Surrogate Modeling and Uncertainty Quantification for Radio Frequency and Optical Applications

Surrogat-Modellierung und Unsicherheitsquantifizierung für hochfrequente und optische Anwendungen

Zur Erlangung des akademischen Grades Doktor-Ingenieur (Dr.-Ing.)
genehmigte Dissertation von Niklas Georg, geboren in Siegen, Deutschland
Fachbereich Elektrotechnik und Informationstechnik, Technische Universität Darmstadt
Tag der Einreichung: 14. September 2021, Tag der Prüfung: 19. November 2021

1. Gutachten: Prof. Dr. Ulrich Römer
2. Gutachten: Prof. Dr. Sebastian Schöps
Darmstadt – D 17

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Fachbereich 18
Computational
Electromagnetics Group

Surrogate Modeling and Uncertainty Quantification for Radio Frequency and Optical Applications
Surrogat-Modellierung und Unsicherheitsquantifizierung für hochfrequente und optische Anwendungen

genehmigte Dissertation von Niklas Georg, geboren in Siegen, Deutschland
Fachbereich Elektrotechnik und Informationstechnik, Technische Universität Darmstadt

1. Gutachten: Prof. Dr. Ulrich Römer
2. Gutachten: Prof. Dr. Sebastian Schöps

Tag der Einreichung: 14. September 2021
Tag der Prüfung: 19. November 2021

Darmstadt – D 17

Bitte zitieren Sie dieses Dokument als:
URN: urn:nbn:de:tuda-tuprints-211498
URL: http://tuprints.ulb.tu-darmstadt.de/21149

Darmstadt, Technische Universität Darmstadt
Jahr der Veröffentlichung der Dissertation auf TUprints: 2022

Erklärungen laut Promotionsordnung

§8 Abs. 1 lit. c PromO

Ich versichere hiermit, dass die elektronische Version meiner Dissertation mit der schriftlichen Version übereinstimmt.

§8 Abs. 1 lit. d PromO

Ich versichere hiermit, dass zu einem vorherigen Zeitpunkt noch keine Promotion versucht wurde. In diesem Fall sind nähere Angaben über Zeitpunkt, Hochschule, Dissertationsthema und Ergebnis dieses Versuchs mitzuteilen.

§9 Abs. 1 PromO

Ich versichere hiermit, dass die vorliegende Dissertation selbstständig und nur unter Verwendung der angegebenen Quellen verfasst wurde.

§9 Abs. 2 PromO

Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Darmstadt, 14. September 2021

Niklas Georg

Zusammenfassung

Diese Dissertation behandelt verschiedene Methoden zur Surrogat-Modellierung und Unsicherheitsquantifizierung (UQ) für parametrisierte/stochastische Formulierungen der Maxwell-Gleichungen. Die Surrogat-Modellierung wird in diesem Zusammenhang verwendet, um den Berechnungsaufwand für die wiederholte Auswertung von numerischen Modellen mit unterschiedlichen Parameterkonfigurationen zu reduzieren. Zunächst wird eine rationale kernbasierte Interpolationsmethode zur effizienten Approximation von Frequenzantworten vorgestellt. Anschließend wird der Einfluss von unsicheren Geometrie- oder Materialparametern betrachtet, die beispielsweise Fertigungstoleranzen oder Messabweichungen unterliegen. Dazu werden verschiedene Techniken zur Konvergenzbeschleunigung von etablierten spektralen UQ-Verfahren, wie generalisiertem polynomialen Chaos oder stochastischer Kollokation, präsentiert. Insbesondere werden mithilfe von konformen Abbildungen transformierte Basisfunktionen konstruiert, welche die Holomorphie-Eigenschaften der zugrundeliegenden Funktionen besser ausnutzen. Um ein effizientes dimensionsadaptives Verfahren zu erhalten, wird eine adjungierte Darstellung des stochastischen Fehlers hergeleitet, welche ebenfalls zur Fehlerkorrektur verwendet wird.

Zusätzlich werden weitere Problemstellungen im Kontext von UQ für hochfrequente und optische Anwendungen behandelt. Es wird eine Prozedur zur effizienten und zuverlässigen Schätzung von Ausfallwahrscheinlichkeiten entwickelt, welche die Auswertungen von einem Surrogat-Modell sowie von Finite-Elemente-Modellen unterschiedlicher Genauigkeit mithilfe von adjungierter Fehlerschätzung kombiniert. Um die Anwendung von spektralen UQ-Techniken für Maxwells Eigenwertproblem zu ermöglichen, wird ein Homotopie-basiertes Eigenwert-Tracking-Verfahren vorgeschlagen, welches eine konsistente Zuordnung der Eigenmoden erreicht. Für quasi-periodische optische Strukturen mit einer endlichen Anzahl von Einheitszellen, die unabhängigen Unsicherheiten unterliegen, wird eine Methode zur entkoppelten Unsicherheitspropagation für individuelle Einheitszellen präsentiert.

Die Methoden werden für verschiedene Beispielprobleme, die sowohl akademische als auch realistische Modelle umfassen, numerisch untersucht und deren Effizienz dargelegt. Abschließend werden umfangreiche UQ- und Sensitivitätsstudien für die supraleitenden 9-Zell TESLA Kavitäten sowie verschiedene nanooptische Strukturen durchgeführt.

Abstract

This thesis addresses surrogate modeling and forward uncertainty propagation for parametric/stochastic versions of Maxwell's source and eigenproblem. Surrogate modeling is employed to reduce the computational complexity of sampling an underlying numerical solver. First, a rational kernel-based interpolation method is developed for the efficient approximation of frequency response functions. Next, the impact of uncertain shape and material parameters is considered, which originate, for instance, in manufacturing tolerances or measurement errors. To this end, several techniques for convergence acceleration of established spectral surrogate modeling techniques, as generalized polynomial chaos or stochastic collocation, are presented. In particular, transformed basis functions are constructed based on conformal maps that suitably transform the region of holomorphy. In addition, an adjoint representation of the stochastic error is employed for an efficient dimension-adaptive scheme as well as error correction.

Several challenges arising in uncertainty quantification for radio frequency and optical components are addressed. A multifidelity scheme for an efficient and reliable yield estimation is presented which comprises sampling of a surrogate model as well as finite element models of different fidelity based on adjoint error estimation. To enable the application of spectral surrogate modeling techniques for Maxwell's eigenproblem with uncertain input data, a homotopy-based eigenvalue tracking method is proposed to ensure a consistent matching of eigenmodes. Quasi-periodic structures of finite size, subject to independent shape uncertainties, are tackled using a decoupled uncertainty propagation procedure on the unit cell level.

The methods are numerically investigated using a number of benchmark problems that encompass academic and real-world models, and their efficiency is demonstrated. Finally, comprehensive uncertainty quantification and sensitivity studies are presented for the 9-cell TESLA cavities as well as different nano-optical structures.

Contents

List of figures

List of tables

1 Introduction

The chapter starts with the motivation for the research. Then, after giving a short overview of related work from the literature, we discuss the contributions of the thesis. An outline is given at the end of the chapter.

1.1 Motivation

The design process of many electromagnetic devices relies on numerical simulations solving an underlying mathematical model. These models describe a physical phenomenon of interest in order to predict the performance of certain designs. Thereby, the input data of the numerical model, such as the geometry of the components and the material properties, are usually assumed to be known exactly. However, in practice, the fabrication process leads to (possibly small) shape variations due to manufacturing tolerances. Furthermore, often also the material data are not known exactly because of measurements errors or limited and indirect observations. Neglecting these uncertainties in the input data might lead to discrepancies between the real phenomenon and the model prediction, which can eventually even cause the fabricated devices to fail the design specifications. In addition, neglecting uncertainties might lead to the selection of disproportionate safety factors, which is, for example, particularly problematic in nanotechnology due to the continuous demand for decreasing structure sizes. Hence, to achieve robust designs and ensure that the manufactured devices fulfill their specifications, this uncertainty should be taken into account in the simulation-based design workflow using uncertainty quantification (UQ) techniques. In Fig. 1.1, the idea of forward uncertainty propagation is illustrated, which aims to endow predictions with confidence intervals, reflecting the uncertainty. To this end, the input data of the numerical model as well as the output data are both described as random variables (RVs). The associated probability density functions (PDFs) of the inputs, describing the uncertainty of the shape or material parameters, are then propagated through the numerical model to the outputs, i.e. the quantities of interest (QoIs). Finally, statistics about the QoI are evaluated, e.g. expected value, standard deviation, or confidence intervals. Alternatively or complementary, the impact of uncertainty may also be quantified in terms of the yield, i.e. the proportion of realizations in a manufacturing process that fulfill predefined performance feature specifications [97].

