviour, e.g. local loss of stiffness. For the sake of illustration, in this case, it is assumed that the model parameter responsible for this behaviour lies in a large interval x_3^I, while this is lower in normal circumstances. These intervals with corresponding locations of the

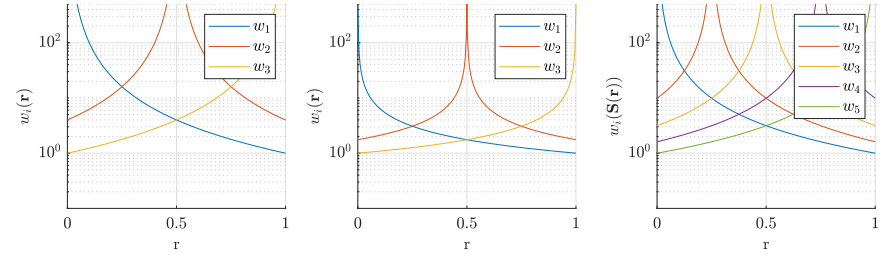
control points are provided in Table 3.1 for this case. The envelope is constructed by means of linear interpolation between all the control points, indicated as a dashed black line in each of the realisations in Figure 3.4. In this figure standard IDW is used in three different configurations: first, using three control points with p set to 2; second, standard IDW where p is determined by means of optimisation, equation (3.9); and third, standard IDW where optimisation is used to determine p using additional control points.

i	1	2	3	4	5
r_i	0	0.25	0.5	0.75	1
$\overline{x}_i^I$	320	320	600	320	320
$\underline{x}_i^I$	290	290	0	290	290

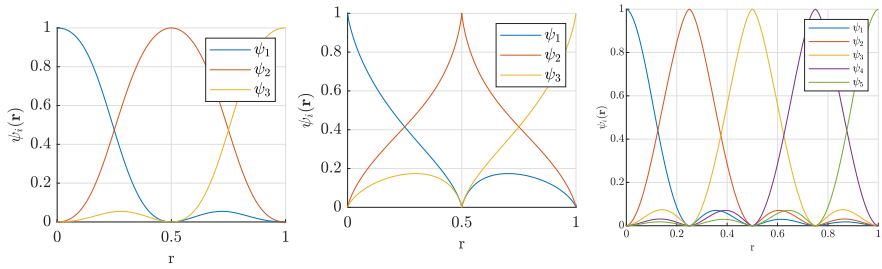
Table 3.1: Intervals and control points used for case 1

Figure 3.4 shows the effects of the power p and adding additional control points on both the basis functions and the weights. It is clear from Figures 3.4a and 3.4b, that lowering the power p localises the influence of an interval around a control point, which also attributes to a higher rate of decline of the weight at this control point. However, due to this rapid decline in weight around the control points, the weight of all intervals increases between the control points, which is best seen by comparing the basis functions of Figure 3.4d and 3.4e. This effect is also responsible for the *inflation* of the realisations of the interval field, illustrated in Figure 3.4i. In the latter case it is clear that the large interval at $r = 0.5$ has an influence on the realisation outside the local area. This effect is mitigated by scaling the basis functions in order to obtain a better agreement between the envelope using a minimal number of control points. This is illustrated in Figure 3.5.

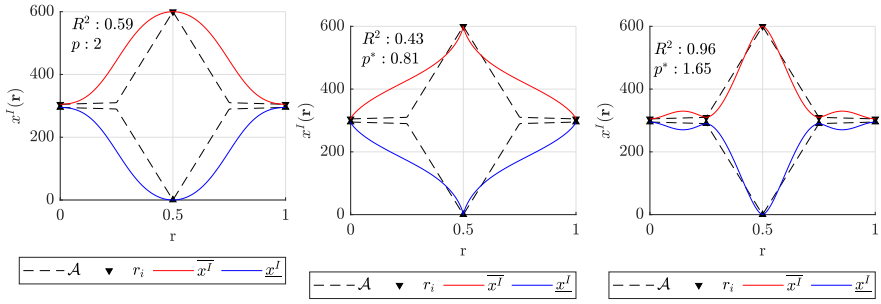
In Figure 3.5 three different configurations of scaled IDW are shown, where in the left column, both p and q are set at an initial value of one, optimisation towards q is performed for the column in the middle, and optimisation towards both p and q is performed in the right column. It is shown in Figure 3.5g that a moderate fit is already obtained by setting q to one, without any means of optimisation. When optimisation is used, a R^2 higher than 0.9 is achieved without the addition of additional control points, as shown in the other configurations, as shown in Figure 3.5h and 3.5i. The effects of using scaled IDW is clearly visible at the weighting functions, which is clearly visible by comparison of the weighting functions in Figure 3.4a and 3.5a, the latter being less *smooth*. This effect is attributed to the sharp transitions of the envelope, which is caused by the use of linear interpolation to construct the envelope, as after this sharp transition, the mapping function provides a sudden large change in distance between points.



(a) Standard weight functions (b) Optimised weight functions (c) Optimised weight functions



(d) Standard basis functions (e) Optimised basis functions (f) Optimised basis functions

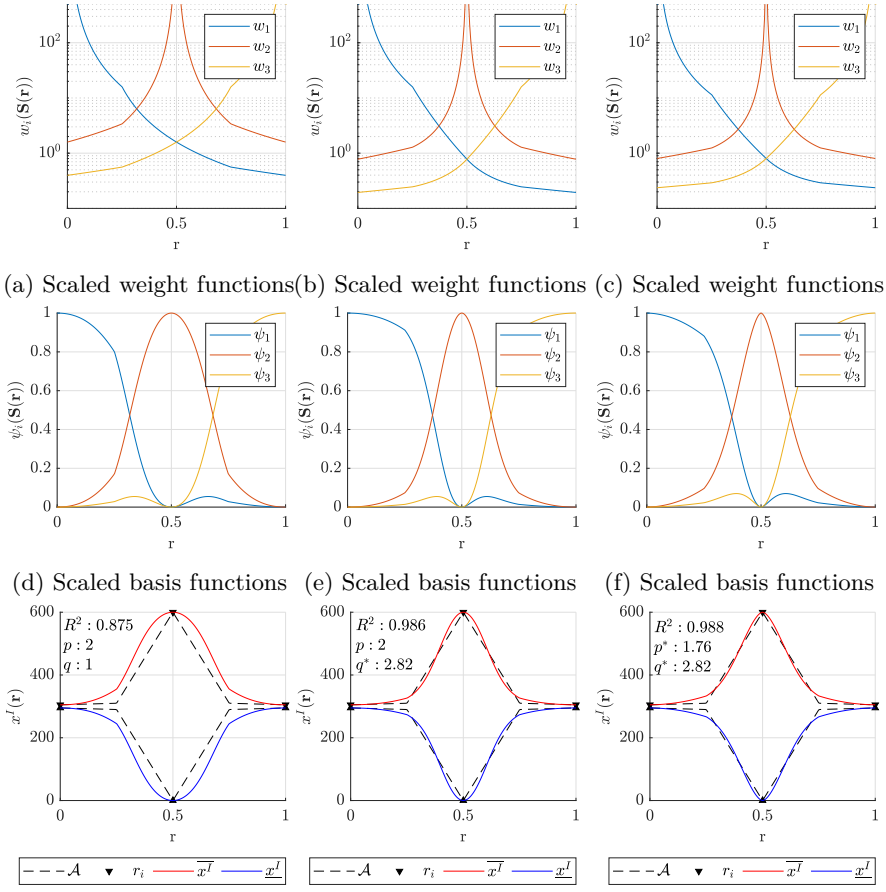


(g) Standard IDW realisations (h) Optimised realisations (i) Optimised realisations

Figure 3.4: Columns representing cases using standard IDW in order to fit the envelope, black dashed line, below with corresponding weight functions, basis functions and the outer realisations indicated in red and blue lines, respectively.

However, as real data in general is smoother, this effect will be less pronounced.

In addition, Figure 3.5a illustrates how scaling of the weights locally decreases the influence where, for instance the weight function w_1 decreases when moving



(a) Scaled weight functions (b) Scaled weight functions (c) Scaled weight functions

(d) Scaled basis functions (e) Scaled basis functions (f) Scaled basis functions

(g) Scaled realisations (h) Scaled realisations (i) Scaled realisations

Figure 3.5: Columns representing cases using scaled IDW in order to fit the envelope, black dashed line, with corresponding weight functions, basis functions and the outer realisations indicated in red and blue lines, respectively.

away from the control point $\mathbf{r}_1$, but when reaching a value of $r = 0.25$ this decrease is amplified. As a result, one can see a difference in the influence of the basis functions in Figure 3.5e, where the influence of $\psi_1(\mathbf{r})$ is almost not decreasing when moving away from the control point $\mathbf{r}_1$, while the influence of $\psi_2(\mathbf{r})$ is more localised, illustrated by the faster decline of the basis function when moving away from the control point $\mathbf{r}_3$. Hence, the realisations obtained

with scaling the weights are in better agreement with the envelope, indicated by the higher R^2 values.

3.6.2 Case 2: varying both the midpoint and radius

For this case, the data consists of different intervals with both a varying interval radius and midpoint. The intervals and corresponding control points used in this case are provided in Table 3.2, which are all used to construct the envelope by means of linear interpolation. The results of this case are illustrated in Figure 3.6, where the left column represents standard IDW with only the control points $i = 1, 2, 5$, scaled IDW is optimised, in the middle and on the right, scaled IDW is used with an additional control point $i = 1, 2, 4, 5$.

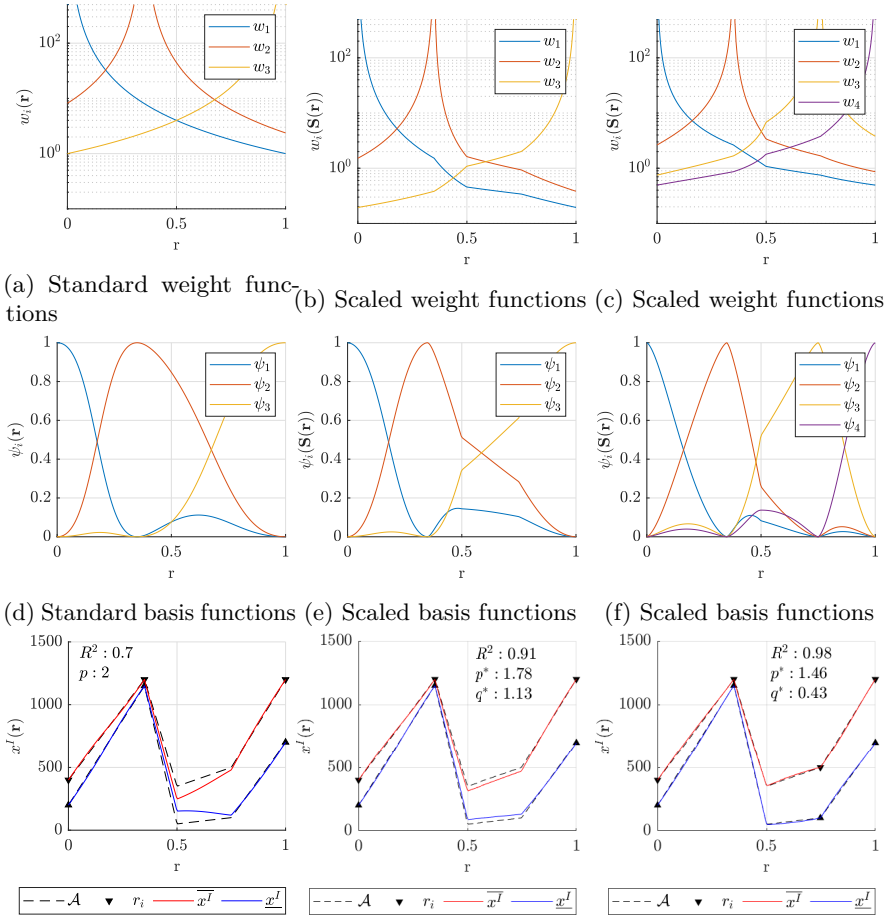
i	1	2	3	4	5
r_i	0	0.35	0.5	0.75	1
$\overline{x}_i^I$	400	1200	350	500	1200
$\underline{x}_i^I$	200	1150	50	100	700

Table 3.2: Intervals and control points used for case 2

This case shows how the information about the radius of the envelope is translated to the weight and basis functions, which is demonstrated by omitting some of the control points that are used to create the envelope from the interval field definition (i.e., $\mathbf{r}_3$ and $\mathbf{r}_4$, are excluded). Here, the envelope is constructed by linear interpolation, and the midpoint of this envelope is used to construct the interval field, according to Equation (2.41). With the use of standard IDW there is a clear discrepancy between the envelope and the interval field at $r = 0.5$, in Figure 3.6g, which can only be mitigated by an additional control point at this location. However, in the second configuration in Figure 3.6h, scaling of the weights based on the midpoint and the radius yields a clear improvement with an R^2 of 0.91. The main reason can be found at the basis functions where a comparison between 3.6a and 3.6b indicates that the influence of the small interval x_2^I is decreased faster, resulting simultaneously in a fast increase of the effect of x_1^I and x_5^I.

In order to increase the coefficient of determination a control point can be added, which also allows the analysis to sample different realisations. However, additional control points will increase the n_b dimensional input space which increases the number of samples to be propagated, i.e., vertex analysis. In addition, for each control point an additional basis function and corresponding weight is needed. This is illustrated in Figure 3.6i, were an additional control point $r_3 = 0.75$ is used to create the interval field. The increased R^2 value of

0.98 clearly shows a large improvement of agreement, and the power of the scaling parameter q is decreased from 1.3 to 0.43 resulting in smoother basis functions as the scaling is less pronounced.



(a) Standard weight functions (b) Scaled weight functions (c) Scaled weight functions

(d) Standard basis functions (e) Scaled basis functions (f) Scaled basis functions

(g) Standard realisations (h) Scaled realisations (i) Scaled realisations

Figure 3.6: Left column representing standard IDW, while the other cases use scaled IDW in order to fit the envelope in the black dashed line below, with corresponding weight functions, basis functions and the outer realisations indicated in red and blue lines, respectively.

3.6.3 Case 3: interval field model for stress-strain measurements

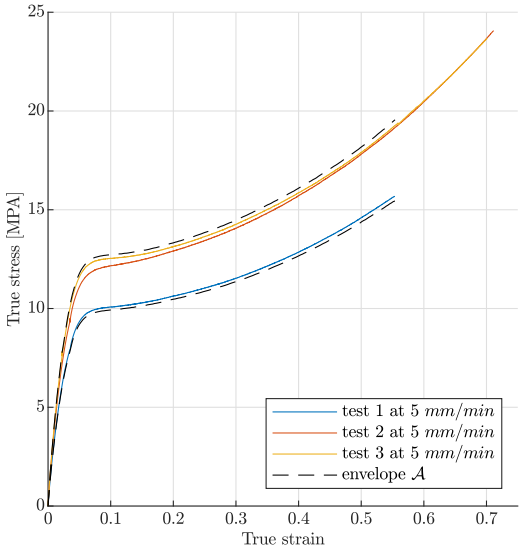


Figure 3.7: Experimentally obtained stress-strain curves with the corresponding envelope, plotted at a 0.1 offset in black dashed lines encapsulating the data

The last case under consideration is based on real stress-strain measurements of tensile tests, performed on dog-bone shaped samples in accordance with ASTM D638 at a speed of 5 mm/min. The samples are produced by means of additive manufacturing, using a FORM 2 stereolithography printer, and are tested on a INSTRON 4204 tensile test machine. The stress-strain curves that were obtained are illustrated in Figure 3.7 with in dashed black lines an envelope $\mathcal{A}^I$, which encloses the stress-strain data. The main goal in this case is now to represent this envelope $\mathcal{A}^I$ by means of an interval field, which was proposed by the authors in [150]. In this case the spatial dependence is provided by the stress-strain curves as they represent the underlying physics of the material that has been tested. Therefore, the envelope $\mathcal{A}^I$ is constructed by the minimum and maximum stress at a specified strain. During the test not all specimen failed at a same amount of strain, therefore, the envelope is constructed up to the strain where the first specimen failed. The envelope is illustrated in Figure 3.7, indicated by the black dashed lines that are plotted at a 10% offset. Note that the underlying assumption is made that these tests are representative

not only for the non-determinism, but also for the spatial nature in the strain domain. The interval midpoint and radius are defined on the envelope trough the following expression:

$$\hat{\sigma}_i = \frac{\underline{\mathcal{A}}(\epsilon_i) + \overline{\mathcal{A}}(\epsilon_i)}{2} \tag{3.12a}$$

$$\Delta\sigma_i = \frac{\overline{\mathcal{A}}(\epsilon_i) - \underline{\mathcal{A}}(\epsilon_i)}{2} \tag{3.12b}$$

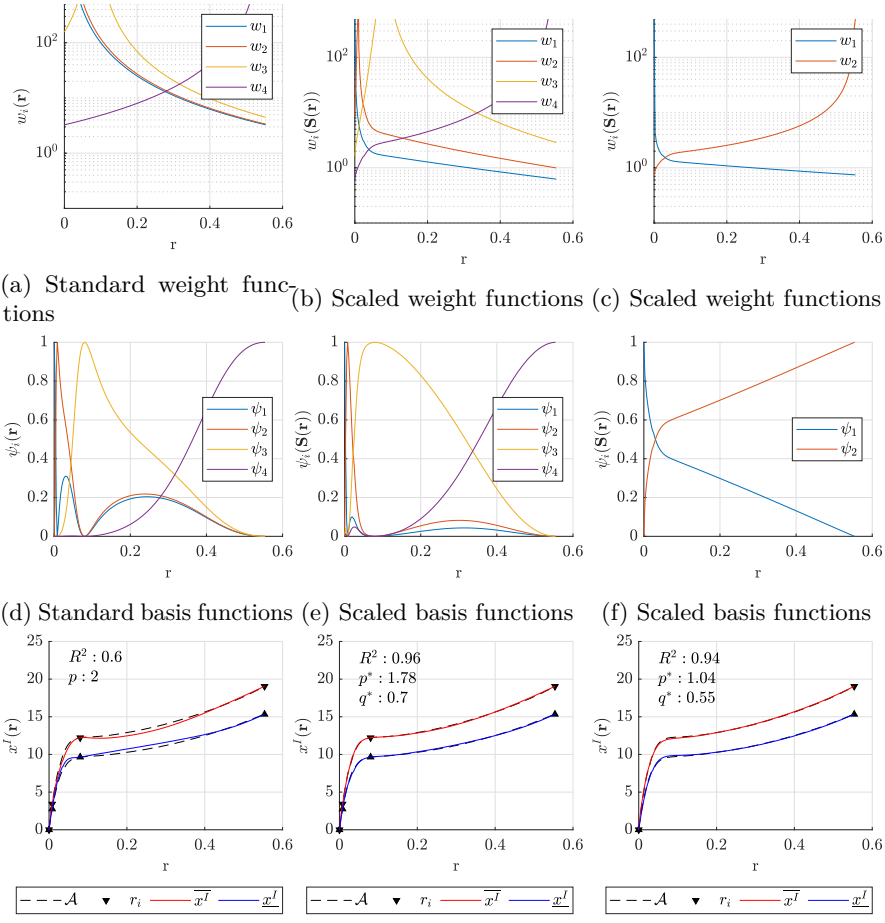
where ϵ denotes the strain, and $\hat{\sigma}$ and $\Delta\sigma$ denote the stress interval midpoint and radius, respectively. Using equation (3.12) the size of the intervals only depends on the location of the control point, denoted by $\mathbf{r}_i$. However, all stress-strain curves should be zero at the origin, which is achieved by keeping the interval $\epsilon = 0$ at zero.

In Figure 3.8, there are three cases illustrated, with on the left, standard IDW using four control points, scaled IDW with four control points in the middle, and finally scaled IDW with only two control points on the right. The specific challenge in this case is the uneven distribution of control points, where the first three control points $\mathbf{r}_{1,2,3}$ are very close to each other. The choice for this configuration is made for mechanical reasons, namely to allow direct control of the initial stiffness and the yield stress. The intervals and control points that were used in this case are provided in table 3.3.

i	1	2	3	4
r_i	0	0.008	0.08	0.75
$\overline{x}_i^I$	0	3.4	12.5	19.3
$\underline{x}_i^I$	0	2.8	9.9	15.7

Table 3.3: Intervals and control points based on the data of case 3

The main problem using standard IDW is that the influence of the intervals at these low strain values remains throughout the domain, which causes the realisations to *deflate*. This *deflation* effect can be traced back to the basis functions where it is clearly seen, in Figure 3.8d, that both ψ_1 and ψ_2 have a large influence between control point $\mathbf{r}_2$ and $\mathbf{r}_3$, and again between control point $\mathbf{r}_3$ and $\mathbf{r}_4$. Especially concerning the latter range, the effect is translated towards the realisations, where the influence of the two small intervals at low strain values *pinches* the outer realisations closer to the midpoint, illustrated in Figure 3.8g. In the second configuration, illustrated in the middle, the scaled basis functions in Figure 3.8e drastically reduce the influence of the first two basis functions, $\psi_1(\mathbf{r})$ and $\psi_2(\mathbf{r})$. Additionally, when the basis functions are



(a) Standard weight functions (b) Scaled weight functions (c) Scaled weight functions

(d) Standard basis functions (e) Scaled basis functions (f) Scaled basis functions

(g) Standard realisations (h) Scaled realisations (i) Scaled realisations

Figure 3.8: Left Column representing standard IDW, while the other cases use scaled IDW in order to fit the envelope of experimental data in the black dashed line below, with corresponding weight functions, basis functions and the outer realisations, indicated in red and blue lines, respectively.

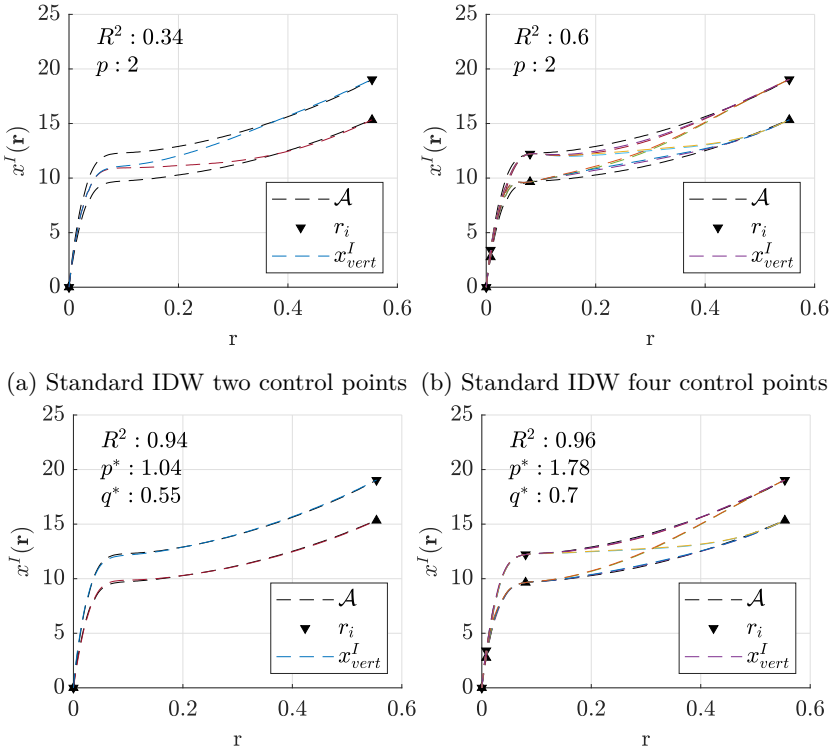
scaled, less control points can be used without loss of fitness, already briefly shown in the previous case study. Here, the number of control points is reduced to two, while still showing a good agreement with the data, as illustrated in Figure 3.8i. In this figure it is illustrated that, compared to the case in

Figure 3.8h with four control points, a similar coefficient of determination, of $R^2 = 0.95$ is obtained, while the optimised parameters are quite different. This is one of the key strengths of this technique where the information of the data is used to increase the agreement with experimental results without using additional control points. Therefore, the size of the n_d dimensional input space can be kept reasonable even for larger models without compromising on spatial resolution. However, with less control points the possible realisations of the interval field reduce as well, where is this case the weight of w_1 diminishes quickly with increasing distance from the control point, and w_2 controls most of the realisations. Therefore, sampling the interval field in Figure 3.8i only yields realisations parallel to the midpoint as the interval field only depends on $\psi_2(\mathbf{r})$, with the first interval kept at zero. This is illustrated in Figure 3.9 where vertex samples of the interval fields in Figure 3.8 are shown in various coloured dashed lines for both standard and scaled IDW. In addition, this figure shows that with identical control points, in the left and middle configuration, the sampled realisations are different, and for the last configuration on the left only two samples are required for a full vertex analysis.

3.7 Conclusion

The framework of interval fields is a both numerically and theoretically convenient concept to represent non-deterministic quantities that have time, space or even space-time dependencies. In addition, the interval field can be used to represent experimental measurements that quantify the non-deterministic value throughout the domain. The general idea behind this concept is that independent interval scalars are expanded though the domain by a priori defined basis functions. Inverse distance Weighting (IDW) provides an intuitive manner of defining these basis functions. However, the realisations obtained through standard IDW generated basis functions are not always satisfactory, even when additional control points are used. Therefore, this chapter introduces a methodology based on the idea that the realisations of the interval field should be limited to an envelope through the domain. This is accomplished by scaling the weight functions such that the corresponding basis functions yield interval fields realisations that always respect this envelope.

The scaling of the distance measure that drives the basis functions is performed by mapping the control points to a higher dimensional feature space, by using an explicit mapping function $\mathcal{S}$, as defined in Equation (3.8). The definition of these mapping functions is either based on interval radius, or the sum of the absolute midpoint with the radius of the envelope. When the mapping is based on the interval radius, interpolation between points in the feature space



(a) Standard IDW two control points (b) Standard IDW four control points

(c) Scaled IDW two control points (d) Scaled IDW four control points

Figure 3.9: Vertex samples of the interval field indicated by $\mathbf{x}^I_{vert}$, for standard IDW in Figure 3.9a and 3.9b and for scaled IDW in Figure 3.9c and 3.9d, both for two and four control points, respectively.

is required, which should be chosen appropriately. The weights of the basis functions are then measured in this feature space.

The envelope that is used to accomplish this scaling can be based on either expert knowledge or a set of experimental measurements. The main concern is that the bounds of the envelope should be representative for the spatial nature of the non-deterministic quantity under consideration, as these bounds are used to scale the basis functions towards the envelope. This is demonstrated in the third case where a set of stress-strain measurements are represented through an interval field. In this case, the envelope clearly represents the underlying non-deterministic physical behaviour of the material at hand. In addition, optimisation is used to increase the agreement between the envelope and the

interval field. This is accomplished by minimising the residual between the envelope and the outer realisations of the interval field, in a least-squares sense. While the method is shown to be effective in reducing the problems associated with classical IDW, the developments are limited to 1-dimensional interval fields only. Further work is aimed at investigating how the mapping to a higher dimensional feature space, as well as the calculation of the necessary distances, can be performed efficiently and effectively.

Chapter 4

Interval field methods with local gradient control

This chapter is a slight improved version of the previously published paper: C. van Mierlo et al. "Interval field methods with local gradient control". In: *Proceedings of the European Congress on Computational Methods in Applied Sciences and Engineering.* event-place: Oslo, Norway. ECCOMAS

4.1 Abstract

This chapter introduces a novel method to create an interval field based on measurement data. Such interval fields are typically used to describe a spatially distributed non-deterministic quantity, e.g., Young's modulus. The interval field is based on a number of measurement points, i.e., control points, expanded throughout the domain by a set of basis functions. At the control point the non-deterministic quantity is known and bounded by an interval. However, at these measurement points information about the gradients might also be available. In addition, the non-deterministic quantity might be described better by estimating the gradients based on the other measurements.

Hence, the proposed interval field method allows to incorporate this gradient information. The method is based on Inverse Distance Weighing (IDW) with an additional set of basis functions: one set of basis functions interpolates the value, and the second set of basis functions controls the gradient at the control

points. The additional basis functions can be determined in two distinct ways: first, the gradients are available or can directly be measured at the control point, and second, a weighted average is taken with respect to all control points within the domain. In general, the proposed interval field provides a more versatile definition of an interval field compared to the standard implementation of inverse distance weighting. The application of the interval field is shown in a number of one-dimensional cases where a comparison with standard inverse distance weighting is made. In addition, a case study with a set of measurement data is used to illustrate the method and how different realisations are obtained.

4.2 Introduction

In common engineering practice the main goal is to provide or validate component designs that should perform under a wide range of circumstances, e.g., extreme weather, impact loads, and sometimes even in space or at other planets. To ensure the performance of these components, engineers often use numerical methods to approximate the set of differential equations governing the physical behavior of the component under investigation. However, this can be a daunting task as the parameters governing these equations are often only known vaguely, as they are inherent variable, or only limited knowledge about these quantities is available, as direct measurement is not possible, or a combination of both [84]. Hence, during the last decades a number of techniques are introduced that aim to quantify these non-deterministic quantities. In general, these techniques are categorized as probabilistic [213] and possibilistic approaches such as: intervals [68], fuzzy sets [100], and imprecise probabilities [19]. Where probabilistic methods describe non-determinism as the likelihood that parameters assume a value via a joint probability density function, possibilistic methods as, i.e. interval methods, consider non-deterministic quantities to be bounded.

In a number of cases these quantities, e.g., wind loads, Young's modulus, dielectric constants, experience a spatial or temporal dependency, which is difficult to consider within the classic interval method as described by R.E. Moore [163]. Therefore, the framework of interval fields was introduced [158], which is capable of providing the spatial or temporal dependency structure by a set of basis functions. Hence this method can be seen as a possibilistic counterpart to random fields [224]. In the last decade a number of researchers have introduced different basis functions that model the dependence structure, which can be based on inverse distance weighting [66, 152], affine arithmetic [207, 208, 211], radial basis functions [112], a spatial averaging method [234], or set-theoretical approaches [116, 176]. The basis functions that are introduced in this

chapter are an extension to the existing technique of Inverse Distance Weighting (IDW), which was introduced by Sheppard [200].

In this work the focus lies on interval fields defined trough IDW, which is a convenient way of constructing an interval field. The idea behind IDW is that the non-deterministic quantity is known or measured at independent locations within the physical domain. The interval field is then constructed based on the assumption that the weight of this information decreases proportional to the inverse of the distance moving further from this measurement point, which is referred to as a control point. Although this technique is successfully applied in a number of cases it has a number of shortcomings [152], and one of such shortcomings is that the maximum value of the interval field can never exceed the maximal value placed at a control point. This property is attributed to the definition of the basis functions, which will always have a zero gradient at each of the control points. Hence, this paper proposes an interval field based on two independent sets of basis functions at each control point, where one will interpolate the value of the control points and the second set of basis functions will control the gradients at the control point.To control the effect of the basis functions for the gradients a scaling parameter v is introduced. In general this information about the gradients may not always be available. Therefore, two methods are proposed to determine the gradients at the control points: the first method uses direct measurement, and the second technique estimates the gradients based on the observed trend of the data. This chapter is structured as follows: in Section 4.3 the interval field is introduced, and in Section 4.4 the application of this interval field is compared with the standard technique. Finally, Section 4.5 illustrates a real case study and conclusions are made in section 4.6.

4.3 Interval field analysis

In this section a brief description of the interval field analysis is provided, for a more detailed description the reader is referred to [66].

4.3.1 Explicit interval fields

The definition of the proposed explicit interval field is given in Equation (4.1) where, opposed to the literature in [158], a second set of basis functions ϕ_i is added. The new interval field consists of the superposition of two times $n_b \in \mathbb{N}$ independent basis functions $\psi_i + \phi_i$. Here, the range of the interval is interpolated by $\psi_i : \Omega \mapsto \mathbb{R}$, and in a similar way the gradients are determined

by $\phi_i : \Omega \mapsto \mathbb{R}$. Both of these basis functions are defined over the geometrical domain $\Omega \subset \mathbb{R}^t$, where t is defined as the physical dimension of the problem. These basis functions describe the spatial nature of the non-deterministic parameter, distributed along the coordinate $\mathbf{r} \in \Omega$. An interval field is created by scaling both these basis functions $\psi_i,\ \phi_i$ with independent interval scalars $\alpha_i^I,\ \beta_i^I \in \mathbb{IR}$. This interval field is formally defined as:

$$\mathbf{x}^I(\mathbf{r}) = \hat{\mathbf{x}} + \sum_{i=1}^{n_b} [\psi_i(\mathbf{r})\alpha_i^I + \phi_i(\mathbf{r})\beta_i^I], \tag{4.1}$$

with $\hat{\mathbf{x}} \in \mathbb{R}$ the midpoint of the interval field. Note that the existing IDW framework for interval fields is a special case where $\beta_i^I = 0$. When Ω is discretised into k finite elements, these base functions ψ_i and ϕ_i interpolate the independent interval scalars α_i^I and β_i^I to dependent intervals for each element in the domain Ω. Hence, the size of the bounded uncertain input space is $2n_b$, which can be reduced when only the range or the gradient at a control point is considered, i.e., $\Delta\mathbf{x}_i = 0$ while the gradient lies between $[0;\ 1]$. Nevertheless, in general this means that the input space dimension can be reduced if $2n_b < k$, which is double the amount compared to the standard method of IDW.

4.3.2 Interval finite element analysis

Omitted to avoid redundancy, see Chapter 2 Section 2.5.2

4.3.3 Definition of the basis functions

The definition of an interval field, as presented in Equation (4.1) takes two basis functions, the first interpolates the range of the interval field from the control points ψ_i, and the second basis function ϕ_i controls the gradients. Through the definition of these basis functions the spatial dependence of the non-deterministic quantity of interest is modeled throughout the domain Ω. An important property of these basis functions is that they should be self-complementary, i.e. $\sum_{i=1}^{n_b} \psi_i(\mathbf{r}_j) = 1\ \forall\ \mathbf{r}_j \in \Omega$. Furthermore, they should behave as unit vectors at the control points to ensure that independent intervals are retained (see [72] for a more thorough discussion). An intuitive definition of basis functions that comply with these requirements is provided by means of IDW interpolation, as applied in [71], which is also used to interpolate the range at the control points to the domain Ω in this chapter.

Basis functions for the range

The first set of basis functions is the standard IDW approach where basis functions are defined for each control point $\mathbf{r}_i$. This is accomplished by a normalisation of weight functions $w_i(\mathbf{r}) \in \Omega$, denoted as:

$$\psi_i(\mathbf{r}) = \frac{w_i(\mathbf{r})}{\sum_{j=1}^{n_b} w_j(\mathbf{r})}, \tag{4.2}$$

with $i = 1, \ldots, n_b$. The weight functions w_i are inversely proportional with the Euclidean distance measure $d(\cdot)$ measured between the control point $\mathbf{r}_i$ and other coordinates $\mathbf{r}$ in the domain:

$$w_i(\mathbf{r}) = \frac{1}{[d(\mathbf{r}_i, \mathbf{r})]^p}. \tag{4.3}$$

Herein, the power $p \in \mathbb{R}^+$ allows the analyst to influence the rate of decay of the weight function. Note that for a power $p < 1$ no derivative of the basis function exists at the control points, while in the case that $p > 2$ the basis functions flatten and higher gradients at the transitions are obtained. Empirical evidence suggests that in general $p = 2$ is a good starting point [71], if no further information about the spatial nature is available. The distance measure $d(\cdot)$ is measured in Euclidean space, defined as:

$$d(\mathbf{r}_i, \mathbf{r}) = \|\mathbf{r}_i - \mathbf{r}\|_2 , \tag{4.4}$$

with $\|\cdot\|_2$ denoting the L_2 norm.

Basis functions for the gradient

The second set of basis functions is constructed in a similar manner as the IDW basis functions and identical weight functions are used, as they assign a higher weight to points in $\mathbf{r}$ closer to a control point $\mathbf{r}_i$. These basis functions are defined as:

$$\phi_i(\mathbf{r}) = \frac{w_i(\mathbf{r})\delta_i(\mathbf{r})}{\sum_{j=1}^{n_b} w_j(\mathbf{r})}, \tag{4.5}$$

with $\delta_i : \mathbb{R} \mapsto \mathbb{R}$ a factor to set the gradients at the control points $\mathbf{r}_i$, which is defined as:

$$\delta_i = A_i(\mathbf{r} - \mathbf{r}_i) \left[\frac{R_i}{R_i + d(\mathbf{r}_i, \mathbf{r})} \right], \tag{4.6}$$

here $A_i \in \mathbb{R}_+$ represents the desired gradient at the control point and the constant $R_i \in \mathbb{R}_+$ is a scaling factor, defined as:

$$R_i = \frac{v(\max(\hat{\mathbf{x}}_i) - \min(\hat{\mathbf{x}}_i))}{A_i}, \tag{4.7}$$

with $v \in \mathbb{R}_+$ a scaling factor that determines the effect of the gradient terms similar to parameter p, which determines the effect of the basis functions for the value. The value $|\delta_i(\mathbf{r})|$ that is added by these basis functions ϕ_i causes the derivatives to be $\frac{\partial \delta_i}{\partial r} = A_i$. This is only valid at the control points as the value of $\phi_i(\mathbf{r})$ for consecutive points moving away from $\mathbf{r}_i$ as the factor $R_i / [R_i + d(\mathbf{r}_i, \mathbf{r})]$ will go from 1 to behaving like $d(\mathbf{r}_i, \mathbf{r})^{-1}$ for large $d(\cdot)$. In addition, note that $\delta_i(\mathbf{r}_i) = 0$ thus keeping the independence of the basis functions at the control points, which can therefore be scaled by independent interval scalars while retaining the self-complementary basis and the independency of the intervals scalars at the control points. However, note that these basis functions are not self complementary beyond the control points.

The only remaining parameter is the constant A_i which will be the gradient at the control point. Depending on the available data two distinct ways of calculating A_i are presented. The first method is to directly calculate the constant based of points close to $\mathbf{r}_i$, defined as:

$$A_i = \frac{(\hat{\mathbf{x}}_j - \hat{\mathbf{x}}_i)(\mathbf{r}_i - \mathbf{r}_j)}{d(\mathbf{r}_i, \mathbf{r}_j)^2}, \tag{4.8}$$

here the index j is given to a neighbouring point $\mathbf{r}_{i\pm j}$ used to calculate the constant A_i. Depending on the side and the distance the result of equation (4.8) can differ, in this case the maximal value is taken. In this case the assumption is made that there is more information available around the control points, which may not be the case in general.

Therefore, a second approach is to determine the constants A_i as a weighted average of the control points $\mathbf{r}_j \in \mathbf{r}_i$, defined as:

$$A_i = \frac{\sum_j^{n_b} w_i \frac{(\hat{\mathbf{x}}_j - \hat{\mathbf{x}}_i)(\mathbf{r}_i - \mathbf{r}_j)}{d(\mathbf{r}_i, \mathbf{r}_j)^2}}{\sum_j^{n_b} w_j}, \tag{4.9}$$

where the weights w_j are defined as in equation (4.3), which assigns less weight to control points far from $\mathbf{r}_i$. Furthermore, in this work A_i is determined based on the midpoint $\hat{\mathbf{x}}$ of the intervals, which will give an average gradient that is acceptable for a large number of cases. However, it is easy to define a case with increasing non-determinism with a zero midpoint, thus depending on the case better results can be obtained by changing this to the interval radius Δx, or the extremes of the interval $\overline{x}_i$, $\underline{x}_i$. Using these basis functions will ensure that the desired derivatives A_i are obtained at each control point $\mathbf{r}_i$.

One of the important things to note here is that the basis functions ϕ_i of the gradients add a value $|\delta_i(\mathbf{r})|$ to the standard IWD basis functions ψ_i. This value of $|\delta_i(\mathbf{r})|$ is zero at the control points, and the effect can be changed between

the control points by setting the parameter v, which changes the value of $|\delta_i|$ as:

$$|\delta_i(\mathbf{r})| \leq v[\max(\Delta\mathbf{x}_{i\pm j}) - \min(\Delta\mathbf{x}_{i\pm j})], \tag{4.10}$$

with the index $j \in \mathbf{r}_i \pm r$. As such, the scaling parameter v controls the effect of the basis for the gradients, which can be local for a small value of v or have an effect far away from the control point for larger values of v. Therefore this parameter can be seen as the gradient counterpart to the parameter p that determines the effect of the basis functions for the value. An interesting implication of using the basis functions ϕ_i is that the maximum value of the interval field is no longer restricted to the location of a control point and can be anywhere within the domain Ω.

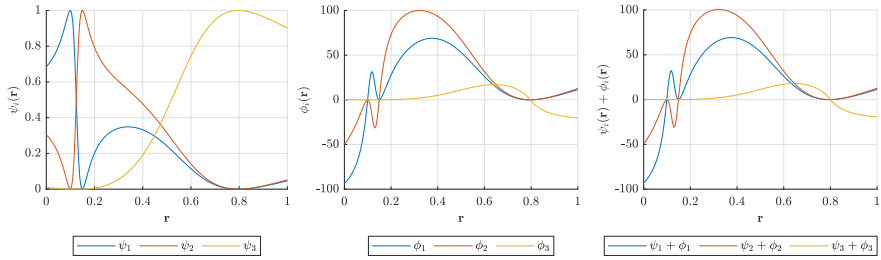
4.4 Illustration of interval fields with local gradient control

The following cases compare and illustrate the use of the novel interval technique. First a comparison is made with the existing IDW technique, second the two different strategies to dermine the gradients are explored, and the third case is about extracting samples and the possibilities towards dependence structures between α^I and β^I.

4.4.1 Comparison between IDW with and without gradient control

To demonstrate the additional value of incorporating gradient information in the basis functions a case is considered where only limited information about a set of measurements is available. This set $\mathcal{A}$ represents the true underlying spatial non-determinism for a parameter, e.g., used in the FE method. The measurements are made at the points [0.1, 0.15, 0.8], which are therefore also used as the control points $\mathbf{r}_i$ of the interval field. In order to calculate the first basis functions that will interpolate the range at the control points ψ_i, the parameter p is set to $p = 2$. Figure 4.1a illustrates the basis functions that are obtained, which are 1 at the location of the control point.

The second set of basis functions ϕ_i provides the desired spatial gradients at the control points, which are calculated exactly from the midpoint $\hat{\mathbf{x}}_i$ of the set $\mathcal{A}$. The motivation is that these points represent a small group of measurements or this can be based on engineering judgement. Hence, the basis functions are calculated following equation (4.5) where the spatial gradients are taken



(a) Basis functions for the interval field value

(b) Basis functions for the derivative

(c) Sum of the basis functions without scalars

Figure 4.1: Basis functions obtained by measuring the gradients, as indicated by equation (4.8)

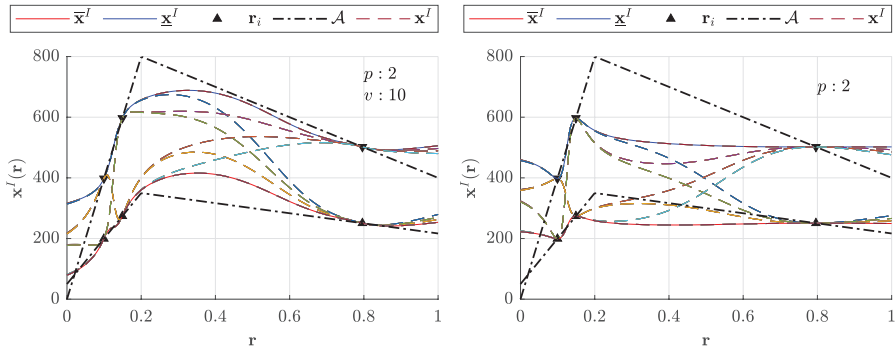
from equation (4.8). In this case the parameter v is set at 10, which is an arbitrary choice. The obtained basis functions are shown in Figure 4.1b, where these are zero at the control points. Note that these basis functions are not self complementary $\sum_1^{nb} \phi_i \neq 1 \; \forall \; i$ outside the control points. In Figure 4.1c the sum of these basis functions is given with the interval scalars set at one $\alpha_i^I, \beta_i^I = 1$.

Figure 4.2a illustrates the outer realisations, in red and blue lines, and the vertex realisations, in dashed lines, of the interval scalar α^I. Hence, we are only considering the combinations of different values without different gradients, as indeed β^I can also vary between $[-1 \; 1]$ causing the gradients to lie within the interval $[-A_i \; A_i]$. Therefore, the illustrations in this case and the following case are limited to $\beta^I = 1$, which will set the gradients at the control points equal to the calculated values of A_i.

In Figure 4.2b a similar plot is made with the standard IDW basis functions, which are identical to ψ_i. A comparison between Figure 4.2a and 4.2b shows that the proposed method is capable of capturing the gradients and thus provides a better estimation of the set $\mathcal{A}$. In addition, one can see that the maximum value of the interval field based on IDW is located at the control point where in Figure 4.2a this is located between $\mathbf{r}_2$ and $\mathbf{r}_3$, which is an intuitive location based on the gradient information at the control points.

4.4.2 Determination of the gradient from other control points

In the previous case the gradients are directly calculated at the control points,

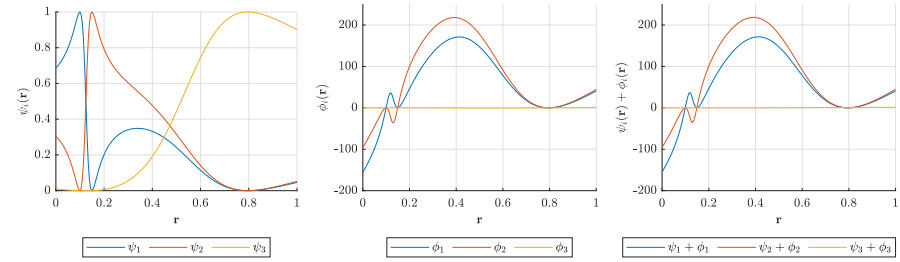


(a) Interval field using the derivatives

(b) Interval field using IDW

Figure 4.2: Vertex realisations of an interval field based on IDW (b) and one using the information about the derivatives (a); the dash-dotted line indicates the unknown underlying uncertainty

which requires additional information at the control points. However this information might not always be available. Therefore, a different approach is considered where the constants A_i are calculated based on a weighted average of the midpoints, as described in equations (4.9). The basis functions that are obtained in this way are given in Figure 4.3 where the first illustration 4.3a is identical to this in 4.1a. However, looking at the basis functions of the gradients ϕ_i shows a gradient close to zero for the third control point $\mathbf{r}_3$, which is attributed to the distance and the small relative change of midpoint at these locations.



(a) Basis functions for the interval field value

(b) Basis functions for the derivative

(c) Sum of the basis functions without scalars

Figure 4.3: Basis functions based on the derivatives calculated from the information at the other control points, as described by (4.9)

In Figure 4.4 the realisations of this interval field are given, which have a higher maximum value compared to the previous case without changing the value of v. This effect is also seen in the basis function in Figure 4.3b, which has a maximal value almost double as high as the previous basis functions. This is caused by the larger difference in the values that are used to calculate the constants, which is described by (4.10) and can be changed by selecting a different value v.

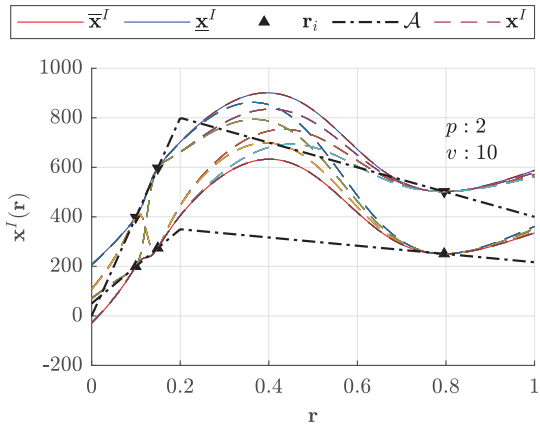
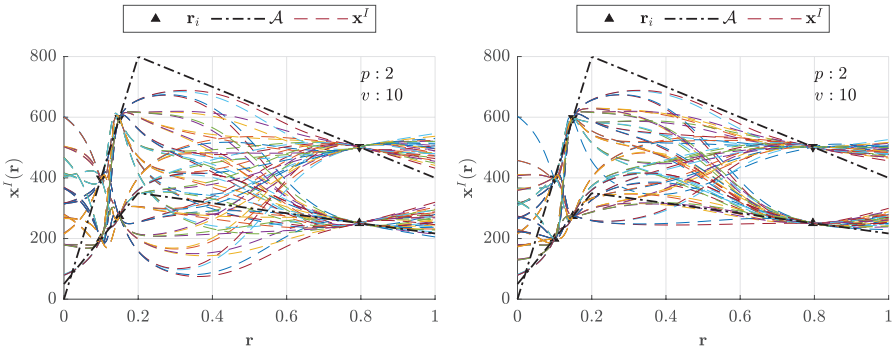


Figure 4.4: Realisations case 1

4.4.3 Interval field realisations

As described in the first case, each of the realisations until here are given with vertex samples from α^I, while $\beta^I = 1$ is kept constant. However, this interval field consists of a two-dimensional uncertain input space at each control point. Thus, a full vertex analysis consists out of 2^{2n_b} samples. In addition, it is up to the analysts to determine the range of the gradients as these can vary between $[-A_i \; A_i]$ illustrated in Figure 4.5a where the sign of the gradient is unknown at the control points and there are a large number of possible realisations. It is also possible to limit the values in β^I from $[0 \; A_i]$ as shown in Figure 4.5b where the gradients are equal to A_i or smaller.

It is clear from Figure 4.5a that the possible values of the gradients should be limited in this case as a large number of realisations lies outside the data $\mathcal{A}$. Although, limiting β^I to lie within an interval $[A_i \; 0]$ it should be noted that more complex dependency structures could be defined, as described in [66]. Nevertheless, one should always define an interval for β^I as keeping β^I at a



(a) Vertex realisation for both $\alpha^I, \beta^I \in [-1\ 1]$

(b) Vertex realisation for $\alpha^I \in [-1\ 1]$ and $\beta^I \in [0\ 1]$

Figure 4.5: Two realisations of an interval field representing the envelope $\mathcal{A}$

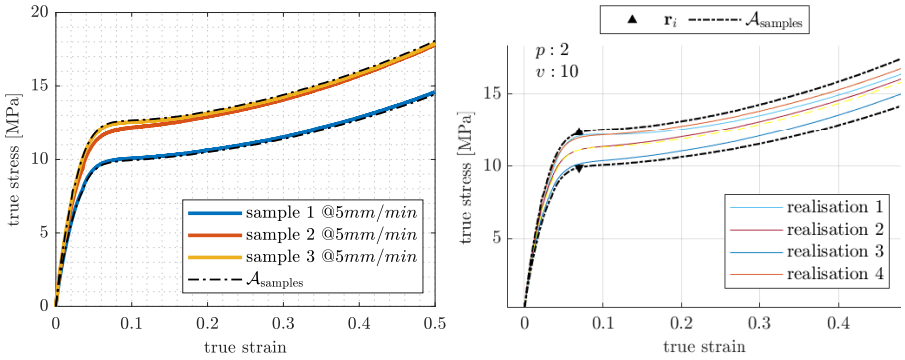
fixed value will fully couple the value at a control point with the gradient, which is not the case in general.

4.5 Case study

In this final case study the method is applied to capture the non-determinism of a set of stress-strain curves. The objective is to represent the set of measurements with an interval field, and each sample of this interval field should represent a feasible stress-strain curve. To obtain the stress-strain curves provided in Figure 4.6a three samples have been printed and tested under uni-axial tension, in accordance with ASTM D638.

To illustrate the additional value of the method only two control points are placed, one at the origin and one at $r = 0.07$. The constant A_i are in this case calculated at each point directly, following Equation (4.8), based on the interval radius Δx of $\mathcal{A}_{\text{samples}}$. The resulting interval field shown in Figure 4.6b is only sampled at $\beta_1^I = 0, -0.5, -0.5, -1$ and $\alpha_2^I = 0.8, 0, -0.8, 0.7$ labeled realisations $1, 2, 3$ and 4, respectively. Thus, the non-deterministic measurement set is represented only using the gradient at the first control point and the value at the second control point, which could be regarded as sampling the initial stiffness of the material and the yield strength.

This case is used as an example where the parameters and control points are set by hand. In a more comprehensive study to find the optimal interval field to



(a) (b)

Figure 4.6: Measured stress-strain curves (a) and the interval field representation (b); the dashed line represents the envelope $\mathcal{A}_{\text{samples}}$

represent a set of measurements optimisation approaches can be used, as in [152]. Furthermore, detailed investigations into the set of admissible realisations need to be made.

4.6 Conclusions

In this chapter an extension to the existing framework of interval fields is presented. This extension allows for the incorporation of information about gradients at the control points. It is shown that this method is better capable in representing a spatially distributed non-deterministic quantity compared with the existing technique. Even without explicit information about the gradients at the control points better agreement is obtained by making an estimation about these gradients. To control the effects, a parameter v is introduced that allows the user to adjust the realisations of the interval field by tuning the influence of the gradients on the final result. In addition, a case study is conducted with a set of real measurements, which could be represented using only the gradient information in one control point and the range at the other control point. Further research will focus on the application of admissible set decomposition, which allows for dependency structures within the interval framework.

Chapter 5

Interval methods for lack-of-knowledge uncertainty in crash analysis

This chapter was previously published as:
C. van Mierlo et al. “Interval methods for lack-of-knowledge uncertainty in crash analysis”. In: *Mechanical Systems and Signal Processing* 168 (2022). Publisher: Elsevier, p. 108574

5.1 Abstract

This chapter deals with lack-of-knowledge uncertainty in complex non-linear simulations on a component level, i.e., a crashbox during frontal impact of a vehicle. Specifically, the focus lies on using interval field techniques to model the uncertain boundary conditions during impact simulations. The uncertainty considered in this work is the unknown mechanical response from the adjacent structure. This uncertainty is considered to be epistemic, representing the case where this adjacent structure is unknown at the time the impact analysis is performed. In practice, this refers to the situation where the adjacent structure is still under development, e.g., at a different department or even outsourced. In addition, the safety critical performance of both, the component and the overall structure should be guaranteed under a wide range of circumstances, which are typically encountered in real-life situations. Typically, car manufacturers use

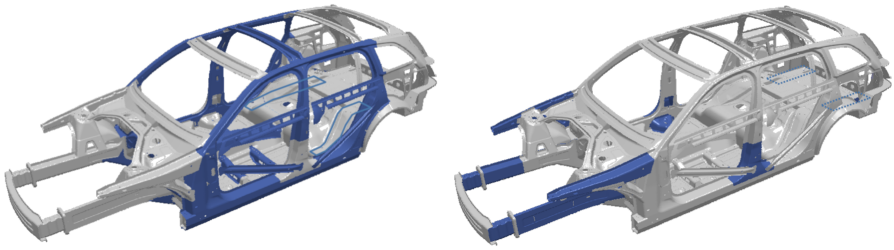
multidisciplinary optimisation to identify component designs that perform best on all requirements in a deterministic sense, while minimising the overall weight. Unfortunately, the results of such optimisation schemes are known to converge to an often non-robust optimum. As a result, the response of the structure may be sensitive to small changes in input parameters or boundary conditions.

As an answer to these challenges, this chapter proposes an interval field approach that accounts for the epistemic, i.e., lack-of-knowledge, uncertainty of the adjacent structures, even in an early design stage. This is accomplished by introducing a spatially varying uncertain mechanical compliance in elements that connect the component to the adjacent structures. These elements have an interval valued stiffness, which is varied along the component following the realisations of an interval field. The bounds on the interval-valued response quantities of interest, i.e., mean force and peak force, are identified using a differential evolution algorithm. This method is demonstrated on four case studies of a full overlap crash analysis of a rectangular crash box, which represents a generic component within the front structure of a vehicle. These case studies demonstrate the applicability and the potential of the proposed method. In addition, in the last case it is shown that the performance of the component can be assessed under an increasing range of uncertainty.

5.2 Introduction

In recent years car manufacturers are changing from traditional test-based design towards more simulation-driven approach due to the ever rising complexity in development and increase of safety requirements. Examples of such safety requirements are, e.g., proposed by the United Nations Economic Commission for Europe (UN-ECE) [60], or based on consumer tests, e.g., those of the New Car Assessment Programmes (NCAP) like Euro NCAP [64] or Global NCAP [92]. These tests represent the relevant accident scenarios while being also sufficiently repeatable to enable controlled vehicle assessments and ratings. In these numerical approaches advanced numerical methods for multi-disciplinary and multi-criteria optimisation are used to identify the appropriate design compromises, see, e.g., [59]. However, even in standard cases, the performance of designs obtained though optimisation are known to be very sensitive to small changes in input parameters. This problem is further amplified when considering highly non-linear phenomena encountered in crashworthiness studies [4]. Furthermore, the robustness - low sensitivity of responses to input variations - as well as the reliability - low probability of constraint violations - have to be considered additionally. This leads to an even higher numerical effort than just needed for a deterministic optimisation.

In addition design criteria for crashworthiness are mostly related to bio-mechanical measures (accelerations, velocities, deformations, forces, and moments) registered by Anthropometric Test Devices (ATDs), also known as "dummies". Examples are the Head Injury Criterion (HIC), see the discussion in [226], or the Neck Injury Criterion (NIC), see [29]. The optimal quantity to use in crashworthiness assessment is changing frequently, see e.g. [145]. However, during the early development process, it is standard to consider mainly structural criteria, i.e., criteria related to the performance of the car structures, as detailed geometrical and material data is not available. Car-body related criteria address either aspects of the safety cage (deformation resistance parts) or aspects of the crumple zones (energy absorbing parts), as illustrated in Figure 5.1. The design of energy absorbing parts remains challenging, criteria as specific energy absorption (SEA), which is the total energy absorption divided by mass, peak force or peak acceleration are commonly used [45, 81].



(a) Deformation resistance parts (blue) (b) Energy absorbing parts (blue)

Figure 5.1: Example of a car body highlighting deformation resistance (left) and energy absorbing parts (right) for a frontal impact [103]

5.2.1 Simulation based car body development

The development of the car body structure is a highly complex task, where all components interact and the required force-deformation behaviour of the components is completely inter-dependent. Even the design of a single component is highly complex due to the high non-linear behaviour in terms of mechanical plasticity, failure, contact, buckling, large deformations, strain- and stress-rate-dependencies. In addition, a high number of different materials have to be modelled, ranging from different steel types and other metals to glass, polymers, foams, composites and bio-materials. The related computational effort is high; standard simulations take several hours and in some cases even days, despite the integration of high performance computing (HPC) in the simulation workflows. Therefore, to reduce this computational cost for the structural design

of energy absorbing components there are two main approaches found in crash related literature, which are listed below and are illustrated in Figure 5.2.

a) **Full vehicle FEM simulations** with models as shown exemplarily in Figure 5.2: the complete structure of the car is modelled and the developer modifies a component (or components) assessing the changes via a complete repeat of the full set of crash simulations (note that a small change will affect the car performance in multiple different crash tests).

b) **Component simulations** with pre-defined boundary and initial conditions. For this, there are three options:

   b1) The energy absorption of a component is assessed under drop test conditions, i.e., a rigid plate or block with a certain mass and initial velocity is hitting the component. Here, the dynamic effects are covered more correctly.

   b2) A similar configuration as for b1) is used but by a crush test where a rigid wall with a prescribed and constant velocity deforms the component in axial direction. Because this is often done in a quasi-static manner, dynamic effects like inertia forces and (strain-)rate dependencies are neglected.

   b3) An alternative can be realised by using a full vehicle simulation and by registering the deformation- or velocity-over-time of the FE nodes at the interface between the component and the complete car structure. Then, the data of the interface nodes is used as constraints in the component simulations.

In some cases, a full-vehicle simulation is used to assess the performance of a single component. The advantage of this approach is that it takes all surrounding parts, as well as their interaction with the component, into account. However, especially during early stages of the development process, properties or design details of neighbouring components are not fully known. Typically, the development is a concurrent process between multiple designers or even departments / companies, where each designer or department is designing an individual component in parallel with activities of the others. Therefore, the full vehicle model at this point may not be available, under construction, or far from the final version. Hence, having a complete vehicle simulation during development may mean that pseudo-accuracy is introduced by the level of detail that is obtained, which neglects the development of other components. Therefore, potential wrong conclusions are made and redesigns at a later stage would be needed to correct for these decisions.

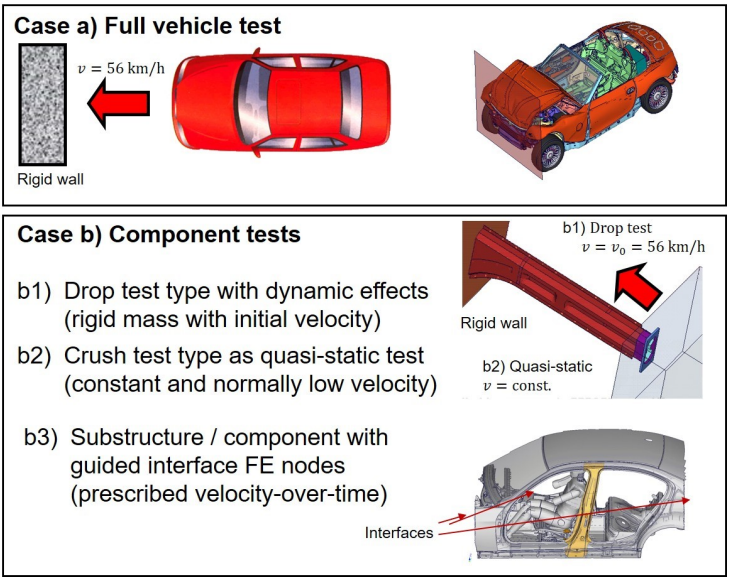


Figure 5.2: Example of full vehicle test (top) and component / sub-structure tests (bottom), after [59]

In addition, the potentially high computational effort of a full-vehicle simulation makes it nearly infeasible to realise a high number of full model simulations, such as needed for optimisation and robustness or reliability assessments. For instance, a single simulation of the Honda Accord model with 1.9 million elements [175] takes 14 hours on eight Intel(R) Core i9-7980XE CPUs. The academic example used in this work on the other hand, as illustrated in case a) in Figure 5.2, requires several minutes to calculate. Moreover, as the full car model consists of multiple parts and materials that interact with each other, the uncertainties about all these parameters should be carefully assessed and quantified. Therefore, detailed investigations should be conducted concerning the range of these parameters as well as the relative likelihood of certain parameter values within this range (as commonly quantified by a distribution function). This is very challenging in general, and especially in an early design stage where many design decisions may still be open. In recent years, robustness studies on full vehicle models have been realised in, e.g., [4, 107]. However, these are rarely embedded in an industrial development and more importantly, the uncertainties considered are far from complete.

5.2.2 Uncertainty in crashworthiness

In the three versions of crashworthiness assessment of a single component illustrated in Figure 5.2, the parts are evaluated in an isolated environment neglecting the influence of possible modifications in the other components. However, from experience, we know that the mechanical response of other components strongly influences the behaviour of the component under consideration in the design study. Therefore, to the opinion of the authors, this - often unknown - difference between fixed boundary conditions and coupled boundary conditions to adjacent parts is of very high relevance, and should be considered in a single component impact performance optimisation under uncertainty. When neglected, the identified optimum may be of questionable value, as robustness or reliability problems on component level may lead to critical performance issues in the global crash performance of the complete vehicle. The advantage of the single component assessment is clearly the computational cost, i.e., a single assessment can be conducted at a fraction of the time it would take to run a full crash model. In addition, validation of these simulations via physical experiments is less complicated as drop-tower tests or quasi-static tests are widely used for single component testing, as opposed to full vehicle tests. The number of scientific papers on the assessment of components is very high, see [2, 247] to give only some of the most recent papers. However, the consideration of uncertainties is rarely undertaken on component level. As an example the reader is referred to [215].

Nevertheless, a range of methods is proposed in recent crash related literature that take these uncertain input parameters into account. Examples include load case and geometrical uncertainties [79, 107], or material uncertainties [122]. In these approaches, variations of the impact angles, locations and velocities are considered. In some cases, these quantities are also combined with the influence of manufacturing tolerances (variations in thickness, material parameters or geometrical features like radii) [59]. Following these numerical approaches, one typically assumes the uncertain input parameters to be independent. Regarding the parameters mentioned above, most of them are direct input parameters of the finite element model except the geometrical changes such as radii. For the latter, parametric shape modelling and mesh morphing tools have been developed [59]. In addition, efforts have been made on reducing the computational cost of uncertainty propagation by a multi-fidelity approach in [138], or adaptive Kriging based approaches in [167].

5.2.3 Complexity of hierarchical development

The application of optimisation with robustness and reliability analyses in an industrial setting remains challenging and time consuming, not only because of the high numerical effort. The main reason is related to the context of systems engineering and the necessity of hierarchical development caused by the high complexity of the product. This means that the different crash types (e.g. front, side, and rear impacts) are treated by different people or even departments and companies. As a consequence, every developer is working on a single component and not on the complete vehicle. Therefore, requirements must be broken down to the component level. Consequently, assessments are done as well on single components rather than on the full vehicle or system. The well-known V-model approach and the more recently developed Component Solution Space methodology [48, 249] enable this hierarchical development. However, following the Component Solution Space approach, it is challenging to include the inter-dependencies of the different components during a crash. As in the original Component Solution Space approach [249], deterministic force-deformation curves are obtained for each of the components, with a range that is maximised for each component until constraints are violated, e.g., order of plastic deformation, or acceleration limits. However, in real incidents, impact angle, speed and impacting object are unknown and the occurring deformations and force levels are uncertain. To resolve this, Component Solution Space methods have been introduced that incorporate epistemic uncertainty: where [47] focuses on uncertainties in force levels and [46] on remaining uncertainties, i.e., deformation lengths, energy to be absorbed, critical acceleration limit. These methods provide bounds on the range in which the component is performing as well as information about the range of uncertainty allowed for by the adjacent structure.

To overcome the issues related to the decoupled development of complex interacting structures, this chapter presents a novel method to consider the interactions of a single component with these adjacent structures. Typically, the design and optimisation of these single components are based on droptower tests, b1 in Figure 5.2, where one typically measures force and deformation of an impacting object on a fixed specimen or component. However, this chapter proposes a novel way to design and optimise a single component by introducing uncertain boundary conditions that account for the unknown behaviour of the adjacent structure, which is neglected in the typical tests. Nevertheless, from experience, we know that the mechanical response of other components strongly influences the behaviour of the component under consideration in the design study. Therefore, to the opinion of the authors, this - often unknown - difference between fixed boundary conditions and coupled boundary conditions to adjacent parts is of very high relevance, and should be considered in a single component

impact performance optimisation under uncertainty. The structure of this chapter is as follows: Section 5.3 gives a detailed description of the proposed implementation of interval fields at the boundary conditions. The difference between deterministic and uncertain boundary conditions is illustrated for a number of case studies in Section 5.4, followed by a discussion of the results in Section 5.5. Final conclusions are drawn in Section 5.6, which also provides an outlook of future challenges.

5.3 Non-deterministic modelling of the adjacent structure

The uncertainty in the proposed modelling strategy stems from the assumptions and abstractions that are made concerning the mechanical behaviour of the adjacent components. Since this uncertainty stems from a lack-of-knowledge about the final components, it is an attribute of the analysis, and hence, epistemic in nature. Therefore, it is proposed to model it using the interval framework. For the sake of argumentation, when one would attempt to model this type of uncertainty using probabilistic methods, subjective information is inserted into the analysis [66], which might give a false sense of accuracy. Applying interval analysis therefore is the most objective approach since it acknowledges that there is no information on the likelihood of relative parameter values within the interval bounds. Furthermore, when limited data about the actual boundary conditions are available, approaches to infer the bounds based either on Bayesian analysis [109] or inverse analysis can be applied [71].

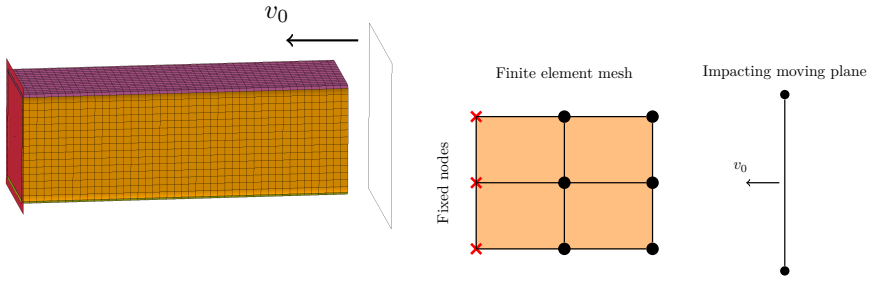
A particular convenient interval technique for parameters that are spatially distributed is the recently introduced framework of interval fields, which can be regarded as a possibilistic counterpart to random fields [224] for quantities that are spatial or time dependent [158]. Following this framework of interval fields, locally defined intervals are expanded through the model domain based on a set of a priori defined basis functions. Multiple definitions of these basis functions can be found in literature, which are based on inverse distance weighting [71], affine arithmetic [203, 207, 211], radial basis functions [112], a spatial averaging method [234], or set-theoretical approaches [116, 176]. A recent overview of interval fields can be found in [66]. The following sections start with a detailed description of the interval field framework, and end with a description how this concept is used to model the epistemic uncertainty about the adjacent structure in a component finite element simulation.

5.3.1 Interval field analysis

Omitted to avoid redundancy, see Chapter 2 Section 2.5

5.3.2 Boundary conditions described by interval fields

The discussion in this chapter is based on a full vehicle crash model and an exemplary component, here a generic crash box, to illustrate the principal ideas. Figure 5.3 illustrates this component and the typical simulation setup, where on the left the finite element model of the component is shown and on the right an illustration of the typical boundary conditions is given where the red crosses indicate the fixed nodes. The case that is considered in this work is a full-width overlap crash test of a passenger car driving against a rigid barrier at 56 km/h, in accordance with the corresponding NCAP test [64]. This type of test set-up is defined in several consumer and regulation tests and is commonly used in scientific studies.



(a) Finite Element Model of the crashbox with a rigid plane attached to the nodes in the back (red) and impacting plane right

(b) Illustration of a general crash set-up, with the red crosses indicating the fixed nodes, and the impacting rigid plane on the right

Figure 5.3: Illustration of the crashbox and the general set-up of a crash analysis as used in this chapter

The interval field concept is not directly applicable to this typical crash simulation, as illustrated in Figure 5.3b. First, a representation of the adjacent structure should be defined. In this case the adjacent structure is modelled at the back of the component between the rigid wall and the fixed nodes, which corresponds well to the physical location of these components within the vehicle. Figure 5.4 illustrates the adjacent structure modelled by a set of connecting elements. The epistemic uncertain lateral stiffness of these elements is represented by the interval field. Note, that the described method also

works for elements placed in front of the component, or a combination of both, although this would require additional considerations about the properties of these elements.

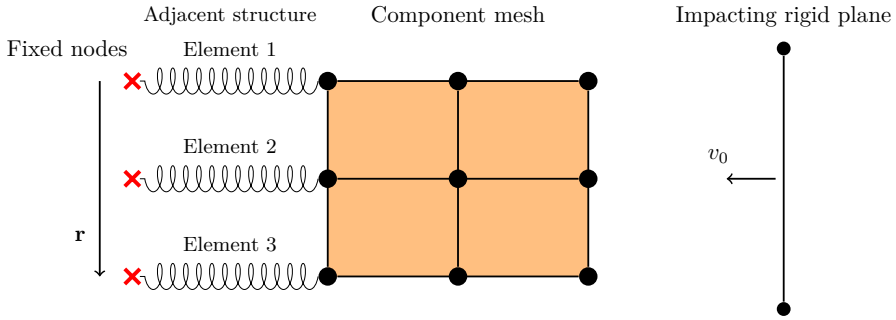


Figure 5.4: Illustration of the connecting elements, with the fixed nodes in red, the component in orange, and the rigid impacting plane on the right

Interval field modelling of the connecting elements

In this work, a novel technique is used to model a one-dimensional interval field on a three-dimensional component. This interval field is defined on the lateral stiffness of the elements connecting the crash-box to the surrounding. As such, the interval field models the uncertain compliance of the structure that is adjacent to the crash box. Specifically, the crash box is modelled as a rectangular shell that is meshed by two-dimensional shell elements. In this case, the nodes of the shell elements describe the circumference of a rectangular shape, as the thickness is considered within the shell formulation. Therefore, the distance measure used in IDW is calculated along the circumference of the rectangular box, which yields a continuous one-dimensional interval field along the circumference of the component. However, since the vector $\mathbf{r}$ describes a position on a continuous rectangular loop, the determination of the distance $d(\cdot)$ from the control point $\mathbf{r}_i$ to the other nodes $\mathbf{r}$ is less trivial as each nodal point can be reached following two distinct paths, i.e., clockwise, or counterclockwise along the circumference of the component. In this case, we consider the shortest distance between two points on the circumference. This can be solved by only using the shortest path between $\mathbf{r}_i$ and $\mathbf{r}$ to determine the distance measure, e.g., using Dijkstra shortest path algorithm [54]. In addition, element lengths can be directly used as weights in these shortest path algorithms and one directly obtains the distance. Note that the application of Dijkstra's algorithm in this case is superfluous since only two possible distances exist. However, in more

general cases, multiple paths may exist. This motivates the application of shortest-path algorithms. In Figure 5.13 an illustration of the nodes is provided

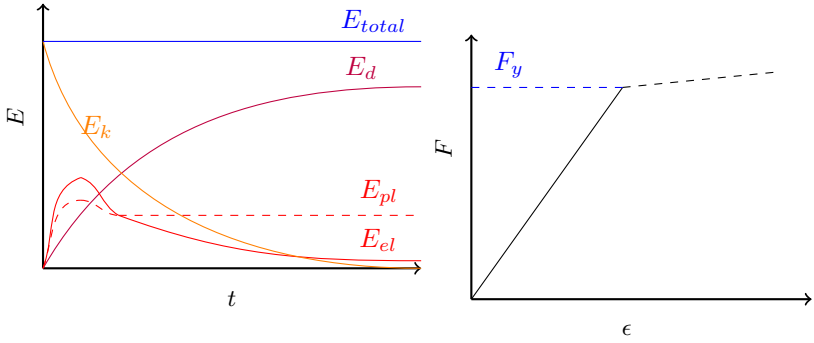
The other parameters to fully determine the interval field as described in Chapter 2.5.2 are: the power p, the interval scalars $\alpha_i^I,\ i, \dots n_b$, and the location of the control points $\mathbf{r}_i,\ i, \dots, n_b$. The influence of each of these parameters is described in detail in different case studies (see Section 5.4. In this chapter, the power p is set at 2, which is reasonable based on prior experience [66]. It is interesting to point out that higher values of p will increase the weight of the closest control point and flatten the realisations near the control points, while a lower value of p decreases the influence of the control points where the realisation are all closer to an average value.

Two illustrations of realisations of the interval field are given in Figure 5.15, where the dashed black lines and red dots illustrate the variation of the normalised lateral stiffness of the elements towards the fixed wall nodes in blue. In addition, the control points $\mathbf{r}_i$ in this figure are shown as blue nodes with a black circle located at the coordinates $\mathbf{r}_i = [30, 30; 30, -30; -30, -30; -30, 30]$, and for some cases the control points are placed between the corner nodes located at the coordinates $\mathbf{r}_i = [2, 30; 30, -2; -2, -30; -30, 2]$.

Modelling the connecting elements

Depending on the analysis there are several ways to model the connecting elements at the back of the component. The appropriate selection of the element type is important, since it influences the energy balance of a crash simulation significantly. Figure 5.5 shows the effect of the connecting elements on the energy balance for two different material models. In a typical crash scenario, the kinetic energy E_k of the moving vehicle is fully translated into elastic and plastic deformation energy $E_d = E_{elastic} + E_{plastic}$, which is stored and dissipated by the deformation of the component. However, an additional energy storing and dissipation element is introduced by the introduction of the connecting elements. The amount of energy stored or dissipated in the elements depends on the interval field realisation and the material model that describes the behaviour of these elements. Figure 5.5b illustrates this behaviour where both, a linear and a bi-linear material model are shown by respectively the full and dashed lines. Here, the plastic deformation of the bi-linear model dissipates a part of the kinetic energy, which will therefore not be translated to the component. Therefore, the crash box is not subjected to the full kinetic energy of the impact. Such situation is undesirable as this biases the comparison of the dissipated energy in the crash box with respect to cases where less energy is dissipated in these connecting elements. Therefore, the connecting elements

are modelled with a linear material behaviour. Further, such linear model also ensures a constant interaction between the component and the adjacent structure. The physical interpretation corresponds to a crash where a certain amount of energy is stored elastically within the complete structure, e.g., front structure of a vehicle, test machine, and this energy is released back from the most rigid components to the deformed components. However, note that when the elements are modelled as linear elastic elements, i.e., beams, forces higher than the yield force of the material can be reached for a short moment of time. Since the failure of these connecting elements are not of interest for the analysis, this is not critical.



(a) Energy balance of one simulation where all the kinetic energy E_k is transformed into deformation energy E_d, and energy in the springs E_s; a part elastic E_{el} an part plastic E_{pl}

(b) Illustration of the force-strain behaviour for a bi-linear model where F_y indicates a yielding point at which the response becomes less stiff

Figure 5.5: Two figures illustrating the impact of different material models on the energy balance of an impact simulation; indicated by the full and dashed lines

5.4 Case studies

In this section, four different approaches to model the lack-of-knowledge uncertainty about the adjacent structure are illustrated on a generic crash example. Specifically, the lateral stiffness of the linear connecting elements is modelled according to following approaches: (1) a deterministic benchmark case, (2) a scalar interval valued model, (3) an interval field approach, and finally (4) an interval field approach with a varying degree of uncertainty, modelled

Material model properties used for the component					
Mass density	ρ	7830 kg/m^3	Strain-rate parameter	P	5
Young's modulus	E	200 GPa	Strain-rate parameter	C	40
Poisson ratio	ν	0.3	Yield stress	σ_0	366 MPa
Equivalent stress	σ_1	424 MPa	Equivalent strain	e_1	0.025
Equivalent stress	σ_2	476 MPa	Equivalent strain	e_2	0.049
Equivalent stress	σ_3	507 MPa	Equivalent strain	e_3	0.072
Equivalent stress	σ_4	529 MPa	Equivalent strain	e_4	0.095
Equivalent stress	σ_5	546 MPa	Equivalent strain	e_5	0.118
Equivalent stress	σ_6	559 MPa	Equivalent strain	e_6	0.14
Equivalent stress	σ_7	584 MPa	Equivalent strain	e_7	0.182

Table 5.1: Material properties used in the piece-wise linear plasticity material model of the component

by changing the interval radius. The reasoning behind each of these cases is different where in the first cases (1-3) the main goal is to quantify the bound on the output given a certain degree of uncertainty, and the final case (4) is an investigation on the effect of different levels of uncertainty.

5.4.1 General setup and quantities of interest

In this section, a detailed investigation of the interactions between the interval field and a generic impact-critical component is conducted under a load case that is defined on the full overlap crash test. The generic component is represented by a rectangular box, which has sides of 60 mm, a total length of 180 mm with a thickness of 2 mm, which is modelled by 2700 four-node shell elements, as illustrated in Figure 5.3a. The properties of the sheet metal used for these components are modelled using a piece-wise linear plastic model [137]. The corresponding parameters are listed in Table 5.1. Following the load case, the component is impacted by a rigid moving wall with a mass of 60 kg. The initial velocity is set to 56 km/h or an equivalent 15.6 m/s. This provides a total kinetic energy of 7300.8 J at the start of the simulation. In engineering practice, it is common to assess the performance of these crash boxes in terms of the peak force and the mean force that are generated during impact. The goal of a general engineering design optimisation for impact is to identify the input parameters, such that an acceptable performance threshold is met. Conventionally, in crash analysis the goal is to achieve a force that is as constant as possible during the deformation.

Peak force

The peak force is a measurement of the highest force that occurs during the impact simulation. In all considered cases, the peak force is measured at the rigid plane located at the back of the springs. The location is also indicated as a red plane in Figure 5.7. The peak force is measured directly from the output data without using any additional filtering:

$$F_{\text{peak}} = \max_{t \in \Delta t} F(t) \tag{5.1}$$

This causes this measurement to be noisy due to numerical instability of the explicit solution scheme. Typically, the peak force is measured just after the component and the rigid wall make contact, which initiates the start of the typical deformation folds. In general, high peak forces are avoided by car manufacturers as these are associated with high accelerations, which impose high forces on the adjacent structure and eventually the passengers leading to more severe injuries.

Mean force

The mean force is an average measurement of the force during impact and provides global information about the performance of a particular design. The mean force is calculated following Equation (5.2) where the total energy of the component E_{comp} is divided by the average final deformation $D(t_{\text{final}})$. In order to omit zero entries, only the force and deformation starting from impact until the kinetic energy is zero are considered $t_{\text{final}} : E_{\text{kinetic}}(t_{\text{final}}) = 0$, neglecting the elastic spring-back of the component.

$$F_{\text{mean}} = \frac{E_{\text{comp}}(t_{\text{final}})}{D(t_{\text{final}})}, \tag{5.2}$$

where $D(t_{\text{final}})$ is calculated as the average displacement between the nodes of the start and the end of the crash box.

5.4.2 Benchmark case

This first case is used as a benchmark where the boundary conditions are applied in a normal way, with two rigid planes in correspondence with the illustration in Figure 5.3b. Therefore, only a single simulation is performed as there are no uncertainties considered in this case. The result of this simulation is provided

by means of a force-displacement graph, shown in Figure 5.6. In this graph, the peak force and mean force measurements are indicated by a blue dot and an orange dashed line, respectively. Note that it is common within industry to filter the results of the numerical simulations of crash scenarios, see e.g., [196]. However, as there are no experimental data to compare to, the results shown in this chapter are provided without the use of any filtering. It is clear that there is a large difference in peak force and mean force, which is not unexpected for a component with this geometry, which is not optimised in any sense. In an industrial environment, one would typically optimise the component such that the peak and mean force are more or less equal to each other, or below an a priori set threshold.