Such forward UQ studies can be conducted using the well-known Monte Carlo (MC) method [103]. However, considering large-scale numerical models, the associated cost might become prohibitive, as often many (possibly millions of) model evaluations might be required to achieve a suitable accuracy. Spectral surrogate modeling techniques for UQ [93, 211] provide an efficient alternative to the MC method, which reduce the computational complexity of sampling the underlying numerical solver in many cases. However, the so-called curse-of-dimensionality [18] poses a major challenge, i.e. the computational effort grows rapidly with respect to the number of uncertain parameters. Hence, sophisticated methods are required to address computationally expensive simulation models with many uncertain parameters.

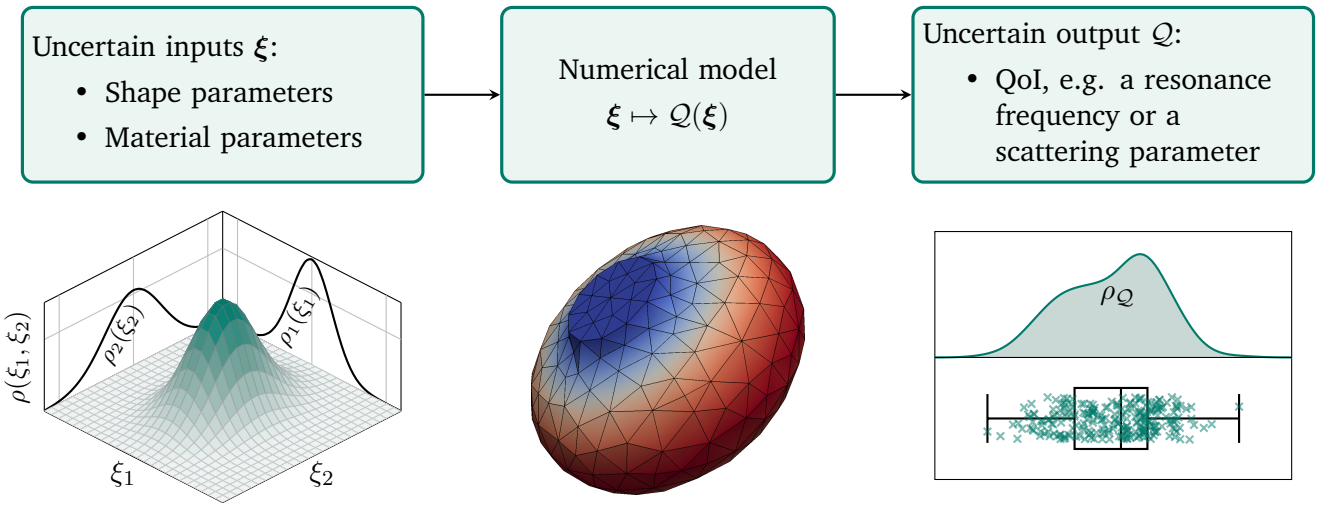


Figure 1.1: Illustration of forward uncertainty propagation: the uncertainty regarding the shape or material data of the component is propagated through the numerical model to obtain statistical information about the respective output distribution.

Surrogate modeling can also be employed in another related context to reduce the computational effort for sampling the underlying numerical solver. In particular, we also address the efficient approximation of frequency response functions (FRFs) based on a few training points, as the repeated evaluation of the FRF for different frequencies can be computationally demanding as well.

In this thesis, we focus on parametric/stochastic problems for radio frequency (RF) and optical applications. In particular, we consider particle accelerator cavities [60, 88] as an RF application example. Superconducting cavities can be used to accelerate particles and are installed in many accelerator facilities, e.g. the European X-ray Free Electron Laser (EXFEL) facility [4]. Due to their demanding performance requirements, e.g. specifications with a relative error margin below 10^{-4}, a dedicated treatment of the shape uncertainties arising in the manufacturing process is necessary [8, 60]. In addition, we consider optical gratings or metasurfaces [89, 167, 181] as application examples. Due to the progress in nanotechnological manufacturing, optical structures have tremendous future potential, such that the 21st century has already been called the "century of the photon" [42]. Nowadays photonic components can be designed by arranging tiny features which are even significantly smaller than the optical wavelength. These nano-scale optical devices are a promising technology, which allow to control light in an almost arbitrary manner, even by creating artificial components with a negative index of refraction, see e.g. [42, Chapter 5.4]. Today, nano-optic components have a wide range of applications [48, 51, 139]. UQ for such optical structures is highly relevant, as the nano-scale manufacturing process leads to relatively large manufacturing imperfections, see e.g. [42, 167]. In this regard, further challenges may arise as certain optical components have a very large *electrical* size, i.e. the physical dimensions relative to the wavelength, which can lead to large-scale simulation models. In addition, quasi-periodic optical structures can be subject to a very high number of uncertain parameters, which makes the application of standard surrogate modeling techniques for UQ as generalized polynomial chaos (GPC) [213] challenging due to the aforementioned curse-of-dimensionality.

1.2 Literature review

Simulation-based UQ is a rapidly evolving research topic in recent years, also in computational electromagnetics, see e.g. [55]. UQ comprises numerous methods for uncertainty propagation, Bayesian inverse problems, and optimal experimental design, among others, see [93, 126, 211] for a detailed background. Note that, although in this work we address forward uncertainty propagation, the surrogate modeling methods could be employed in an inverse problem setting as well. Different surrogate modeling techniques for UQ have already been developed and analyzed, in particular GPC and stochastic collocation (SC), i.e. higher-order spectral polynomial methods [11, 93, 140, 155, 210, 211]. They have been successfully applied in many different application areas and we refer to [1, 9, 50, 83, 204] regarding electromagnetic applications. In order to address complex models with a large number of uncertain parameters, which still represent a computational challenge, dimension-adaptive extensions [23, 53, 149] of the polynomials methods have been proposed to delay the curse-of-dimensionality. Alternative approaches, which are not in the scope of the present thesis, employ for instance model order reduction (MOR) [20, 77, 207], active subspaces [57] or low-rank tensor expansions [100, 136]. If only small uncertainties are considered, perturbation methods [69, 96, 185] can also provide a very efficient alternative. However, as perturbation methods only employ local measures and might not yield reliable results if the QoIs depend strongly on the uncertain parameters, they are not considered here. Finally, we note that the spectral UQ methods crucially depend on the smoothness of the underlying model, in particular on a smooth map from input to output parameters. For a number of different problem classes holomorphy results were shown in [52, 178]. If this assumption is not fulfilled, MC methods are still the preferred approach, in particular their multilevel and multifidelity extensions [94, 165].

1.3 Contribution

In this thesis, we propose different methods for parametric/stochastic problems in RF and optical applications. A major part is concerned with efficient surrogate modeling for parametric problems, in order to reduce the computational complexity of sampling a corresponding numerical solver. In this respect, we distinguish between the approximation of FRFs and spectral approximations w.r.t. the uncertain shape and material parameters. In the first case, we propose a rational kernel-based interpolation (RKI) method [86] for complex-valued FRFs as the frequency variable usually features a large parameter range and eventually limited smoothness. In the second case, we investigate techniques for convergence acceleration of spectral UQ methods, such as SC and GPC. In particular, to address high parametric sensitivities, we propose different schemes [89, 90] based on transformed basis functions, which are derived using conformal maps in order to suitably transform the region of holomorphy. To consider a large number of parameters, we suggest an adaptive collocation method [89] based on mapped Leja interpolation points [148] where the dimension-adaptivity is efficiently steered based on an adjoint representation of the stochastic error [43, 111]. Moreover, the adjoint-based error indicator is employed for error correction, in order to further enhance the convergence.

Next, we investigate an efficient and reliable yield estimation based on the adjoint-based surrogate approximation. Estimating probabilities based on sampling the surrogate model solely can lead to arbitrarily large errors, see [131] for details. Hence, we suggest a yield estimation scheme [81] in the spirit of multifidelity methods which takes all relevant error sources into account.