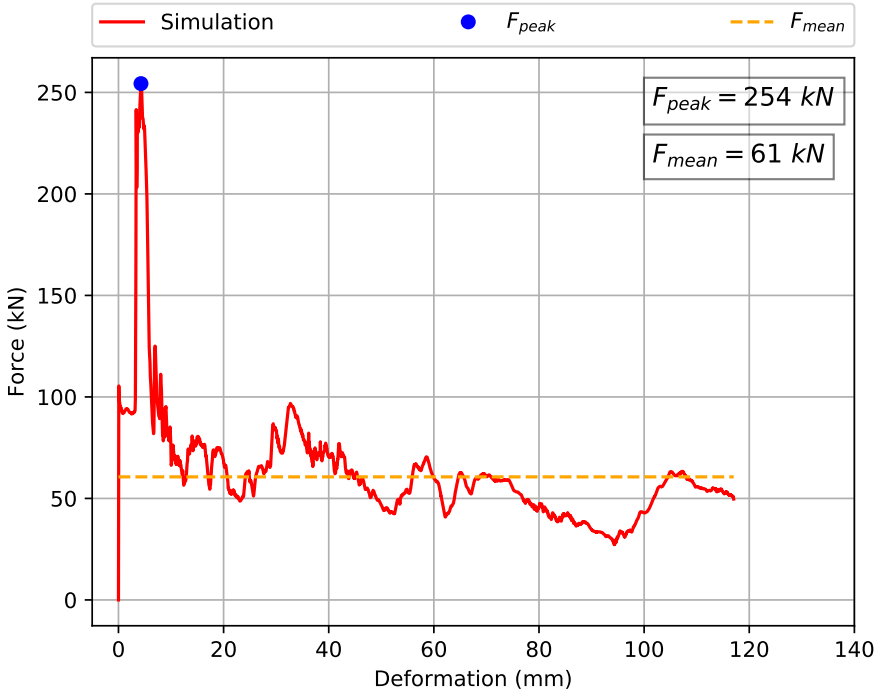
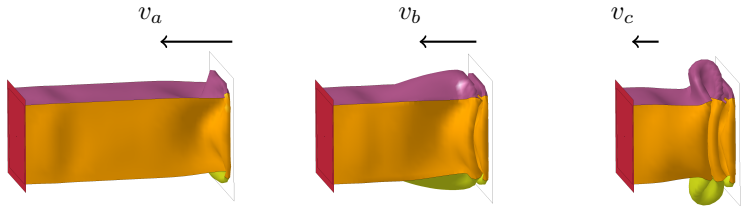


Figure 5.6: Force-deformation curve of the benchmark case without filtering, red; location of the peak force, blue dot; and the mean force, orange dashed line

The multiple peaks that are seen in Figure 5.6 are located at times where the force has built up until reaching a threshold before the next fold is initiated. This corresponds perfectly with the observed folding pattern, as illustrated in Figure 5.7. In this figure, the red plane is fixed and the white plane on

the right is impacting the structure, in correspondence with Figure 5.3b. The folding pattern shows that three folds are created during the first 12 ms of the impact, which is a local buckling mode starting at the impacting plane. The computational time for this simulation is approximately 2 minutes on two cores of an Intel(R) Xeon(R) CPU E5-2695 v3 @ 2.3GHz processor, which is reasonable in comparison to performing simulations on a full vehicle model.



(a) time = 3 ms (b) time = 6.5 ms (c) time = 12 ms

Figure 5.7: Deformation of the benchmark case with fixed boundary conditions at different time steps, with the fixed plane in red, and the impacting plane in white

5.4.3 Interval valued non-deterministic modelling of the adjacent structure

In the following case studies, the non-deterministic effects of the adjacent structure are modelled by a set of springs at the back of the component. These springs are illustrated in Figure 5.4. The lateral stiffness of each element is determined following a discretisation of the interval field. This interval field with IDW basis functions is used to model the spatial dependency of the element stiffness, which corresponds to the physical reality where the point-wise deformation of two points in an adjacent component is also dependent on the neighbouring areas in this component. As such, the interval field models the spatial distribution that is in a full-scale analysis provided by adjacent connecting elements. Additionally, it is shown that it is needed to optimise the input parameters of the interval field to obtain the worst case response of the structure, i.e., a response that results in structural failure defined as higher accelerations.

Interval valued stiffness of the adjacent structure

For this case, the bounds of the lateral stiffness interval are considered to be given as $k_t^I = [200;\ 330]$ MPa. Furthermore, it is assumed that all elements take the same stiffness value. This assumption will not provide the worst-case bounds on the response as this would require the use of optimisation, which is used in general for non-monotonic problems [61, 160], and as will be applied in the latter case studies in this chapter. Nonetheless, the analysis is performed with these assumptions to illustrate the effect the elements have on the overall performance of the crashbox. The results of this case study are illustrated in Figure 5.8. In this figure, the force deformation curves of this case are compared with those of the benchmark case. Figure 5.8 shows that both the obtained mean force as well as the peak forces are lower than those of the benchmark case, by 5.4 kN and 15 kN respectively. In addition, it also shows that the peak force is reached at a lower deformation in both cases. This behaviour is explained by the elements that absorb, and therefore deform, a part of the kinetic energy especially at the start of the impact, which is shown in Figure 5.9. This figure shows that the time to absorb the kinetic energy is both higher and lower depending whether the lower or upper bound is used. Hence, the time to build up the force and initiate the first folds is increased.

Figure 5.9 shows that in the final stages of the impact event the elements set at the lower value of the stiffness accumulate more energy than the stiffer elements (indicated in red), which causes the total kinetic energy to be absorbed sooner. Therefore, the elements influence the time in which the kinetic energy is absorbed by the component and the amount of kinetic energy, as a part remains within the elements. The latter is of course an undesired effect as these components are designed to dissipate a certain amount of kinetic energy. Therefore, care should be taken to limit the amount that is elastically stored within the springs. Moreover, this figure also shows the energy accumulated by the deformation of the component E_{comp}, the kinetic energy E_{kin}, hourglass energy E_{hg}, and the total energy E_{tot} of the simulation, which are truncated at the time all kinetic energy is dissipated.

Finally, Figure 5.10 illustrates the deformation pattern at specific time steps where the results of the upper row are set at the upper limit of the interval and the second row is set at the lower interval value. It is clear from this figure that the deformation in both runs is quite similar but with a small time delay for the lower limit, which is also less deformed at the end of the impact after 18 ms.

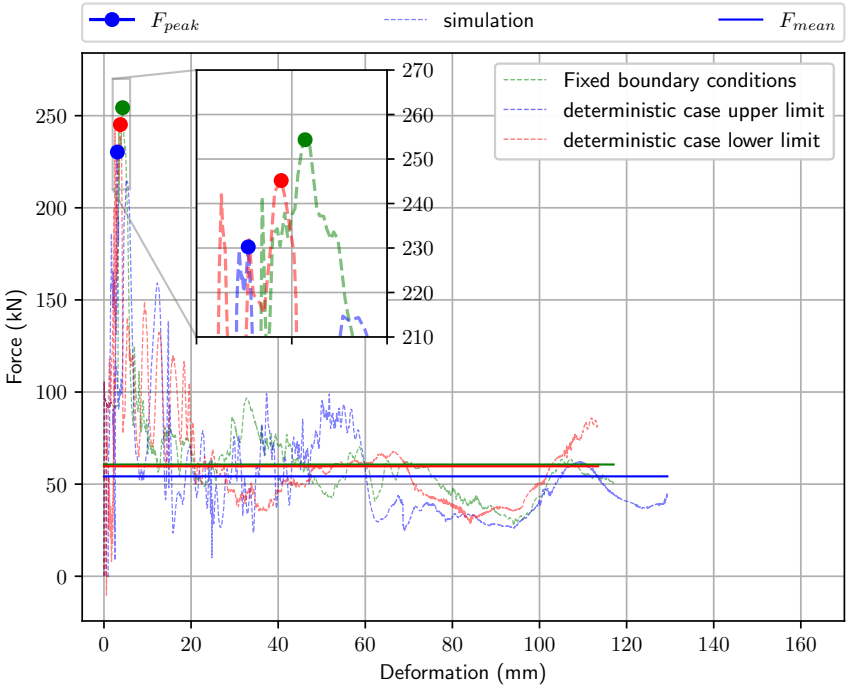


Figure 5.8: Force-deformation curves of the interval valued case with the minimal and maximal peak- and mean force indicated by arrows and solid lines, in blue and red

Interval valued spatial uncertain stiffness controlled at the corners

In this case study, the interval valued stiffness for the elements is assumed to be spatially coupled, while the size of the interval is identical to this of the previous case. As explained earlier, this corresponds to the physical presence of the adjacent structure. The stiffness values of the elements are coupled by means of an interval field. In this interval field, a set of discrete control points are placed at the corner nodes of the crashbox. Further, rather than modelling the stiffness of each of the 60 elements separately, only 4 parameters are required. This is advantageous from a computational standpoint. The interval field used in this case is defined in section 2.5.2 with basis functions that are based on IDW with $p = 2$, and the interval of the lateral stiffness is assumed to have a midpoint of $\hat{x} = 265$ MPa with a radius of $\Delta x = 65$ MPa, which corresponds to the interval used in previous case $k_t^I = [200;\ 330]$ MPa. The bounds of the response

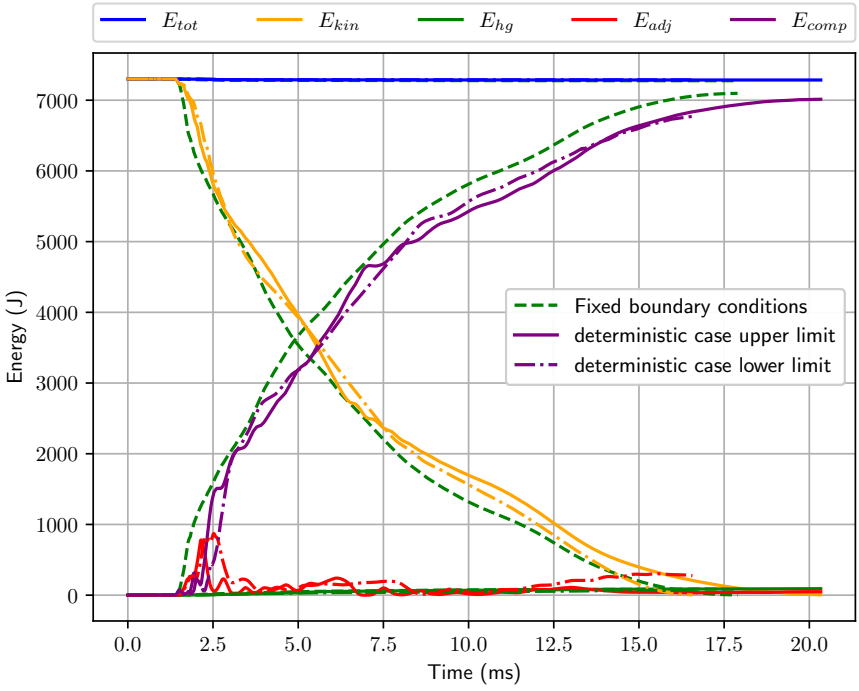
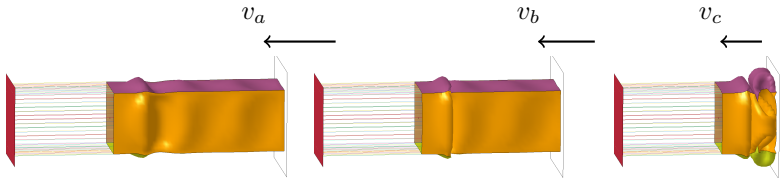


Figure 5.9: Energy balance of the interval valued case with: the total energy E_{tot} blue, the kinetic energy E_{kin} yellow, hourglass energy E_{hg} green, the energy of the adjacent structure E_{adj} red, and the energy of the component E_{comp} purple

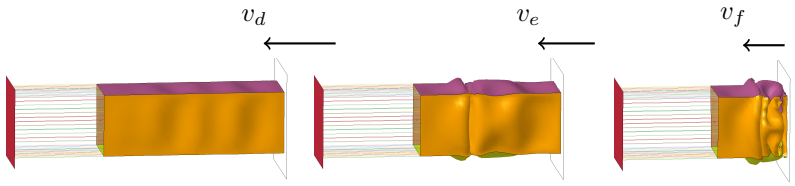
are in this case estimated by global optimisation using a differential evolution algorithm (DE). DE uses different populations for each generation within the input space to actively search for the global minimum. The results and the settings for the optimisation algorithm are summarised in Table 5.2, where the interval scalars are denoted with an $*$ when obtained through optimisation $\alpha^* = \max m_i(\mathbf{x}^I)$.

From the summary in Table 5.2 and the corresponding force-deformation curves in Figure 5.11, it is clear that when optimisation is used to actively search for the bounds, a larger interval is found for both, the mean force and the peak force. Especially in comparison with the previous case, it is clear that variation of the stiffness between elements yields larger bounds on the response for both quantities of interest. Figure 5.11 further illustrates that both the minimal



(a) time = 2.5 ms (b) time = 5 ms (c) time = 18 ms

Deformation obtained by propagation of upper limit



(d) time = 2.5 ms (e) time = 5 ms (f) time = 18 ms

Deformation obtained by propagation of lower limit

Figure 5.10: Deformation of the interval valued case at identical time steps, with the fixed plane in red, and the impacting plane in white

results of DE using control points at the corner nodes										
value	pop	rec	mut	nfal	nit	α_1^*	α_2^*	α_3^*	α_4^*	Optimisation
224.4	20	0.3	[0.9 1.7]	1155	56	0.0634	0.9663	0.1508	0.2748	min F_{peak}
261.7	20	0.3	[0.9 1.7]	220	8	0.4767	0.6706	0.6464	0.3902	max F_{peak}
42.2	20	0.3	[0.9 1.7]	1325	59	0.5267	0.9036	0.5703	0.0222	min F_{mean}
61.4	20	0.3	[0.9 1.7]	9850	257	0.7300	0.1332	0.8988	0.1077	max F_{mean}

Table 5.2: Results of the case study with control points at the corner nodes, including the DE parameters: population size (pop), recombination constant (rec), mutation constant (mut), with the number of evaluations (nfal) and iterations (nit) needed to identify the optimal interval scalar parameter α^* for the different optimisation runs

mean force as well as the minimal peak force are very low in the region between 30 and 80 mm of deformation, before starting to increase again. The cause of this effect can be seen in the deformation pattern in Figure 5.12, where it is clear that a global buckling mode is activated. This causes the crashbox to "fold" and lose its structural integrity. The force is only going up after 140 mm of deformation because the collapsed structure is still between the two rigid planes and is starting to get further compressed. Hence, it is argued that from this level of uncertainty realisations are possible where the performance of the

component is no longer guaranteed, as the global buckling mode prevents the dissipation of the kinetic energy.

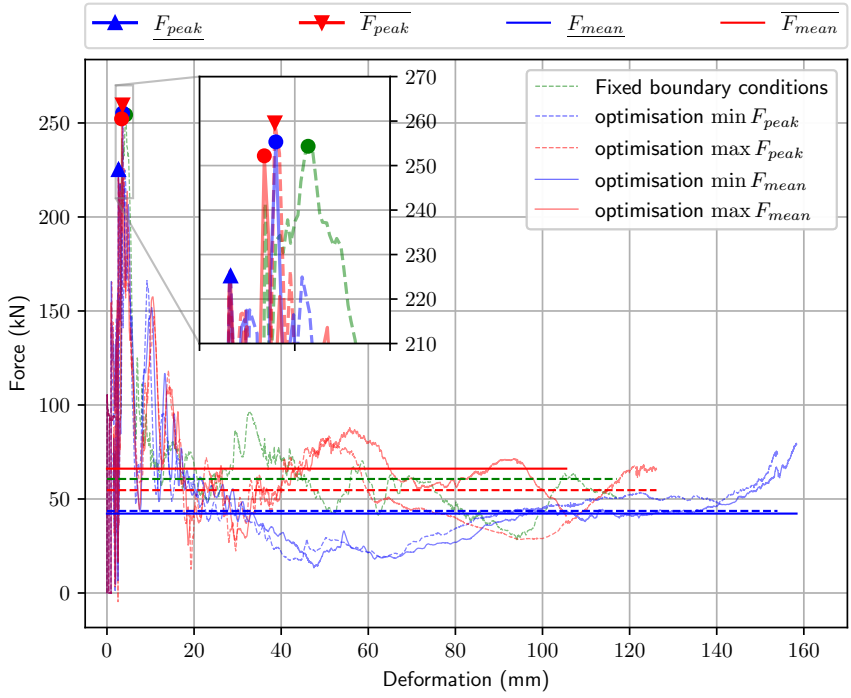
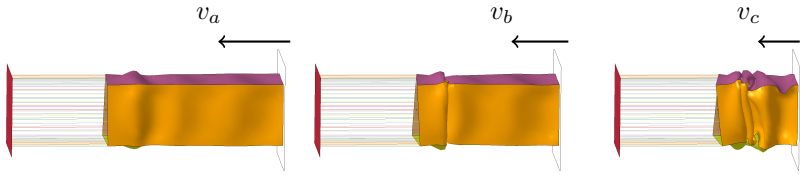


Figure 5.11: Force-deformation curves for the case with four control points at the corners obtained through optimisation; with the minimal and maximal peak- and mean force indicated by arrows and solid lines, in blue and red

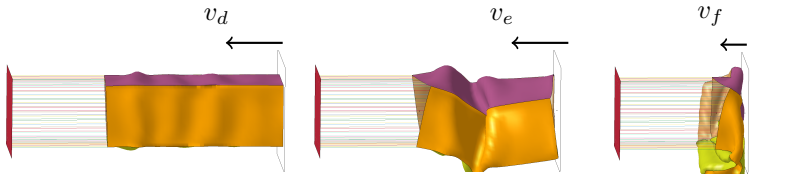
Figure 5.11 also shows that the optimisation procedure yielded a mean force which is higher than the mean force that was obtained in the benchmark case. This is illustrated in green colour. The reason for a higher mean force is found in Figure 5.12. Based on this figure, the higher mean force is attributed to a more dense folding pattern. Because of this denser folding pattern, the total deformation of the crashbox is also shorter than for the benchmark case, which can also be seen in Figure 5.11. This indicates that, for an equal kinetic energy, interactions between the component and the adjacent structure can result in mean forces both higher and lower than these identified with the benchmark case.

To gain a better understanding of these interactions, it is also useful to look



(a) time = 2.5 ms (b) time = 5 ms (c) time = 18 ms

Deformation resulting in the maximal mean force



(d) time = 2.5 ms (e) time = 5 ms (f) time = 18 ms

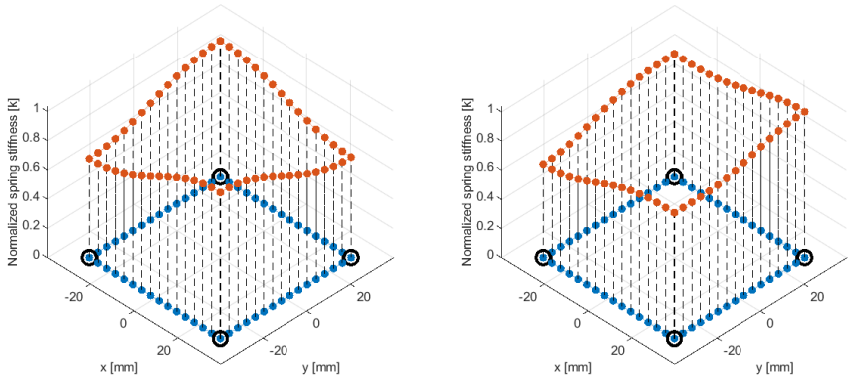
Deformation resulting in the minimal mean force

Figure 5.12: Deformation of the case with four control points at the corners at identical time steps, with the fixed plane in red, and the impacting plane in white

at the realisations of the interval field. These realisations are illustrated in Figure 5.13 with the left hand realisation corresponding to the maximum peak force and the right hand configuration to the minimum mean force. The control points of the interval field are indicated by a black circle in this figure and the normalised stiffness of the elements is indicated by the relative length of the black dashed lines. It is clear from this figure that the global buckling mode is obtained by a realisation that resembles a plane which is placed at an angle, while the maximum mean force is obtained by making differences between opposite corners. These realisations are not only interesting from the point of UQ as they can also assist in the way these components are manufactured and joined together, which initiates relative changes of stiffness.

Interval valued spatial uncertain stiffness controlled between the corner nodes

For this case, the locations of the control points are changed, which directly influences the possible realisations of the interval field. A summary of the results obtained through optimisation using a differential evolution algorithm is given in Table 5.3. In a comparison with the previous case it is noticed that there is a



(a) Realisation of the interval field according to $\alpha^* = \max m_i(\mathbf{x}^I)$

(b) Realisation of the interval field according to $\alpha^* = \min m_i(\mathbf{x}^I)$

Figure 5.13: Realisations of the interval field with four control points, resulting in the minimal and maximum mean force; control points are indicated by a black circle and the length of the dashed line indicates the normalised stiffness value

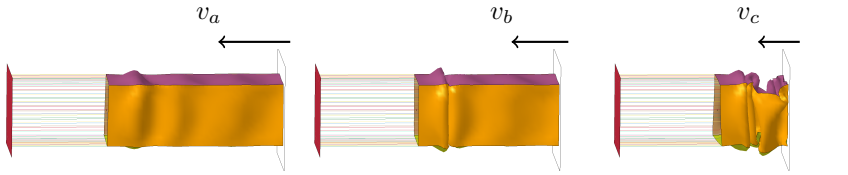
results of DE using control points between the corner nodes										
value	pop	rec	mut	nfal	nit	α_1^*	α_2^*	α_3^*	α_4^*	Optimisation
224.3	20	0.3	[0.9 1.7]	1140	55	0.0391	0.2447	0.8808	0.3498	$\min F_{\text{peak}}$
264.0	20	0.3	[0.9 1.7]	320	14	0.5126	0.7282	0.3702	0.6913	$\max F_{\text{peak}}$
41.9	20	0.3	[0.9 1.7]	5125	250	0.3034	0.9278	0.9953	0.1621	$\min F_{\text{mean}}$
62.8	20	0.3	[0.9 1.7]	975	44	0.5595	0.5065	0.9198	0.2090	$\max F_{\text{mean}}$

Table 5.3: Results of the case with control points between the corner nodes, including the DE parameters: population size(pop), recombination constant (rec), mutation constant (mut), with the number of evaluations (nfal) and iterations (nit) needed to identify the optimal interval scalar parameter α^* for the different optimisation runs

change in the upper limit of the peak force and the mean force, which indicates that this configuration allows for a different interaction with the elements.

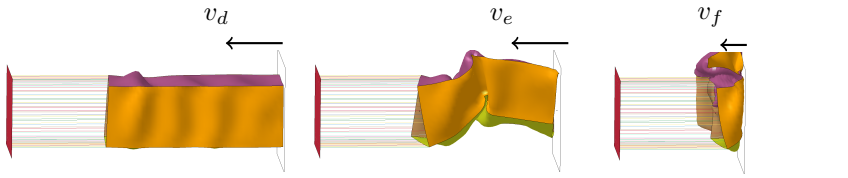
Figure 5.14 shows the deformation pattern that yielded the minimal and maximal mean force at different time steps. Compared to the previous case, these deformation patterns look quite different at a first glance, nevertheless when a closer look is taken it seems that these are more familiar to the previous cases,

seen from a different viewpoint. This could be the case as the configuration of the interval field is not unique, which causes the component to buckle in a different direction when the control points are rotated. This is not true in general as in this case the box is a simple symmetric geometry, which is not true in the presence of holes and fold initiators.



(a) time = 2.5 ms (b) time = 5 ms (c) time = 18 ms

Deformation resulting in the maximal mean force



(d) time = 2.5 ms (e) time = 5 ms (f) time = 18 ms

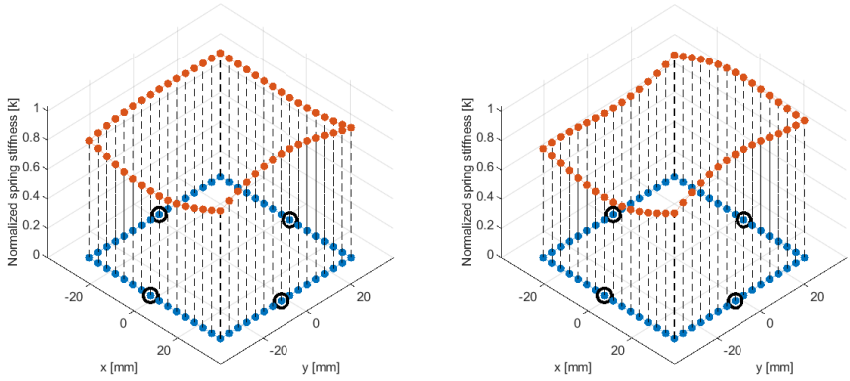
Deformation resulting in the maximal mean force

Figure 5.14: Deformation of the case with four control points at the corners at identical time steps, with the fixed plane in red, and the impacting plane in white

In addition, the realisations of the interval field are provided in Figure 5.15 where the different location of the control points are indicated by the black circle. The non-uniqueness in this case can be seen as rotating the interval realisation by 90 degrees, which yields the same results. It is also visible in this figure that the minimal mean force is obtained by a similar realisation as seen in the previous case. For the maximal mean force a different realisation is responsible for the observed differences.

Increased degree of freedom by placing additional control points

In this case, the degrees of freedom of the interval field are increased by placing additional control points, which allows the realisations of the interval field to have a more complex shape. Hence, this case represents a combination of the previous cases constructed by placing control points at both, the corner nodes



(a) Realisation of the interval field according to $\alpha^* = \max m_i(\mathbf{x}^I)$

(b) Realisation of the interval field according to $\alpha^* = \min m_i(\mathbf{x}^I)$

Figure 5.15: Realisations of the interval field with four control points, resulting in the minimal and maximum mean force; control points are indicated by a black circle and the length of the dashed line indicates the normalised stiffness value

results of DE using eight control points												
value	pop	nfal	nit	α_1^*	α_2^*	α_3^*	α_4^*	α_5^*	α_6^*	α_7^*	α_8^*	Optimisation
224.1	26	5652	350	0.08	0.05	0.79	0.55	0.65	0.68	0.19	0.02	$\min F_{\text{peak}}$
263.4	26	864	32	0.25	0.85	0.33	0.55	0.66	0.59	0.46	0.89	$\max F_{\text{peak}}$
41.3	26	5832	350	0.95	0.90	0.56	0.04	0.04	0.53	0.20	0.92	$\min F_{\text{mean}}$
64.7	26	5831	350	0.53	0.67	0.05	0.96	0.46	0.26	0.98	0.35	$\max F_{\text{mean}}$

Table 5.4: Results of the case using eight control points, here the DE parameters: recombination constant (rec), mutation constant (mut) are identical to the previous case, while the population size(pop), number of evaluations (nfal) and iterations (nit) needed to identify the optimal interval scalar parameter α^* for the different optimisation runs are provided

and between them. The results of this case are summarised in Table 5.4, which indicates that in general the bounds on the response have increased.

These increasing bounds are expected as the additional control points increase the dimension of the input space, which also results in an increased time to perform the optimisation. It is noticed that the optimisation algorithm quickly identifies realisations that result in a high or low mean force, and starts

optimising the elements to have the lowest stiffness that still initiates the global buckling mode. The mean reason is that after buckling of the component the moving rigid plane starts impacting the elements, which provide a lower force if they have a lower stiffness. This is observed by the fast increase in force in Figure 5.16 while the energy in the springs Figure 5.17 is not increasing. Hence, it can be argued that the component is not capable of dissipating all kinetic energy under this amount of uncertainty.

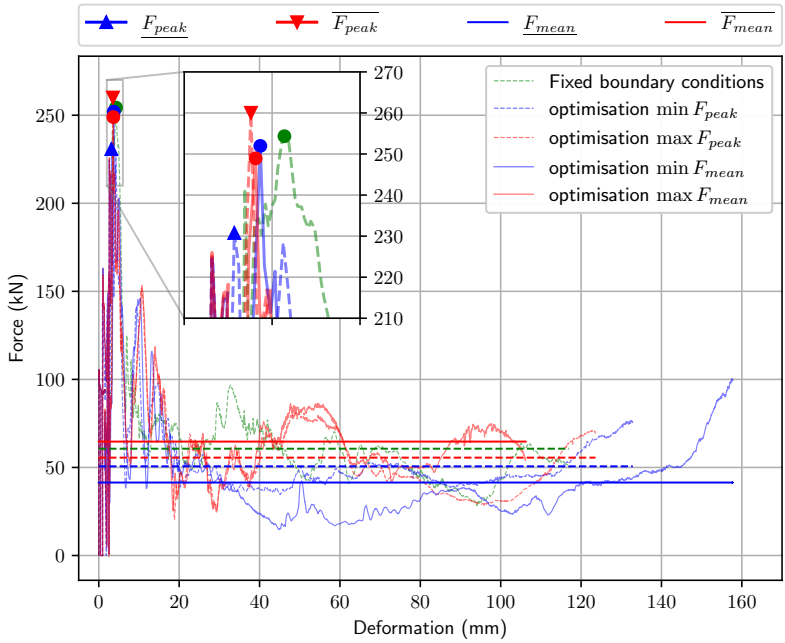


Figure 5.16: Force-deformation curves for the case with eight control points obtained through optimisation; with the minimal and maximal peak- and mean force indicated by arrows and solid lines, in blue and red

The realisations of the interval field with eight control points are shown in Figure 5.18b where the realisation yielding the maximum mean force is shown on the left and the minimal mean force on the right. As in the previous cases, the realisation that yields the minimal mean force is quite similar and the maximum mean force is a result of a more complex interaction with the elements at the end of the component.

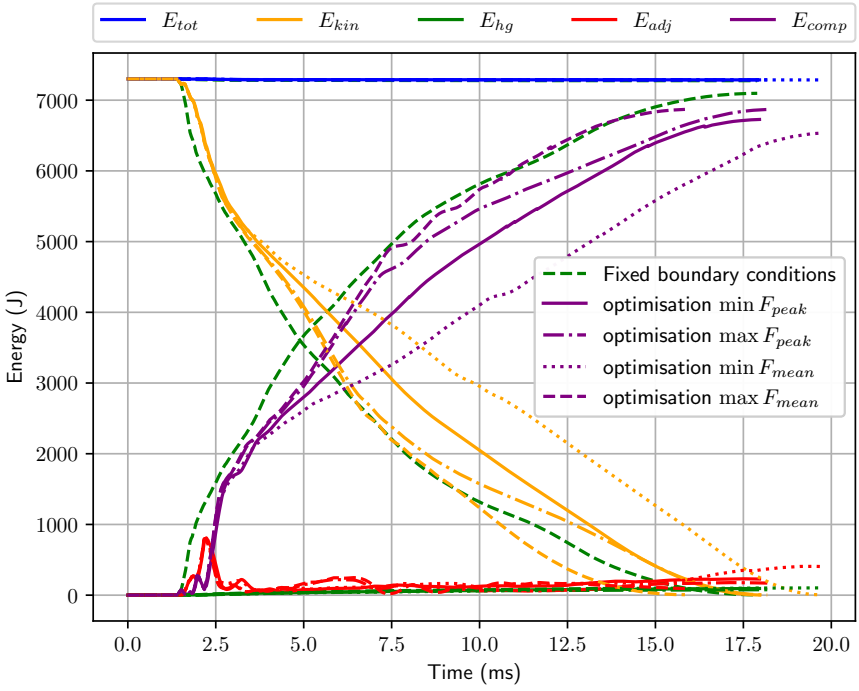
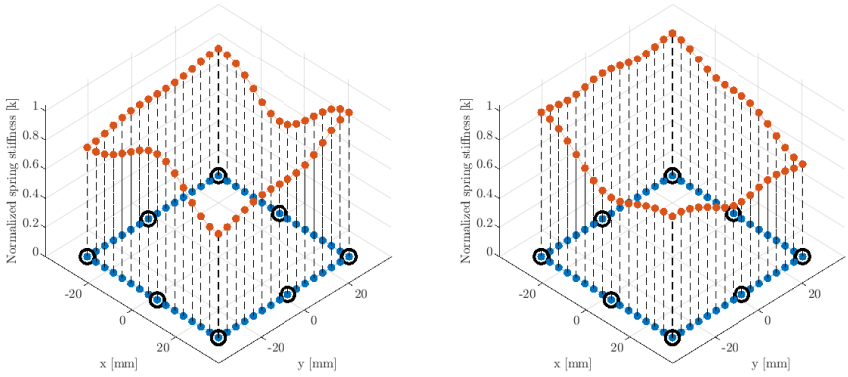


Figure 5.17: Energy balance of the interval valued case whit: the total energy E_{tot} blue, the kinetic energy E_{kin} yellow, hourglass energy E_{hg} green, the energy of the adjacent structure E_{adj} red, and the energy of the component E_{comp} purple

5.4.4 Interval field with increasing uncertainty

In this case the uncertainty in the interval field model, quantified by the width of the bounds, is varied by changing the radius of the interval Δx. This study is aimed at identifying the performance of the component under different levels of uncertainty. In this case, the level of uncertainty that allows to deform the component by a global buckling mode is of main concern as this prevents the component from fully dissipating the kinetic energy, which is the main purpose of this component. The corresponding level of uncertainty is identified by running a set of optimisations using the same settings as in Section 5.4.3 while the interval radius Δx is varied. The results of all these individual optimisations are provided in Figure 5.19 where each of the optimisation runs are identified by a marker for the upper and lower bound. Figure 5.19 shows that with an



(a) Realisation of the interval field according to $\alpha^* = \max m_i(\mathbf{x}^I)$

(b) Realisation of the interval field according to $\alpha^* = \min m_i(\mathbf{x}^I)$

Figure 5.18: Realisations of the interval field with eight control points, resulting in the minimal and maximum mean force; control points are indicated by a black circle and the length of the dashed line indicates the normalised stiffness value

increase of the interval radius a non-monotonic increase of the bounds on the output, indicated in red and blue, is obtained. Especially the large step made by the lower bound of the mean force between $\Delta x = 32.5$ MPa and $\Delta x = 34.5$ MPa is of interest as this indicates the transition between a folding pattern towards the global buckling mode, which is regarded as a failure. This is also observed in the deformation patterns, in the same figure, at a single time step of 5 ms, which illustrate the transition in the observed deformation pattern. This information can be used in a component optimisation where a better design is performing better under a wider range of uncertainty, which would make it more robust. This robustness is not limited to the component alone as it translates to the complete structure, which will meet the requirements under a wider range of circumstances.

In addition, the markers in Figure 5.19 show that the DE algorithm was unable to identify the same minimum that was obtained in another optimisation run, which yields the adjusted bounds identified by a circle. The adjustments of the bound for each of these circles was about ten times smaller in absolute value compared to the step that is observed at the minimum mean force bound. Although, this step occurs at a seemingly arbitrary value of $\Delta x = 32.5$ MPa, the important lessons are the different worst-case deformation patterns that

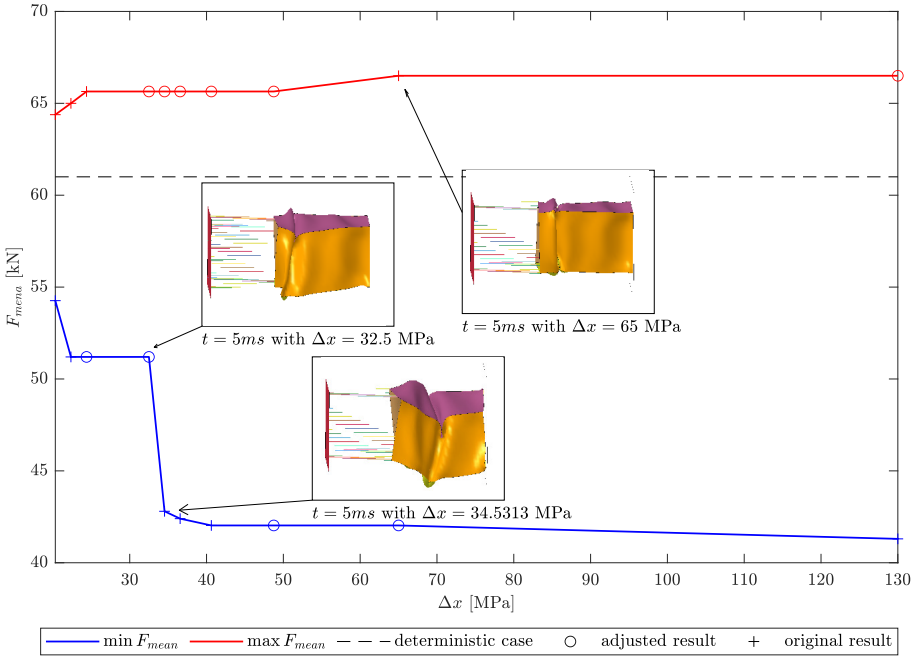


Figure 5.19: Bounds of the mean force identified through global optimisation for different values of interval radius Δx, with the corresponding deformation at identical times

are occurring. Moreover, this value of Δx can be seen as a measure of the robustness of the component with respect to the uncertain input. Hence, the robustness is interpreted as the ability of the component to perform within certain limits for a range of uncertainty.

5.5 Discussion

In the previous section, a number of cases are shown starting from a benchmark case, an interval valued case, interval field analysis and finally an interval field approach with increasing uncertainty. These cases illustrate the use and additional value of using non-deterministic modelling strategies in crash simulation. However, a number of important findings are further elaborated on in this section that allow for a more general discussion about the results.

The first finding is that the elements at the back of the component are also dissipating kinetic energy, which is a direct result of the stiffness of each spring and the reaction force of the component. This effect is first shown in the benchmark case, Section 5.4.2, where at the start of the impact energy is stored at the springs, which is released later. Nevertheless, in Section 5.4.3 it is also shown that the optimisation algorithm converges to a configuration of the elements that ensures failure of the component while maximising the amount of elastic energy stored within the elements. This configuration leads to lowest mean force after the component lost structural rigidity, which can be interpreted as failure. Therefore, energy storage at the end of the impact event stored within these elements is undesired and should be limited or accounted for within the optimisation, as this limits the kinetic energy dissipated by the component.

The second point that stands out in this analysis is related to the optimisation algorithm that is used to obtain the bounds on the output quantities. It is noted that the DE algorithm experiences some difficulties to reach a converged solution for some of the optimisation runs. Since, these simulations are quite time consuming, a limit on the maximum number of iterations of the DE solver has to be placed for practical reasons. Specifically, this bound was set at 350 iterations, which corresponds to about 5700 deterministic crash simulations. Figure 5.20 shows the convergence of the best candidate point at each iteration for the minimisation of the mean force, for the cases in Section 5.4.3. For each of these optimisation runs, the best candidate point is not improved for the last 50 iterations before reaching the maximum number of allowed iterations. Hence, this point is accepted as the global minimum with the knowledge that with a large number of additional iterations a better candidate point might be identified. In the authors opinion this is not justified by the additional computational cost that would be required. Note that it is not possible in general to prove that the global minimum is identified using global optimisation approaches in combination with non-convex functions.

In addition, Figure 5.19 shows that for the multiple independent simulation runs different global minima were identified, which were not always lower than the minima found at a lower interval radius. The bounds of these optimisation runs, marked by a circle, are adjusted to the previously identified minima. One of the reasons for these difficulties is of course the heavy non-linear response of the crash model with respect to the uncertain input parameters. To illustrate this, Figure 5.21a shows the function evaluation of 5000 samples, for which $X_{1,2}$ are generated by a Latin hyper-cube and $X_{3,4}$ are set at zero. This figure shows that optimisation of this function is not trivial as there are multiple local minima and maxima, which means that small perturbations of the input parameters can easily lead to a different result. In addition, Figure 5.21b shows the same data as in Figure 5.21a represented as a two-dimensional colour

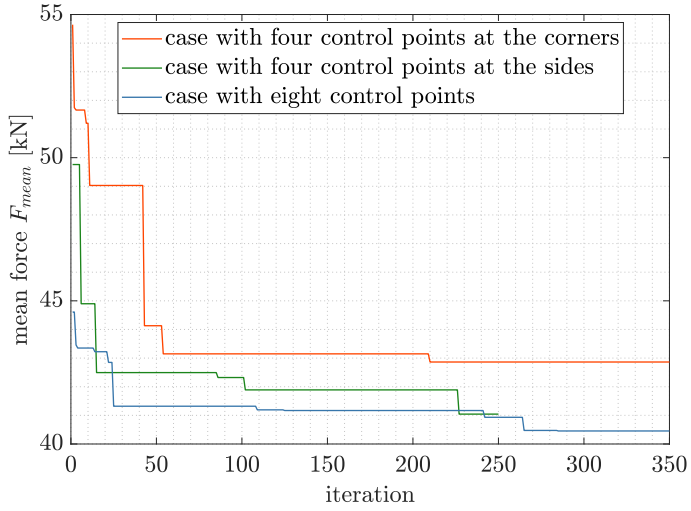
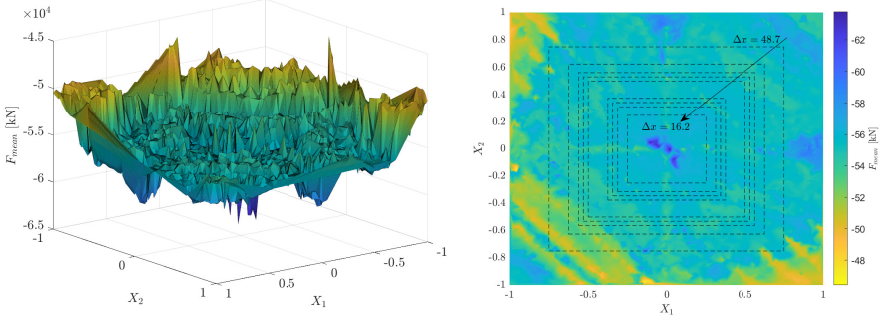


Figure 5.20: Convergence of the differential evolution algorithm for the minimisation of the mean force, described in section 4.3.2., 4.3.3., and 4.3.4.

plot. The rectangles in Figure 5.21b represent the input space dimensions that were used in the optimisation runs of Figure 5.19 with the edge of the figure representing $\Delta x = 65$ MPa. This figure shows the symmetry that exists between the interval field and the rectangular crash box, and some of the local minima and maxima. However, note that because of the interpolation used to create this colour plot some of these local effects are not well-represented. With these challenges in global optimisation of this function in mind it can be argued that the differences between the three cases in Sections 5.4.3, 5.4.3 and 5.4.3 are not that significant. This is especially interesting towards the case with eight control points, in Section 5.4.3 where the dimension of the search space doubled resulting in a much larger computational cost. This case demonstrates that there is a dependence between the number of control points of the interval field and the performance of the component. Hence, it is worthwhile to investigate this dependence as in a more complex case the presence of small triggers, e.g., holes, can lead to bifurcations.

A final point of discussion relates to the use of the peak force and the mean force measure for anti-optimisation of crash structures under uncertainty. Although these measures have a profound physical background and are widely used within the crash community, it is illustrated in Section 5.4.4 that optimisation on these responses is very hard. In addition, in Section 5.4.3 it is shown that over time the optimisation is more focused on storing energy within the elements than



(a) three-dimensional representation of the input and output

(b) two-dimensional representation of the slice

Figure 5.21: Slice of the input space with the mean force as a response, constructed with 5000 Latin hyper-cube samples for X_1 and X_2 while $X_{3,4} = 0$ the surface is then created by linear interpolation

it is at identifying bifurcation modes for the component. Therefore, further investigations should be made to a measure that captures the performance of the overall system with an output that is less prone to small bifurcations in optimisation.

5.6 Conclusions

This chapter introduces a new framework for modelling and evaluating the crashworthiness of a single component in an early development stage under epistemic uncertainty. This is accomplished by modelling the behaviour of the impacted adjacent structure as unknown but spatially coupled uncertain element stiffnesses. The interval valued performance of the structure is obtained using a global optimisation approach, which is shown to be challenging yet feasible for interval field analysis applied to crash simulation. Based on an academic case study, the results obtained by this innovative framework are demonstrated. The focus is on the sensitivity of the typical main quantities of interest, i.e., mean and peak force during impact, towards the uncertainty included in the interval field modelling strategy. In the presented case study, three dominant deformation modes are identified for the considered range of uncertainty, one of them a global buckling mode. The results indicate that even limited uncertainty in the adjacent structure can affect the deformation mode significantly, resulting in fundamentally different conclusions. In addition, by

investigating the realisations of the interval field, the cause of these deformation modes can be further analysed.

Although this work is aimed specifically at crash analysis, this technique can be applied to impact simulations in general. Especially in cases that typically consider fixed boundary conditions while the actual conditions are unknown, the interval field proves to be a powerful concept that allows to tackle uncertainty efficiently. Hence, in future work the combination with the Component Solution Spaces for early stage crash component design is further investigated, which allows for faster design of complex structures in a large and decentralised design process while guaranteeing overall system performance from an early design stage.

Chapter 6

Robust design optimisation under lack-of-knowledge uncertainty

This chapter was previously published as:
C. van Mierlo et al. “Robust design optimisation under lack-of-knowledge uncertainty”. In: *Computers & Structures* 275 (2023). Publisher: Elsevier, p. 106910

6.1 Abstract

Design optimization is common practice in engineering where the goal is to find the optimal combination of design parameters under prescribed constraints. However, some parameters may be impossible to define in a deterministic sense and may only be known with significant uncertainty. This limitation has led to an alternative definition of design optimality called robustness, where attention is payed to the variation around the optimal performance. Straightforward methods to solve robust optimization problems are usually limited in two ways: (1) the computation burden of the so-called ‘double-loop’ optimization problem hinders application to realistic models, and (2) the formalisms are typically limited to probabilistic descriptions of the uncertainty. This chapter presents a formulation of the robust optimization problem under interval uncertainty and proposes a new approach taking advantage of the so-called adaptive Gaussian

processes to solve it efficiently. The proposed surrogate approach mitigates the computational burden of the resolution, and a dedicated learning function is proposed to ensure iterative minimization of the surrogate modelling error and convergence towards the robust optimum. The algorithm uses a stopping criterion related to the level of confidence associated with the optimality of the solution. The approach is illustrated on six analytical and engineering benchmark problems.

6.2 Introduction

Current engineering practice involves the development and design of products that span an ever growing field of applications, while the performance of these products should also be guaranteed under a wide range of circumstances. In other words, the performance of a product should be only minimally affected by, e.g., load variations, changing environments, boundary conditions. The idea of products and processes that are insensitive to variations, e.g., in manufacturing, was pioneered by *Genichi Taguchi* who first applied his methodology on electrical circuits [218, 219]. However, the description of these variations, including the details about their underlying probability density functions (PDF's), is in general a challenging task. The main reasons for this are that the corresponding quantities are inherently variable, e.g. wind loads, there is incomplete knowledge about the quantity, e.g. direct measurement is challenging, or the designer is faced with a combination of both [84]. Additionally, in an early design stage, where the fundamental design decisions are made, only rough estimations of the quantities influencing the performance might exists. Historically, in engineering practice uncertainties are covered by safety factors. Although this approach is very straightforward, these safety factors will not provide information about the actual conservatism in the design. Therefore, numerous techniques for uncertainty quantification have been introduced during the last decades to account for these uncertainties. Typically, these techniques are categorised as probabilistic and possibilistic approaches [213]. The latter includes techniques as: interval [68, 163], fuzzy sets [100], information gap methods [22], and imprecise probabilities [19, 77]. In general, probabilistic methods are best suited for aleatory uncertainties as they describe non-determinism via random variables defined by their joint probability distributions, while possibilistic approaches are usually well suited to cover both aleatory and epistemic uncertainties.

In addition to the variety of possibilistic methods, different definitions of the robustness are proposed in literature; the relevance of which depends on, e.g., the application and the available information. For a review of different robustness measures under probabilistic uncertainty the reader is referred

to the work in [166, 179, 246]. In the context of possibilistic uncertainties, robustness definitions have been introduced in the framework of information gap theory [130, 192], convex models [8], and for fuzzy sets [20]. The definitions in these works are mainly based on two criteria: the first is minimising the variation of the output [20], and the second is to optimise simultaneously both the output (e.g. performance) and its variance around the optimal value [106, 157, 244]. In addition to the definition of robustness, a range of methods have been developed for its evaluation, with sampling strategies for most mixed uncertainty problems [20, 135], forward or inverse propagation [219], meta-model assisted methods [191], and fully decoupled methods for reliability based design optimisation [78].

This work focuses on developing a meta-model assisted method to determine the robustness at different design points. The meta-model that is used is a Gaussian Process (GP) model also known as Kriging [129, 194], which is used in this context as an emulator of the physical model. After calibration of the GP-model on a set of evaluated points, i.e., Design of Experiments (DOE), the model is fast to evaluate. Based on this easy to evaluate GP model, fast approximations can be made about the underlying problem, i.e., numerical model, and this approximation can be improved by increasing the calibration points in the DOE. The well-known framework of Efficient Global Optimisation (EGO) [118] successfully exploits the GP mean and variance to select additional calibration points and improve on the predicted minimum. In the specific case of interval uncertainties, the GP is used to estimate the interval width in un-sampled regions, including the confidence bounds about this estimate. Hence, the GP estimate can be used in place of the actual model for the optimization problem. The estimation will be affected by a modeling error but can be bounded by a confidence interval. An improvement function is proposed that finds the next point to evaluate as a compromise between its estimated robustness and the uncertainty regarding its estimation (high GP variance). The improvement function in this work is based on the work of M. De Munck et al. [52]. However, in this work some adaptions are proposed to the improvement function to efficiently perform the robust optimisation. To solve the robustness optimisation efficiently two improvement functions are combined: first an improvement of the interval width throughout the domain and second an improvement towards the most robust design. The combination of these two improvement functions provides a powerful improvement function that refines the GP model both globally and locally around the most robust design point. The proposed Robustness under Lack-of-Knowledge method is abbreviated as RULOK.

This chapter is structured as follows: Section 6.3 the robustness measure under interval uncertainty is introduced, while Section 6.4 provides the details

towards the Gaussian Process model that is used. Section 6.5 describes the adaptive sampling strategy that is used to calibrate the Gaussian process and the performance of this method is demonstrated on a number of cases in Section 6.6. Finally, in Section 6.7 a discussion about the results is held before conclusions are drawn in Section 6.8.

6.3 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. In practice, this kind of uncertainty is encountered when the best estimate of a parameter is limited to a range of possible values, even when it is based on all available data and/or knowledge. The *real* value of the quantity, be it deterministic or variable, is in this case represented by the bounds between which it is deemed to lie. Precisely, an interval is defined as:

$$\mathbf{x}^I = [\underline{\mathbf{x}}; \overline{\mathbf{x}}] = \{\mathbf{x} \in \mathbb{R}^{n_x} \mid \underline{\mathbf{x}} \leq \mathbf{x} \leq \overline{\mathbf{x}}\}, \tag{6.1}$$

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.

6.3.1 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 consists out of 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 $y \in \mathbb{Y} \subset \mathbb{R}$, with the set of admissible model parameters $\mathbb{Y}$, defined as:

$$m : y = m(\boldsymbol{\theta}). \tag{6.2}$$

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 [66]:

$$\tilde{y}_s = \left\{ y_j \mid y_j = m(\boldsymbol{\theta}_j); \mathbf{x}_j \in \mathbf{x}^I; j = 1, \ldots, n_q \right\}. \tag{6.3}$$

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 x_j such that $\tilde{y}_s$ is a 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. 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 corresponding optimisation problem is defined as:

$$\begin{aligned} \underline{y} &= \min_{\mathbf{x} \in \mathbf{x}^I} m(\boldsymbol{\theta}), \\ \overline{y} &= \max_{\mathbf{x} \in \mathbf{x}^I} m(\boldsymbol{\theta}), \end{aligned} \tag{6.4}$$

where $y^I = [\underline{y}; \overline{y}]$ is the solution interval. When a global minimum or maximum is found through optimisation, the exact output set bounds are obtained. However, it should be noted that the behaviour of the goal function with respect to the uncertain parameters is unpredictable in the case of strongly non-linear problems, which makes the computational effort highly problem dependent [159].

There is a special case for monotonic problems, where the vertices of the hyper-cubic input space are sampled, called the vertex method, introduced by Dong and Shah [55]. Following this method the output set is determined exactly within 2^{n_x} evaluations. However, the underlying assumption is that the model output behaves monotonically with respect to the input parameters, which is not true in general. Other approaches are intrusive methods to solve interval problems, which have been proposed in [172], and interval arithmetic methods as proposed in [210].

6.3.2 Defining robustness in the case of interval valued uncertainty

As mentioned in the introduction, multiple definitions of robustness exist, depending on the context and application. In this work, it is proposed to define robustness as the design with minimum variation in the performance given a well-defined input uncertainty. Following this definition, robustness can be

defined as the ratio of input uncertainty to the output uncertainty. However, quantifying this uncertainty is non-trivial in general. Therefore, the focus lies on the interval radius as a measure for the uncertainty. In this way, this robustness measure can be regarded as an interval counterpart to robustness measures that minimize the variance of the performance. For a case with one interval valued input parameter, the input and output uncertainty are represented respectively by the scalar interval radius Δx and the associated scalar output interval radius Δy. The output radius is a function of the design parameter $\mathbf{z}$ and should be evaluated for multiple designs $\mathbf{z} \in \mathcal{Z}$. The robustness for this case is defined as:

$$R(\mathbf{z}) = \frac{\Delta x}{\Delta y(\mathbf{z})} = \frac{\overline{x} - \underline{x}}{\overline{y}(\mathbf{z}) - \underline{y}(\mathbf{z})}, \tag{6.5}$$

Since Δx is independent of the design $\mathbf{z}$, finding the most robust design $\mathbf{z}^*$ is reformulated to the minimisation of the output uncertainty, which can be evaluated for multidimensional cases, defined by:

$$\mathbf{z}^* = \underset{\mathbf{z}\in\mathcal{Z}}{\operatorname{argmin}}[\overline{y} - \underline{y}] = \underset{\mathbf{z}\in\mathcal{Z}}{\operatorname{argmin}}[\max_{\mathbf{x}\in\mathbf{x}^I} m(\boldsymbol{\theta}) - \min_{\mathbf{x}\in\mathbf{x}^I} m(\boldsymbol{\theta})]. \tag{6.6}$$

Figure 6.1 illustrates the proposed robustness measure R for a point $\mathbf{z}^*$ and shows the associated upper bound $\overline{y}(\mathbf{z})$ and lower bound $\underline{y}(\mathbf{z})$, in red and blue. The point $\mathbf{z}^*$ is also the point with the maximum robustness R, indicated in orange. As suggested from Equation (6.6) finding the robustness of just one design involves a global optimisation to construct the conservative approximation of the solution set $\tilde{y}$, which should be repeated for each of the design points in $\mathcal{Z}$. Thus, crude optimisation of the problem described in Equation (6.6) involves two other optimisation problems: first an optimisation that actively looks for the upper-bound $\overline{y}$, and second, an optimisation that searches the lower bound $\underline{y}$, both for the same design $\mathbf{z}$. Therefore, crude optimisation is a time consuming effort, as this would involve a large number of evaluations of the model m under consideration. In an attempt to alleviate this problem, the next section discusses the use of a well-designed Gaussian process model $\mathcal{G}$ that could be used in place of the model m.

6.4 Gaussian process model for robustness under interval uncertainty

This section provides a short theoretical summary of GP models or Kriging [129][132], an introduction with examples is also available in [190]. A GP

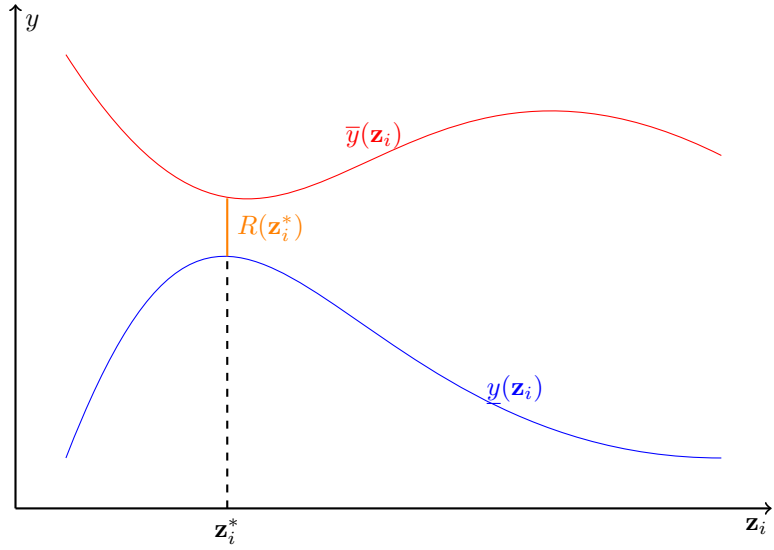


Figure 6.1: Illustration of the optimal robust design points $R(\mathbf{z}^*)$ (orange) for the upper and lower bounds $\overline{y}$ and $\underline{y}$ for a specific design parameter $\mathbf{z}_i$.

model is a stochastic meta-model that assumes $m(\boldsymbol{\theta})$ to be a realisation of a Gaussian process, which is defined as [198]:

$$\mathcal{G} = \beta^T \mathbf{f}(\alpha) + \sigma^2 \mathcal{F}(\mathbf{x}, \Omega), \tag{6.7}$$

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 $\mathcal{F}(\mathbf{x}, \Omega)$ 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 [3] for details about different covariance functions in Gaussian processes. In this chapter two well-known covariance functions are used: The Gaussian kernel (also known as squared-exponential covariance function) and the Matérn $\frac{5}{2}$ kernel.

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 [198]:

$$\mu_{gp} = \mathbf{f}(\mathbf{x})^T \hat{\beta} + \mathbf{r}(\mathbf{x})^T K^{-1}(\mathbf{y}_{DOE} - \mathbf{F}\hat{\beta}), \tag{6.8}$$

$$\sigma^2_{gp} = \sigma^2 \left(1 - \mathbf{r}^T(\mathbf{x}) + \mathbf{u}^T(\mathbf{x})(\mathbf{F}^T\mathbf{K}^{-1}\mathbf{F})^{-1}\mathbf{u}(\mathbf{x})\right), \tag{6.9}$$

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

$$\hat{\beta} = (\mathbf{F}^T\mathbf{K}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{K}^{-1}\mathbf{y}_{DOE}, \tag{6.10}$$

the general least-squares estimate of β and

$$\mathbf{u}(\mathbf{x}) = \mathbf{F}^T\mathbf{K}^{-1}\mathbf{r}(\mathbf{x}) - \mathbf{f}(\mathbf{x}). \tag{6.11}$$

Equations (6.8) and (6.9) are referred to as the mean and variance of the GP predictor, respectively. The GP that is used in this work is an interpolating GP, which means that the prediction of the variance at an experimental point $\mathbf{x} \in \mathbf{x}_{DOE}$ tends to zero.

6.4.1 Predicting interval bounds with a Gaussian Process model

In this work a GP-model is used to predict the output of the model m with as input $\boldsymbol{\theta}$ the set of uncertain and design parameters. To this end, μ_{gp} is considered to be the best GP-estimate and σ_{gp} is the confidence over this estimate. For the specific application of estimating an output interval based on the GP-model the main interest goes to the maximum and the minimum response over the complete range of uncertainty. Therefore, the bounds of the response are estimated by:

$$\overline{y}_{gp}(\mathbf{z}) = \overline{\mu}_{gp}(\mathbf{z}) = \max_{\mathbf{x}\in\mathbf{x}^I} \mu_{gp}(\boldsymbol{\theta}), \tag{6.12}$$

$$\underline{y}_{gp}(\mathbf{z}) = \underline{\mu}_{gp}(\mathbf{z}) = \min_{\mathbf{x}\in\mathbf{x}^I} \mu_{gp}(\boldsymbol{\theta}). \tag{6.13}$$

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

$$\overline{\delta}_{\mu+\sigma}(\mathbf{z}) = \max_{\mathbf{x}\in\mathbf{x}^I} (\mu_{gp}(\boldsymbol{\theta}) + c\sigma_{gp}(\boldsymbol{\theta})), \tag{6.14}$$

$$\underline{\delta}_{\mu+\sigma}(\mathbf{z}) = \min_{\mathbf{x}\in\mathbf{x}^I} (\mu_{gp}(\boldsymbol{\theta}) + c\sigma_{gp}(\boldsymbol{\theta})), \tag{6.15}$$

$$\overline{\delta}_{\mu-\sigma}(\mathbf{z}) = \max_{\mathbf{x}\in\mathbf{x}^I} (\mu_{gp}(\boldsymbol{\theta}) - c\sigma_{gp}(\boldsymbol{\theta})), \tag{6.16}$$

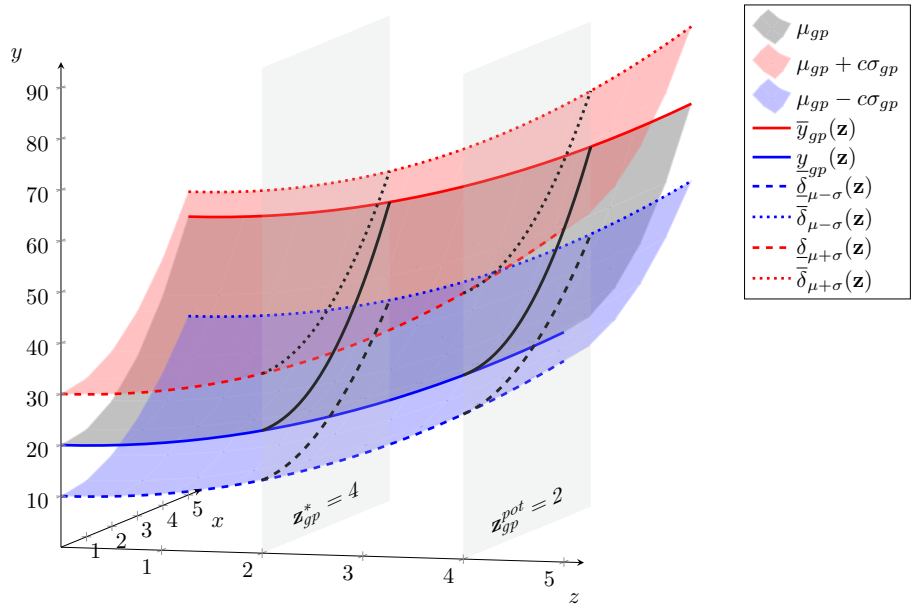
$$\underline{\delta}_{\mu-\sigma}(\mathbf{z}) = \min_{\mathbf{x}\in\mathbf{x}^I} (\mu_{gp}(\boldsymbol{\theta}) - c\sigma_{gp}(\boldsymbol{\theta})), \tag{6.17}$$

with $c\sigma$ confidence bounds. The bounds of the response are estimated for each design point $\mathbf{z}$, based on Equations (6.12-6.17). Figure 6.2a illustrates a simplification of the GP-model output for a single uncertain parameter $\mathbf{x} \in \mathbf{x}^I$ and a single design variable $\mathbf{z} \in \mathcal{Z}$. The upper bound of the output interval is determined by Eq. (6.12), indicated by the red line, and using Eq. (6.13) the lower bound is found, indicated by the blue line. In addition, the bounds based on the mean plus variance $\mu_{gp} + c\sigma_{gp}$ are predicted by Equations (6.14) and (6.15), indicated by the red dotted and dashed lines. Similarly, the bounds based on the mean minus the variance $\mu_{gp} - c\sigma_{gp}$ are given by Equations (6.16) and (6.17) are indicated by the blue dotted and dashed lines. Moreover, two designs $\mathbf{z}^*_{gp}$ and $\mathbf{z}^{pot}_{gp}$ are shown, illustrating the predicted behaviour along the uncertain parameter $\mathbf{x}$. Note that in general, for one specific design, e.g., $\mathbf{z}^*_{gp}$, the location of $\mathbf{x}$ for the predicted upper bound $\overline{y}_{gp}(\mathbf{z}^*_{gp})$ and the location of $\mathbf{x}$ of the maximum of the CI for the upper bound $\overline{\delta}_{\mu+\sigma}$ are different.

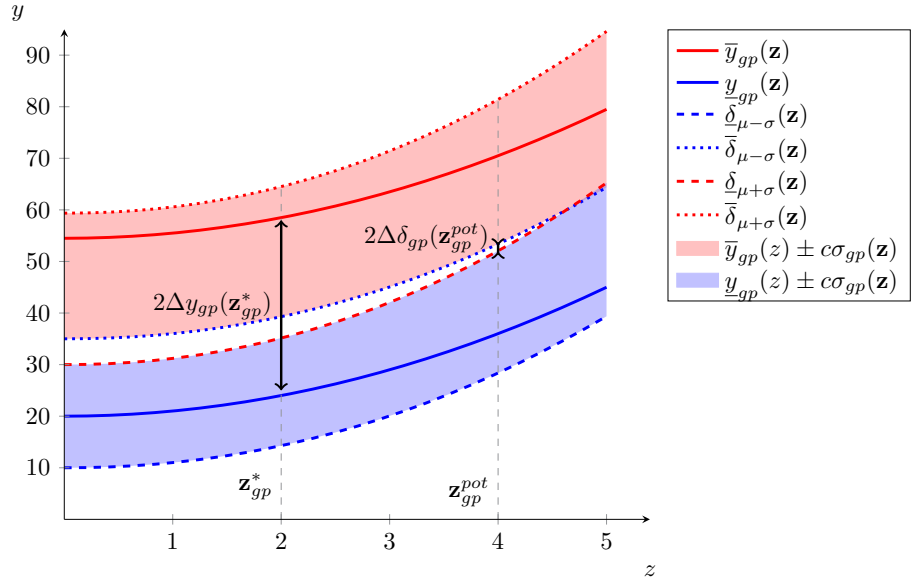
In the second illustration, Figure 6.2b, the estimated interval bounds are shown by the red and blue lines, with the CI about these estimates indicated by the red area for the upper bound, and blue area for the lower bound. Note that the red area is drawn between the upper bound of the minimum prediction and the upper bound of the maximum prediction by the GP-model. In addition, two designs $\mathbf{z}^*_{gp} = 2$ and $\mathbf{z}^{pot}_{gp} = 4$ are highlighted to illustrate the robustness measure. The robustness in Eq. (6.5) can be calculated based on these bounds given by the GP-model. Specifically, for the design $\mathbf{z}^*_{gp}$ the robustness is given by:

$$R(\mathbf{z}^*_{gp}) = \frac{\overline{x} - \underline{x}}{\overline{y}(\mathbf{z}^*_{gp}) - \underline{y}(\mathbf{z}^*_{gp})} \tag{6.18}$$

with $\overline{y}(\mathbf{z}^*_{gp}) - \underline{y}(\mathbf{z}^*_{gp})$ the estimated interval width, which corresponds to $2\Delta y(\mathbf{z}^*_{gp})$. Moreover, based on the CI it is also possible to estimate the potential interval width for $\mathbf{z}^{pot}_{gp}$, which would potentially have a higher robustness. To make this



(a) Illustration of the Gaussian Process model spanning the uncertain $\mathbf{x}$, design $\mathbf{z}$ and output $\mathbf{y}$ space, with the GP mean prediction μ_{gp} in black and the $\mu \pm c\sigma$ CI in red and blue surfaces; Two potential design points $\mathbf{z}^*_{gp}$ and $\mathbf{z}^{pot}_{gp}$ are shown as a slice.



(b) Illustration of the $(\mathbf{z})$-(y) perspective where the Confidence bounds around the upper-bound $\overline{y}_{gp}(\mathbf{z})$ and lower-bound $\underline{y}_{gp}(\mathbf{z})$ are given by the red and blue areas; and the two designs $\mathbf{z}^*_{gp}$ and $\mathbf{z}^{pot}_{gp}$ show the difference in mean bound prediction $\Delta y_{gp}(\mathbf{z}^*_{gp})$and the lowest CI prediction $\Delta\delta_{gp}(\mathbf{z}^{pot}_{gp})$ of the bounds

Figure 6.2: Illustration of the domain to determine the robustness based on the GP-model predictions

estimate the confidence bounds about the mean prediction are used:

$$R(\mathbf{z}_{gp}^{pot}) = \frac{\overline{x} - \underline{x}}{\overline{\delta}_{\mu-\sigma}(\mathbf{z}_{gp}^{pot}) - \underline{\delta}_{\mu+\sigma}(\mathbf{z}_{gp}^{pot})} \tag{6.19}$$

with $\overline{\delta}_{\mu-\sigma}(\mathbf{z}_{gp}^{pot}) - \underline{\delta}_{\mu+\sigma}(\mathbf{z}_{gp}^{pot})$ the estimated interval width, which corresponds to $2\Delta\delta_{gp}(\mathbf{z}_{gp}^{pot})$. The difference between these two robustness measures is that $R(\mathbf{z}_{gp}^{*})$ is estimated on the mean and the potential robustness $R(\mathbf{z}_{gp}^{pot})$ is estimated using the CI. Hence, the learning function introduced in Section 5 will exploit this difference, to search for designs with a potential higher robustness. Note that changing the constant c in Equations (6.14) to (6.17) from, e.g., 2σ to 3σ will enlarge the distance between red and blue surfaces.

In general, identifying the minimum and maximum as stated in Equations (6.12) until (6.17) is not trivial and involves numerous calls to the GP model. In addition, the mean and variance of the GP model are hard to use for optimisation as in the general case the problem is non-convex. Therefore, using a GP model is challenging for global optimisation methods. However, a number of successful strategies have been proposed to efficiently optimise such problems e.g., using branch and bound algorithms as proposed in [118]. In this work, the continuous problem is discretised over a grid with a fixed number of points. In that case, the complex problem of identifying the maximum and minimum reduces to identifying the highest value in a set of candidates. Note that this only works efficiently with a low number of parameters, as the computational burden increases exponentially $\mathcal{O}(n^d)$ with the d-dimensions of the problem for a full grid. In addition, an associated disadvantage is the finite accuracy achievable by the discretisation of the problem, with a finer discretisation causing a higher computational burden. The effects of discretisation can be mitigated in low dimensional problems by using a high number of grid points and changing the number of points to check the dependency of the solution on the discretisation.

6.5 Adaptive refinement of the Gaussian process model

To identify the robust design point in a limited number of evaluations of the model m the GP-model is adaptively refined with the specific goal of identifying the most robust design point. Therefore, the GP-model itself is used to identify regions of interest based on two criteria related to the famous compromise between exploration (low prediction confidence) and exploitation (identified areas of possible optimum). The learning function to achieve this is described in this section, starting first with an introduction of the maximum improvement function.

6.5.1 Maximum improvement function

The learning function introduced in this chapter is based on the maximum improvement function, which was introduced in [52]. Before applying this idea to the robustness problem as stated in the previous section, the general idea is briefly summarised. The goal of the learning function is to determine which sample is the best candidate to enrich the set of calibration samples for the GP. This effectively means improving the precision of the GP around the selected sample. Here this is illustrated on a general continuous function $f(\mathbf{u}) : \mathbb{R}^{n_u} \mapsto \mathbb{R}$, which is approximated by a GP-model $g(\mathbf{u})$. Using the learning function as defined in (6.20), a compromise is made between improving the calibration around the expected minimum using the GP mean (exploitation) and in areas of high prediction uncertainty based on the GP variance where a better minimum could be found (exploration). By iteratively enriching the calibration of the GP with the best sample improves the estimation of the minimum until a stopping criterion is eventually reached. The learning function is defined as [52]:

$$MI(\mathbf{u}) = \frac{\min(\mu_g(\mathbf{u})) - (\mu_g(\mathbf{u}) - c\sigma_g(\mathbf{u}))}{\min(\mu_g(\mathbf{u}))}, \tag{6.20}$$

with $\mu_g(\mathbf{u})$ the GP model prediction at $\mathbf{u}$, $\min(\mu_g(\mathbf{u}))$ the current minimum, and $c\sigma_g(\mathbf{u})$ represents the variance around the prediction of $\mathbf{u}$. Here, the variance is truncated at a certain confidence bound with c in Eq. (6.14). Hence, when the confidence bounds are based on, e.g., 3σ, more effort is dedicated to reducing the uncertainty about the approximation. Contrarily, lower confidence bounds, e.g., 2σ, reduce the confidence interval and favour improving approximately found maxima or minima. To identify the new candidate point $\mathbf{u}_{\text{new}}$ the maximum MI is identified over the domain $\mathbf{u} \in \mathcal{U}$ found by:

$$\mathbf{u}_{\text{new}} = \underset{\mathbf{u}\in\mathcal{U}}{\operatorname{argmax}} \left(\frac{\min(\mu_g(\mathbf{u})) - (\mu_g(\mathbf{u}) - c\sigma_g(\mathbf{u}))}{\min(\mu_g(\mathbf{u}))} \right). \tag{6.21}$$

Figure 6.3 shows the true function $f(\mathbf{u})$ in red and the GP based approximation $g(\mathbf{u})$ in black. The black dot is a point that is part of the DOE used to calibrate the GP-model. Furthermore, this figure shows how the learning function in Eq. (6.20) is used to evaluate the point $\mathbf{u}_{new} \in \mathcal{U}$ to determine which point should be added to the DOE. When the GP-model is re-calibrated using the newly evaluated point $\mathbf{u}_{\text{new}}$, the minimum of $f(\mathbf{u})$ is further approximated. If it is unlikely that a point $\mathbf{u}_{\text{candidate}}$ provides a minimum of $f(\mathbf{u})$ lower than the current $\min g(\mathbf{u})$, a negative improvement is obtained.

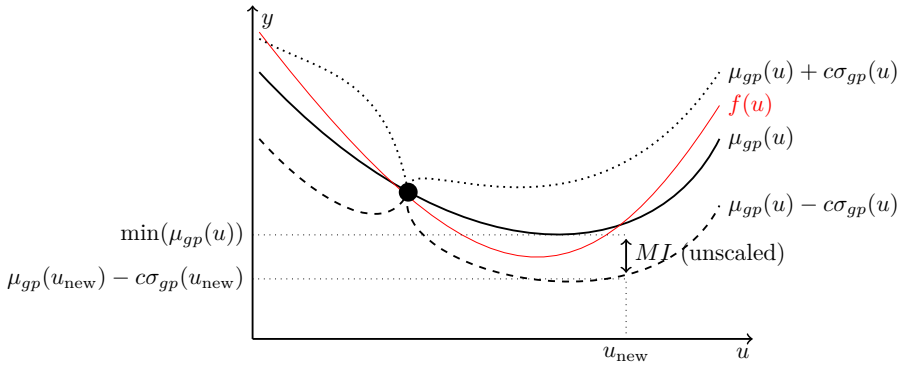


Figure 6.3: Illustration of the learning function where the goal is to approximate the minimum of the true function $f(u)$ in red, by the GP-model prediction $g(u)$; the black dot is the point that is part of the DOE and the next point u_{new} is selected by the learning function, adapted from [52]

6.5.2 Maximum improvement of the robustness

After the introduction of the maximum improvement, the remainder of section 5 describes how this is used in this specific case of robustness. 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, introduced in Eq. (6.20), to work directly on the minimum interval width. Specifically, it is adapted to:

$$MI_z(\mathbf{z}) = \frac{\min_{\mathbf{z}\in\mathbf{z}^I}\left(\overline{y}_{gp}(\mathbf{z}) - \underline{y}_{gp}(\mathbf{z})\right) - \left(\overline{\delta}_{\mu-\sigma}(\mathbf{z}) - \underline{\delta}_{\mu+\sigma}(\mathbf{z})\right)}{\min_{\mathbf{z}\in\mathbf{z}^I}\left(\overline{y}_{gp}(\mathbf{z}) - \underline{y}_{gp}(\mathbf{z})\right)}, \tag{6.22}$$

with $\overline{\delta}_{\mu-\sigma}(\mathbf{z}) - \underline{\delta}_{\mu+\sigma}(\mathbf{z})$ the predicted minimum bound $2\Delta\delta(\mathbf{z})$ with a confidence interval 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 minimum bound predicted by the mean estimate. Note that the mean estimated bounds correspond to $2\Delta y_{gp}(\mathbf{z}^*_{gp})$ in Figure 6.2b, and $\overline{\delta}_{\mu-\sigma}(\mathbf{z}) - \underline{\delta}_{\mu+\sigma}(\mathbf{z})$ to $2\Delta\delta_{gp}(\mathbf{z}^{pot}_{gp})$ in the same figure. By reaching a $MI_z(\mathbf{z}) \leq 0$, when the two intervals are equal, 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 $\mathbf{z} \in \mathcal{Z}$. Figure 6.4 illustrates in the top graph the improvement function where $\Delta\delta$ provides a possible smaller bound for the interval Δy. In the graph below the value for $MI_z(\mathbf{z})$ is given, illustrating that it is likely to improve the robustness at $\min \Delta\delta$.

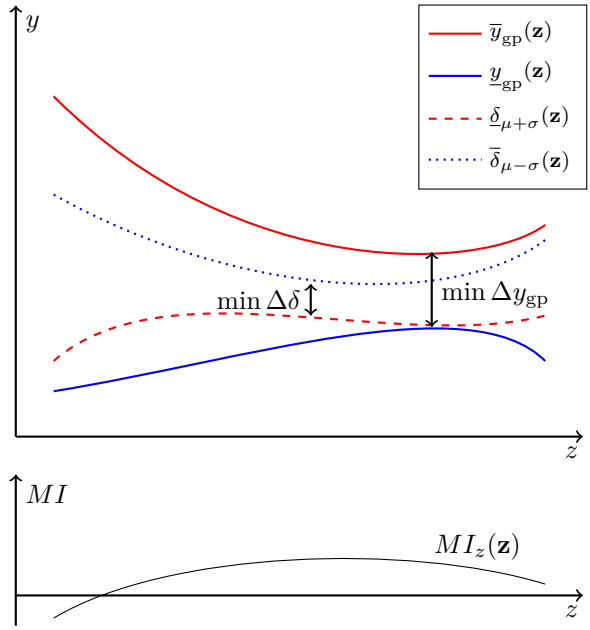


Figure 6.4: Illustration of the predicted mean bound $\Delta y_{\mathrm{gp}}(\mathbf{z}) = \overline{y}_{\mathrm{gp}}(\mathbf{z}) - \underline{y}_{\mathrm{gp}}(\mathbf{z})$ and the minimum bound based on the confidence interval $\Delta\delta(\mathbf{z}) = \underline{\delta}_{\mu+\sigma}(\mathbf{z}) - \overline{\delta}_{\mu-\sigma}(\mathbf{z})$

6.5.3 Maximum improvement of the predicted bounds

The previously introduced improvement function Eq. (6.22) finds a promising design point, based on the estimates of the GP-model. However, to estimate promising design points the overall GP-model must be refined as well, especially around these promising design points. Therefore, a second improvement function is introduced to increase the confidence of the predicted bounds. Here the maximum improvement Eq. (6.20) is adapted to obtain a best estimate of the minimal interval width $\Delta\delta_{\min}$, which depends on both the upper and lower bound. Figure 6.5 illustrates the idea behind the improvement function used here. In general, the goal is to approximate the output set $\tilde{y}_s$ for each design $\mathbf{z} \in \mathcal{Z}$. The point that provides the largest improvement of the lower bound of this interval is given as:

$$MI_{\min}(\boldsymbol{\theta}) = \min_{\mathbf{x}\in\mathbf{x}^I} \left[\mu_{gp}(\boldsymbol{\theta}) + c\sigma_{gp}(\boldsymbol{\theta})\right] - \mu_{gp}(\boldsymbol{\theta}), \tag{6.23}$$

and the improvement of the upper bound is given as:

$$MI_{\max}(\boldsymbol{\theta}) = \mu_{gp}(\boldsymbol{\theta}) - \max_{\mathbf{x}\in\mathbf{x}^I} \left[\mu_{gp}(\boldsymbol{\theta}) - c\sigma_{gp}(\boldsymbol{\theta})\right]. \tag{6.24}$$

Note that unlike the improvement functions in Equations (6.20) and (6.22) the one given in (6.23) and (6.24) are not normalized and calculated for each design in $\mathcal{Z}$. Hence, there is a guaranteed possible improvement even if the global minimum and maximum are identified. The improvement function is illustrated in Figure 6.5 for a single point $\mathbf{x}^* \in \mathbf{x}^I$. In the illustrated case, the improvement of the minimum bound $MI_{\min}$ is unlikely (negative value) while it seems likely to improve the upper limit $MI_{\max}$. In the end, only one candidate point can be chosen to be added to the design of experiments. Therefore, for each evaluated point the highest improvement value is used, which can either improve the lower bound or the upper bound:

$$MI_x = \max(MI_{\min}, MI_{\max}). \tag{6.25}$$

This means that for the illustration in Figure 6.5 only the value of MI_{max} is saved for the point $\mathbf{x}^*$.

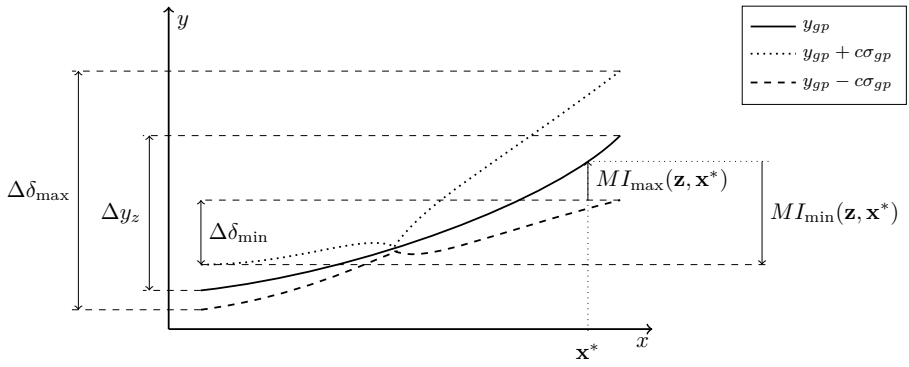


Figure 6.5: Illustration of the learning function for a candidate point x^*, showing the MI of the lower and upper bound; here the improvement of the lower bound is negative

Finally, the candidate point that performs best over the sum of the two improvement functions Eq. (6.22) and Eq. (6.25) is selected. Hence, the next candidate point $\boldsymbol{\theta}_{\text{candidate}}$ is obtained by:

$$\boldsymbol{\theta}_{\text{candidate}} = \underset{\mathbf{z}\in\mathcal{Z}\ \mathbf{x}\in\mathbf{x}^I}{\operatorname{argmax}} \left[MI_z(\mathbf{z}) + MI_x(\boldsymbol{\theta})\right]. \tag{6.26}$$

Note that it is possible here to assign weighting factors to the two functions. However, to the author's knowledge no advantage is gained in this regard. Hence, these weights are not used in this work.