We also address a number of further challenges arising in UQ for RF and periodic optical applications: In general, global UQ techniques cannot be straightforwardly applied using standard eigenvalue solvers for Maxwell's eigenproblem with uncertain input data, e.g. to evaluate the uncertainty of eigenmodes in accelerator cavities subject to shape variabilities. In particular, eigenvalues crossing with respect to shape parameter changes can occur and, hence, a proper matching procedure of the respective eigenmodes for different deformed geometries is required. Note that a classification of eigenmodes in post-processing based on the field solution is cumbersome, see e.g. [34]. Hence, in this thesis, we suggest an eigenvalue tracking method [88] that can be employed for UQ based on homotopies between collocation points in the parameter space. Furthermore, we investigate UQ for different periodic optical structures where one has to distinguish between global and local uncertainties, i.e. either preserving or violating the periodicity of unit cells. In particular, for the case of a finite size quasi-periodic structure subject to independent deformations, we suggest a decoupled uncertainty propagation procedure [87]. It is based on a surrogate approximation on the unit cell level and a scattering matrix approach (SMA) [13, 95], which is embedded into a multifidelity Monte Carlo (MFMC) framework.

Numerical investigations for both academic examples and real-world models are carried out to verify the improved convergence or computational efficiency gains of the different methods. Finally, the methods are applied to conduct comprehensive UQ studies. This involves, in particular, a stochastic modeling of different uncertainty sources in the TeV-Energy Superconducting Linear Accelerator (TESLA) cavity manufacturing process as well as a proper treatment of the eventually present correlations of the uncertain shape parameters [60, 88]. Furthermore, we quantify the impact of material and fabrication uncertainties on different optical gratings [89, 181]. These investigations provide important insights on sensitivities and on the overall expected average performances for a large number of devices.

1.4 Structure of this treatise

This work is structured as follows: Chapter 2 introduces parametric versions of Maxwell's source and eigenproblem, i.e. the partial differential equations (PDEs) describing the physical phenomena considered in this thesis. Furthermore, we discuss their finite element (FE) approximation and define some key notions for UQ. In Chapter 3 we first recall some standard methods for the approximation of FRFs before presenting the suggested RKI method. In Chapter 4 the forward uncertainty propagation problem is addressed. In particular, we first review several existing UQ methods before suggesting the aforementioned improved techniques. In Chapter 5 the proposed surrogate modeling and UQ techniques are applied and investigated for a number of benchmark problems. Furthermore, comprehensive UQ studies for TESLA cavities as well as different periodic nano-optical structures are presented. Finally, in Chapter 6 the thesis is concluded and possible extensions are suggested for further research.

2 Fundamentals

The chapter starts with recalling Maxwell's eigenproblem as well as the corresponding source problems where periodic unit cell models and rectangular waveguides are addressed in particular. Suitable boundary conditions as well as their FE approximation will be discussed. Then parametric versions of the different models are considered and the basic setting for UQ is introduced. The content and structure of this chapter are based on our works [60, 81, 88, 89].

2.1 Maxwell's equations in frequency domain

We are interested in high-frequency electromagnetic wave phenomena, e.g. in the context of microwave or optical waveguides, resonant cavities, or scattering problems. All such macroscopic electromagnetic phenomena are described by Maxwell's equations [110, 141], which can, in the time-harmonic case, be expressed as

$$\nabla \times \mathbf{E} = -\mathrm{i}\omega\mu\mathbf{H} \qquad \text{in } D, \tag{2.1a}$$
$$\nabla \times \mathbf{H} = \mathbf{J}_\mathrm{s} + \mathrm{i}\omega\varepsilon\mathbf{E} \qquad \text{in } D, \tag{2.1b}$$
$$\nabla \cdot (\varepsilon\mathbf{E}) = \varrho \qquad \text{in } D, \tag{2.1c}$$
$$\nabla \cdot (\mu\mathbf{H}) = 0 \qquad \text{in } D, \tag{2.1d}$$

where $\mathbf{E}, \mathbf{H}, \mathbf{J}_\mathrm{s} : D \to \mathbb{C}^3$ denote the phasors for the electric field, the magnetic field and the source current, respectively. Furthermore $\omega \in \Omega \subset \mathbb{R}^+$ refers to the angular frequency, ϱ to the charge density phasor, ε to the complex-valued permittivity and μ to the magnetic permeability. Note that the complex vector functions $\mathbf{E}, \mathbf{H}, \mathbf{J}_\mathrm{s}$ depend on a spatial coordinate $\mathbf{r} = [x, y, z]^\top \in D \subset \mathbb{R}^3$ and are related to the respective time-domain quantities as

$$\mathbf{E}^\mathrm{time}(\mathbf{r}, t) = \Re[\mathbf{E}(\mathbf{r})e^{\mathrm{i}\omega t}], \qquad \mathbf{H}^\mathrm{time}(\mathbf{r}, t) = \Re[\mathbf{H}(\mathbf{r})e^{\mathrm{i}\omega t}], \qquad \mathbf{J}_\mathrm{s}^\mathrm{time}(\mathbf{r}, t) = \Re[\mathbf{J}_\mathrm{s}(\mathbf{r})e^{\mathrm{i}\omega t}], \tag{2.2}$$

see, e.g. [144, Section 1.2]. It can be seen that the time-harmonic form of Maxwell's equation (2.1) assumes all field quantities to oscillate with a single frequency ω. Hence, they can be employed to describe the steady state of a sinusoidally excited system but are also often employed for a Fourier analysis of electromagnetic phenomena in a certain frequency range Ω.

In this thesis, we assume the (computational) domain D to be bounded, simply connected and to have a Lipschitz continuous boundary ∂D. Furthermore, assuming absence of charges and source currents, i.e. $\varrho = 0$ and $\mathbf{J}_\mathrm{s} = 0$, one can derive the so-called curl-curl equation by eliminating $\mathbf{H}$ in (2.1)

$$\nabla \times \left(\mu_\mathrm{r}^{-1} \nabla \times \mathbf{E}\right) - \omega^2 \varepsilon \mu_0 \mathbf{E} = 0 \qquad \text{in } D, \tag{2.3}$$

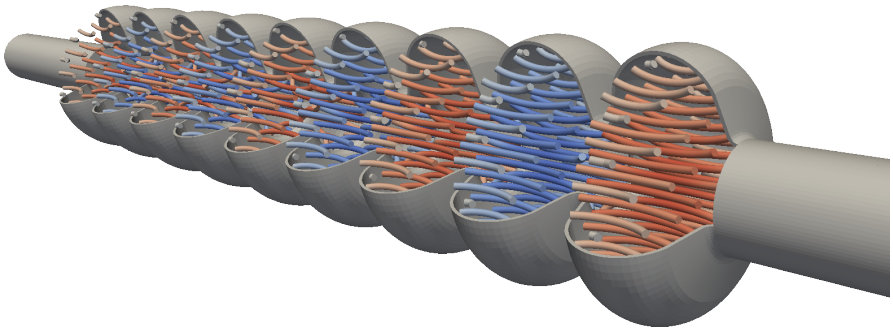


Figure 2.1: Electric field lines of the accelerating eigenmode in the 9-cell TESLA cavity [8], taken from [88, Figure 2a].

where μ_r and μ_0 represent the relative and vacuum permeability, respectively, s.t. $\mu = \mu_\mathrm{r}\mu_0$. We further assume a dispersive linear material behavior for ε and μ, i.e. the material parameters do not depend on the field solution $\mathbf{E}$ but may depend on the angular frequency ω.

The corresponding boundary conditions then distinguish the different types of boundary value problems (BVPs) which are mainly considered in this thesis and discussed in the following. In particular, Maxwell's eigenproblem in a cavity is recalled before the source problems for waveguide models as well as unit cells of periodic structures are addressed. To this end, for brevity of notation, the operators

$$\pi_\mathrm{t}[\mathbf{u}] \coloneqq \mathbf{e}_z \times \mathbf{u}, \tag{2.4a}$$
$$\pi_\mathrm{T}[\mathbf{u}] \coloneqq (\mathbf{e}_z \times \mathbf{u}) \times \mathbf{e}_z, \tag{2.4b}$$

along with the traces

$$\mathbf{u}_\mathrm{t} \coloneqq \mathbf{n}_{\partial D} \times \mathbf{u}|_{\partial D}, \tag{2.5a}$$
$$\mathbf{u}_\mathrm{T} \coloneqq (\mathbf{n}_{\partial D} \times \mathbf{u}|_{\partial D}) \times \mathbf{n}_{\partial D}, \tag{2.5b}$$

are introduced, where $\mathbf{e}_z$ refers to the unit vector in z-direction, ∂D to the boundary of D and $\mathbf{n}_{\partial D}$ to its outer unit normal.