6.5.4 Stopping criterion for adaptive refinement

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 $\min_{\mathbf{z}\in\mathbf{z}^I}\left(\overline{y}_{gp}(\mathbf{z}) - \underline{y}_{gp}(\mathbf{z})\right)$. This point is identified with $c\sigma$ confidence when the maximum improvement $MI_z \leq 0$. However, this is only achieved when the GP-model variance at location $\mathbf{z}^*$ reduces to zero. Although possible in theory, this is highly unlikely to be achieved in practice. Hence a small error term ϵ is defined, which assures that when:

$$MI_z \leq \epsilon, \tag{6.27}$$

there is with 95% confidence no point R within the domain smaller than $R(1+\epsilon)$. Unless explicitly specified otherwise, the default value for $\epsilon = 1\cdot 10^{-3}$ throughout this work.

6.5.5 Overview of the method

In Figure 6.6, a flowchart of the method is provided. The flowchart describes in detail the steps needed to perform the optimisation as proposed in this chapter. The method starts at the initialisation where all parameters are selected by the user, i.e., correlation function, 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 adaptive refinement starts with finding new potential robust designs points based on the learning function in Section 5. 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 Eq. (6.27) is met. Finally, after finishing the optimisation, it is highly recommended to validate and verify the results of the GP. A good starting point to check the accuracy of the GP-model is to perform Leave-one-Out (LOO) cross-validation with the points already in the Design of experiments.

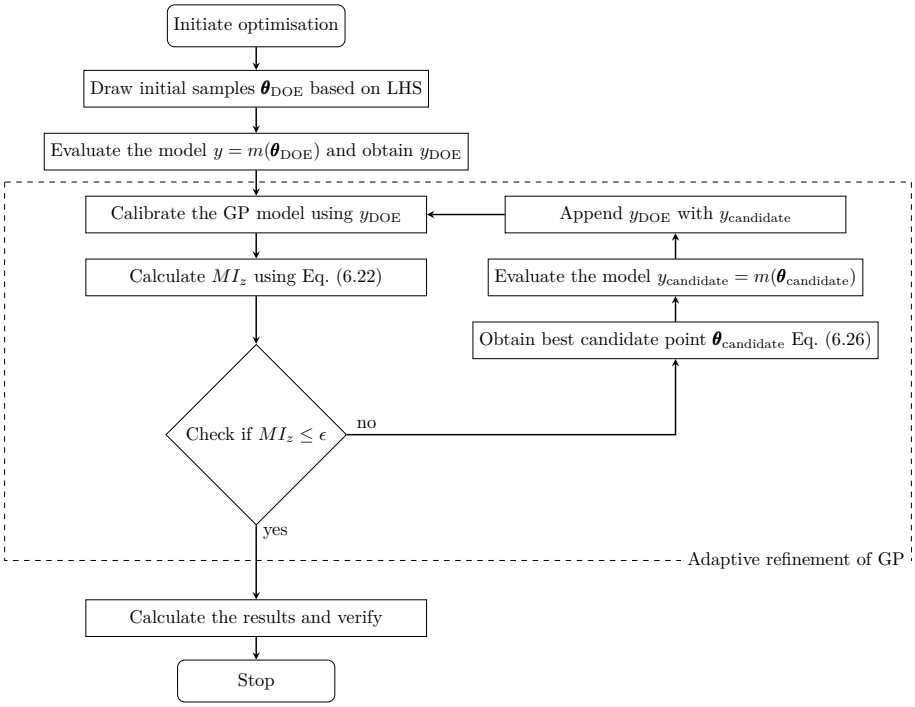


Figure 6.6: Flowchart of the robustness under lack-of-knowledge method

6.6 Case studies

In this section the RULOK technique is tested and validated for different problems, which start with a set of analytical functions and build up to higher dimensional engineering examples. For the first analytical cases a comparison is made with classical optimisation techniques, which require direct evaluations of the numerical model for each of the sample points. Moreover as the robustness measure in Eq. (6.6) requires a double-loop optimisation approach, where the outer-loop is focused on the next design point and the inner loop identifies the upper and lower bound of the response for a given design $\mathbf{z} \in \mathbb{Z}$. This optimisation directly uses the expensive to evaluate numerical model. Thus, the efficiency is measured in the amount of required function evaluations.

6.6.1 Analytical test functions

To study the basic properties of the proposed method a set of analytical test functions is used. Each of the three test functions presents a different challenge in terms of optimisation, starting form a convex and smooth function and progressing to non-convex problems. The analytical test functions are defined as:

$$f_a(x_1, x_2) = x_1^2 x_2 - x_2^2, \tag{6.28}$$

$$f_b(x_1, x_2) = x_2 x_1 - \sin(x_1) x_2^2 + x_1^2, \tag{6.29}$$

$$f_c(x_1, x_2) = \cos(4\pi x_1) - \sin(x_1 x_2) + x_2, \tag{6.30}$$

with $x_1 \in [-5, 5]$ the design parameter and $x_2^I = [-5, 5]$ the uncertain parameter. The goal of the optimisation is to identify the value for x_1 at which the bounds on Δf are minimal for each $x_2 \in x_2^I$. This optimisation is defined as:

$$\max R(x_1) = \min_{x_1 \in x_1^I} \left(\max_{x_2 \in x_2^I} f_n(x_1, x_2) - \min_{x_2 \in x_2^I} f_n(x_1, x_2) \right), \tag{6.31}$$

with n indicating the three functions f_a, f_b, f_c. In these particular cases, without the need for optimisation, one can determine that the minimum of the functions f_a, f_b and f_c lies at $x_1 = 0; \forall x_2 \in x_2^I$. Nevertheless to demonstrate the additional value of the proposed method two well-known optimisation algorithms are used in a comparison. These two optimisation approaches used in this work are: Unconstrained Optimisation (UO) where the minimum of a function is searched using a quasi-Newton algorithm; another strategy is to use a Genetic Algorithm (GA) to solve the outer-loop where the bounds of the response in the inner-loop are identified using UO. The population for the GA is set to a default value of 20.

The results of the method and those oobtained by classical optimisation approaches are compared in Table 6.1. It is noticed that the proposed method outperforms the brute optimisation approaches, which is expected with the use of a meta-model. The table also shows that depending on the level of confidence the number of iterations increases. Note that the amount of iterations needed to obtain a result is difficult to estimate a priori as this depends on the underlying problem and the correctness of the GP model at each iteration. The error term in the table refers to the discretization error introduced by using a fixed grid to sample the meta model. For both function f_a and f_b the optimal point is part of the samples in the grid using $n_{\text{samples}} = 501$. However, for function

f_c this dependence is checked and the optimal point is not part of the grid points $n_{\text{samples}} = 200$ or $n_{\text{samples}} = 500$. Therefore, the analysis returns the next best point, which is the closest to the optimal point. Using a larger number of grid-points will therefore increase the accuracy of the estimation at a higher computational cost.

Function	Method	Optimum	Iterations	Evaluations	Error*	confidence
f_a	analytic	0	-	-	-	-
f_a	RULOK	0	36	38	≈ 0.02	$1,96\sigma$
f_a	RULOK	0	42	44	≈ 0.02	3σ
f_a	UO	-7e-6	3	184	-	-
f_b	analytic	0	-	-	-	-
f_b	RULOK	0	28	30	≈ 0.02	$1,96\sigma$
f_b	RULOK	0	49	51	≈ 0.02	3σ
f_b	GA	0.099	77	50680	-	-
f_c	analytic	0	-	-	-	-
f_c	RULOK	-2.5e-2	279	281	0.05	$1,96\sigma$
f_c	RULOK	1e-2	172	174	0.02	$1,96\sigma$
f_c	RULOK	0	214	216	≈ 0.02	$1,96\sigma$
f_c	RULOK	-1e-2	242	244	0.02	3σ
f_c	GA	4.4e-5	30	2760857	-	-

* the discretization error of the grid is determined by $\Delta x/n_{points} = 10/501$ for f_a and f_b.

Table 6.1: Results of the analytic test functions

To further illustrate how the method works Figure 6.7 shows the function value for all three functions f_a,f_b and f_c at each design point x_1. For each function the true bounds are given by the black dashed lines, the evaluated points are indicated with a green cross, and the predicted upper- and lower-bound are given in red and blue, including their 95% confidence intervals, and the optimal design point is indicated by a circle. Starting at the top of Figure 6.7 function f_a is shown where the gradient decreases when moving towards the robust design point $x_1 = 0$. The middle sub-figure illustrates the function f_b with larger confidence bounds around the predicted optimum, shown by the red and blue areas. It is also shown that the CI about the upper bound is larger than the CI of the lower bound, which is exactly the goal during optimisation. Finally, the bottom graph of Figure 6.7 shows the more complex function f_c with the optimum at $x_1 = 0$. This figure illustrates the additional function evaluations needed to ensure the global minimum was found, and not one of the many local minima. Note that for this case the optimal robust point $x_1 = 0$ is not part of the grid as the grid is discretized by an even number of samples, which include the end and start point.

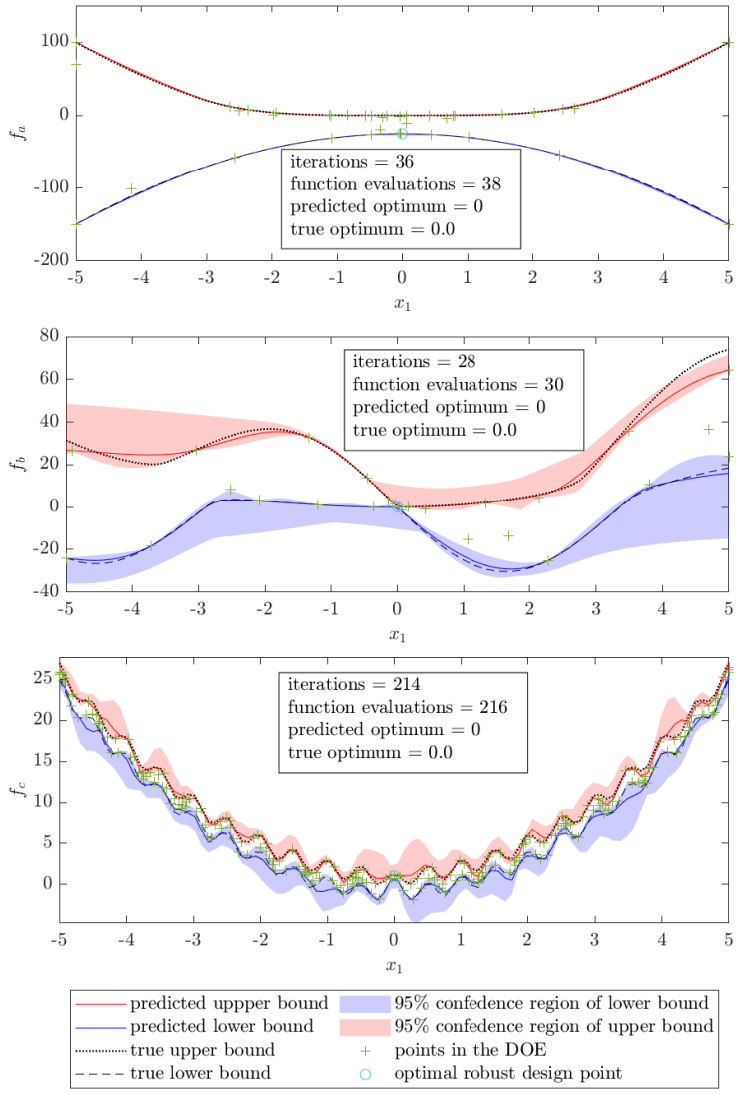


Figure 6.7: The GP predicted bounds of the interval valued uncertainty including the 95% confidence intervals; for from the top to the bottom function f_a,f_b and f_c, respectively.

In Figure 6.8 the meta-model of function f_a is plotted for all points within the domain with the black surface the mean response, the red and blue surfaces the lower- and upper-bound of the 95% confidence intervals, and the green dots are the points used to calibrate the GP-model. This figure illustrates the dispersion of the evaluation points at the edges of the domain and concentration of points around the optimal point, which reduces the variance of the GP-model at this location. Hence, the distance between the bounds increases in locations that are further from the optimal point since there are considerably less points evaluated here. Nevertheless, it is possible to use the GP-model further to analyse the problem at hand. However, one should be aware that due to the selection of training points an overall agreement between the GP-model and the underlying problem is not guaranteed.

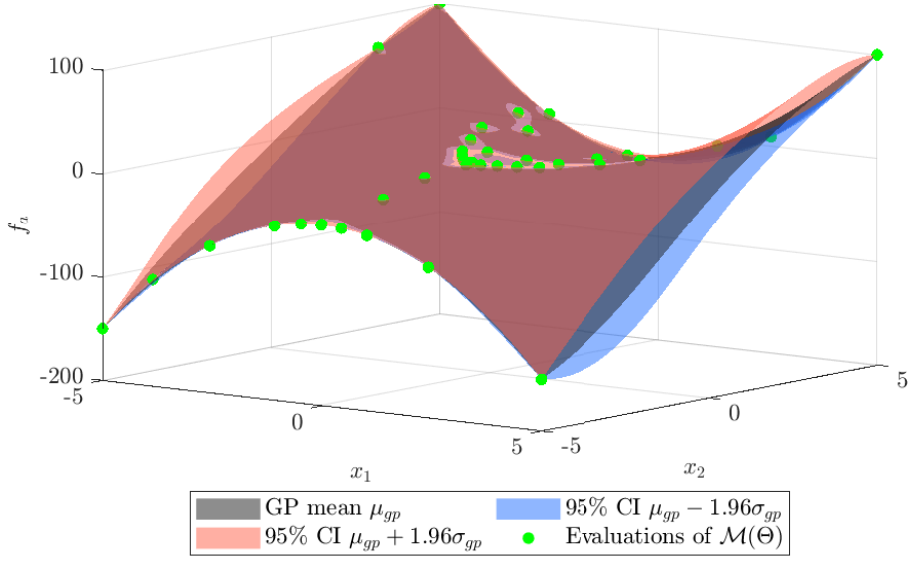


Figure 6.8: GP-model prediction of function A with the black surface the mean response, the red and blue surfaces the 95% CI on the mean prediction, and the green points indicating the evaluated points

6.6.2 Plate subjected to a point load

In this case study, the thickness of a plate with two equal sides of 100mm is chosen within the interval $t \in [3,\ 6]$mm. The uncertain parameter is the

Young's modulus of the material, which is known to be bounded by the interval: $E^I = [110,\ 280]$GPa. All degrees of freedom of one side of the plate are completely fixed and a point load of 100N is applied to one of the opposite corners. The performance of this design is measured by the displacement of the corner node that is subjected to the load. The analysis is performed by a FE model using 1000 shell elements.

At the start of the analysis two initial points are evaluated based on Latin Hyper-Cube sampling. Figure 6.9 shows the results that are obtained after just 9 function evaluations, with the true bounds of the model in dashed black lines, the GP-model prediction of the upper- and lower-bound in red and blue, and the confidence intervals as colored areas. The optimal design point $t_{\text{robust}} = 6$mm is as expected, the thickest plate. The rationale behind this simple example is that the thickest plate will bend less than a thinner plate under identical uncertainty of the Young's modulus. However, Figure 6.9 presents an illustration of the refinement around the optimal point, with only two evaluation points lower than 4.5mm. The order of the points that are added is further highlighted by the numbers next to the crosses in the plot starting with the initial evaluations 1 and 2, up to 9, the final point.

Although the physical interpretation of the problem explains the identified optimum a double-loop approach is used to validate this result. Here using CO a total of 82 evaluations of the numerical model were needed to identify the optimum $t_{\text{robust}} = 6$mm, which is identical. However, with this classical optimisation no additional information is obtained regarding the problem that is studied.

6.6.3 The borehole function

The second engineering example is the Borehole function [217], which is a typical test case for computer experiments. The borehole function describes the water flow f_{borehole} though a borehole between two underground aquifers by the flow rate of the water m^3/year:

$$f_{\text{borehole}} = \frac{2\pi T_u (H_u - H_l)}{\ln\left(\frac{r}{r_w}\right)\left(1 + \frac{2LT_u}{\ln\left(\frac{r}{r_w}\right) r_w^2 K_w + \frac{T_u}{T_l}}\right)}. \tag{6.32}$$

It is assumed that the diameter $r_w \in [0.05, 0.15]$ m, the length of the borehole $L \in [1120, 1680]$m of the borehole can be controlled and are therefore the design parameters. All other parameters are listed in Table 6.2. Two cases are considered with this example, first a case where only two parameters are

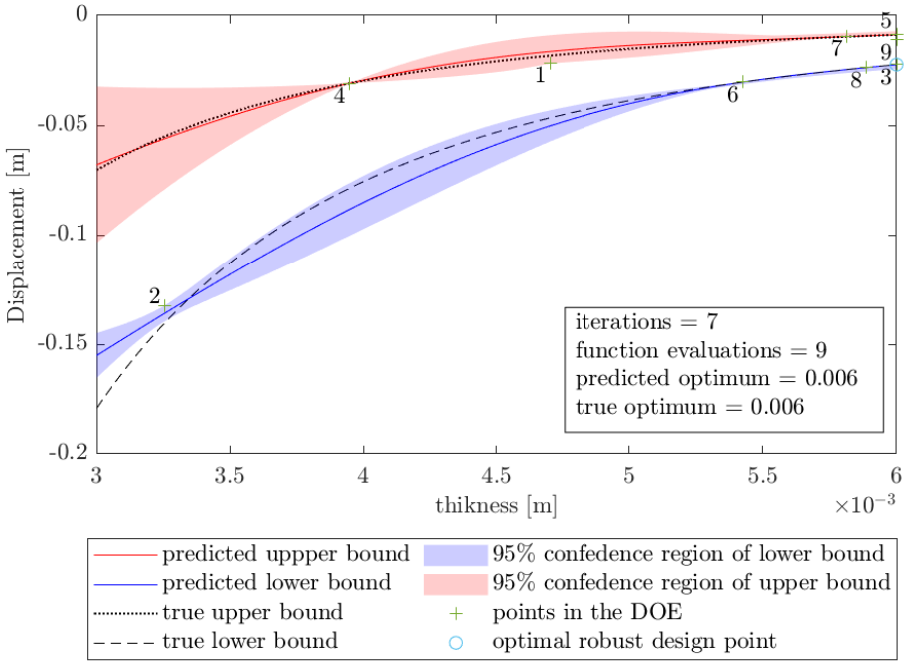


Figure 6.9: GP predicted bounds of the interval valued uncertainty including the order in which the points within the DOE where evaluated; including the 95% confidence intervals, evaluated points indicated by crosses, and the optimal design point indicated by a circle.

uncertain and the others are taken at the midpoint, second a case where all parameters are considered uncertain.

Borehole function with two uncertain parameters

In this first case only the potentiometric head of the upper aquifer H_u and the hydraulic conductivity K_w are regarded as uncertain. The remaining uncertain parameters are taken at the midpoint of their interval. The results of the analysis are shown in Figure 6.10, which shows a contour plot of the true interval width on the top, the predicted interval width based on the mean of the GP-model below, and the minimal interval width based on the 95% CI next to it. In all contour plots of Figure 6.10 the red circle and green dot indicate the location

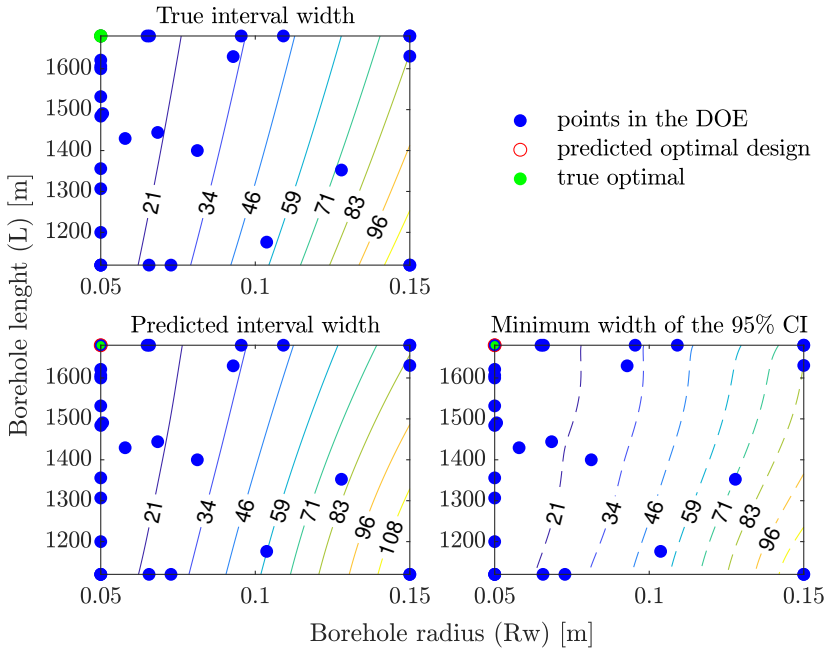


Figure 6.10: Top: contour plot of the true interval width in function of the design parameters, borehole radius r_w and the borehole length L and only the potentiometric head of the upper aquifer H_u and hydraulic conductivity K_w are regarded uncertain; bottom: mean GP prediction of the interval width (left) and the minimal interval width by 95% CI (right)

of the robust design point, located at the lower-bound of the diameter and the upper-limit of the length of the borehole. In addition the blue dots indicate the points where the original function was sampled. The physical interpretation of the location of the robust point is that a borehole with a smaller diameter limits the possible flow through the borehole. However, for the length of the borehole this observation is not obvious. The results in Figure 6.10 are obtained with a total of 35 evaluations of the borehole function including the four initial evaluations.

Borehole function with six uncertain parameters

In this case all six uncertain- and the two design-parameters of the previously discussed borehole function Eq. (6.32) are considered within the ranges as

parameter		$\underline{x}$	$\overline{x}$	$\hat{x}$	unit
radius of influence	r^I	100	50 000	2550	m
transmissivity of upper aquifer	T_u^I	63 070	50 000	56 535	m^2/year
potentiometric head of the upper aquifer	H_u^I	990	1110	1050	m
transmissivity of lower aquifer	T_l^I	63.1	116	89.55	m^2/year
the potentiometric head of the lower aquifer	H_l^I	700	820	760	m
hydraulic conductivity of the borehole	K_w^I	9855	12 045	10 950	m^2/year

Table 6.2: Parameters of the borehole function

defined in Table 6.2. The results of the analysis are shown in Figure 6.11, which shows the true interval width on the top and the GP prediction on the bottom left and the interval width based on the 95% CI on the right. The number of evaluations to obtain these results has only increased slightly to 64, which includes 8 initial evaluations, while the complexity of the problem is increased by four additional uncertain parameters. The location of the robust design point remained at the lower-bound of the diameter and the upper-bound of the borehole length. The physical reason for this difference is not directly clear from the formulation of the borehole function. However, the additional parameters seem to have little effect to the overall behaviour of the function while the width of the interval has increased slightly, which can be seen by comparing Figure 6.10 and Figure 6.11. To better understand the effect of the additional parameters the interval sensitivities are investigated. The reader is referred to [161] for a thorough discussion about interval sensitivities. However, note that the fundamental difference between the classical sensitivity studies and interval sensitivities is that the latter is valid over the full range of the interval, while the former focuses on local sensitivities, which are not valid over the full range of the interval. The interval sensitivities for the borehole function with six uncertain parameters are provided in Figure 6.12, which shows that the radius of influence r, transmissivity of the upper aquifer T_u, and lower aquifer T_l have an negligible effect on the output interval. Moreover, this figure shows that all parameters behave the least sensitive around the robust design point. The latter means that with a relative change of input interval width only a minimal change in output interval width happens.

Although the obtained results are convincing and could be compared with the true solution, this is not always possible especially with the use of complex numerical models. However, one can validate the GP-model based on the points that were evaluated in the Design-of-Experiments, which provides an indication about the *correctness* to capture the underlying physical behaviour. This validation is accomplished by a number of tests shown in Figure 6.13 which are based on the Leave-One-Out prediction of the points within the DOE. Note that this is a conservative choice as the prediction is now made with a GP

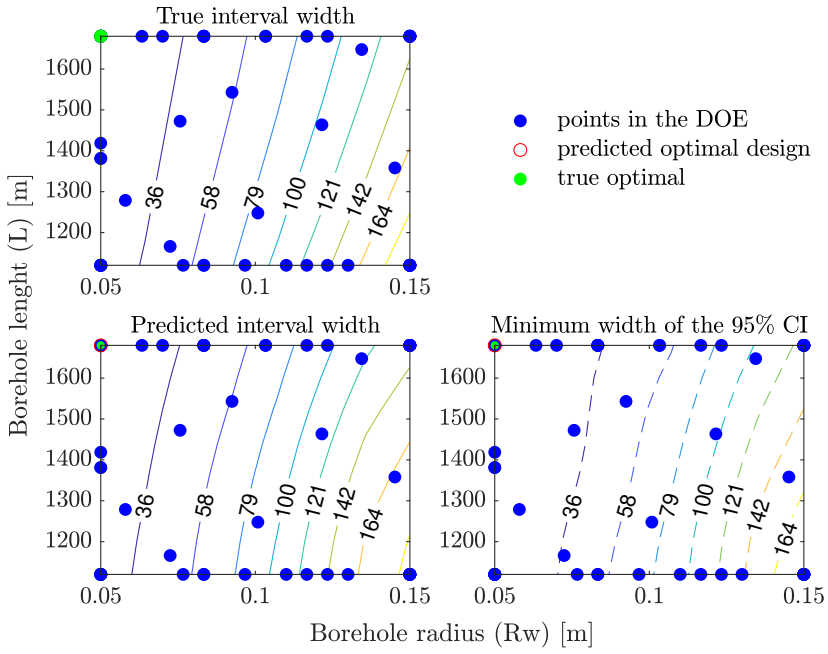


Figure 6.11: Top: contour plot of the true interval width in function of the design parameters: borehole radius r_w, and the borehole length L; with uncertain parameters: the radius of influence r, transmissivity of upper aquifer T_u, the potentiometric head of the upper aquifer H_u, transmissivity of lower aquifer T_l, the potentiometric head of the lower aquifer H_l, the hydraulic conductivity of the borehole K_w; bottom: mean GP prediction of the interval width (left) and the minimal interval width by 95% CI (right)

containing $n-1$ training points, which is especially conservative with a low number of training points. In Figure 6.13a the true function response and the Leave-One-Out (LOO) response are shown including the R^2 value. Figure 6.13 illustrates that most of the points in the DOE are located at the lower-bound of the function output and an increasing error towards the upper-bound of the output. The latter is a direct effect of the selection of points that are added to the DOE, which results in a GP model that is especially good in a specific region. The second Figure 6.13b the true function value and the standardized LOO residual are shown with the two red lines indicating the 95% CI. A similar conclusion can be made where the model in correct at lower output values but misses the true function at higher output predictions. Finally in Figure 6.13c

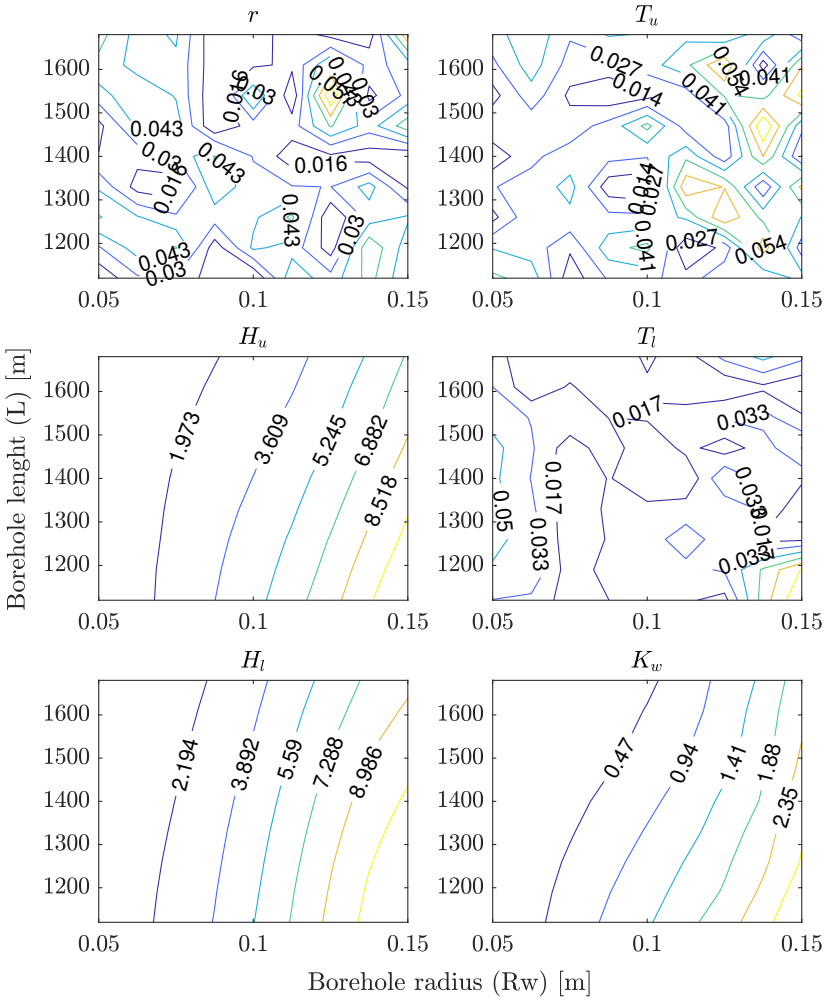


Figure 6.12: Relative interval sensitivity of the uncertain parameters r, T_u, H_u, T_l, H_l and K_w in function of the borehole diameter r_w and length L.

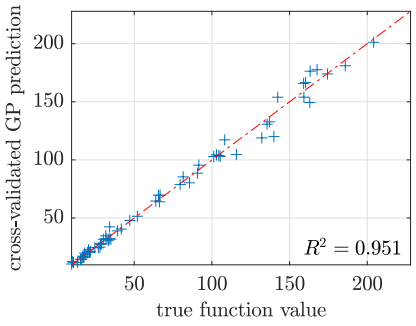
the true model quantiles versus the predicted quantiles are shown. We can conclude that the GP-model performs well at low flow rates, with an error that increases at higher flow rates.

6.7 Discussion

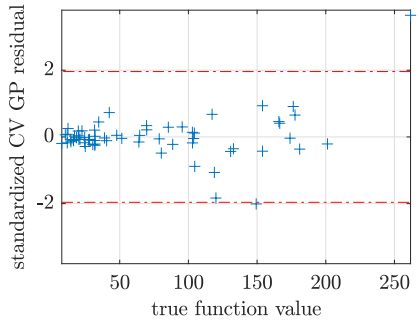
In general the results of the presented method are convincing and show that this method is capable of identifying the robust design point with only a limited amount of evaluations of the underlying expensive function, which is demonstrated in a number of case studies. Nevertheless, a few things are noted by the authors that should be addressed for further research and implementation. As mentioned before, the obtained results are based on the GP model as implemented in UQlab [140] for all case studies. It is noted by the authors that using different implementations of the GP can lead to an increase in the number of iterations before convergence is reached. This is attributed to the use of a noise parameter in the GP, which is set at a minimum of $1e-4$ for the Matlab built-in implementation [142]. Hence, the error term ϵ in Equation (6.27) should increase to reflect this.

The number of samples in the initial DOE can affect the convergence and in this chapter, as a rule of thumb, the number of initial evaluations is kept at the total amount of uncertain- and design-parameters. Quantifying the effect of the initial population size on the rate on convergence is challenging as this depends on the underlying problem, i.e., that what is resembled by the GP model. This rule of thumb is regarded as the minimal amount of initial evaluations needed by the GP to make a first estimation. Nevertheless, the number of iterations is difficult to determine a priori, as this depends on the complexity of the response surface, the added value of the point added at each iteration, and the calibration error of the GP model.

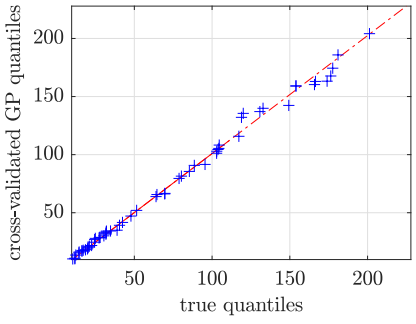
Finally, as the improvement function is evaluated on a fixed number of equally spaced grid points, a limited precision is reached. Although using a large number of grid points the precision increased, the computational cost to evaluate all these points increases exponential in d-dimensions $\mathcal{O}(n^{-d})$ for a full grid. Hence, in high dimensional cases this becomes a bottleneck without sacrificing the resolution of the grid.



(a)



(b)



(c)

Figure 6.13: Validation tests for the borehole function with two uncertain parameters: (a) shows the cross-validated prediction vs. the true function value, (b) shows the standard normalized residuals of the cross-validated GP model within the 95% bounds in red, and (c) shows the cross validated quantiles vs. the true quantiles

6.8 Conclusion

This chapter introduces a novel method to design robust structures in an early stage of development under lack-of-knowledge uncertainty. The presented method uses an adaptively refined GP-model to perform the global optimisation of the robustness and locate the most promising designs, which are the least sensitive to the modelled sources of uncertainty. Based on a set of analytical test functions the effectiveness and efficiency of the proposed method is demonstrated and compared with typical well-known optimisation algorithms. It is shown that the proposed method efficiently solves the double-loop problem, which is typically associated with robustness-based optimization methods. In addition, three additional case studies: a plate in bending, and two times the borehole function are included to demonstrate the applicability to both industrial problems and problems in moderately high dimensions. For all of these examples the results are obtained with a reasonable number of evaluations of the underlying function or numerical model. Future research is aimed at enlarging the application domain of the proposed method, specifically for time-dependent problems.

Chapter 7

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

This chapter was previously published as:
C. van Mierlo et al. "Robust design optimization of expensive stochastic simulators under lack-of-knowledge". In: *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg* 0.0 (2023). Publisher: American Society of Mechanical Engineers Digital Collection, p. 0

7.1 Abstract

Robust design optimisation of stochastic black-box functions is a challenging task in engineering practice. Crashworthiness optimisation qualifies as such problem especially with regards to the high computational costs. Moreover, in early design phases, there may be significant uncertainty about the numerical model parameters. Therefore, this chapter 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.

7.2 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 [218, 219]. 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., [18, 20, 154], while the second is aimed at optimizing of both the objective function and the variance associated with this optimum, see, e.g., [106, 157, 244]. 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 approaches [213]. 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 [68], fuzzy sets [100], information gap methods [22], and imprecise probabilities [19, 77].

The authors of this work recently introduced the Robustness Under Lack-Of-Knowledge method (RULOK) [154]. 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, such 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 represent 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) [44, 222] 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., [59] 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 FE 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 deterministic simulations might exhibit *numerical inadequacies*, i.e., dynamic- and numerical instabilities that can cause a small (infinitesimal) change in the input to produce a major change in the output [4, 124, 148]. In addition to these *numerical inadequacies*, in the specific case of explicit dynamic analysis there is a small but progressively increasing *numerical error* accumulation [14, 16, 59, 223] over the total duration of the simulation. The accumulated error term of both the *numerical inadequacies* and *numerical error* makes the deterministic simulation act 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 [129, 194, 202] 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 adaptively 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 [118]. In this chapter, 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 chapter 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., [125, 181] 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 [154] 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 7.3 describes the measure of robustness under lack-of-knowledge. In Section 7.4 the details about noisy GP's are provided, while Section 7.5 describes the new stopping criterion and provides an overview of the RULOK method. In Section 7.6, the method is tested on three noise contaminated analytical functions and in Section 7.7 an example about crashworthiness optimisation of a crashbox is given. Finally, in section 7.8 a discussion about the results is presented before conclusions are drawn in Section 7.9.

7.3 Robustness under lack-of-knowledge uncertainty

Omitted to avoid redundancy with Section 2.5.2 in Chapter 2.

7.3.1 Propagation of interval valued uncertainty

Omitted to avoid redundancy with Section 2.5.2 in Chapter 2.

7.3.2 Robustness for interval analysis

Robustness under lack-of-knowledge uncertainty is defined in [154] 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:

$$R(\mathbf{z}) = \frac{\Delta x}{\Delta y(\mathbf{z})} = \frac{\overline{x} - \underline{x}}{\overline{y}(\mathbf{z}) - \underline{y}(\mathbf{z})}. \tag{7.1}$$

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:

$$\mathbf{z}^* = \underset{\mathbf{z}\in\mathcal{Z}}{\operatorname{argmin}}[\overline{y} - \underline{y}] = \underset{\mathbf{z}\in\mathcal{Z}}{\operatorname{argmin}}[\max_{\mathbf{x}\in\mathbf{x}^I} m(\boldsymbol{\theta}) - \min_{\mathbf{x}\in\mathbf{x}^I} m(\boldsymbol{\theta})], \tag{7.2}$$

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 (7.2) 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 7.1 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 in the noise variance of the GP.

7.4 Gaussian process model for noisy responses

This section provides a short theoretical summary of GP models or Kriging [129][132], an introduction with examples is also available in [190]. 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 [198]:

$$\mathcal{G} = \beta^T \mathbf{f}(\alpha) + \sigma^2 \digamma(\boldsymbol{\theta}, \Omega), \tag{7.3}$$

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 $\digamma(\boldsymbol{\theta}, \Omega)$ 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 $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$ is defined by the covariance function $K(\boldsymbol{\theta}, \boldsymbol{\theta}', l_c)$, 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 [3] for details about different covariance functions in Gaussian processes. In this chapter the Matérn $\frac{5}{2}$ kernel is used. Note that there is no general framework

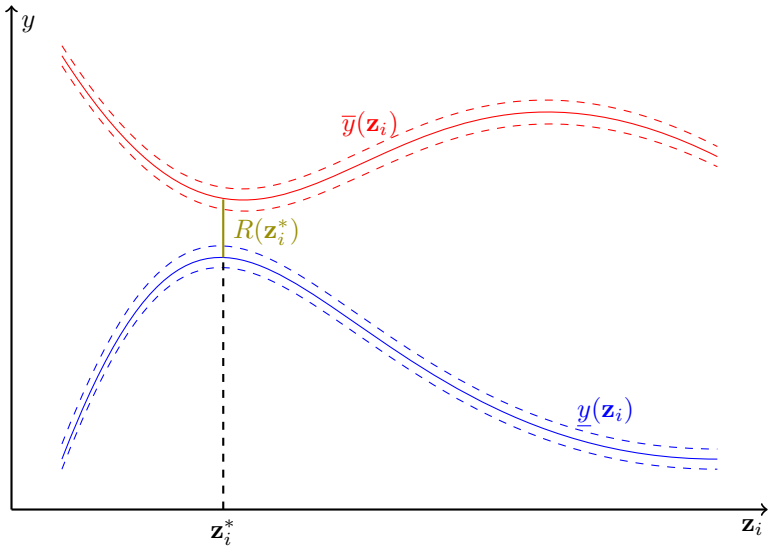


Figure 7.1: Illustration of the optimal robust design points $R(\mathbf{z}^*)$ (orange) for the with noise contaminated upper- and lower bounds $\overline{y}$ and $\underline{y}$ for a specific design parameter $\mathbf{z}_i$, adapted from [154]

at this moment to a priori select the appropriate covariance function. The choice of this particular kernel is based on the fact that this function is 3 times differentiable opposed to an infinite time as seen in the squared exponential covariance function [76].

7.4.1 Predictions made by a noisy Gaussian process

Omitted to avoid redundancy with Section 2.7.3 in Chapter 2.

7.4.2 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_{gp}^2 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:

$$\overline{y}_{gp}(\mathbf{z}) = \overline{\mu}_{gp}(\mathbf{z}) = \max_{\mathbf{x}\in\mathbf{x}^I} \mu_{gp}(\boldsymbol{\theta}), \tag{7.4}$$

$$\underline{y}_{gp}(\mathbf{z}) = \underline{\mu}_{gp}(\mathbf{z}) = \min_{\mathbf{x}\in\mathbf{x}^I} \mu_{gp}(\boldsymbol{\theta}). \tag{7.5}$$

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

$$\overline{\delta}_{\mu+\sigma_\mu}(\mathbf{z}) = \max_{\mathbf{x}\in\mathbf{x}^I} (\mu_{gp}(\boldsymbol{\theta}) + c\sigma_\mu(\boldsymbol{\theta})), \tag{7.6}$$

$$\underline{\delta}_{\mu+\sigma_\mu}(\mathbf{z}) = \min_{\mathbf{x}\in\mathbf{x}^I} (\mu_{gp}(\boldsymbol{\theta}) + c\sigma_\mu(\boldsymbol{\theta})), \tag{7.7}$$

$$\overline{\delta}_{\mu-\sigma_\mu}(\mathbf{z}) = \max_{\mathbf{x}\in\mathbf{x}^I} (\mu_{gp}(\boldsymbol{\theta}) - c\sigma_\mu(\boldsymbol{\theta})), \tag{7.8}$$

$$\underline{\delta}_{\mu-\sigma_\mu}(\mathbf{z}) = \min_{\mathbf{x}\in\mathbf{x}^I} (\mu_{gp}(\boldsymbol{\theta}) - c\sigma_\mu(\boldsymbol{\theta})), \tag{7.9}$$

with $c\sigma$ confidence bounds. The bounds of the response are estimated for each design point $\mathbf{z}$, based on Equations (7.4-7.9). Note that although the GP is cheap to evaluate finding the minimum and maximum response as in Equations (7.4) until (7.9) 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 [118]. 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.