2.1.1 Maxwell's eigenproblem: strong formulation

In this subsection, we address RF resonators in which at certain eigenfrequencies the electromagnetic field oscillates with a higher magnitude. Such devices are, for instance, employed to construct dedicated filter components or to amplify the magnitude of an excitation. In particular, in this thesis, we mainly focus on accelerator cavities which are specifically designed resonators for the acceleration of charged particles. One example is given by the 9-cell TESLA cavity [8] which is illustrated in Fig. 2.1 and will be described in more detail in Section 5.2. In this case, D in (2.3) refers to the inner domain of the cavity. Applying so-called perfect electric conducting (PEC) boundary conditions, one then obtains the eigenvalue formulation: find $\omega \in \mathbb{R}^+$ and $\mathbf{E} \neq 0$ such that

$$\nabla \times \left(\mu_\mathrm{r}^{-1} \nabla \times \mathbf{E}\right) = \omega^2 \mu_0 \varepsilon \mathbf{E} \qquad \text{in } D, \tag{2.6a}$$
$$\mathbf{E}_\mathrm{t} = 0 \qquad \text{on } \partial D. \tag{2.6b}$$

The PEC boundary conditions can be justified by the superconducting walls, however, different (more advanced) boundary conditions are also possible, e.g. a surface impedance boundary condition [144, Chapter 1]. Note that there is an infinite set of eigenmodes $\mathbf{E}_n$, $n = 0, 1, \ldots$ with corresponding eigenfrequencies s.t. $0 \leq \omega_0 = 2\pi f_0 \leq \omega_1 = 2\pi f_1 \leq \ldots < \infty$ which are the solutions of (2.6), cf. [144, Theorem 4.18].

In a TESLA cavity the particle beam is usually accelerated in longitudinal direction by a fundamental transverse magnetic (TM) eigenmode, i.e. one particular solution of (2.6) which is illustrated in Fig. 2.1. In a 9-cell TESLA cavity there are 9 eigenmodes of this type with eigenfrequencies $f_0 < f_1 < \ldots < f_8$ with different phase-shift in the electromagnetic fields between neighboring cells which vary from 0 to π radians. The accelerating mode is the so-called π-mode with frequency f_8 such that, given the chosen length of the cells [159], the field direction reverses as the particles traverse a cell and, hence, the particles are consistently accelerated in the same direction. For details on linear particle accelerators, we refer to [201].

Finally, we introduce two relevant figures of merit for the 9-cell TESLA cavity design which are based on the eigensolution of (2.6). The so-called cell-to-cell coupling coefficient

$$k_{\text{cc}} \coloneqq 2\frac{f_8 - f_0}{f_8 + f_0}, \tag{2.7}$$

quantifies the spread of the fundamental modes [19]. Note that a sufficiently large cell-to-cell coupling coefficient is desirable since there is a risk of exciting other modes than f_8 with the RF generator if their eigenfrequencies are too close. Another relevant quantity is the field flatness which is here defined as

$$\text{FF} \coloneqq \frac{\min_{i=1,\ldots,9} E^{(i)}_{\text{ax,max}}}{\max_{i=1,\ldots,9} E^{(i)}_{\text{ax,max}}}, \tag{2.8}$$

where $E^{(i)}_{\text{ax,max}}$ denotes the maximum magnitude of the longitudinal electric field component along the longitudinal axis in the i-th cell. A high field flatness, corresponding to a uniform field distribution, is of practical interest as it increases the accelerating voltage [160, p.129] and reduces the peak surface fields and, thus, the risk of field emissions.

2.1.2 Unit cell problem: strong formulation

Next, we address the scattering and transmission behavior of electromagnetic waves on periodic structures. This setting is here motivated by the optical gratings, which will be discussed in Chapter 5 where tiny sub-wavelength features are arranged in a periodic manner using modern nano-scale fabrication techniques in order to design innovative optical components [167, 181]. To this end, in this subsection, an infinitely periodic structure and a periodic excitation is assumed, such that Floquet's Theorem [115, Chapter 13] can be applied. It allows to confine the computational domain D to a single unit cell of the periodic structure. In particular, we assume periodicity in the x and y directions, whereas ∂D_{z^+} and ∂D_{z^-} denote the boundaries in the non-periodic direction. The corresponding unit cell is depicted in Fig. 2.2. Vacuum material properties ε_0, μ_0 are assumed for $z \geq z_{\max}$ and the structure is excited from the top, i.e. at ∂D_{z^+}, by an incident plane wave

$$\mathbf{E}^{\text{inc}} = \mathbf{E}_0 e^{-\mathrm{i}\mathbf{k}^{\text{inc}}\cdot\mathbf{r}}, \qquad \mathbf{k}^{\text{inc}} = \begin{bmatrix} k_x^{\text{inc}} \\ k_y^{\text{inc}} \\ k_z^{\text{inc}} \end{bmatrix} = -k_0 \begin{bmatrix} \sin\theta^{\text{inc}} \cos\phi^{\text{inc}} \\ \sin\theta^{\text{inc}} \sin\phi^{\text{inc}} \\ \cos\theta^{\text{inc}} \end{bmatrix}, \tag{2.9}$$

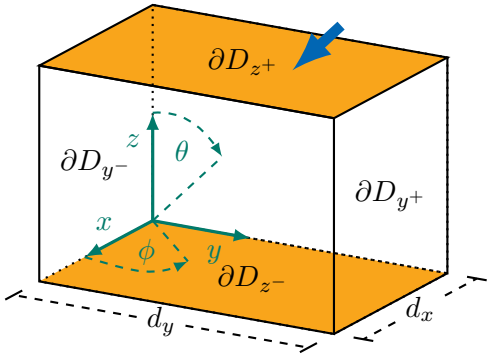


Figure 2.2: The computational domain D is represented by a rectangular unit cell. The incident wave vector is depicted by an blue arrow. The orange boundaries truncate the structure in the non-periodic directions. The illustration is based on [89, Fig. 8] and [90, Fig. 1b].

where $\theta^{\text{inc}},\ \phi^{\text{inc}}$ are the angles of incidence, $k_0 = \omega\sqrt{\mu_0\varepsilon_0}$ is the vacuum wave number and $\mathbf{E}_0 \in \mathbb{C}^3$. Depending on the values of $\theta^{\text{inc}},\ \phi^{\text{inc}}$, the excitation may have a different periodicity than the geometrical periodicity of the structure. Hence, periodic phase-shift boundary conditions need to be enforced on $\partial D_{x^+} \cup \partial D_{x^-}$ and on $\partial D_{y^+} \cup \partial D_{y^-}$, i.e.

$$\mathbf{E}|_{\partial D_{x^+}} = \mathbf{E}|_{\partial D_{x^-}} e^{-\mathrm{i}k_x^{\text{inc}} d_x}, \tag{2.10a}$$

$$\mathbf{E}|_{\partial D_{y^+}} = \mathbf{E}|_{\partial D_{y^-}} e^{-\mathrm{i}k_y^{\text{inc}} d_y}, \tag{2.10b}$$

where $d_x,\ d_y$ denote the dimensions of the unit cell.