7.5 Adaptive refinement of the noisy Gaussian process model

In this section the learning function introduced in [154] 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 [52] to work directly on the minimum interval width:

$$MI_z(\mathbf{z}) = \frac{\min_{\mathbf{z}\in\mathbf{z}^I}\left(\overline{y}_{gp}(\mathbf{z}) - \underline{y}_{gp}(\mathbf{z})\right) - \left(\overline{\delta}_{\mu-\sigma}(\mathbf{z}) - \underline{\delta}_{\mu+\sigma}(\mathbf{z})\right)}{\min_{\mathbf{z}\in\mathbf{z}^I}\left(\overline{y}_{gp}(\mathbf{z}) - \underline{y}_{gp}(\mathbf{z})\right)}, \tag{7.10}$$

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}. Figure 7.2 illustrates the learning function in Equation (7.10) with on the top illustration the GP predicted upper- and lower-bound of the model and on the lower illustration the learning function MI_z. The illustration shows that a design point z^* at $\min 2\Delta\delta$ is possible more robust than the current optimum z^{opt} at $\min \Delta y_{\text{gp}}$. This is also apparent from 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 $\mathbf{z} \in \mathcal{Z}$.

7.5.1 Maximum improvement of the predicted bounds

The learning function in Equation (7.10) 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 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 [52] 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:

$$MI_{\min}(\boldsymbol{\theta}) = \min_{\mathbf{x}\in\mathbf{x}^I}\left[\mu_{gp}(\boldsymbol{\theta}) + c\sigma_{\mu}(\boldsymbol{\theta})\right] - \mu_{gp}(\boldsymbol{\theta}), \tag{7.11}$$

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

$$MI_{\max}(\boldsymbol{\theta}) = \mu_{gp}(\boldsymbol{\theta}) - \max_{\mathbf{x}\in\mathbf{x}^I}\left[\mu_{gp}(\boldsymbol{\theta}) - c\sigma_{\mu}(\boldsymbol{\theta})\right], \tag{7.12}$$

both of which are not normalised as seen in Equation (7.10). This to guarantee a possible improvement even if the global minimum and maximum are

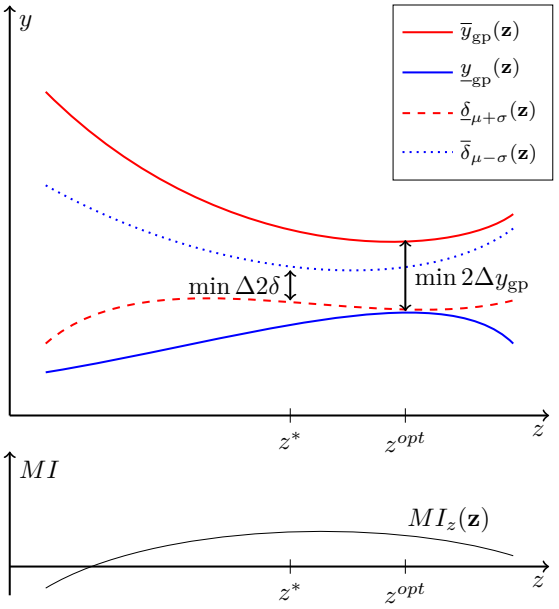


Figure 7.2: Illustration of the predicted mean bound $\Delta y_{\mathrm{gp}}(\mathbf{z}) = \overline{y}_{\mathrm{gp}}(\mathbf{z}) - \underline{y}_{\mathrm{gp}}(\mathbf{z})$ and the minimum bound based on the confidence interval $\Delta\delta(\mathbf{z}) = \underline{\delta}_{\mu+\sigma}(\mathbf{z}) - \overline{\delta}_{\mu-\sigma}(\mathbf{z})$, adapted from [154]

identified. This learning function is illustrated in Figure 7.3, where 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:

$$MI_x = \max(MI_{\min}, MI_{\max}). \tag{7.13}$$

This means that for the illustration in Figure 7.3 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 (7.10) and Equation (7.13) is selected. Hence, the next candidate point $\boldsymbol{\theta}_{\mathrm{candidate}}$ is obtained by:

$$\boldsymbol{\theta}_{\mathrm{candidate}} = \underset{\mathbf{z}\in\mathcal{Z}\ \mathbf{x}\in\mathbf{x}^I}{\operatorname{argmax}} \left[MI_z(\mathbf{z}) + MI_x(\boldsymbol{\theta})\right]. \tag{7.14}$$

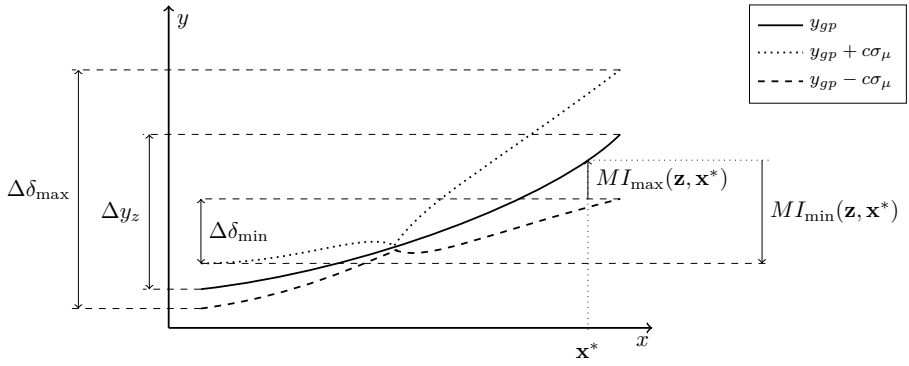


Figure 7.3: Illustration of the learning function for a candidate point x^*, showing the MI of the lower and upper bound; here the improvement of the lower bound is negative [154]

Note that MI_z and MI_x have a different magnitude as MI_x is not normalised. Hence, it is possible at this point to scale MI_z and MI_x. However, to the knowledge of the authors no significant improvements have been achieved in this regard [154].

7.5.2 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 7.4. 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 $\sigma_{\text{gp}}^2 + \sigma^2$. In accordance with the learning function in Equation (7.10) 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

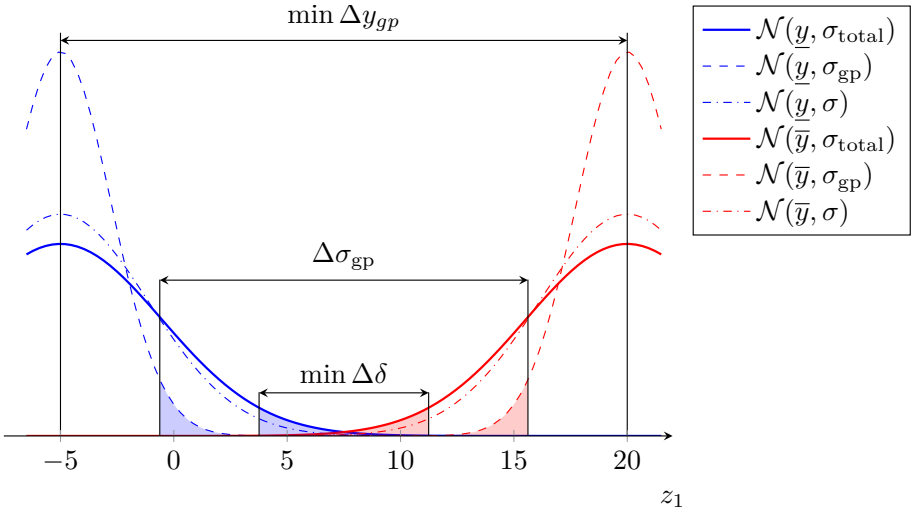


Figure 7.4: Illustration of the stopping criteria for a GP with noise; The illustration shows that the $\Delta\delta$ canFor both the upper and lower bound three normal distributions are drawn $\mathcal{N}(\overline{y}, \sigma_n)$,$\mathcal{N}(\overline{y}, \sigma_{gp})$, $\mathcal{N}(\overline{y}, \sigma_{\text{total}})$ indicated by the dashed, dash-dotted and full lines, respectively

homoscedastic noise the stopping criteria is defined as:

$$MI_s = \frac{2c\sigma_{\text{gp}}}{\min_{z\in\mathbf{z}^I}\left(\overline{y}_{gp}(z) - \underline{y}_{gp}(z)\right)} + \epsilon, \tag{7.15}$$

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 than the maximal improvement given the noise of the GP:

$$MI_z \leq MI_s. \tag{7.16}$$

By the end of the adaptive refinement one can state that according 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 (7.16) would not work in the case of *heteroscedastic* noise. Hence, the remainder of this chapter focuses on homogeneous or *homoscedastic* noise.

7.5.3 Overview of the method

In Figure 7.5, a flowchart of the method is provided. The flowchart describes in detail the steps needed to perform the optimisation as proposed in this chapter. The method starts at the 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 design point based on the learning function in Section 7.5. 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 (7.16) is met two consecutive times. Finally, after finishing the optimisation, it is considered good practise to validate the results of the GP.

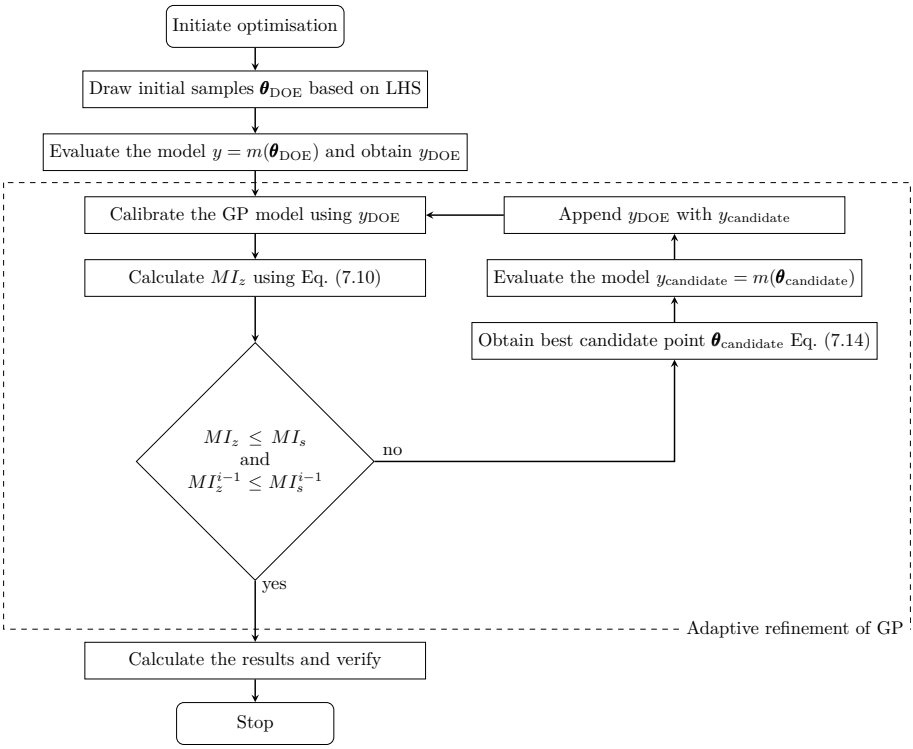


Figure 7.5: Flowchart of the robustness under lack-of-knowledge method for noisy functions

7.6 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 [154]. However, in this work a random noise term is added. The analytical test functions with noise are defined as:

$$f_a(z_1, x_1) = z_1^2 x_1 - x_1^2 + \zeta_{\text{fn}}, \tag{7.17}$$

$$f_b(z_1, x_1) = x_1 z_1 - \sin(z_1) x_1^2 + z_1^2 + \zeta_{\text{fn}}, \tag{7.18}$$

$$f_c(z_1, x_1) = \cos(4\pi z_1) - \sin(z_1 x_1) + x_1 + \zeta_{\text{fn}}, \tag{7.19}$$

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 7.6 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 cases under both known- and unknown-homogeneous noise.

7.6.1 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 terms on both the analytical function σ_{fn}^2 and GP σ_{gp}^2 is demonstrated. The proposed approach is stochastic in nature since it depends on the noise-affected realisations of the system. Therefore, each of the cases shown in this section are repeated ten times, and the mean and envelope of all runs are shown. The first case illustrated in Figure 7.7 shows the effect of increasing the imposed noise variance σ_{fn}^2 for function f_a Equation (7.17). The figure on the left shows an increasing error for an increased imposed noise variance σ_{fn}^2, indicated with the mean relative error in a blue line and the blue area showing the minimum and maximum error 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_{gp}^2 is set at different values, while the imposed noise variance is kept at $\sigma_{\text{fn}}^2 = 10$, again for function f_a Equation (7.17). The results are shown in Figure 7.8, which is identical in setup to the previous case. It is clear that with an increase of GP variance

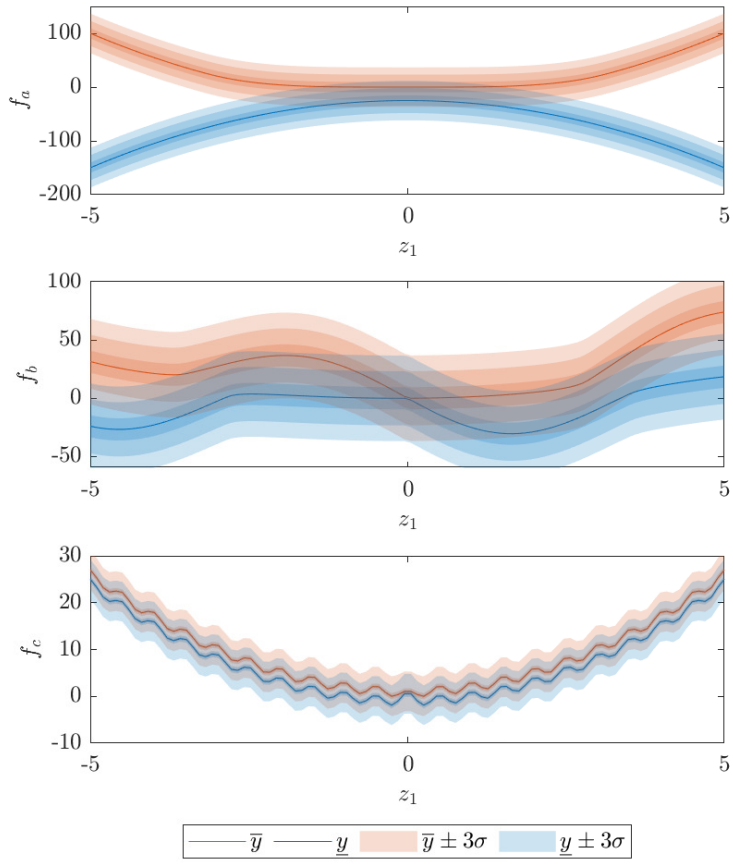


Figure 7.6: Illustration of the effect of noise on function f_a,f_b,f_c; The red and blue lines indicate the upper- and lower-bound of the functions, while the red and blue areas represent the effect of noise on the upper- and lower-bound illustrated by the 3σ CI for three noise variances $\sigma^2_{\text{fn}} = 10, 60, 150$ for f_a,f_b and $\sigma^2_{\text{fn}} = 0.01, 0.05, 2$ for f_a

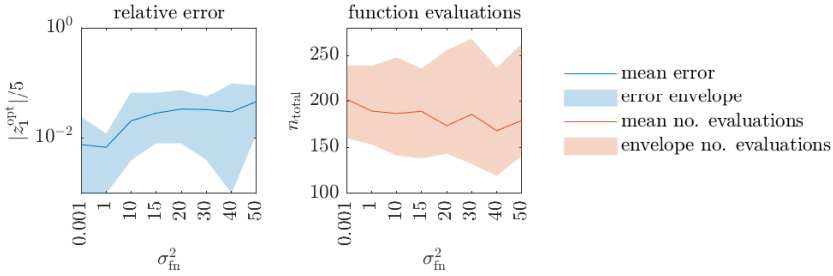


Figure 7.7: The mean and envelope of ten runs for function f_a Equation (7.17) with an increased imposed noise variance σ^2_{fn}

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 high variance, a large number of evaluations is required to reach the desired accuracy.

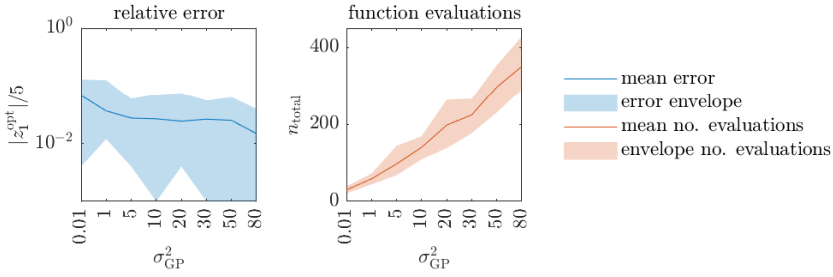


Figure 7.8: The mean and envelope of ten runs for function f_a Equation (7.17) with a increased set GP variance σ^2_{GP}

For the third analytical case, function f_b Equation (7.18) is used following a similar approach. The results of this case are shown in Figure 7.9 where the noise imposed on the function is increased and the variance of the GP model is kept at $\sigma^2_{\text{GP}} = 10$. These results are a bit different than 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 7.7. 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 be seen in Figure 7.6. One possible interpretation of that result is that high noise levels mask the local minima of function f_b.

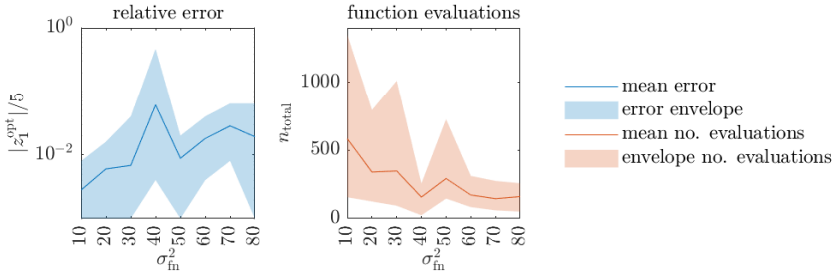


Figure 7.9: The mean and envelope of ten runs for function f_b Equation (7.18) with a increased imposed noise variance σ^2_{fn}

Finally, Figure 7.10 shows the results of function f_c Equation (7.19) 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 7.7 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 possesses too much of a challenge when noisy responses are considered.

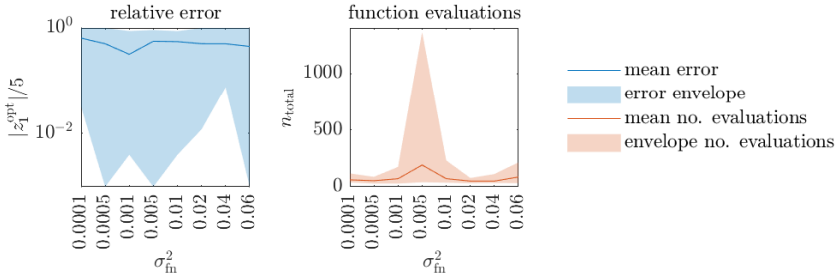


Figure 7.10: The mean and envelope of ten runs for function f_c Equation (7.19) with increasing imposed noise variance σ^2_{fn}

7.6.2 Analytical functions with unknown homogeneous noise

This section focusses 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 sufficiently large number of observations. Hence, in the first case shown in Figure 7.11 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 too fast, as with the limited number of evaluations no correct estimates of the imposed noise variance σ^2_{fn} are made. This behaviour 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.

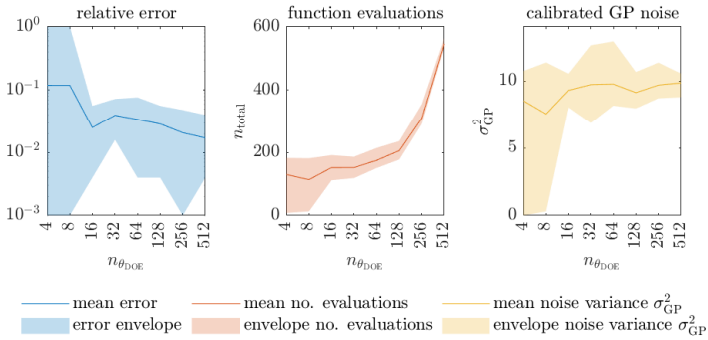


Figure 7.11: Results of function f_a Equation (7.17) with on the left the relative error, in the middle the total number of function evaluations including the initial evaluations, and on the right the calibrated noise of the GP in a full yellow line.

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 7.12, with again the same structure as before. The decrease of ϵ and the 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 relatively low. However, it can be seen that for very low values the number of evaluations starts to increase. 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.

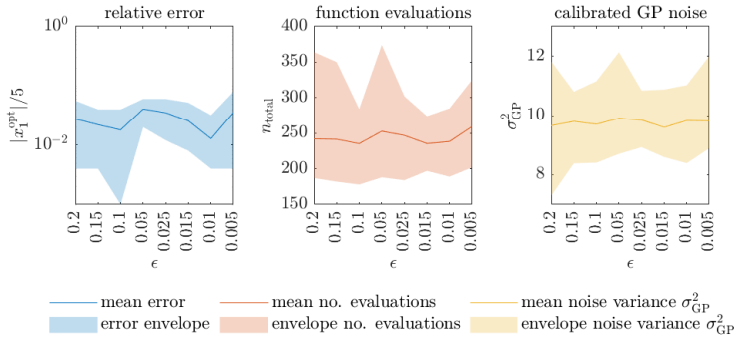


Figure 7.12: Results of Function f_a Equation (7.17) for eight cases with increasing values for Epsilon, with on the left the relative error, in the middle the total number of function evaluations including the initial evaluations, and on the right the calibrated noise of the GP in a full yellow line.

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 7.13 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 7.13 that the method is capable of tracking these high imposed variances without increasing the function evaluations.

7.6.3 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

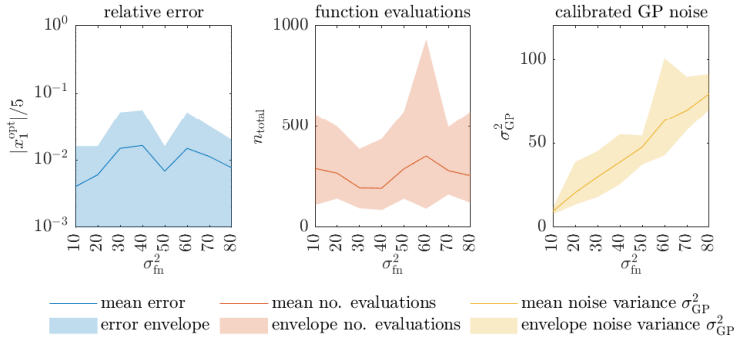


Figure 7.13: Results of function f_b Equation (7.18) for eight cases with increasing imposed noise variance σ^2_{fn}, with on the left the relative error, in the middle the total number of function evaluations including the initial evaluations, and on the right the calibrated noise of the GP in a full yellow line.

noise calibrated as part of the GP maximum likelihood estimation. Both methods are capable of providing satisfactory results for the analytical functions defined in this section. However, general conclusions are not easily drawn on the basis of the results obtained, as the 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 fewer evaluations, were obtained by a priori estimated noise.

7.7 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 7.14, 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 amount of energy during frontal impact, and to prevent further structural

Table 7.1: 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

damage at low-speed impact events. The numerical model to represent the crashbox is taken from the publicly available Toyota Yaris model, downloaded from [175], and consists 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 7.14, 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 7.1. More details can be found in the publicly available numerical model [156]. Note that T_1 refers to the thickness of the green plate in Figure 7.14, which has PID 2000121, and T_2 to the blue plate in the back with PID 2000142.

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 7.15 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 7.1.

7.7.1 Crashbox with uncertain spotweld diameter

In this section, robust optimisation is performed with T_1 as design parameter, and T_{sw} the uncertain interval valued parameter as described in Table 7.1, while $T_2 = 1.8$ mm is fixed. As a reference, the existing RULOK method using

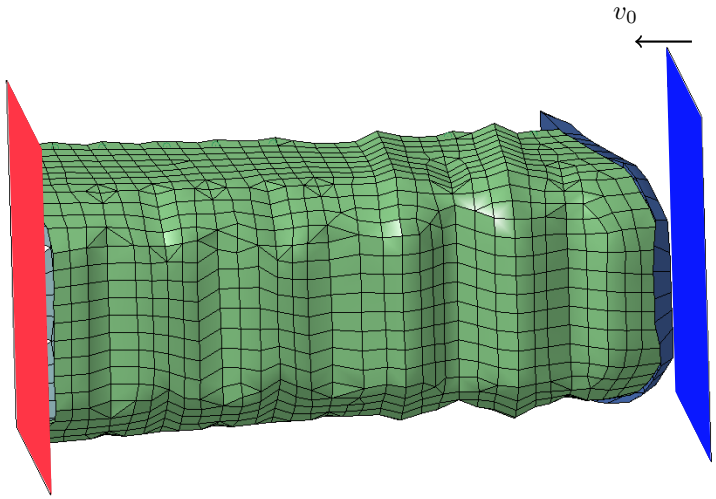


Figure 7.14: Finite Element Model of the crashbox with a rigid plane attached to the nodes in the back (red) and impacting plane right (blue); adapted from the Toyota Yaris model [175]

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 7.16. 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 7.16 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 $\sigma^2_{\text{GP}} = 5\text{kN}$. 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 (C.O.V.) of about 0.045 on average within the domain. Therefore, 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 C.O.V.. In Figure 7.17 the results are shown for different assumed C.O.V.'s and a mean response of the 20 initial evaluations of 110kN. The top figure shows the robust design point for each of the cases with the blue line indicating the mean of the 10 evaluations and the blue area the envelope. On

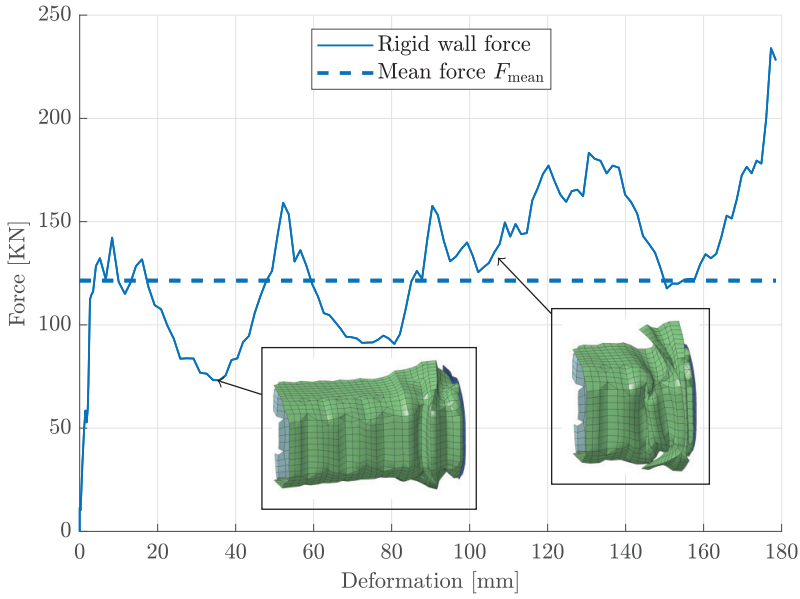


Figure 7.15: Typical force-displacement curve, obtained from the numerical simulation, with the mean-force F_{mean} as a dashed line; in addition, two deformed states of the crashbox are provided

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 results obtained 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 C.O.V.'s the number of evaluations start increasing, and sometimes wrong optima are identified, while the mean predicted optimum is always in the correct range.

7.7.2 Crashbox with uncertain spotwelds and plate thickness

To demonstrate the proposed method in 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 listed in Table 7.1. When an additional uncertain parameter is introduced, the location of the robust design point has also changed. Therefore, a reference is created based on 1000 Latin-Hyper-Cube (LHS) samples, before initiating the optimisation

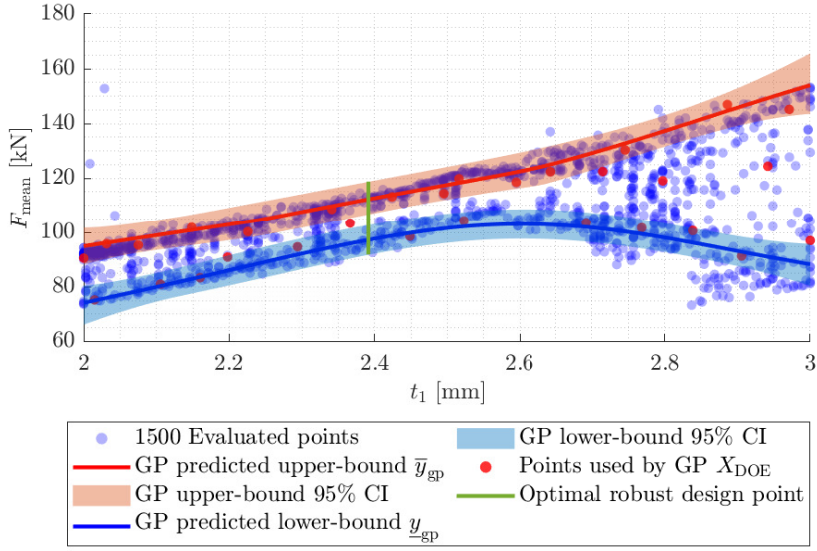


Figure 7.16: Evaluations by the RULOK method without noise kernel, in blue vs. the evaluations, in red, and prediction of the method with a noise kernel

by RULOK. The results obtained are shown in Figure 7.18 with the LHS samples 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 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, appears correct based on the LHS samples.

7.8 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 [154] where a Genetic

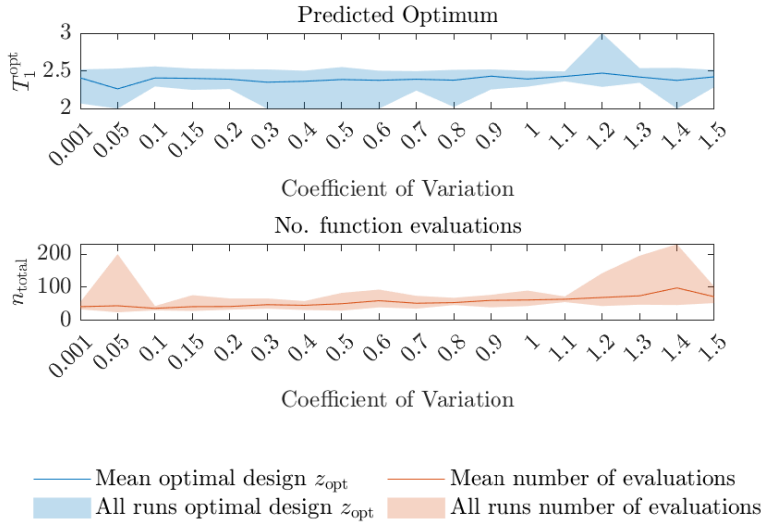


Figure 7.17: Obtained results for the crashbox with the GP noise depending on the COV of the initial 20 evaluations

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 7.16 and Figure 7.18 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 work of the authors [13, 148], the optimisation of similar crash cases took a minimum of about 160 evaluations, and sometimes more than 3000 function evaluations for a single design. Reducing this to only about 50 evaluations for a range of designs is a huge improvement in terms of efficiency. To determine the number of initial evaluations ten times the dimension of the input space $10n_\theta$ can be taken as a rule of thumb. However, the method is capable of achieving result

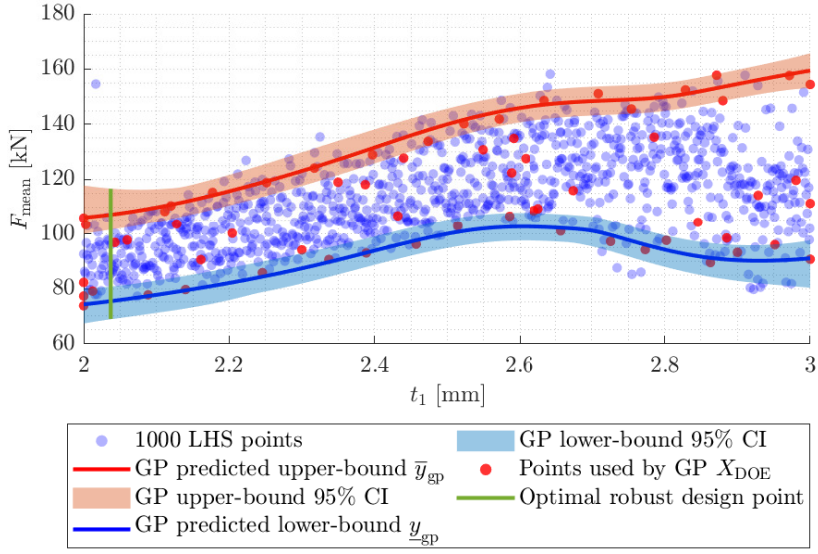


Figure 7.18: 1000 LHS samples of the numerical model vs. the GP prediction of the upper- and lower bound based on only 62 function evaluations

with far fewer initial evaluations as shown in Section 7.6

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 implemented in UQlab [140] for all case studies. However, using the stopping criterion proposed in this chapter 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 [142], which can be taken into account as Gaussian noise. Finally, it should be noted that, as with the original RULOK method, a structured grid is used. Therefore, the computational cost to evaluate all points on this grid increases exponentially 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.

7.9 Conclusion

In this chapter 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 of identifying a robust design point, with fewer model evaluations than what would be expected from a general optimisation algorithm.

Chapter 8

Valorisation

This chapter is a requirement within the Faculty of Engineering Technology, which is aimed at the valorisation potential of the methods and developments presented in the thesis. In this chapter, two independent valorisation trajectories are explored and described in detail in the sections below. The valorisation is mainly found in licensing implementations of state-of-the-art techniques such that industry's R&D departments can make more efficient use of these techniques, and second in, expert roles providing assistance to industry to use advanced methods and consultation in impact modelling and testing.

8.1 Uncertainty Quantification software licensing

FIRST, abbreviated from Fuzzy, Interval Ready Software Technology, is a software tool used and developed by the Reliable and Robust Design research group at the KU Leuven. This tool is specifically aimed to be intuitive in use and is equipped with a recently developed graphical user interface. **FIRST** allows direct coupling with multiple commercially available Finite Element packages, i.e., Siemens NX, Abaqus. The main goal of this software package is to lower the threshold experienced by industry to use advanced uncertainty quantification tools, in particular field techniques, i.e., random fields, and interval fields. Therefore, **FIRST** can be used by the research & development teams at these companies to: (1) increase and understand the *validity of numerical models* by including the relevant uncertainty from an early conceptual stage, (2) assist in the design of appropriate *experimental campaigns*, and (3) assist in analysing *manufacturing variability* and the impact hereof on product

reliability. The analyst is assisted in this regard using the intuitive graphical interface of **FIRST** to enable advanced uncertainty quantification techniques, i.e., interval fields, random fields. Based on the developments in this thesis, the interval field capabilities were extended with the techniques described in Chapter 3 and 4. Furthermore, based on the developments in this work the main contribution would be an additional material calibration module. This module can be used to calibrate material models that can be used in various FE simulations. In addition, the module would allow the user to quantify the effect of scattered material data on the product performance, following the developments in Chapter 3.

The material calibration module would allow the analyst to upload a set of experimental data, as shown in Figure 3.7, and capture the non-determinism using an interval field. Realisations of this fitted interval field are then directly usable in numerical simulation and other UQ techniques, e.g., sensitivity analysis, robustness based design optimisation. This methodology was first tested and implemented by the author in [149], and is also used for the optimisation results shown later in this section. The material calibration module would be licensed as a separate add-on for **FIRST** following the Software As A Service (SAAS) model. The license costs of this module would be in the range of 5K€ for an annual licence. This license fee is in line with currently available material calibration tools (Mcalibration), which do not offer non-deterministic modelling strategies. However, first a test phase would be conducted to gather customer feedback, which will be mainly from the industrial partners and license holders who already have an additional service package for **FIRST**. The reason for this testing phase is of course to ensure that the package is working as intended and that the instructions are clear to new users. The typical costumers for this add-on would be:

- **Machine building companies** aiming to increase the reliability of the machines, and reduce the variance in produced products. Using the material calibration module these companies can quantify effects of scattered material properties on the production process of the machines.
- **Product development and design companies** where FE methods are used to determine the performance of the final product. Often, significant uncertainty exists about the exact material behaviour, especially at an early development stage. Using the calibration module the material models can be calibrated and used for UQ

Based on the growing number of workshops, learning platforms and conferences industry's interest is increasing to use non-deterministic modelling, uncertainty quantification, and model Verification and Validation (V&V) techniques, which is

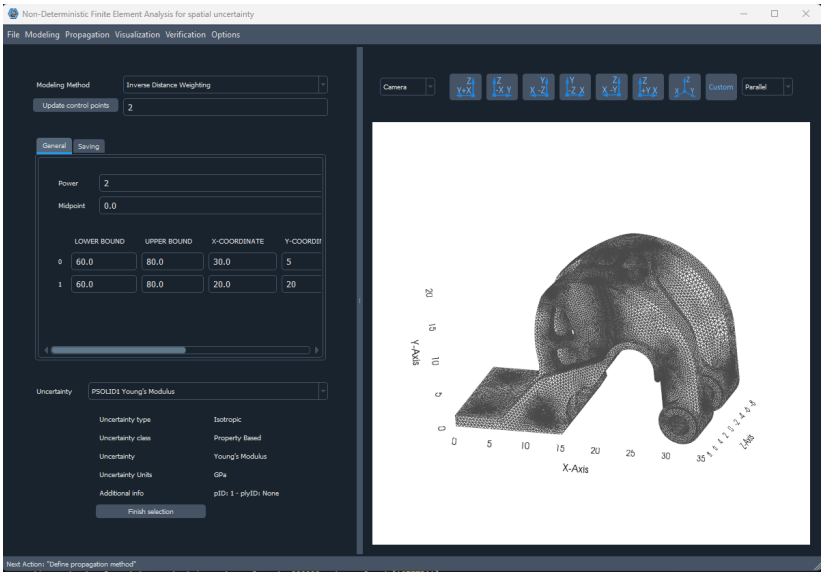


Figure 8.1: Screenshot of the graphical user interface of FIRST

further supported by the increasing active scientific community in this field. This indicates the economic potential. Moreover, renown standardisation institutes, i.e., ASME, NASA, ISO, are increasingly describing the use of advanced methods to perform these activities. Furthermore, the future of crashworthiness assessment in automotive industry is moving towards a simulation driven approach, with other industries following these developments. However, at the moment only a few tools are commercially available and suited to perform advanced UQ analysis, e.g., Noesis Optimus, COSSAN-X, UQ lab, UQpy. At this moment, only COSSAN-x offers random field capabilities with an intuitive graphical user interface. Moreover, none of these packages have a non-deterministic material calibration tool. Therefore, **FIRST** including the material calibration tool provides a lot of additional value for the customers that use commercially available simulation packages. Figure 8.1 shows a snapshot of the graphical user interface of **FIRST**. This figure shows how large three-dimensional models can be displayed. On the left, the field parameters are set by the user who can directly see these changes appear on the model.

8.2 Consulting and providing expert knowledge to industry

In this research the focus was put on impact and crash analysis. Hence, over the years a certain amount of expertise regarding modelling and testing of these events is obtained. In addition, from the start of this research development and in-house construction of a Digital Image Correlation (DIC) [119] capable impact tower started. Furthermore, an accompanying set of High Speed cameras where acquired that could be used during impact experiments. These unique capabilities make our research group well positioned to provide consulting to industry, which is also a valid valorisation strategy as the developments in this research project, i.e., capturing material uncertainty, account for uncertain adjacent components, and a robust design optimisation technique are challenging to combine in a single product. Therefore, the gained expertise of this research project can be provided to industry in the form of consulting services and the creation of unique devices to assess the crashworthiness of components. An example of such a device is the in house built impact tower, which is compatible with multiple high-speed cameras, which have a clear view on the specimen.

The need for these consulting services is clear as multiple companies reached out to our research group with specific questions regarding the design and simulation of impact critical components. The main challenge faced in industry is usually the validation of impact or crash simulations. An example of these such question would be: how should the design be adapted to accommodate a new material, while keeping identical impact performance? In these cases a more detailed investigation is required and a software solution alone is clearly insufficient. In another project, detailed investigations in crack propagation was needed to check that the material and production process could be used to produce safety critical components. Based on the results obtained with HS-DIC new computational models could be created, which predicted the failure mechanisms in higher detail. Furthermore, multiple companies contacted the research group about concerns with changing safety regulations, which increases the use of simulation based design as large scale testing is simply to expensive.

8.2.1 In-house built impact tower

Based on the high need for validation of impact tests in the projects conducted with industry, an impact setup capable to use with High-Speed cameras was developed. Although, plenty of impact and drop towers are commercially available, and some are already operational in the Material Science labs (MTM) of KU Leuven, none of these were designed to be used with a camera setup.