2.1.2.1 Floquet absorbing boundary condition

Next, the truncation of the structure at ∂D_{z^+} is discussed. To this end, a Floquet absorbing boundary condition is derived by splitting the electric field in the unbounded region $z \geq z^+$ into the known incident field $\mathbf{E}^{\text{inc}}$ and the unknown scattered field $\mathbf{E}^{\text{sc}}$

$$\mathbf{E} = \mathbf{E}^{\text{inc}} + \mathbf{E}^{\text{sc}}. \tag{2.11}$$

The scattered field $\mathbf{E}^{\text{sc}}$ can be decomposed into an infinite series of Floquet modes [22, Chapter 3],[216, Chapter 12.2.1],

$$\mathbf{E}^{\text{sc}} = \sum_{\substack{m,n\in\mathbb{Z} \\ \alpha\in\{\text{TE,TM}\}}} c_{\alpha,mn} \mathbf{E}^{\alpha,mn} e^{-\mathrm{i}\kappa_{mn}(z-z^+)}, \tag{2.12}$$

where

$$\mathbf{E}^{\text{TE},mn} := \frac{e^{-\mathrm{i}(k_{xm}x+k_{yn}y)}\,(k_{yn}\mathbf{e}_x - k_{xm}\mathbf{e}_y)}{\sqrt{d_x d_y}\sqrt{k_{xm}^2+k_{yn}^2}}, \tag{2.13a}$$

$$\mathbf{E}^{\text{TM},mn} := \frac{e^{-\mathrm{i}(k_{xm}x+k_{yn}y)}\left(k_{xm}\mathbf{e}_x + k_{yn}\mathbf{e}_y - \frac{k_{xm}^2+k_{yn}^2}{\kappa_{mn}}\mathbf{e}_z\right)}{\sqrt{d_x d_y}\sqrt{k_{xm}^2+k_{yn}^2}}, \tag{2.13b}$$

with

$$k_{xm} \coloneqq k_x^{\mathrm{inc}} + \frac{2\pi m}{d_x}, \quad k_{yn} \coloneqq k_y^{\mathrm{inc}} + \frac{2\pi n}{d_y}, \quad \kappa_{mn} \coloneqq \sqrt{k_0^2 - k_{xm}^2 - k_{yn}^2}. \tag{2.14}$$

Note that the transverse electric (TE) modes $\mathbf{E}^{\mathrm{TE},mn}$ and the TM modes $\mathbf{E}^{\mathrm{TM},mn}$ fulfil $\mathbf{E} \perp \mathbf{e}_z$ and $\mathbf{H} \perp \mathbf{e}_z$, respectively. It can be seen that the finite number of propagating modes with $\kappa_{mn} \in \mathbb{R}$, depends only on a few parameters, i.e. the wavenumber k_0, the unit cell dimensions $d_x,\ d_y$ and the angles of incidence $\theta^{\mathrm{inc}},\ \phi^{\mathrm{inc}}$.

We then introduce the modal admittances

$$Y_{\alpha,mn} \coloneqq \begin{cases} \frac{\kappa_{mn}}{\omega\mu} & \text{for } \alpha = \mathrm{TE}, \\ \frac{\omega\varepsilon}{\kappa_{mn}} & \text{for } \alpha = \mathrm{TM}, \end{cases} \tag{2.15}$$

such that

$$\pi_{\mathrm{t}}\left[\mathbf{H}^{\alpha,mn} e^{-\mathrm{i}\kappa_{mn}(z-z^+)}\right] = \pi_{\mathrm{t}}\left[\frac{\mathrm{i}}{\omega\mu}\nabla \times \left(\mathbf{E}^{\alpha,mn} e^{-\mathrm{i}\kappa_{mn}(z-z^+)}\right)\right] = -Y_{\alpha,mn}\pi_{\mathrm{T}}\left[\mathbf{E}^{\alpha,mn} e^{-\mathrm{i}\kappa_{mn}(z-z^+)}\right]. \tag{2.16}$$

The incident plane wave $\mathbf{E}^{\mathrm{inc}}$, see (2.9), corresponds to the lowest order Floquet modes $\mathbf{E}^{\alpha,00}$ with modal admittance Y^{inc}

$$\pi_{\mathrm{t}}\left[\mathbf{H}^{\mathrm{inc}}\right] = Y^{\mathrm{inc}}\pi_{\mathrm{T}}\left[\mathbf{E}^{\mathrm{inc}}\right], \quad Y^{\mathrm{inc}} \coloneqq \begin{cases} \frac{\sqrt{\varepsilon}\cos(\theta^{\mathrm{inc}})}{\sqrt{\mu}} & \text{for } \alpha = \mathrm{TE}, \\ \frac{\sqrt{\varepsilon}}{\sqrt{\mu}\cos(\theta^{\mathrm{inc}})} & \text{for } \alpha = \mathrm{TM}. \end{cases} \tag{2.17}$$

The magnetic field for $z \geq z_{\max}$ is then obtained by applying the operator $\pi_{\mathrm{t}}[\frac{\mathrm{i}}{\omega\mu}\nabla \times (\cdot)]$ to (2.11) which leads to

$$\pi_{\mathrm{t}}\left[\mathbf{H}\right] + \sum_{\substack{m,n\in\mathbb{Z} \\ \alpha\in\{\mathrm{TE},\mathrm{TM}\}}} c_{\alpha,mn} Y_{\alpha,mn} \pi_{\mathrm{T}}[\mathbf{E}^{\alpha,mn} e^{-\mathrm{i}\kappa_{mn}(z-z^+)}] = Y^{\mathrm{inc}}\pi_{\mathrm{T}}\left[\mathbf{E}^{\mathrm{inc}}\right]. \tag{2.18}$$

Introducing the inner product notation

$$(\mathbf{u}, \mathbf{v})_{\partial D_{z^+}} \coloneqq \int_{\partial D_{z^+}} \mathbf{u} \cdot \overline{\mathbf{v}}\, \mathrm{d}\mathbf{x}, \tag{2.19}$$

where $\overline{\mathbf{v}}$ denotes the complex conjugate of $\mathbf{v}$, and employing the orthogonality of the modal basis,

$$\left(\mathbf{E}_{\mathrm{T}}^{\mathrm{TE},mn}, \mathbf{E}_{\mathrm{T}}^{\mathrm{TE},ij}\right)_{\partial D_{z^+}} = \delta_{mi}\delta_{nj}, \quad \left(\mathbf{E}_{\mathrm{T}}^{\mathrm{TM},mn}, \mathbf{E}_{\mathrm{T}}^{\mathrm{TM},ij}\right)_{\partial D_{z^+}} = \delta_{mi}\delta_{nj}, \quad \left(\mathbf{E}_{\mathrm{T}}^{\mathrm{TE},mn}, \mathbf{E}_{\mathrm{T}}^{\mathrm{TM},ij}\right)_{\partial D_{z^+}} = 0, \tag{2.20}$$

where δ denotes the Kronecker delta, the unknown coefficients $c_{\alpha,mn} \in \mathbb{C}$ are obtained as

$$c_{\alpha,mn} = \left(\mathbf{E}_{\mathrm{T}}^{\mathrm{sc}}, \mathbf{E}_{\mathrm{T}}^{\alpha,mn}\right)_{\partial D_{z^+}} = \left(\mathbf{E}_{\mathrm{T}}, \mathbf{E}_{\mathrm{T}}^{\alpha,mn}\right)_{\partial D_{z^+}} - \left(\mathbf{E}_{\mathrm{T}}^{\mathrm{inc}}, \mathbf{E}_{\mathrm{T}}^{\alpha,mn}\right)_{\partial D_{z^+}}. \tag{2.21}$$

Inserting (2.21) in (2.18) then yields the boundary condition on ∂D_{z^+}

$$\mathbf{H}_{\mathrm{t}} + \sum_{\substack{m,n\in\mathbb{Z} \\ \alpha\in\{\mathrm{TE},\mathrm{TM}\}}} \left(\mathbf{E}_{\mathrm{T}}, \mathbf{E}_{\mathrm{T}}^{\alpha,mn}\right)_{\partial D_{z^+}} Y_{\alpha,mn} \mathbf{E}_{\mathrm{T}}^{\alpha,mn} = 2Y^{\mathrm{inc}}\mathbf{E}_{\mathrm{T}}^{\mathrm{inc}}. \tag{2.22}$$

For practical computations the infinite sum is usually truncated to $-m^{\max} \leq m \leq m^{\max}$, $-n^{\max} \leq n \leq n^{\max}$. The boundary condition can be further simplified if d_x, d_y are small enough, such that only the fundamental modes $\mathbf{E}_{\alpha,00}$ propagate. In that case, one can place the boundary ∂D_{z^+} sufficiently far away from the structure, ensuring that all evanescent higher order modes (HOMs) have a negligible magnitude at ∂D_{z^+} and can hence be omitted. In this case, the first-order absorbing boundary condition [115, Chapter 13.4.1]

$$\mathbf{H}_\mathrm{t} - \frac{\mathbf{k}_\mathrm{t}^\mathrm{inc}}{\omega\mu_0 k_z^\mathrm{inc}} \left(\mathbf{k}_\mathrm{t}^\mathrm{inc} \cdot \mathbf{E}_\mathrm{T}\right) - \frac{k_z^\mathrm{inc}}{\omega\mu_0}\mathbf{E}_\mathrm{T} = 2Y^\mathrm{inc}\mathbf{E}_\mathrm{T}^\mathrm{inc} \qquad \text{on } \partial D_{z^+}, \tag{2.23}$$

can be employed which will be mainly considered in this work.