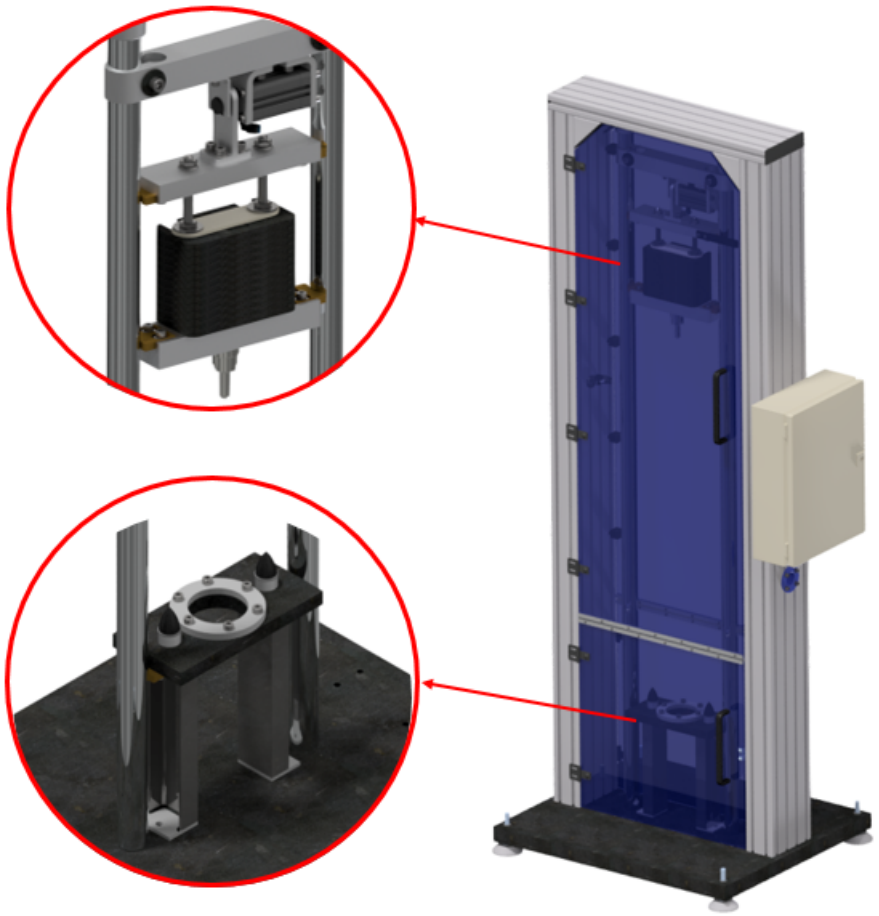


Figure 8.2: Special build droptower to be used with High-Speed Digital image correlation

Hence, a flexible in-house setup was designed that could be used for multiple experiments without blocking the line-of-sight of the cameras. Figure 8.2 shows a render of the design as envisioned at the start of the research. On the left top the release mechanism and falling weights are shown, and on the left bottom the sample holder is shown. The idea with raising the sample holder is to enable the cameras to be directly pointed at the bottom of the sample. Figure 8.3 shows the built drop tower created mainly out of ITEM-profiles. On this picture a puncture test following the ASTM 3763 [7] standard, which is comparable to the

ISO 6603 standard [114]. The controller enabling safe operations is located in the white box shown on the right side, and the height adjustable falling weights and the raised sample holder are shown in the detailed views. By adjusting the fall height and weight, impact energies within a range of about 2J up to 255J are realized, as shown in Table 8.1. It should be pointed out here that only the potential energy of the weight can be used, i.e., there is no additional source of energy. Therefore, a suitable compromise should be found between impact speed, drop weight and energy. Note that the minimum impact energy is mostly determined by the weight of the dropweight assembly, shown in the top detail of Figure 8.3. The lightest version, weighing only 2.13 kg, was created to test the landing gear of a drone [150].

parameter	minimum	maximum
Height [m]	0.1	1.3
weight [kg]	1.8	20
energy [J]	1.96	255

Table 8.1: ranges of the parameters used to design the impact tower

The main quantities of interest during an impact test are the displacement and force during impact, which are typically used in standards and are also common output quantities of numerical simulations. The setup was designed to be used with various types of sensors and data-acquisition tools. Initially an accelerometer and piezo-electric load-cell were used, in later setups multiple load-cells, and laser vibrometers were used depending on the specific needs. An important aspect is the synchronisation of the cameras time-frame with the data acquisition, to ensure that image and force-displacement measurements are taken at identical times. In some cases the sampling rate of the data acquisition could be so high that interpolation could easily be performed and a common trigger would be sufficient for synchronisation.

The main novelty with respect to existing drop-towers is the possibility to use a set of high-speed cameras during the tests, which is the reason for the raised sample holder. A picture of the setup is shown in 8.4 where the cameras are placed under the sample holder pointing at the sample under an angle of about 30-degrees. The exact setup varies depending on the test, materials, and impact energy. To enable the use of DIC high quality images are required of both cameras having similar Field Of View (FOV), taken at exactly the same time, the area of interest is in focus, and the lighting conditions are even in all images. The latter is especially troublesome in the case of high speed imaging as this usually requires high power light sources, which have a tendency to create reflections. Additionally, large deformations during impact change the angle of the incident light, which can cause reflections and changes in light intensity.

Figure 8.3: Picture of the as built impact tower with the controller in the white box on the right side, adjustable weights and sample holder in the middle between two finely machined rods used to guide the falling weights

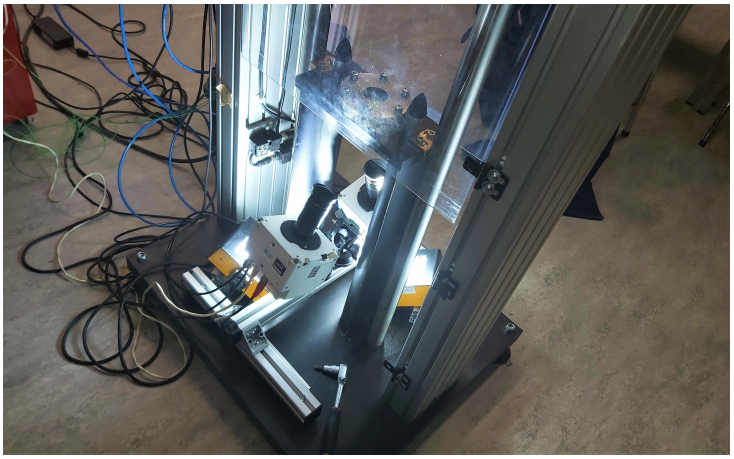


Figure 8.4: Picture of the experimental setup to perform puncture tests ISO 6603

8.2.2 High speed digital image correlation

Digital Image Correlation DIC is an optical measuring technique where patterns of a reference and deformed image are being matched. These images are usually captured using quite common CCD/CMOS cameras, and more advanced techniques can be found in literature, e.g., electron microscopes [121], or x-ray imaging [98]. Depending on the number of cameras being used different terms are found in literature: 2D-DIC when there is only one camera, stereo-DIC when two cameras are used, and multi-camera DIC in the case where more than two cameras are being used [119, 199]. The pattern that is used to correlate subsequent images is usually applied on the surface of the sample, which in some cases is not needed as the surface provides enough features. The introduction of DIC has led to a revolution in experimental mechanics where novel test procedures are proposed to obtain constitutive material parameters [183]. The main benefit is that a measurement is not limited to a single point, or location, as the complete FOV can be used, which in practice usually reduces to a certain area of interest. This is shown in Figure 8.5 by the coloured area. Note that this area is a lot smaller than the complete FOV, which is required as the area should be within the FOV under large deformations.

In the specific case of High-Speed Digital Image Correlation (HS-DIC) the main change is the type of cameras that are being used. Typical standard cameras work on a frame-rate between $[24 - 90]$ fps with high-speed cameras capable of reaching frame-rates of $2 * 10^6$ fps and beyond. Due to advancements in

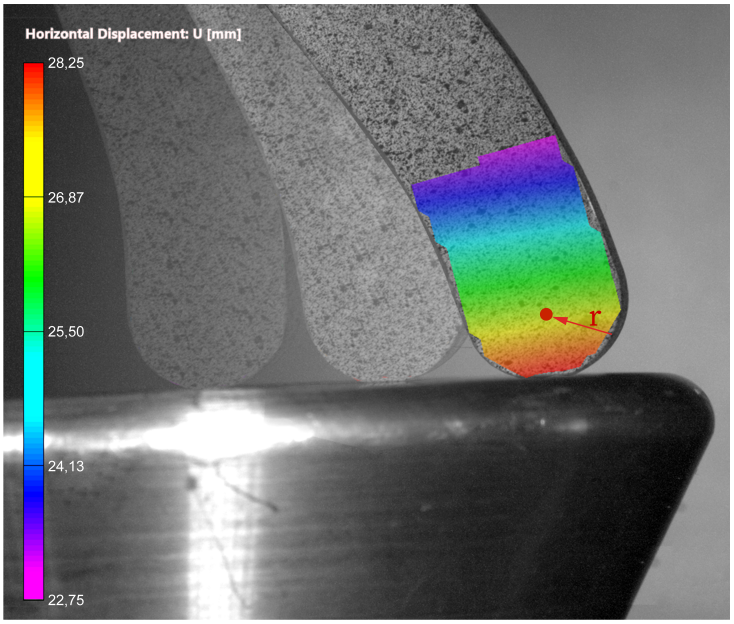


Figure 8.5: Result of HS-DIC on the compliant drone leg at multiple time steps

sensor technologies the frame-rates have increased almost exponentially over the last decades, at lower costs than ever before. Nevertheless, High-speed cameras are not cheap by any means and most experiments with high speed cameras use only one [86, 240] camera. In addition, multiple attempts are made to find alternative setups that use a mirror based setup [36, 239] to increase the capabilities of a single camera. Especially, as the price-point is largely determined by the frame-rate and increases exponentially.

Performing validation on impact structures remains a challenge requiring both specialised hard- and software [12, 193, 228]. The general idea is given in Figure 8.6 where on the bottom the numerical procedure is shown and on the top the experimental procedure. These two are not directly comparable as the influence of capturing the image and the processing settings should be accounted for. Using the simulation results to numerically deform images from the experiment is an option to cancel out these effects, as these images undergo the identical steps as the images used in the experiment. An overview of the main considerations is given below:

- **Boundary conditions** used in the experiment should correspond with these used in simulation. These should be checked before the experimental

campaign as due to manufacturing, assembly and operational procedures the intended, e.g., dimensions, features might not correspond.

- **Camera settings** that are used to capture the behaviour directly impact the final result. The most important features are:
  - **Frame rate** determines the sampling rate of the system. Simulation output results should also be available at these approximate sampling times.
  - **Resolution** determines the amount of pixels that are captured within the area of interest. At a higher resolution smaller spatial features can be detected, while the noise floor increases at the same time.
  - **Optics**, such as, the lenses will cause distortions in the image, which directly impacts the results [133]. Lenses of optically lower quality will have more distortions within the image and therefore increase the noise floor. In addition, on older sets of lenses there might be small artefacts that influence the performance, e.g., scratches.
  - **Light sources** are imported as there should be sufficient light that the sensor can capture. Especially where the frame rate is high and therefore the shutter times are very short there should be sufficient light reaching the imaging sensor.
  - **Pattern** that is applied to the surface, should remain applied to the surface. When the pattern is moving independently of the component the measurements are wrong. Generally this happens when there is insufficient bonding between the paint and the substrate. In the specific case of impact experiments this might also happen in areas of high strain rates, near a crack or breakage of the substrate, or after large deformations. The best solution for these problems is to let the paint only partially dry, to prevent the paint from fully curing and becoming very brittle. There are also special types of paint available that might mitigate these issues.
- **DIC settings** refer to the algorithmic settings used to correlate the reference to the deformed image. These settings are typically subset size, step size and the correlation algorithm, among other more advanced settings. A typical setting that should be considered for HS-DIC is to update the reference image or not, depending on the strains that are measured in subsequent images.

The illustration in Figure 8.5 is a good example of the application and challenges in using HS-DIC where the deformed state of the compliant drone leg is shown at multiple time instances. It should be pointed out that, the position of the

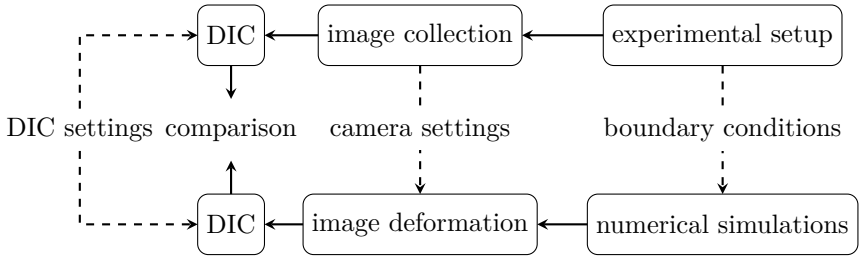


Figure 8.6: Flowchart of the validation strategy using digital image correlation

impacting metal surface is only at the correct location for the last deformed state. This metal plate is actually moving up in the current image layout, and the drone leg position is adapted to create the figure. The main challenges in this case are the large displacements of the drone leg, where at the start of the experiment the tip of the drone leg was located in the left corner and at the end would move out of the FOV.

8.2.3 Case study: validation of a compliant drone leg

In the remainder of this section, a validation of a numerical impact simulation is discussed, where uncertain hyper-elastic material behaviour is modelled based on the framework discussed in Chapter 3. This section only focuses on the validation strategy, so only limited information about the modelling strategy is provided. The model under consideration is this of a compliant drone leg, which is designed to dissipate impact energy and protect sensitive camera equipment mounted on the drone. To validate the numerical model a secondary model was created using the exact boundary conditions of the lab tests. The idea is that if the numerical model is valid in lab conditions, results of the numerical model in other scenarios are also correct, as long as the underlying conditions are not changed, i.e., similar deformation rate. In this case the lab experiments were limited by the minimum falling mass weight of 2.13 kg, which was used to create a kinetic energy at impact of 2.98 J. This impact energy is identical to this of the drone falling from one meter. In the numerical simulation the upper part of the drone leg is fully constrained while at the bottom a rigid surface is moving up with a kinetic energy of 2.98 J. Important factors in this simulation are the material parameters and the friction model used between the drone leg and the rigid surface. Time discretisation was performed by means of an explicit scheme until the total simulation time of 20ms was reached. Here, eight material curves were generated by vertex samples from the interval field that was used to describe the material uncertainty, see [150] for details. These

realisations are the result of using four control points with one kept at zero and creating a realisation for all other combinations. These realisations were used to generate eight different results as shown by the red lines in Figure 8.7.

High-speed Digital Image Correlation	
technique used	Stereovision DIC
Noise	
Camera 0	0,5 %
Camera 1	0,6 %
Correlation	
Pre-filtering	Gaussian - Kernel 5
Subset	27
Step	11
Correlation criterion	ZNSSD
Process history	Temporal
Shape function	Quadratic
Interpolation function	Bicubic Splines
Measurement points	347
Total number of images	113 (3 KHz)
Displacement	
In-plane resolution	2 μm
Out-of-plane resolution	6 μm

Table 8.2: HS-DIC parameters that were used for this experiment and the obtained resolution

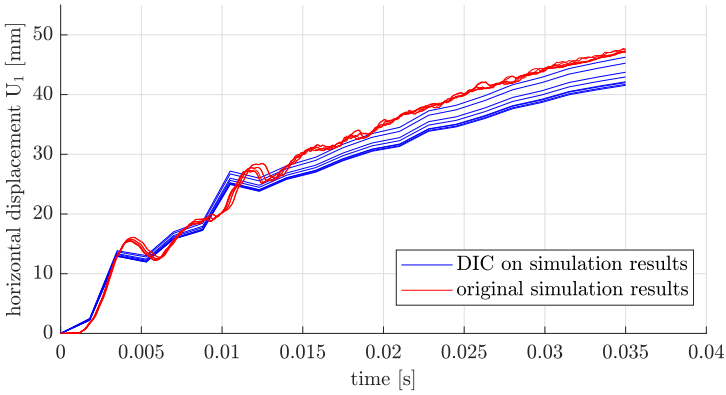


Figure 8.7: Horizontal displacement u_1 for the drone leg as predicted by the explicit numerical model in red and the results of using DIC on the numerically deformed images in blue

The blue lines in Figure 8.7 are the results of performing DIC on the numerically deformed images using the HS-DIC parameters as listed in Table 8.2. For this the Region of Interest of the images as shown by the colored region in Figure 8.5 is selected and numerically deformed using the nodal displacements of the simulation. Due to the low sampling rate it is not entirely clear but the DIC algorithm acts as a low pass filter. In addition, the figure shows that the spread on the blue curves increases over time where the red results of the simulation keep a similar spread in time, which can probably be attributed to an error caused by temporal updating the reference image. However, no further investigations have been conducted to prove this.

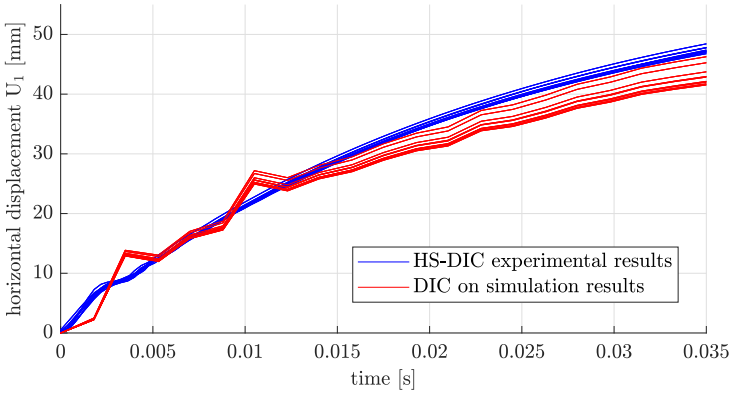


Figure 8.8: Comparison of horizontal displacement u_1 for the drone leg between simulation results in red lines and HS-DIC results in blue lines

Quantitative results from the DIC measurements were obtained by extracting the horizontal u_1 displacements at a specific point, i.e., the tip of the drone leg. However, before the quantities of interest can be obtained the DIC algorithm has to be run, which used the parameters listed in Table 8.2. It should be pointed out here that the displacement resolution in Table 8.2 is quite high, and represents a noise floor of the first images. Hereafter, the resolution will degrade due to other errors, e.g., updated references, changing conditions. The results of the DIC measurements are shown in Figure 8.8 where they are directly compared with the results from the simulation. It is clear from Figure 8.8 that there is quite a difference between the results from simulation and test. In a follow-up investigation model updating techniques could be used to try and mitigate this difference. This is not a trivial task and a detailed analysis should be conducted to further explore the possibilities of model updating. This example highlights some important factors to validate impact models and the challenges that are associated with these factors.

This example case illustrates the increased capabilities of the impact tower that is developed in combination with state-of-the-art full field material identification techniques. The combination with the developed field techniques strengthens these capabilities further, and leaves room for further explorations of more comprehensive material identification techniques and validation strategies. Furthermore, due to wide range of operations and flexible construction the droptower can easily be adjusted to facilitate the wide variety of impact experiments.

Chapter 9

Conclusions

9.1 General conclusions

In this work, the challenges in Systems Engineering are explored and various ways to efficiently and effectively design complex components are proposed. The main contribution is a robust design methodology, which is demonstrated to achieve good results using only a very limited amount of simulations. Therefore this work contributes to a simulation driven design approach starting from an early stage of development, reducing the number of redesigns and tests that are needed.

In Chapter 3 an interval field method is proposed to evaluate and assess the effect of scattered material curves on the overall design, accomplished by bounding the realisations of the interval field to an envelope encompassing a set of experimental results. This is achieved by using an explicit mapping function that scales the distance measure of the basis functions. As such, the intuitive properties of the IDW framework are maintained while more complex functional relations are represented without the need for additional control points. Furthermore, in Chapter 4 a second manner to achieve more complex functional realisations is to enhance the IDW basis functions with an interval field method that controls the gradients at a control point. The proposed interval field uses the value, gradient, or both pieces of information at a control point, as demonstrated on a set of simple example problems.

In Chapter 5 a new framework for modelling and evaluating the crashworthiness of a single component in an early development stage under epistemic uncertainty is proposed. This is accomplished by modelling the behaviour of the

impacted adjacent structure as unknown but spatially coupled uncertain element stiffnesses. The interval valued performance of the structure is obtained using a global optimisation approach, which is shown to be challenging yet feasible for interval field analysis applied to crash simulation. In a follow-up research [13] this method was corroborated by isolating a single component from a full scale crash simulation. The results from this research show that the component-level simulations following this framework are closer to the component performance within the full system. Therefore, the framework enables a first step towards distributed development of complex interacting structures, such as the front structure of a body-in-white.

Finally, to obtain a meaningful design under various types of epistemic parameters, a robustness-based optimisation technique is proposed in Chapter 6, and applied in the context of stochastic functions in Chapter 7. The central idea is that the performance of a design should be robust towards the influence of uncertain parameters. Therefore, a robustness measure is introduced as the ratio between input interval width over output interval width. Based on an efficient global optimisation strategy, using predictions made by a Gaussian Process, the most robust design point is identified with a certain confidence and a minimum amount of function evaluations. In various cases ranging from analytical cases, engineering examples and even on crash simulations, the benefits and performance are benchmarked against state-of-the-art optimisation algorithms. Moreover, using a GP that accounts for homoscedatic noise in the underlying function, a certain negligence towards the numerical inadequacies is achieved. Therefore, the method is able to obtain approximate optimal results in cases that are extremely challenging for most optimisation techniques. Furthermore initial tests have shown that both methods also work in the case of interval field uncertainties, in which case the number of control points determines the number of uncertain parameters from the interval field. Hence reducing the dimension of the interval field reduces the dimension of the underlying GP, which performance deceases drastically for higher dimensions.

The main contributions of this thesis can be summarised as follows:

- A novel interval field basis function is developed that allows to control gradients at the control points, following the intuitive inverse distance weighting method.
- A method is introduced to capture and propagate material uncertainty beyond elasticity, while respecting the underlying physics of the material.
- A novel robustness-based design method is proposed and benchmarked on real numerical impact simulations, taking into account both numerical inadequacies and parameter lack-of-knowledge.

- A dedicated learning function to identify robust design points based on the estimates of a Gaussian process model, is developed.
- Introduction and verification of a framework to deal with uncertain boundary conditions of adjacent components.

9.2 Recommendations for future work

To conclude, in this section a few recommendations and ideas are given that could be investigated for future work.

Improving multi-core capabilities of RULOK The robustness under lack-of-knowledge framework is based on an adaptive refinement of the underlying Gaussian process. The set of samples that is used to calibrate this GP and the subsequently added adaptive samples, are obtained from non-linear numerical models. It is well-known that these models take quite some time to solve, even with sophisticated parallelisation methods. However, the method itself could also benefit from parallelisation, not only for the initial samples, but the adaptively selected points as well. These additional samples would not be optimum in the single-core strategy. However, they could be obtained within the same time-frame and provide additional information. Therefore, reducing the number of iterations and the total time of the optimisation method. A first proposed method in this direction can be found in [44].

Experimental validation using High-Speed DIC All hardware requirements to enable this are provided and various small tests have already been conducted. The main challenge is the design of the experiment, which should still be in-line with the validation requirements. This experiment must be designed to limit all factors of noise and contamination from the experiment. Furthermore, the framerate of the high-speed cameras dictates the speed at which the experiment could be run. It should be noted here, that at impact of the component a (shock) wave is propagating through the medium. Experiments using DIC that use wave propagation have already been proposed [86]. If the setup is unable to capture this wave, the subsequent reflections and harmonics might cast undesired artefacts.

Combination of the component solution spaces to create a general design framework The method proposed in Chapter 5 could be extended to include the component solution spaces. In this regard, more unknown components and

their effects could be modelled, i.e., each designer is provided with bounds on the performance and a modelling framework in which these bounds are tested.

Bibliography

[1] K. Abdel-Tawab and A. K. Noor. "Uncertainty analysis of welding residual stress fields". In: *Computer methods in applied mechanics and engineering* 179.3-4 (1999). Publisher: Elsevier, pp. 327–344.

[2] E. İ. Albak. "Multi-objective crashworthiness optimization of thin-walled multi-cell tubes with different wall lengths". In: *International Journal of Crashworthiness* (2020). Publisher: Taylor & Francis, pp. 1–18.

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

[4] N. Andricevic, F. Duddeck, and S. Hiermaier. "A novel approach for the assessment of robustness of vehicle structures under crash". In: *International Journal of Crashworthiness* 21.2 (2016). Publisher: Taylor & Francis, pp. 89–103. ISSN: 17542111.

[5] ASME. *V & V 10-2019: Standard for Verification and Validation in Computational Solid Mechanics*. 2019.

[6] ASME. *V & V 20-2009: Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer: An American National Standard*. 2009.

[7] ASTM. *D3763-18: Standard Test Method for High Speed Puncture Properties of Plastics Using Load and Displacement Sensors*. 2018.

[8] F. Au, Y. Cheng, L. Tham, and G. Zeng. "Robust design of structures using convex models". In: *Computers & structures* 81.28-29 (2003). Publisher: Elsevier, pp. 2611–2619.

[9] S. Avril, M. Bonnet, A.-s. B. Michel, F. Hild, P. Ienny, F. Latourte, D. Lemosse, S. Pagano, E. Pagnacco, and F. Pierron. "Overview of Identification Methods of Mechanical Parameters Based on Full-field Measurements". In: (2008), pp. 381–402. DOI: 10.1007/s11340-008-9148-y.

[10] F. Bachoc. “Cross validation and maximum likelihood estimations of hyper-parameters of Gaussian processes with model misspecification”. In: *Computational Statistics & Data Analysis* 66 (2013), pp. 55–69.

[11] E. Baker, P. Barbillon, A. Fadikar, R. B. Gramacy, R. Herbei, D. Higdon, J. Huang, L. R. Johnson, P. Ma, A. Mondal, B. Pires, J. Sacks, and V. Sokolov. “Analyzing Stochastic Computer Models: A Review with Opportunities”. eng. In: *Statistical science* 37.1 (2022). Place: Hayward Publisher: Institute of Mathematical Statistics, pp. 64–89. ISSN: 0883-4237.

[12] R. Balcaen, P. Reu, P. Lava, and D. Debruyne. “Stereo-DIC uncertainty quantification based on simulated images”. In: *Experimental Mechanics* 57.6 (2017). Publisher: Springer, pp. 939–951.

[13] R. Barzanooni, C. van Mierlo, M. Pabst, C. Boegle, M. Faes, D. Moens, and F. Duddeck. “Evaluation of uncertain boundary conditions for analysis of structural components with respect to crashworthiness”. In: *Online proceedings ISMA2022-USD2022*. event-place: Leuven, Belgium. 2022.

[14] K. Bathe and E. Wilson. “Stability and accuracy analysis of direct integration methods”. In: *Earthquake Engineering & Structural Dynamics* 1.3 (1972). Publisher: Wiley Online Library, pp. 283–291.

[15] K.-J. Bathe. “Finite element method”. In: *Wiley encyclopedia of computer science and engineering* (2007). Publisher: Wiley Online Library, pp. 1–12.

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

[17] M. J. Bayarri, J. O. Berger, M. C. Kennedy, A. Kottas, R. Paulo, J. Sacks, J. A. Cafeo, C.-H. Lin, and J. Tu. “Predicting vehicle crashworthiness: Validation of computer models for functional and hierarchical data”. In: *Journal of the American Statistical Association* 104.487 (2009). Publisher: Taylor & Francis, pp. 929–943.

[18] A. T. Beck, W. J. Gomes, and F. A. Bazán. “On the robustness of structural risk optimization with respect to epistemic uncertainties”. In: *International Journal for Uncertainty Quantification* 2.1 (2012). Publisher: Begel House Inc.

[19] M. Beer, S. Ferson, and V. Kreinovich. “Imprecise probabilities in engineering analyses”. In: *Mechanical systems and signal processing* 37.1-2 (2013). Publisher: Elsevier, pp. 4–29.

[20] M. Beer and M. Liebscher. “Designing robust structures–a nonlinear simulation based approach”. In: *Computers & Structures* 86.10 (2008). Publisher: Elsevier, pp. 1102–1122.

[21] Y. Ben-Haim. “Info-gap decision theory (IG)”. In: *Decision Making under Deep Uncertainty*. Springer, Cham, 2019, pp. 93–115.

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

[23] Y. Ben-Haim, C. C. Dacso, and N. M. Zetola. “Info-gap management of public health Policy for TB with HIV-prevalence and epidemiological uncertainty”. In: *BMC Public Health* 12.1 (2012). Publisher: BioMed Central, pp. 1–17.

[24] A. Bergek, C. Berggren, K. R. Group, et al. “The impact of environmental policy instruments on innovation: A review of energy and automotive industry studies”. In: *Ecological Economics* 106 (2014). Publisher: Elsevier, pp. 112–123.

[25] W. Betz, I. Papaioannou, and D. Straub. “Numerical methods for the discretization of random fields by means of the Karhunen–Loéve expansion”. In: *Computer Methods in Applied Mechanics and Engineering* 271 (2014), pp. 109–129. ISSN: 0045-7825. DOI: https://doi.org/10.1016/j.cma.2013.12.010. URL: http://www.sciencedirect.com/science/article/pii/S0045782513003502.

[26] H.-G. Beyer and B. Sendhoff. “Robust optimization – A comprehensive survey”. In: *Computer Methods in Applied Mechanics and Engineering* 196.33 (2007), pp. 3190–3218. ISSN: 0045-7825. DOI: https://doi.org/10.1016/j.cma.2007.03.003. URL: https://www.sciencedirect.com/science/article/pii/S0045782507001259.

[27] R. Blumhardt. “Numerische Optimierung des Crashverhaltens von Fahrzeug-Strukturen und-komponenten”. PhD Thesis. Technische Universität München, 2001.

[28] L. Bogaerts, M. Faes, and 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.

[29] O. Boström, M. Y. Svensson, B. Aldman, H.-A. Hansson, Y. Håland, P. Lövsund, T. Seeman, A. Säljö, and T. örtengren. “A new neck injury criterion candidate-based on injury findings in the cervical spinal ganglia after experimental neck extension trauma”. In: *Proceedings of the 1996 International Ircobi Conference on the Biomechanics of Impact, September 11-13, Dublin, Ireland*. 1996, pp. 123–136.

[30] D. L. Bourell, T. J. Watt, D. K. Leigh, and B. Fulcher. “Performance limitations in polymer laser sintering”. In: *Physics Procedia* 56 (2014). Publisher: Elsevier, pp. 147–156.

[31] M. Broggi, M. Faes, E. Patelli, Y. Govers, D. Moens, and M. Beer. "Comparison of Bayesian and interval uncertainty quantification: Application to the AIRMOD test structure". In: *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2017, pp. 1–8.

[32] M. Bujny, N. Aulig, M. Olhofer, and F. Duddeck. "Evolutionary crashworthiness topology optimization of thin-walled structures". In: *ASMO UK, Munich, Germany* (2016).

[33] R. R. Callens, M. G. Faes, and D. Moens. "Local explicit interval fields for non-stationary uncertainty modelling in finite element models". In: *Computer Methods in Applied Mechanics and Engineering* 379 (2021). Publisher: Elsevier, p. 113735.

[34] L. Catallo. "Genetic anti-optimization for reliability structural assessment of precast concrete structures". In: *Computers & structures* 82.13-14 (2004). Publisher: Elsevier, pp. 1053–1065.

[35] Q. Chang, C. Zhou, M. A. Valdebenito, H. Liu, and Z. Yue. "A novel sensitivity index for analyzing the response of numerical models with interval inputs". In: *Computer Methods in Applied Mechanics and Engineering* 400 (2022). Publisher: Elsevier, p. 115509.

[36] B. Chen and B. Pan. "Mirror-assisted multi-view digital image correlation: Principles, applications and implementations". In: *Optics and Lasers in Engineering* 149 (2022). Publisher: Elsevier, p. 106786.

[37] L. Chen and S. Rao. "Fuzzy finite-element approach for the vibration analysis of imprecisely-defined systems". In: *Finite elements in analysis and design* 27.1 (1997). Publisher: Elsevier, pp. 69–83.

[38] S. Chen, L. Liu, S. Zhou, Z. Chen, and Z. Yue. "Multi-domain substructures synthesis with general joints for the dynamics of large structures". In: *Aip Advances* 7.10 (2017). Publisher: AIP Publishing LLC, p. 105007.

[39] Z.-Y. Chen, M. Imholz, L. Li, M. Faes, and D. Moens. "Transient landing dynamics analysis for a lunar lander with random and interval fields". In: *Applied Mathematical Modelling* 88 (2020). Publisher: Elsevier, pp. 827–851.

[40] J. L. D. Comba and J. Stol. "A ne arithmetic and its applications to computer graphics". In: *Proceedings of VI SIBGRAPI (Brazilian Symposium on Computer Graphics and Image Processing)*. Citeseer, 1993, pp. 9–18.

[41] J. A. Cottrell, T. J. Hughes, and Y. Bazilevs. *Isogeometric analysis: toward integration of CAD and FEA*. John Wiley & Sons, 2009.

[42] L. G. Crespo, D. P. Giesy, and 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.

[43] C. Dang, P. Wei, M. G. Faes, M. A. Valdebenito, and M. Beer. “Interval uncertainty propagation by a parallel Bayesian global optimization method”. In: *Applied Mathematical Modelling* 108 (2022). Publisher: Elsevier, pp. 220–235.

[44] C. Dang, P. Wei, M. G. Faes, M. A. Valdebenito, and M. Beer. “Parallel adaptive Bayesian quadrature for rare event estimation”. In: *Reliability Engineering & System Safety* (2022). Publisher: Elsevier, p. 108621.

[45] M. Daub. “Optimizing Flexibility for Component Design in Systems Engineering under Epistemic Uncertainty”. PhD thesis. Munich, Germany: Technische Universität München, 2020.

[46] M. Daub and F. Duddeck. “A decoupled design approach for complex systems under lack-of-knowledge uncertainty”. In: *International Journal of Approximate Reasoning* 119 (2020). Publisher: Elsevier, pp. 408–420.

[47] M. Daub and F. Duddeck. “Maximizing Flexibility for Complex Systems Design to Compensate Lack-of-Knowledge Uncertainty”. In: *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg* 5.4 (2019). Publisher: American Society of Mechanical Engineers Digital Collection, p. 041008.

[48] M. Daub, F. Duddeck, and M. Zimmermann. “Optimizing component solution spaces for systems design”. In: *Structural and Multidisciplinary Optimization* 61 (2020). Publisher: Springer, pp. 2097–2109.

[49] A. De Luca and S. Termini. “A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory”. eng. In: *Information and control* 20.4 (1972). Publisher: Elsevier B.V, pp. 301–312. ISSN: 0019-9958.

[50] C. A. De Moura and C. S. Kubrusly. “The courant–friedrichs–lewy (cfl) condition”. In: *AMC* 10.12 (2013). Publisher: Springer.

[51] W. De Mulder, D. Moens, and D. Vandepitte. “Modeling uncertainty in the context of finite element models with distance-based interpolation”. In: *Proceedings of the 1st international symposium on uncertainty quantification and stochastic modeling*. 2012.

[52] M. De Munck, D. Moens, W. Desmet, and D. Vandepitte. “An efficient response surface based optimisation method for non-deterministic harmonic and transient dynamic analysis”. In: *Computer Modeling in Engineering & Sciences* 47.2 (2009). Publisher: Tech Science Press, pp. 119–166.

[53] D. Degrauwe, G. Lombaert, and G. De Roeck. "Improving interval analysis in finite element calculations by means of affine arithmetic". In: *Computers & structures* 88.3-4 (2010). Publisher: Elsevier, pp. 247–254.

[54] E. W. Dijkstra. "A note on two problems in connexion with graphs". In: *Numerische mathematik* 1.1 (1959), pp. 269–271.

[55] W. Dong and H. C. Shah. "Vertex method for computing functions of fuzzy variables". In: *Fuzzy sets and Systems* 24.1 (1987). Publisher: Elsevier, pp. 65–78.

[56] D. K. Dries. *Tissue-Level Tolerance Criteria for Crash-Related Head Injuries: A Combined Experimental and Numerical Approach*. eng. Place: Leuven. 2018.

[57] D. Dubois and H. Prade. "Possibility theory: qualitative and quantitative aspects". In: *Quantified representation of uncertainty and imprecision*. Springer, 1998, pp. 169–226.

[58] V. Dubourg. "Adaptive surrogate models for reliability analysis and reliability-based design optimization". PhD Thesis. Université Blaise Pascal-Clermont-Ferrand II, 2011.

[59] F. Duddeck. "Multidisciplinary optimization of car bodies". In: *Structural and Multidisciplinary Optimization* 35.4 (2008). Publisher: Springer, pp. 375–389.

[60] UN-ECE. *United Nations Economic Commision for Europe*. Assessed on: 14-05-2020. URL: `https://www.unece.org/trans/main/welcwp29.html`.

[61] I. Elishakoff, R. T. Haftka, and J. Fang. "Structural design under bounded uncertainty—Optimization with anti-optimization". In: *Computers & Structures* 53.6 (1994), pp. 1401–1405. ISSN: 0045-7949.

[62] I. Elishakoff et al. "A comparison of stochastic and interval finite elements applied to shear frames with uncertain stiffness properties". In: *Computers & structures* 67.1-3 (1998). Publisher: Elsevier, pp. 91–98.

[63] I. Elishakoff and Y. Miglis. "Novel parameterized intervals may lead to sharp bounds". In: *Mechanics Research Communications* 44 (2012). Publisher: Elsevier, pp. 1–8.

[64] EURO-NCAP. *Euro New Car Assessment Programme*. Assessed on: 14-05-2020. URL: `https://www.euroncap.com`.

[65] M. Faes and D. Moens. "On auto- and cross-interdependence in interval field finite element analysis". In: *International Journal for Numerical Methods in Engineering* 121.9 (2020). Publisher: John Wiley and Sons Ltd, pp. 2033–2050. ISSN: 00295981. DOI: `10.1002/nme.6297`.

[66] M. Faes and D. Moens. “Recent Trends in the Modeling and Quantification of Non-probabilistic Uncertainty”. In: *Archives of Computational Methods in Engineering* 27 (2019). Publisher: Springer (part of Springer Nature), pp. 633–671. ISSN: 1134-3060.

[67] M. Faes. “Interval methods for the identification and quantification of inhomogeneous uncertainty in Finite Element models”. In: (2017).

[68] M. Faes, M. Broggi, E. Patelli, Y. Govers, J. Mottershead, M. Beer, and D. Moens. “A multivariate interval approach for inverse uncertainty quantification with limited experimental data”. In: *Mechanical Systems and Signal Processing* 118 (2019), pp. 534–548. ISSN: 0888-3270.

[69] M. Faes, R. Callens, and D. Moens. “Adaptive sparse grid approximation for high dimensional interval field construction”. In: *AIAA Scitech 2020 Forum*. 2020, p. 1418.

[70] M. Faes, M. Imholz, D. Vandepitte, and D. Moens. “A Review of Interval Field Approaches for Uncertainty Quantification in Numerical Models”. In: *Modern Trends in Structural and Solid Mechanics 3: Non-deterministic Mechanics* (2021). Publisher: Wiley Online Library, pp. 95–110.

[71] M. Faes and D. Moens. “Identification and quantification of spatial interval uncertainty in numerical models”. In: *Computers & Structures* 192 (2017). Publisher: Elsevier, pp. 16–33.

[72] M. Faes and D. Moens. “On auto- and cross-interdependence in interval field finite element analysis”. In: *International Journal for Numerical Methods in Engineering* February (Jan. 2020), nme.6297.

[73] M. Faes, G. D. Sabyasachi, and D. Moens. “Hybrid spatial uncertainty analysis for the estimation of imprecise failure probabilities in Laser Sintered PA-12 parts”. In: *Computers & Mathematics with Applications* 78.7 (2019). Publisher: Elsevier, pp. 2395–2406.

[74] M. Faes, J. Sadeghi, M. Broggi, M. De Angelis, E. Patelli, M. Beer, and D. Moens. “On the robust estimation of small failure probabilities for strong nonlinear models”. In: *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg* 5.4 (2019). Publisher: American Society of Mechanical Engineers Digital Collection.

[75] M. Faes, Y. Wang, P. Lava, and D. Moens. “Variability, heterogeneity, and anisotropy in the quasi-static response of laser sintered PA12 components”. In: *Strain* 53.2 (2017). Publisher: Wiley Online Library, e12219.

[76] M. G. Faes, M. Broggi, P. D. Spanos, and M. Beer. "Elucidating appealing features of differentiable auto-correlation functions: A study on the modified exponential kernel". In: *Probabilistic Engineering Mechanics* 69 (2022). Publisher: Elsevier, p. 103269.

[77] M. G. Faes, M. Daub, S. Marelli, E. Patelli, and M. Beer. "Engineering analysis with probability boxes: a review on computational methods". In: *Structural Safety* 93 (2021). Publisher: Elsevier, p. 102092.

[78] M. G. Faes and M. A. Valdebenito. "Fully decoupled reliability-based design optimization of structural systems subject to uncertain loads". In: *Computer Methods in Applied Mechanics and Engineering* 371 (2020). Publisher: Elsevier, p. 113313.

[79] L. Farkas, D. Moens, S. Donders, and D. Vandepitte. "Optimisation study of a vehicle bumper subsystem with fuzzy parameters". In: *Mechanical systems and signal processing* 32 (2012). Publisher: Elsevier, pp. 59–68.

[80] P. E. Farrell, D. A. Ham, S. W. Funke, and M. E. Rognes. "Automated derivation of the adjoint of high-level transient finite element programs". In: *SIAM Journal on Scientific Computing* 35.4 (2013). Publisher: SIAM, pp. C369–C393.