2.1.2.2 Boundary value problem

Finally, there are various choices for the boundary condition at ∂D_{z^-}, for instance, again a Floquet absorbing boundary condition or perfectly matched layer (PML) [115]. Here, a PEC boundary condition is applied, for simplicity. In summary, the BVP for the optical grating coupler, which will be introduced in Section 5.3.1 reads

$$\begin{aligned}
\nabla \times \left(\mu_\mathrm{r}^{-1}\nabla \times \mathbf{E}\right) - \omega^2\varepsilon\mu_0\mathbf{E} &= 0 && \text{in } D, && \text{(2.24a)}\\
\mathbf{E}|_{\partial D_{x^+}} e^{\mathrm{i}k_x^\mathrm{inc}d_x} &= \mathbf{E}|_{\partial D_{x^-}} && \text{on } \partial D_{x^+} \cup \partial D_{x^-}, && \text{(2.24b)}\\
\mathbf{E}|_{\partial D_{y^+}} e^{\mathrm{i}k_y^\mathrm{inc}d_y} &= \mathbf{E}|_{\partial D_{y^-}} && \text{on } \partial D_{y^+} \cup \partial D_{y^-}, && \text{(2.24c)}\\
\mathbf{E}_\mathrm{t} &= 0 && \text{on } \partial D_{z^-}, && \text{(2.24d)}\\
\mathbf{H}_\mathrm{t} - \frac{\mathbf{k}_\mathrm{t}^\mathrm{inc}}{\omega\mu_0 k_z^\mathrm{inc}} \left(\mathbf{k}_\mathrm{t}^\mathrm{inc} \cdot \mathbf{E}_\mathrm{T}\right) - \frac{k_z^\mathrm{inc}}{\omega\mu_0}\mathbf{E}_\mathrm{T} &= 2Y^\mathrm{inc}\mathbf{E}_\mathrm{T}^\mathrm{inc} && \text{on } \partial D_{z^+}. && \text{(2.24e)}
\end{aligned}$$

In addition to the electric field $\mathbf{E}$, one is, in practice, often also interested in scattering parameters (S-parameters), i.e. reflection and transmission coefficients. These scattering parameters at ∂D_{z^+} can here be defined as (affine-) linear functionals of $\mathbf{E}$, as

$$S^{\alpha,mn} := \left(\mathbf{E}_\mathrm{T} - \mathbf{E}_\mathrm{T}^\mathrm{inc}, \mathbf{E}_\mathrm{T}^{\alpha,mn}\right)_{\partial D_{z^+}} \in \mathbb{C}, \quad \text{where } \alpha \in \{\mathrm{TE}, \mathrm{TM}\}, m \in \mathbb{Z}, n \in \mathbb{Z}, \tag{2.25}$$

which project the scattered electric field on the Floquet modes $\mathbf{E}^{\alpha,mn}$ defined in (2.13).

2.1.3 Waveguide problem: strong formulation

In this subsection, we consider again a rectangular domain which now describes the model of an electric waveguide, as illustrated in Fig. 2.3. The boundary consists of three parts $\partial D = \overline{\partial D_\mathrm{PEC}} \cup \overline{\partial D_\mathrm{P1}} \cup \overline{\partial D_\mathrm{P2}}$ where $\partial D_\mathrm{P1}, \partial D_\mathrm{P2}$ refer to the two ports and PEC boundary conditions are employed at the waveguide walls ∂D_PEC. We assume that the waveguide is excited at ∂D_P1 by an incident TE_{10} wave

$$\mathbf{E}^{\mathrm{inc,wg}} = E_0\mathbf{E}_{10}^\mathrm{TE}e^{-\mathrm{i}k_{z10}z} \quad \text{with } \mathbf{E}_{10}^\mathrm{TE} := \sin\left(\frac{\pi x}{a}\right)\mathbf{e}_y, \tag{2.26}$$

where E_0 denotes the wave amplitude, $k_{z10} = \sqrt{\omega^2\mu\varepsilon_0 - \frac{\pi}{a^2}}$ the propagation constant and a the width of the waveguide. Note that the mode indices are now added as a subscript in order to distinguish waveguide and

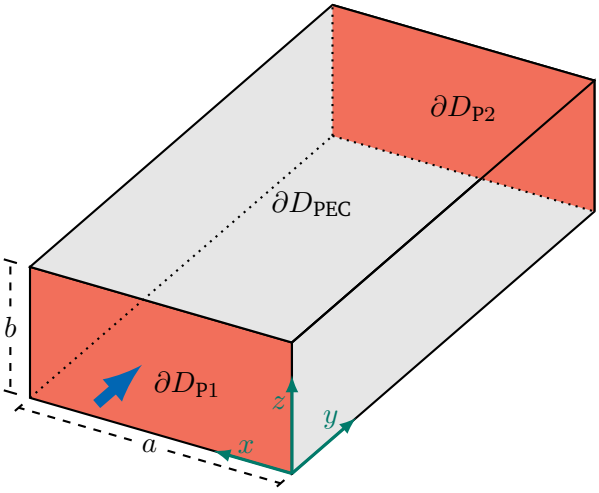


Figure 2.3: The computational domain D is represented by a rectangular waveguide. The incident wave vector is depicted by a blue arrow. The waveguide ports are shown in red while the waveguide walls are depicted in grey.

Floquet modes. Additionally, we assume vacuum material properties ε_0, μ_0 at the waveguide ports $\partial D_{\mathrm{P1}} \cup \partial D_{\mathrm{P2}}$ as well as suitable waveguide dimensions such that all modes, except for the TE_{10} mode, are evanescent. Then, further assuming that the ports are placed with sufficient distance from any possible obstacles inside the waveguide that could excite higher-order modes, lowest order waveguide boundary conditions can be derived. The derivation follows a similar procedure as for the Floquet boundary conditions derived in the last subsection and is hence omitted, see [115, Chapter 8.5] for details. In summary, we are concerned with the BVP

$$\begin{aligned}
\nabla \times \left(\mu_{\mathrm{r}}^{-1} \nabla \times \mathbf{E}\right) - \omega^2 \varepsilon \mu_0 \mathbf{E} &= 0 && \text{in } D, && (2.27a)\\
(\nabla \times \mathbf{E})_{\mathrm{t}} - \mathrm{i} k_{z10} \mathbf{E}_{\mathrm{T}} &= -2\mathrm{i} k_{z10} \mathbf{E}^{\mathrm{inc,wg}} && \text{on } \partial D_{\mathrm{P1}}, && (2.27b)\\
(\nabla \times \mathbf{E})_{\mathrm{t}} - \mathrm{i} k_{z10} \mathbf{E}_{\mathrm{T}} &= 0 && \text{on } \partial D_{\mathrm{P2}}, && (2.27c)\\
\mathbf{E}_{\mathrm{t}} &= 0 && \text{on } \partial D_{\mathrm{PEC}}. && (2.27d)
\end{aligned}$$

Scattering parameters can then be computed in post-processing, for instance, the fundamental scattering parameter of the TE_{10}-mode on ∂D_{P1} is given as

$$S_{10}^{\mathrm{TE}} := \frac{2}{E_0 a b} \left(\mathbf{E} - \mathbf{E}^{\mathrm{inc,wg}}, \mathbf{E}_{10}^{\mathrm{TE}}\right)_{\partial D_{\mathrm{P1}}}, \tag{2.28}$$

if one assumes (without loss of generality) that ∂D_{P1} is located at $z = 0$.

2.2 Weak formulations and discretization

In practice, Maxwell's equations are usually solved with a suitable numerical method. To this end, different methods from computational electromagnetics are available. For example, we mention the finite element method (FEM) [115, 144] in the frequency domain as well as the boundary element method (BEM) [37, 177]. Furthermore, particularly for the respective time-domain problems, discontinuous Galerkin methods [115,

182] as well as the finite difference method [124, 195] or the finite integration technique (FIT) [202] are well established. In this work, we mainly rely on the FEM, as it can easily handle inhomogeneous material distributions and complex geometries.

In order to solve the different BVPs (2.6), (2.24), (2.27) with the FEM, the corresponding weak formulations are derived in the following. To this end, we first recall some standard function spaces. Let $\left(L^2\left(D\right)\right)^3$ denote the complex vector function space of square integrable functions on D, i.e.

$$\left(L^2\left(D\right)\right)^3 \coloneqq \left\{\mathbf{u} : (\mathbf{u},\mathbf{u})_D < \infty\right\}, \tag{2.29}$$

where

$$(\mathbf{u},\mathbf{v})_D = \int_D \mathbf{u}\cdot\overline{\mathbf{v}}\,\mathrm{d}\mathbf{r}, \tag{2.30}$$

denotes the inner product. Furthermore, for a proper treatment of Maxwell's equations, we introduce the space

$$H\left(\mathrm{curl};D\right) \coloneqq \left\{\mathbf{u}\in\left(L^2\left(D\right)\right)^3 : (\nabla\times\mathbf{u},\nabla\times\mathbf{u})_D < \infty\right\}, \tag{2.31}$$

which defines the space of finite-energy solutions for the electric field, see [144] for details. Finally, we define $H_0\left(\mathrm{curl};D\right)$ as the subspace of $H\left(\mathrm{curl};D\right)$ with vanishing tangential components on the boundary ∂D, i.e.

$$H_0\left(\mathrm{curl};D\right) \coloneqq \left\{\mathbf{u}\in H\left(\mathrm{curl};D\right) : \mathbf{u}_\mathrm{t} = 0 \text{ on } \partial D\right\}. \tag{2.32}$$

For further details on the function spaces related to Maxwell's equations as well as their discrete counterparts, which can be obtained using discrete differential forms and an analysis of their properties, we refer to [6, 32, 105].

We then build the inner products of (2.3) with test function $\mathbf{E}'\in V$ and integrate by parts using [144, Theorem 3.31], which yields

$$\left(\mu_\mathrm{r}^{-1}\nabla\times\mathbf{E},\nabla\times\mathbf{E}'\right)_D - \omega^2\mu_0\left(\varepsilon\mathbf{E},\mathbf{E}'\right)_D + < \left(\mu_\mathrm{r}^{-1}\nabla\times\mathbf{E}\right)_\mathrm{t}, \mathbf{E}'_\mathrm{T} >_{\partial D} = 0, \tag{2.33}$$

where $<\cdot,\cdot>$ denotes the duality brackets and V is a subspace of $H\left(\mathrm{curl};D\right)$ such that the boundary term in (2.33) is well-defined. Next, the boundary term can be simplified, depending on the particular boundary conditions of the model problem.

2.2.1 Maxwell's eigenproblem

For Maxwell's eigenproblem, in view of the PEC boundary conditions (2.6b), we employ $H_0\left(\mathrm{curl};D\right)$ for test and ansatz functions which immediately yields the standard variational formulation of (2.6): find $\omega\in\mathbb{R}^+$ with $\mathbf{E}\in H_0\left(\mathrm{curl};D\right)$ such that

$$\left(\mu_\mathrm{r}^{-1}\nabla\times\mathbf{E},\nabla\times\mathbf{v}\right) = \omega^2\mu_0\left(\varepsilon\mathbf{E},\mathbf{v}\right), \quad \forall\mathbf{v}\in H_0\left(\mathrm{curl};D\right). \tag{2.34}$$

Note that the boundary integral in (2.33) vanishes due to the PEC boundary conditions which are enforced for the test functions $\mathbf{v}$ as well. Next, we devise the discrete counterpart of (2.34) by following a Galerkin procedure. To this end, we employ a sequence of finite-dimensional spaces $V_h\subset H_0\left(\mathrm{curl};D\right)$ and correspondingly restrict (2.34): find $\omega_h\in\mathbb{R}$ with $\mathbf{E}_h\in V_h$ such that

$$\left(\mu_\mathrm{r}^{-1}\nabla\times\mathbf{E}_h,\nabla\times\mathbf{v}\right) = \omega_h^2\mu_0\left(\varepsilon\mathbf{E}_h,\mathbf{v}\right) \quad \forall\mathbf{v}\in V_h. \tag{2.35}$$

Note that there are different choices for the space V_h which yield a proper discretization of $H\,(\mathrm{curl}; D)$, see, e.g. [24, 144, 150] for Nédélec edge elements or [38] for an isogeometric approach, however, in any case $\mathbf{E}_h$ can then be represented in terms of basis functions $(\mathbf{w}_j)_{j=1}^{N_h} \subset V_h$, where $\mathbf{w}_j : D \to \mathbb{R}^3$, as

$$\mathbf{E}_h(\mathbf{r}) = \sum_{j=1}^{N_h} e_j \mathbf{w}_j(\mathbf{r}). \tag{2.36}$$

The discrete solution $\mathbf{e} \in \mathbb{C}^{N_h}$ can then be computed by solving the generalized eigenvalue problem

$$\mathbf{K}\mathbf{e} = \omega_h^2 \mathbf{M}\mathbf{e}, \tag{2.37}$$

where the stiffness matrix $\mathbf{K}$ and the mass matrix $\mathbf{M}$ for vacuum are given by

$$\mathbf{K}_{ij} = \left(\mu_\mathrm{r}^{-1} \nabla \times \mathbf{w}_j, \nabla \times \mathbf{w}_i\right)_D, \quad \mathbf{M}_{ij} = \mu_0 \left(\varepsilon \mathbf{w}_j, \mathbf{w}_i\right)_D. \tag{2.38}$$

2.2.2 Unit cell problem

For the unit cell model (2.24), we can simplify the boundary term in (2.33) because the contributions on ∂D_{x^+}, ∂D_{y^+} cancel the contributions on ∂D_{x^-}, ∂D_{y^-}, respectively, due to the periodic phase-shift boundary conditions (2.24b), (2.24c) which are imposed for trial and test functions. Additionally, the part of the integral on ∂D_{z^-} can be eliminated by demanding that the PEC boundary condition (2.24d) is fulfilled by the test functions $\mathbf{E}'$ as well. The appropriate function space V^uc, where uc is short for unit cell, then reads

$$\begin{aligned} V^\mathrm{uc} := \{\mathbf{u} \in H\,(\mathrm{curl}; D) :\ & \mathbf{u}_\mathrm{T}|_{\partial D_{z^-}} = 0 \wedge\ \mathbf{u}_\mathrm{T}|_{\partial D_{x^+}} e^{\mathrm{i}k_x^\mathrm{inc} d_x} = -\mathbf{u}_\mathrm{T}|_{\partial D_{x^-}} \\ & \wedge \mathbf{u}_\mathrm{T}|_{\partial D_{y^+}} e^{\mathrm{i}k_y^\mathrm{inc} d_y} = -\mathbf{u}_\mathrm{T}|_{\partial D_{y^-}} \wedge\ \mathbf{u}_\mathrm{T}|_{\partial D_{z^+}} \in \left(L^2(\partial D_{z^+})\right)^3\}, \end{aligned} \tag{2.39}$$

where the condition $\mathbf{u}_\mathrm{T}|_{\partial D_{z^+}} \in \left(L^2(\partial D_{z^+})\right)^3$ ensures a well-defined boundary integral in (2.33). Inserting the Floquet absorbing boundary condition (2.24e) then leads to the weak formulation: find $\mathbf{E} \in V^\mathrm{uc}$ s.t.

$$\begin{aligned} \left(\mu_\mathrm{r}^{-1} \nabla \times \mathbf{E}, \nabla \times \mathbf{E}'\right)_D - \omega^2 \mu_0 \left(\varepsilon \mathbf{E}, \mathbf{E}'\right)_D - \frac{\mathrm{i}}{k_z^\mathrm{inc}} \left(\mathbf{k}_\mathrm{t}^\mathrm{inc} \cdot \mathbf{E}_\mathrm{T}, \mathbf{k}_\mathrm{t}^\mathrm{inc} \cdot \mathbf{E}'_\mathrm{T}\right)_{\partial D_{z^+}} - \mathrm{i} k_z^\mathrm{inc} \left(\mathbf{E}_\mathrm{T}, \mathbf{E}'_\mathrm{T}\right)_{\partial D_{z^+}} \\ = 2\mathrm{i}\omega\mu_0 Y^\mathrm{inc} \left(\mathbf{E}_\mathrm{T}^\mathrm{inc}, \mathbf{E}'_\mathrm{T}\right)_{\partial D_{z^+}} \quad \forall \mathbf{E}' \in V^\mathrm{uc}. \end{aligned} \tag{2.40}$$