[81] J. Fender, F. Duddeck, and M. Zimmermann. "Direct computation of solution spaces". In: *Structural and Multidisciplinary Optimization* 55.5 (2017). Publisher: Springer, pp. 1787–1796.

[82] J. Fender, F. Duddeck, and M. Zimmermann. "On the calibration of simplified vehicle crash models". In: *Structural and Multidisciplinary Optimization* 49.3 (2014). Publisher: Springer, pp. 455–469.

[83] C. Feng, M. Faes, M. Broggi, C. Dang, J. Yang, Z. Zheng, and M. Beer. "Application of interval field method to the stability analysis of slopes in presence of uncertainties". In: *Computers and Geotechnics* 153 (2023). Publisher: Elsevier, p. 105060.

[84] S. Ferson and L. R. Ginzburg. "Different methods are needed to propagate ignorance and variability". In: *Reliability Engineering and System Safety* 54.2-3 (1996), pp. 133–144. ISSN: 09518320. DOI: 10.1016/S0951-8320(96)00071-3.

[85] F. Fleissner, T. Haag, M. Hanss, and P. Eberhard. "Analysis of granular chute flow based on a particle model including uncertainties". In: *Trends in computational contact mechanics* (2011). Publisher: Springer, pp. 121–134.

[86] L. Fletcher and F. Pierron. "An image-based inertial impact (IBII) test for tungsten carbide cermets". In: *Journal of Dynamic Behavior of Materials* 4.4 (2018). Publisher: Springer, pp. 481–504.

[87] J. K. Foster, J. O. Kortge, and M. J. Wolanin. “Hybrid III—a biomechanically-based crash test dummy”. In: *SAE Transactions* (1977). Publisher: JSTOR, pp. 3268–3283.

[88] K. Fujita and I. Takewaki. “An efficient methodology for robustness evaluation by advanced interval analysis using updated second-order Taylor series expansion”. In: *Engineering Structures* 33.12 (2011). Publisher: Elsevier, pp. 3299–3310.

[89] F. Gao, J. Sacks, and W. J. Welch. “Predicting urban ozone levels and trends with semiparametric modeling”. In: *Journal of Agricultural, Biological, and Environmental Statistics* (1996). Publisher: JSTOR, pp. 404–425.

[90] M. B. Giles, M. C. Duta, J.-D. Muller, and N. A. Pierce. “Algorithm developments for discrete adjoint methods”. In: *AIAA journal* 41.2 (2003), pp. 198–205.

[91] P. E. Gill, W. Murray, and M. A. Saunders. “SNOPT: An SQP algorithm for large-scale constrained optimization”. In: *SIAM review* 47.1 (2005). Publisher: SIAM, pp. 99–131.

[92] GLOBAL-NCAP. *Global New Car Assessment Programme*. Assessed on: 14-05-2020. URL: `http://www.globalncap.org`.

[93] R. Gooding. *Supercomputers in crash simulation*. Publication Title: Automotive testing technology international. Mar. 2021. URL: `https://www.automotivetestingtechnologyinternational.com/features/supercomputers-in-crash-simulation.html`.

[94] R. Goodridge, C. Tuck, and R. Hague. “Laser sintering of polyamides and other polymers”. In: *Progress in Materials science* 57.2 (2012). Publisher: Elsevier, pp. 229–267.

[95] M. Guerder, A. Duval, T. Elguedj, P. Feliot, and J. Touzeau. “Isogeometric shape optimisation of volumetric blades for aircraft engines”. In: *Structural and Multidisciplinary Optimization* 65.3 (2022). Publisher: Springer, pp. 1–21.

[96] T. Haag, J. Herrmann, and M. Hanss. “Identification procedure for epistemic uncertainties using inverse fuzzy arithmetic”. In: *Mechanical Systems and Signal Processing* 24.7 (2010). Publisher: Elsevier, pp. 2021–2034.

[97] R. Haberfellner, O. De Weck, E. Fricke, and S. Vössner. *Systems engineering: fundamentals and applications*. Springer, 2019.

[98] S. A. Hall, M. Bornert, J. Desrues, Y. Pannier, N. Lenoir, G. Viggiani, and P. Bésuelle. “Discrete and continuum analysis of localised deformation in sand using X-ray $/mu$CT and volumetric digital image correlation”. In: *Géotechnique* 60.5 (2010). Publisher: Thomas Telford Ltd, pp. 315–322.

[99] M. Hanss and S. Turrin. “A fuzzy-based approach to comprehensive modeling and analysis of systems with epistemic uncertainties”. In: *Structural Safety* 32.6 (2010). Publisher: Elsevier, pp. 433–441.

[100] M. Hanss. *Applied fuzzy arithmetic*. Springer, 2005.

[101] M. Hanss. “The transformation method for the simulation and analysis of systems with uncertain parameters”. In: *Fuzzy Sets and systems* 130.3 (2002). Publisher: Elsevier, pp. 277–289.

[102] C. Haskins, K. Forsberg, M. Krueger, D. Walden, and D. Hamelin. “Systems engineering handbook”. In: *INCOSE*. Vol. 9. International Council on Systems Engineering Seattle, 2006, pp. 13–16.

[103] E. Hilfrich and D. Seidner. “Crash safety with high strength steels, presented at International Automotive Congress, November 1st”. In: *Shengyang, China: Verlag Heinrich Vogel* (2008).

[104] R. Horst and P. M. Pardalos. *Handbook of global optimization*. Vol. 2. Springer Science & Business Media, 2013.

[105] H. Hu, Y. Wu, A. Batou, and H. Ouyang. “B-spline based interval field decomposition method”. In: *Computers & Structures* 272 (Nov. 2022), p. 106874. DOI: 10.1016/j.compstruc.2022.106874.

[106] N. Hu and B. Duan. “An efficient robust optimization method with random and interval uncertainties”. In: *Structural and Multidisciplinary Optimization* 58.1 (2018). Publisher: Springer, pp. 229–243.

[107] S. Hunkeler, F. Duddeck, M. Rayamajhi, and H. Zimmer. “Shape optimisation for crashworthiness followed by a robustness analysis with respect to shape variables”. In: *Structural and Multidisciplinary Optimization* 48.2 (2013). Publisher: Springer, pp. 367–378.

[108] M. Imholz, M. Faes, J. Cerneels, D. Vandepitte, and D. Moens. “On the comparison of two novel interval field formulations for the representation of spatial uncertainty”. In: *Proceedings of the 7th international workshop on reliable engineering computing*. Vol. 7. University of Bochum; Bochum, 2016, pp. 367–378.

[109] M. Imholz, M. Faes, D. Vandepitte, and D. Moens. “Robust uncertainty quantification in structural dynamics under scarse experimental modal data: A Bayesian-interval approach”. In: *Journal of Sound and Vibration* 467 (Feb. 2020). Publisher: Elsevier Ltd, p. 114983. ISSN: 0022460X. DOI: 10.1016/j.jsv.2019.114983.

[110] M. Imholz, D. Vandepitte, and D. Moens. “Analysis of the effect of uncertain clamping stiffness on the dynamical behaviour of structures using interval field methods”. In: *Applied mechanics and materials*. Vol. 807. Trans Tech Publ, 2015, pp. 195–204.

[111] M. Imholz, D. Vandepitte, and D. Moens. “Application of interval fields to fit experimental data on deepdrawn components”. In: *Proceedings of the joint ICVRAM ISUMA UNCERTAINTIES conference*. University of Sao Paolo, Structural Engineering Department, 2018.

[112] M. Imholz, D. Vandepitte, and D. Moens. “Derivation of an input interval field decomposition based on expert knowledge using locally defined basis functions”. In: *1st ECCOMAS Thematic conference on international conference on uncertainty quantification in computational sciences and engineering*. 2015, pp. 1–19.

[113] T. A. S. Institute. *Robust Design, Technical Report*. Publication Title: The American Supplier Institute. 2005.

[114] *ISO 6603-1:2000: Plastics - Determination of puncture impact behaviour of rigid plastics - Part 1: Non-instrumented impact testing*. Place: Geneva, CH Type: Standard Volume: 2000. Mar. 2000.

[115] J. Janusevskis and R. Le Riche. “Simultaneous kriging-based estimation and optimization of mean response”. In: *Journal of Global Optimization* 55.2 (2013). Publisher: Springer, pp. 313–336.

[116] C. Jiang, B. Ni, N. Liu, X. Han, and J. Liu. “Interval process model and non-random vibration analysis”. In: *Journal of Sound and Vibration* 373 (2016). Publisher: Elsevier, pp. 104–131.

[117] C. Johnsson, A. Laureshyn, and T. De Ceunynck. “In search of surrogate safety indicators for vulnerable road users: a review of surrogate safety indicators”. In: *Transport Reviews* 38.6 (2018). Publisher: Taylor & Francis, pp. 765–785.

[118] D. R. Jones, M. Schonlau, and W. J. Welch. “Efficient global optimization of expensive black-box functions”. In: *Journal of Global optimization* 13.4 (1998). Publisher: Springer, pp. 455–492.

[119] E. M. C. Jones and M. A. Ladicola. *A Good Practices Guide for Digital Image Correlation*. 2018. DOI: 10.32720/idics/gpg.ed1.

[120] I. Kallina, F. Zeidler, K. Baumann, and D. Scheunert. “The offset crash against a deformable barrier, a more realistic frontal impact”. In: *Proceedings: International Technical Conference on the Enhanced Safety of Vehicles*. Vol. 1995. National Highway Traffic Safety Administration, 1995, pp. 1300–1304.

[121] A. D. Kammers and S. Daly. “Digital image correlation under scanning electron microscopy: methodology and validation”. In: *Experimental Mechanics* 53.9 (2013). Publisher: Springer, pp. 1743–1761.

[122] A. Khakhali, N. Nariman-Zadeh, A. Darvizeh, A. Masoumi, and B. Notghi. "Reliability-based robust multi-objective crashworthiness optimisation of S-shaped box beams with parametric uncertainties". In: *International Journal of Crashworthiness* 15.4 (2010). Publisher: Taylor & Francis, pp. 443–456.

[123] H. H. Khodaparast, J. E. Mottershead, and K. J. Badcock. "Interval model updating with irreducible uncertainty using the Kriging predictor". In: *Mechanical Systems and Signal Processing* 25.4 (2011). Publisher: Elsevier, pp. 1204–1226.

[124] J. P. Kleijnen. "Kriging metamodeling in simulation: A review". In: *European journal of operational research* 192.3 (2009). Publisher: Elsevier, pp. 707–716.

[125] J. P. Kleijnen. "Regression and Kriging metamodels with their experimental designs in simulation: a review". In: *European Journal of Operational Research* 256.1 (2017). Publisher: Elsevier, pp. 1–16.

[126] S. Kleiven. "Predictors for traumatic brain injuries evaluated through accident reconstructions". In: *Stapp car crash J* 51.1 (2007), pp. 81–114.

[127] T. Knoke. "Mixed forests and finance—Methodological approaches". In: *Ecological Economics* 65.3 (2008). Publisher: Elsevier, pp. 590–601.

[128] A. Kossiakoff, W. N. Sweet, S. J. Seymour, and S. M. Biemer. *Systems engineering principles and practice*. Vol. 83. John Wiley & Sons, 2011.

[129] D. G. Krige. "A statistical approach to some basic mine valuation problems on the Witwatersrand". In: *Journal of the Southern African Institute of Mining and Metallurgy* 52.6 (1951). Publisher: Southern African Institute of Mining and Metallurgy, pp. 119–139.

[130] A. Kuczkowiak, S. Cogan, M. Ouisse, E. Foltête, and M. Corus. "Experimental Validation of an Info-Gap Uncertainty Model for a Robustness Analysis of Structural Responses". In: *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg* 6.3 (May 2020). ISSN: 2332-9017. DOI: 10.1115/1.4047096.

[131] P. Lamberti and V. Tucci. "Interval approach to robust design". In: *COMPEL-The international journal for computation and mathematics in electrical and electronic engineering* 26.2 (2007). Publisher: Emerald Group Publishing Limited, pp. 280–292.

[132] C. Lataniotis, D. Wicaksono, S. Marelli, and B. Sudret. *UQLab user manual – Kriging (Gaussian process modeling)*. Tech. rep. Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich,Switzerland, 2021.

[133] P. Lava, W. Van Paepegem, S. Coppieters, I. De Baere, Y. Wang, and D. Debruyne. "Impact of lens distortions on strain measurements obtained with 2D digital image correlation". In: *Optics and Lasers in Engineering* 51.5 (2013). Publisher: Elsevier, pp. 576–584.

[134] R. Le Riche, V. Picheny, A. Meyer, N.-H. Kim, and D. Ginsbourger. "Gears design with shape uncertainties using controlled monte carlo simulations and kriging". In: *50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference 17th AIAA/ASME/AHS Adaptive Structures Conference 11th AIAA No.* 2009, p. 2257.

[135] D.-S. Lee, L. F. Gonzalez, J. Periaux, and K. Srinivas. "Robust design optimisation using multi-objective evolutionary algorithms". In: *Computers & Fluids* 37.5 (2008). Publisher: Elsevier, pp. 565–583.

[136] C. Li, B. Chen, H. Peng, and S. Zhang. "Sparse regression Chebyshev polynomial interval method for nonlinear dynamic systems under uncertainty". In: *Applied Mathematical Modelling* 51 (2017). Publisher: Elsevier, pp. 505–525.

[137] Livermore Software Technology Corporation (LSTC). *Keyword Users Manual Volume 2: Material Models*. Vol. II. 2014. ISBN: 0-9778540-3-5.

[138] M. Mäck and M. Hanss. "Efficient Possibilistic Uncertainty Analysis of a Car Crash Scenario Using a Multifidelity Approach". In: *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg* 5.4 (2019). Publisher: American Society of Mechanical Engineers Digital Collection.

[139] G. Manson. "Calculating frequency response functions for uncertain systems using complex affine analysis". In: *Journal of Sound and Vibration* 288.3 (2005). Publisher: Elsevier, pp. 487–521.

[140] S. Marelli and B. Sudret. "UQLab: A Framework for Uncertainty Quantification in Matlab". In: *Vulnerability, Uncertainty, and Risk*. _eprint: https://ascelibrary.org/doi/pdf/10.1061/9780784413609.257, pp. 2554–2563. DOI: 10.1061/9780784413609.257. URL: https://ascelibrary.org/doi/abs/10.1061/9780784413609.257.

[141] B. Matérn. *Spatial variation*. Vol. 36. Springer Science & Business Media, 1960.

[142] *Matlab and Statistics and Machine Learning Toolbox release 2020b*. Publication Title: Fit a Gaussian process regression (GPR) model. URL: https://nl.mathworks.com/help/stats/fitrgp.html.

[143] A. Mayyas, A. Qattawi, M. Omar, and D. Shan. "Design for sustainability in automotive industry: A comprehensive review". In: *Renewable and sustainable energy reviews* 16.4 (2012). Publisher: Elsevier, pp. 1845–1862.

[144] J. P. McDonald, T. Shams, N. Rangarajan, D. Beach, et al. “Design and development of a THOR based small female crash test dummy”. In: *Stapp car crash journal* 47 (2003). Publisher: The Stapp Association, p. 551.

[145] MDPVS. *Measured Data Processing Vehicle Safety Workgroup*. Assessed on: 14-05-2020. URL: `http://mdvfs.org`.

[146] M. Meßmer, L. F. Leidinger, S. Hartmann, F. Bauer, F. Duddeck, R. Wüchner, and K.-U. Bletzinger. “Isogeometric analysis on trimmed solids: A B-spline-based approach focusing on explicit dynamics”. In: *Proceedings of the 13th European LS-DYNA Conference*. 2021.

[147] C. van Mierlo, L. Burmberger, M. Daub, F. Duddeck, M. Faes, and D. Moens. “Component-level impact performance assessment under spatially uncertain boundary conditions”. In: *Proceedings of International Conference on Uncertainty in Structural Dynamics*. International Conference on Uncertainty in Structural Dynamics, 2020, pp. 1–15.

[148] C. van Mierlo, L. Burmberger, M. Daub, F. Duddeck, M. G. Faes, and D. Moens. “Interval methods for lack-of-knowledge uncertainty in crash analysis”. In: *Mechanical Systems and Signal Processing* 168 (2022). Publisher: Elsevier, p. 108574.

[149] C. van Mierlo, W. Dirix, M. Faes, and D. Moens. “Impact modeling of hyper-elastic am compliant mechanisms using high-speed digital image correlation”. In: *Exploring the Design Freedom of AdditiveManufacturing through Simulation, Date: 2018/12/10-2018/12/11, Location: Helsinki, Finland*. 2018, pp. 72–72.

[150] C. van Mierlo, M. Faes, and D. Moens. “Identification of visco-plastic material model parameters using interval fields”. In: *Proceedings of the 3rd International Conference on Uncertainty Quantification in Computational sciences and Engineering*. ECCOMAS, 2019.

[151] C. van Mierlo, M. Faes, and D. Moens. “Robustness of component performance under interval valued uncertainty”. In: *Proceedings of the 4rd International Conference on Uncertainty Quantification in Computational Sciences and Engineering*. ECCOMAS, 2021.

[152] C. van Mierlo, M. G. Faes, and D. Moens. “Inhomogeneous interval fields based on scaled inverse distance weighting interpolation”. In: *Computer Methods in Applied Mechanics and Engineering* 373 (2021). Publisher: Elsevier, p. 113542.

[153] C. van Mierlo, M. G. Faes, and D. Moens. “Interval field methods with local gradient control”. In: *Proceedings of the European Congress on Computational Methods in Applied Sciences and Engineering*. event-place: Oslo, Norway. ECCOMAS.

[154] C. van Mierlo, A. Persoons, M. G. Faes, and D. Moens. "Robust design optimisation under lack-of-knowledge uncertainty". In: *Computers & Structures* 275 (2023). Publisher: Elsevier, p. 106910.

[155] C. van Mierlo, A. Persoons, M. G. Faes, and D. Moens. "Robust design optimization of expensive stochastic simulators under lack-of-knowledge". In: *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg* 0.0 (2023). Publisher: American Society of Mechanical Engineers Digital Collection, p. 0.

[156] K. van Mierlo, A. Persoons, M. Faes, and D. Moens. *Replication Data for: Robust Design Optimization of Expensive Stochastic Simulators Under Lack-of-Knowledge*. Version V1. 2023. DOI: 10.48804/A3V0Z4. URL: https://doi.org/10.48804/A3V0Z4.

[157] I. P. Mitseas, I. Kougioumtzoglou, M. Beer, E. Patelli, and J. Mottershead. "Robust Design Optimization of Structural Systems Under Evolutionary Stochastic Seismic Excitation". In: June 2014, pp. 215–224. ISBN: 978-0-7844-1360-9. DOI: 10.1061/9780784413609.022.

[158] D. Moens, M. De Munck, W. Desmet, and D. Vandepitte. "Numerical dynamic analysis of uncertain mechanical structures based on interval fields". In: *IUTAM Symposium on the Vibration Analysis of Structures with Uncertainties*. Ed. by A. K. Belyaev and R. S. Langley. Dordrecht: Springer Netherlands, 2011, pp. 71–83. ISBN: 978-94-007-0289-9.

[159] D. Moens and M. Hanss. "Non-probabilistic finite element analysis for parametric uncertainty treatment in applied mechanics: Recent advances". eng. In: *Finite Elements in Analysis & Design* 47.1 (2011). Publisher: Elsevier B.V, pp. 4–16. ISSN: 0168-874X.

[160] D. Moens and D. Vandepitte. "A survey of non-probabilistic uncertainty treatment in finite element analysis". In: *Computer Methods in Applied Mechanics and Engineering* 194.12 (2005), pp. 1527–1555. ISSN: 0045-7825.

[161] D. Moens and D. Vandepitte. "Interval sensitivity theory and its application to frequency response envelope analysis of uncertain structures". In: *Computer methods in applied mechanics and engineering* 196.21-24 (2007). Publisher: Elsevier, pp. 2486–2496.

[162] B. Möller, W. Graf, and M. Beer. "Fuzzy structural analysis using α-level optimization". In: *Computational mechanics* 26.6 (2000). Publisher: Springer, pp. 547–565.

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

[164] R. E. Moore. *Interval arithmetic and automatic error analysis in digital computing*. Tech. rep. Stanford Univ Calif Applied Mathematics and Statistics Labs, 1962.

[165] R. E. Moore. *Methods and applications of interval analysis*. SIAM, 1979.

[166] S. Moritz Göhler, T. Eifler, and T. J. Howard. "Robustness Metrics: Consolidating the Multiple Approaches to Quantify Robustness". In: *Journal of Mechanical Design* 138.11 (Sept. 2016).

[167] M. Moustapha, B. Sudret, J.-M. Bourinet, and B. Guillaume. "Adaptive Kriging reliability-based design optimization of an automotive body structure under crashworthiness constraints". In: International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP) (12th : 2015). July 2015. DOI: http://dx.doi.org/10.14288/1.0076165.

[168] M. Moustapha, B. Sudret, J.-M. Bourinet, and B. Guillaume. "Quantile-based optimization under uncertainties using adaptive Kriging surrogate models". In: *Structural and multidisciplinary optimization* 54.6 (2016). Publisher: Springer, pp. 1403–1421.

[169] R. L. Muhanna and R. L. Mullen. "Uncertainty in mechanics problems—interval–based approach". In: *Journal of Engineering Mechanics* 127.6 (2001). Publisher: American Society of Civil Engineers, pp. 557–566.

[170] R. L. Muhanna, H. Zhang, and R. L. Mullen. "Combined axial and bending stiffness in interval finite-element methods". In: *Journal of Structural Engineering* 133.12 (2007). Publisher: American Society of Civil Engineers, pp. 1700–1709.

[171] G. Muscolino and A. Sofi. "Bounds for the stationary stochastic response of truss structures with uncertain-but-bounded parameters". In: *Mechanical Systems and Signal Processing* 37.1-2 (2013). Publisher: Elsevier, pp. 163–181.

[172] G. Muscolino and A. Sofi. "Stochastic analysis of structures with uncertain-but-bounded parameters via improved interval analysis". In: *Probabilistic Engineering Mechanics* 28 (2012). Publisher: Elsevier, pp. 152–163.

[173] J. A. Newman. "A generalized acceleration model for brain injury threshold (GAMBIT)". In: *Proceedings of International IRCOBI Conference, 1986*. 1986.

[174] J. A. Newman and N. Shewchenko. *A proposed new biomechanical head injury assessment function-the maximum power index*. Tech. rep. SAE Technical Paper, 2000.

[175] NHTSA. *National Highway Traffic Safety Administration (NHTSA), Crash Simulation Vehicle Models*. Assessed on: 14-05-2020. URL: https://www.nhtsa.gov/crash-simulation-vehicle-models.

[176] B. Ni and C. Jiang. "Interval field model and interval finite element analysis". In: *Computer Methods in Applied Mechanics and Engineering* 360 (2020), p. 112713. ISSN: 0045-7825. DOI: https://doi.org/10.1016/j.cma.2019.112713. URL: http://www.sciencedirect.com/science/article/pii/S0045782519306012.

[177] B. M. Nicolai, J. A. Egea, N. Scheerlinck, J. R. Banga, and A. K. Datta. "Fuzzy finite element analysis of heat conduction problems with uncertain parameters". In: *Journal of Food Engineering* 103.1 (2011). Publisher: Elsevier, pp. 38–46.

[178] W. H. Organization et al. "Road safety". In: (2020). Publisher: OECD.

[179] G.-J. Park, T.-H. Lee, K. H. Lee, and K.-H. Hwang. "Robust design: an overview". In: *AIAA journal* 44.1 (2006), pp. 181–191.

[180] M. Pavan, M. Faes, D. Strobbe, B. Van Hooreweder, T. Craeghs, D. Moens, and W. Dewulf. "On the influence of inter-layer time and energy density on selected critical-to-quality properties of PA12 parts produced via laser sintering". In: *Polymer testing* 61 (2017). Publisher: Elsevier, pp. 386–395.

[181] V. Picheny and D. Ginsbourger. "Noisy kriging-based optimization methods: a unified implementation within the DiceOptim package". In: *Computational Statistics & Data Analysis* 71 (2014). Publisher: Elsevier, pp. 1035–1053.

[182] L. Piegl and W. Tiller. *The NURBS book*. Springer Science & Business Media, 1996.

[183] F. Pierron and M. Grédiac. "Towards Material Testing 2.0. A review of test design for identification of constitutive parameters from full-field measurements". In: *Strain* 57.1 (2021). Publisher: Wiley Online Library, e12370.

[184] F. Pierron and M. Grédiac. *The virtual fields method: extracting constitutive mechanical parameters from full-field deformation measurements*. Springer Science & Business Media, 2012.

[185] K. Price, R. M. Storn, and J. A. Lampinen. *Differential evolution: a practical approach to global optimization*. Springer Science & Business Media, 2006.

[186] Z. Qiu and X. Wang. "Vertex solution theorem for the upper and lower bounds on the dynamic response of structures with uncertain-but-bounded parameters". In: *Acta Mechanica Sinica* 25.3 (2009). Publisher: Springer, pp. 367–379.

[187] S. S. Rao and L. Berke. “Analysis of uncertain structural systems using interval analysis”. In: *AIAA journal* 35.4 (1997), pp. 727–735.

[188] S. S. Rao and L. Chen. “Numerical solution of fuzzy linear equations in engineering analysis”. In: *International Journal for Numerical Methods in Engineering* 42.5 (1998). Publisher: Wiley Online Library, pp. 829–846.

[189] S. S. Rao and J. P. Sawyer. “Fuzzy finite element approach for analysis of imprecisely defined systems”. In: *AIAA journal* 33.12 (1995), pp. 2364–2370.

[190] C. E. Rasmussen. “Gaussian processes in machine learning”. In: *Summer school on machine learning*. Springer, 2003, pp. 63–71.

[191] S. ur Rehman, M. Langelaar, and F. van Keulen. “Efficient Kriging-based robust optimization of unconstrained problems”. In: *Journal of Computational Science* 5.6 (2014). Publisher: Elsevier, pp. 872–881.

[192] T. Roach, Z. Kapelan, R. Ledbetter, and M. Ledbetter. “Comparison of robust optimization and info-gap methods for water resource management under deep uncertainty”. In: American Society of Civil Engineers, 2016.

[193] M. Rossi, P. Lava, F. Pierron, D. Debruyne, and M. Sasso. “Effect of DIC spatial resolution, noise and interpolation error on identification results with the VFM”. In: *Strain* 51.3 (2015). Publisher: Wiley Online Library, pp. 206–222.

[194] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn. “Design and analysis of computer experiments”. In: *Statistical science* 4.4 (1989). Publisher: Institute of Mathematical Statistics, pp. 409–423.

[195] J. Sadeghi, M. De Angelis, and E. Patelli. “Robust propagation of probability boxes by interval predictor models”. In: *Structural Safety* 82 (2020). Publisher: Elsevier, p. 101889.

[196] SAE. “J211-1.1995”. In: 552 (2010). ISBN: 0784406243. DOI: 10.1520/G0154-12A.

[197] M. Sánchez and D. Abellán. “Development of new deformable barriers for testing vehicle performance in different crash configurations”. In: *International journal of crashworthiness* 20.4 (2015). Publisher: Taylor & Francis, pp. 370–386.

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

[199] H. Schreier, J.-J. Orteu, M. A. Sutton, et al. *Image correlation for shape, motion and deformation measurements: Basic concepts, theory and applications*. Vol. 1. Springer, 2009.

[200] D. Shepard. "A Two-Dimensional Interpolation Function for Irregularly-Spaced Data". In: *Proceedings of the 1968 23rd ACM National Conference*. ACM '68. New York, NY, USA: Association for Computing Machinery, 1968, pp. 517–524. ISBN: 978-1-4503-7486-6. DOI: 10.1145/800186.810616.

[201] T. M. Shortell. "INCOSE systems engineering handbook: a guide for system life cycle processes and activities". In: *John Wiley& Sons,* (2015).

[202] T. W. Simpson, T. M. Mauery, J. J. Korte, and F. Mistree. "Kriging models for global approximation in simulation-based multidisciplinary design optimization". In: *AIAA journal* 39.12 (2001), pp. 2233–2241.

[203] A. Sofi and G. Muscolino. "Static analysis of Euler–Bernoulli beams with interval Young's modulus". In: *Computers & Structures* 156 (2015). Publisher: Elsevier, pp. 72–82.

[204] A. Sofi, G. Muscolino, and I. Elishakoff. "Natural frequencies of structures with interval parameters". In: *Journal of Sound and Vibration* 347 (2015). Publisher: Elsevier, pp. 79–95.

[205] A. Sofi and E. Romeo. "A unified response surface framework for the interval and stochastic finite element analysis of structures with uncertain parameters". In: *Probabilistic Engineering Mechanics* 54 (2018). Publisher: Elsevier, pp. 25–36.

[206] A. Sofi. "Euler–Bernoulli interval finite element with spatially varying uncertain properties". In: *Acta Mechanica* 228.11 (2017). Publisher: Springer, pp. 3771–3787.

[207] A. Sofi. "Structural response variability under spatially dependent uncertainty: Stochastic versus interval model". In: *Probabilistic Engineering Mechanics* 42 (2015), pp. 78–86. ISSN: 0266-8920.

[208] A. Sofi, G. Muscolino, and I. Elishakoff. "Static response bounds of Timoshenko beams with spatially varying interval uncertainties". In: *Acta mechanica* 226 (2015), pp. 3737–3748. DOI: 10.1007/s00707-015-1400-9.

[209] A. Sofi, G. Muscolino, and F. Giunta. "Propagation of uncertain structural properties described by imprecise Probability Density Functions via response surface method". In: *Probabilistic Engineering Mechanics* 60 (2020). Publisher: Elsevier, p. 103020.

[210] A. Sofi and E. Romeo. "A novel interval finite element method based on the improved interval analysis". In: *Computer Methods in Applied Mechanics and Engineering* 311 (2016). Publisher: Elsevier, pp. 671–697.

[211] A. Sofi, E. Romeo, O. Barrera, and A. Cocks. “An interval finite element method for the analysis of structures with spatially varying uncertainties”. In: *Advances in Engineering Software* 128 (2018). Publisher: Advances in Engineering Software, pp. 1–19.

[212] P. D. Spanos and R. Ghanem. “Stochastic finite element expansion for random media”. In: *Journal of engineering mechanics* 115.5 (1989). Publisher: American Society of Civil Engineers, pp. 1035–1053.

[213] G. Stefanou. “The stochastic finite element method: Past, present and future”. In: *Computer Methods in Applied Mechanics and Engineering* 198.9-12 (2009). ISBN: 0045-7825 Publisher: Elsevier B.V., pp. 1031–1051. ISSN: 00457825. DOI: 10.1016/j.cma.2008.11.007.

[214] R. Storn and K. Price. “Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces”. In: *Journal of global optimization* 11.4 (1997). Publisher: Springer, pp. 341–359.

[215] G. Sun, H. Zhang, R. Wang, X. Lv, and Q. Li. “Multiobjective reliability-based optimization for crashworthy structures coupled with metal forming process”. In: *Structural and multidisciplinary optimization* 56.6 (2017). Publisher: Springer, pp. 1571–1587.

[216] T. Sunaga. “Theory of an interval algebra and its application to numerical analysis”. In: *RAAG memoirs* 2 (1958), pp. 29–46.

[217] S. Surjanovic and D. Bingham. *Virtual Library of Simulation Experiments: Test Functions and Datasets*. Published: Retrieved August 19, 2021, from http://www.sfu.ca/ ssurjano.

[218] G. Taguchi. “Performance analysis design”. eng. In: *International journal of production research* 16.6 (1978). Publisher: Taylor & Francis Group, pp. 521–530. ISSN: 0020-7543.

[219] G. Taguchi. “Quality engineering (Taguchi methods) for the development of electronic circuit technology”. eng. In: *IEEE transactions on reliability* 44.2 (1995). Place: NEW YORK Publisher: IEEE, pp. 225–229. ISSN: 0018-9529.

[220] G. Taguchi, S. Chowdhury, and Y. Wu. *Taguchi’s quality engineering handbook*. John Wiley & Sons, 2004.

[221] K. F. Tee, A. U. Ebenuwa, and Y. Zhang. “Fuzzy-based robustness assessment of buried pipelines”. In: *Journal of Pipeline Systems Engineering and Practice* 9.1 (2018). Publisher: American Society of Civil Engineers, p. 06017007.

[222] R. Teixeira, M. Nogal, and A. O’Connor. “Adaptive approaches in metamodel-based reliability analysis: A review”. In: *Structural Safety* 89 (2021). Publisher: Elsevier, p. 102019.

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

[224] E. Vanmarcke. *Random fields: analysis and synthesis*. World Scientific, 2010.

[225] A. V. Vecchia. "Estimation and model identification for continuous spatial processes". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 50.2 (1988). Publisher: Wiley Online Library, pp. 297–312.

[226] J. Versace. *A review of the severity index*. Tech. rep. SAE Technical Paper, 1971.

[227] L. Wang, C. Xiong, X. Wang, M. Xu, and Y. Li. "A dimension-wise method and its improvement for multidisciplinary interval uncertainty analysis". In: *Applied Mathematical Modelling* 59 (2018). Publisher: Elsevier, pp. 680–695.

[228] Y. Wang, P. Lava, S. Coppieters, M. De Strycker, P. Van Houtte, and D. Debruyne. "Investigation of the uncertainty of DIC under heterogeneous strain states with numerical tests". In: *Strain* 48.6 (2012). Publisher: Wiley Online Library, pp. 453–462.

[229] Z. Wang, C. Jiang, B. Ni, C. Wang, J. Zhong, and T. Fang. "An interval finite element method for electromagnetic problems with spatially uncertain parameters". In: *Science China Technological Sciences* 63.1 (2020). Publisher: Springer, pp. 25–43.

[230] M. Warmus. "Calculus of approximations". In: *Bulletin de l'Academie Polonaise de Sciences* 4.5 (1956), pp. 253–257.

[231] T. M. Wasfy and A. K. Noor. "Multibody dynamic simulation of the next generation space telescope using finite elements and fuzzy sets". In: *Computer Methods in Applied Mechanics and Engineering* 190.5-7 (2000). Publisher: Elsevier, pp. 803–824.

[232] A. Wegner and G. Witt. "Correlation of process parameters and part properties in laser sintering using response surface modeling". In: *Physics Procedia* 39 (2012). Publisher: Elsevier, pp. 480–490.

[233] M. Wright. "The interior-point revolution in optimization: history, recent developments, and lasting consequences". In: *Bulletin of the American mathematical society* 42.1 (2005), pp. 39–56.

[234] D. Wu and W. Gao. "Hybrid uncertain static analysis with random and interval fields". In: *Computer Methods in Applied Mechanics and Engineering* 315 (2017). Publisher: Elsevier, pp. 222–246.

[235] D. Wu and W. Gao. “Uncertain static plane stress analysis with interval fields”. In: *International Journal for Numerical Methods in Engineering* 110.13 (2017). Publisher: Wiley Online Library, pp. 1272–1300.

[236] J. Wu, Z. Luo, Y. Zhang, N. Zhang, and L. Chen. “Interval uncertain method for multibody mechanical systems using Chebyshev inclusion functions”. In: *International Journal for Numerical Methods in Engineering* 95.7 (2013). Publisher: Wiley Online Library, pp. 608–630.

[237] Y. Xiang and Z. Shi. “Interval Analysis of Vibro-Acoustic Systems by the Enclosing Interval Finite-Element Method”. In: *Applied Sciences* 12.6 (2022). Publisher: MDPI, p. 3061.

[238] M. Xu and Z. Qiu. “A dimension-wise method for the static analysis of structures with interval parameters”. In: *SCIENCE CHINA Physics, Mechanics & Astronomy* 57.10 (2014). Publisher: Springer, pp. 1934–1945.

[239] L. Yu and B. Pan. “Full-frame, high-speed 3D shape and deformation measurements using stereo-digital image correlation and a single color high-speed camera”. In: *Optics and Lasers in Engineering* 95 (2017). Publisher: Elsevier, pp. 17–25.

[240] L. Yu and B. Pan. “Single-camera high-speed stereo-digital image correlation for full-field vibration measurement”. In: *Mechanical Systems and Signal Processing* 94 (2017). Publisher: Elsevier, pp. 374–383.

[241] L. Zadeh. “Calculus of fuzzy restrictions. Electronics Research Laboratory”. In: *University of California, Berkeley* (1975).

[242] L. A. Zadeh. “Fuzzy sets”. In: *Information and control* 8.3 (1965). Publisher: Elsevier, pp. 338–353.

[243] L. A. Zadeh. “The concept of a linguistic variable and its application to approximate reasoning—I”. In: *Information sciences* 8.3 (1975). Publisher: Elsevier, pp. 199–249.

[244] C. Zang, M. Friswell, and J. Mottershead. “A review of robust optimal design and its application in dynamics”. In: *Computers & structures* 83.4-5 (2005). Publisher: Elsevier, pp. 315–326.

[245] L. W. Zaseck, A. C. Bonifas, C. S. Miller, N. R. Orton, M. P. Reed, C. K. Demetropoulos, K. A. Ott, C. J. Dooley, N. P. Kuo, L. M. Strohsnitter, et al. “Kinematic and biomechanical response of post-mortem human subjects under various pre-impact postures to high-rate vertical loading conditions”. In: *Stapp car crash journal* 63 (2019). Publisher: The Stapp Association, pp. 235–266.

[246] M. Q. Zhang, M. Beer, C. G. Koh, and H. A. Jensen. “Nuanced Robustness Analysis with Limited Information”. In: *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* 2.3 (2016), B4015001.

[247] X. Zhao, G. Zhu, C. Zhou, and Q. Yu. “Crashworthiness analysis and design of composite tapered tubes under multiple load cases”. In: *Composite Structures* 222 (2019). Publisher: Elsevier, p. 110920.

[248] H.-J. Zimmermann. *Fuzzy set theory—and its applications*. Springer Science & Business Media, 2011.

[249] M. Zimmermann and J. E. von Hoessle. “Computing solution spaces for robust design”. In: *International Journal for Numerical Methods in Engineering* 94.3 (2013). Publisher: Wiley Online Library, pp. 290–307.

[250] M. Zimmermann, S. Königs, C. Niemeyer, J. Fender, C. Zeherbauer, R. Vitale, and M. Wahle. “On the design of large systems subject to uncertainty”. In: *Journal of Engineering Design* 28.4 (2017). Publisher: Taylor & Francis, pp. 233–254.

List of publications

Journal articles

Publications in international peer reviewed journals:

1. C. van Mierlo et al. “Inhomogeneous interval fields based on scaled inverse distance weighting interpolation”. In: *Computer Methods in Applied Mechanics and Engineering* 373 (2021). Publisher: Elsevier, p. 113542
2. C. van Mierlo et al. “Interval methods for lack-of-knowledge uncertainty in crash analysis”. In: *Mechanical Systems and Signal Processing* 168 (2022). Publisher: Elsevier, p. 108574
3. C. van Mierlo et al. “Robust design optimisation under lack-of-knowledge uncertainty”. In: *Computers & Structures* 275 (2023). Publisher: Elsevier, p. 106910
4. C. van Mierlo et al. “Robust design optimization of expensive stochastic simulators under lack-of-knowledge”. In: *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg* 0.0 (2023). Publisher: American Society of Mechanical Engineers Digital Collection, p. 0

Conference proceedings

Work that has been presented at an international conference and published within the conference proceedings:

1. C. van Mierlo et al. “Impact modeling of hyper-elastic am compliant mechanisms using high-speed digital image correlation”. In: *Exploring*

the Design Freedom of AdditiveManufacturing through Simulation, Date: 2018/12/10-2018/12/11, Location: Helsinki, Finland. 2018, pp. 72–72

2. C. van Mierlo et al. "Identification of visco-plastic material model parameters using interval fields". In: *Proceedings of the 3rd International Conference on Uncertainty Quantification in Computational sciences and Engineering.* ECCOMAS, 2019

3. C. van Mierlo et al. "Component-level impact performance assessment under spatially uncertain boundary conditions". In: *Proceedings of International Conference on Uncertainty in Structural Dynamics.* International Conference on Uncertainty in Structural Dynamics, 2020, pp. 1–15

4. C. van Mierlo et al. "Robustness of component performance under interval valued uncertainty". In: *Proceedings of the 4rd International Conference on Uncertainty Quantification in Computational Sciences and Engineering.* ECCOMAS, 2021

5. C. van Mierlo et al. "Interval field methods with local gradient control". In: *Proceedings of the European Congress on Computational Methods in Applied Sciences and Engineering.* event-place: Oslo, Norway. ECCOMAS

6. R. Barzanooni et al. "Evaluation of uncertain boundary conditions for analysis of structural components with respect to crashworthiness". In: *Online proceedings ISMA2022-USD2022.* event-place: Leuven, Belgium. 2022

Outreach activities

1. **STEM - workshop** Thin[k]shape: A workshop aimed at high-school students highlighting the importance and capabilities of topology optimisation, 3D-printing and validation brought together in an interactive manner. The workshop has been given multiple times a year for groups typically ranging from eight to twenty students.

2. **R2D network event** An event organized to inform and connect with industrial partners of the research group.

FACULTY OF ENGINEERING TECHNOLOGY
DEPARTMENT OF MECHANICAL ENGINEERING
ROBUST AND RELIABLE DESIGN
Jan Pieter de Nayerlaan 5, B-2860 Sint-Katelijne-Waver
B-3001 Leuven