Eq. (2.40) is then discretized, following a similar procedure as outlined in the last subsection, i.e. using a finite-dimensional subspace of V^uc and approximating the electric field in the form of (2.36). In this case the curl-conforming basis functions $\mathbf{w}_j$ are chosen based on first kind, second order Nédélec basis functions [144, 150] defined on a tetrahedral mesh of D with a periodic surface mesh on ∂D_{x^+} and ∂D_{x^-} as well as on ∂D_{y^+} and ∂D_{y^-}. The phase-shift boundary conditions are then incorporated as explained in [115, Chapter 13], [89, Appendix B], which finally leads to a discrete system in the form of

$$\underbrace{\left(\mathbf{K} - \omega^2 \mathbf{M} + \mathrm{i}\mathbf{D}^\mathrm{uc}\right)}_{\mathbf{A}^\mathrm{uc}} \mathbf{e} = \mathbf{f}^\mathrm{uc}, \tag{2.41}$$

where $\mathbf{A}^{\mathrm{uc}} \in \mathbb{C}^{N_h \times N_h}$ denotes the system matrix and $\mathbf{f}^{\mathrm{uc}} \in \mathbb{C}^{N_h}$ the right-hand side (RHS). Note that $\mathbf{K}$ and $\mathbf{M}$ are defined in (2.38) and

$$\mathbf{D}_{ij}^{\mathrm{uc}} = -\frac{1}{k_z^{\mathrm{inc}}} \left(\mathbf{k}_{\mathrm{t}}^{\mathrm{inc}} \cdot (\mathbf{w}_j)_{\mathrm{T}}, \mathbf{k}_{\mathrm{t}}^{\mathrm{inc}} \cdot (\mathbf{w}_i)_{\mathrm{T}}\right)_{\partial D_{z^+}} - k_z^{\mathrm{inc}} \left((\mathbf{w}_j)_{\mathrm{T}}, (\mathbf{w}_i)_{\mathrm{T}}\right)_{\partial D_{z^+}}, \tag{2.42}$$

$$\mathbf{f}_i^{\mathrm{uc}} = 2\mathrm{i}\omega\mu_0 Y^{\mathrm{inc}} \left(\mathbf{E}_{\mathrm{T}}^{\mathrm{inc}}, (\mathbf{w}_i)_{\mathrm{T}}\right)_{\partial D_{z^+}}. \tag{2.43}$$

The discrete counterpart of (2.25) is given as

$$S_h^{\alpha,mn} = \mathbf{j}^{\alpha,mn} \cdot (\mathbf{e} - \mathbf{e}^{\mathrm{inc,uc}}) \in \mathbb{C}, \quad \text{where } \alpha \in \{\mathrm{TE}, \mathrm{TM}\}, m \in \mathbb{Z}, n \in \mathbb{Z}, \tag{2.44}$$

and $\mathbf{e}^{\mathrm{inc,uc}} \in \mathbb{C}^{N_h}$ denotes the vector of coefficients obtained after projecting the incident wave (2.9) on the basis (2.36) and the vector $\mathbf{q}^{\alpha,mn}$ is given as

$$q_i^{\alpha,mn} = \left((\mathbf{w}_i)_{\mathrm{T}}, \mathbf{E}_{\mathrm{T}}^{\alpha,mn}\right)_{\partial D_{z^+}} \in \mathbb{C}^{N_h}. \tag{2.45}$$

2.2.3 Waveguide problem

For the waveguide problem (2.27), we start again from (2.33), eliminate the portion of the boundary integral on $\partial D_{\mathrm{PEC}}$, due to the PEC boundary conditions (2.27d) and insert the waveguide boundary conditions (2.27b),(2.27c), which yields the variational formulation: find $\mathbf{E} \in V^{\mathrm{wg}}$ s.t.

$$\begin{aligned} \left(\mu_{\mathrm{r}}{}^{-1} \nabla \times \mathbf{E}, \nabla \times \mathbf{E}'\right)_D - \omega^2 \mu_0 \left(\varepsilon\, \mathbf{E}, \mathbf{E}'\right)_D + \mathrm{i} k_{z10} \left(\mathbf{E}_{\mathrm{T}}, \mathbf{E}_{\mathrm{T}}\right)_{\partial D_{\mathrm{P1}} \cup \partial D_{\mathrm{P2}}} \\ = 2\mathrm{i} k_{z10} \left(\mathbf{E}^{\mathrm{inc,wg}}, \mathbf{E}'_{\mathrm{T}}\right)_{\partial D_{\mathrm{P2}}} \qquad \forall \mathbf{E}' \in V^{\mathrm{wg}}, \end{aligned} \tag{2.46}$$

where

$$V^{\mathrm{wg}} := \left\{ \mathbf{u} \in H(\mathrm{curl}; D) : \ \mathbf{u}_{\mathrm{T}}\big|_{\partial D_{\mathrm{P1}}} \in \left(L^2(\partial D_{\mathrm{P1}})\right)^3 \ \wedge \ \mathbf{u}_{\mathrm{T}}\big|_{\partial D_{\mathrm{P2}}} \in \left(L^2(\partial D_{\mathrm{P2}})\right)^3 \ \wedge \ \mathbf{u}_{\mathrm{T}}\big|_{\partial D_{\mathrm{PEC}}} = 0 \right\}. \tag{2.47}$$

After discretization using, again, Nédélec basis functions on a tetrahedral mesh, one can then compute the discrete solution $\mathbf{e}$ by solving the linear system

$$\underbrace{\left(\mathbf{K} - \omega^2 \mathbf{M} + \mathrm{i}\mathbf{D}^{\mathrm{wg}}\right)}_{\mathbf{A}^{\mathrm{wg}}(\omega)} \mathbf{e} = \mathbf{f}^{\mathrm{wg}}, \tag{2.48}$$

where $\mathbf{K}$ and $\mathbf{M}$ are defined in (2.38) and

$$\mathbf{D}_{ij}^{\mathrm{wg}} = k_{z10} \left((\mathbf{w}_j)_{\mathrm{T}}, (\mathbf{w}_i)_{\mathrm{T}}\right)_{\partial D_{\mathrm{P1}} \cup \partial D_{\mathrm{P2}}}, \qquad \mathbf{f}_i^{\mathrm{wg}} = 2\mathrm{i} k_{z10} \left(\mathbf{E}^{\mathrm{inc,wg}}, (\mathbf{w}_i)_{\mathrm{T}}\right)_{\partial D_{\mathrm{P1}}}. \tag{2.49}$$

The discrete counterpart of (2.28) is given as

$$S_{10,h}^{\mathrm{TE}} = \mathbf{j}_{10}^{\mathrm{TE}} \cdot (\mathbf{e} - \mathbf{e}^{\mathrm{inc,wg}}) \in \mathbb{C}, \tag{2.50}$$

where $\mathbf{e}^{\mathrm{inc,wg}}$ is obtained by projecting (2.26) on the basis (2.36) and

$$(\mathbf{j}_{10}^{\mathrm{TE}})_i = \frac{2}{E_0 ab} \left(\mathbf{w}_i, \mathbf{E}_{10}^{\mathrm{TE}}\right)_{\partial D_{\mathrm{P1}}}. \tag{2.51}$$

2.3 Parametric problems

The previous sections of this chapter discussed the numerical solution of the different problem types for particular choices of the computational domain D and the material distributions. Note that the source problems lead to parametric systems (2.41) and (2.48) with respect to the angular frequency ω. These will be discussed in Chapter 3 in more detail. However, in Chapter 4 we will then address another type of parametric problem. In particular, we will consider parametric variations which change the shape or material parameters of the considered model. These parameters can then be used to model uncertainties, e.g. due to manufacturing tolerances or measurement uncertainty. Note that many RF components are sensitive to their shape parameters, e.g. the final accelerator cavity shape determines, to a large extent, their performance. It has also been reported in the context of nano-optical structures that significant variations with respect to certain shape parameters can be observed, see for instance [167] where the considered grating coupler is particularly sensitive to the grating depth. In addition, material uncertainties can have a significant impact, for example it has been pointed out in [114] that the measurement errors in the available data sets, for the optical properties of silver, significantly affect the accuracy of simulation predictions. We note that uncertainties may also arise in the boundary conditions or in the excitation of the source problems, however, these are not considered in this thesis. In the following, we will introduce the respective parameterized versions of the different problems.

2.3.1 Parameterized eigenvalue problem

For the eigenproblem (2.6), we will consider a parametric domain $D(\boldsymbol{\xi})$ with a parametric boundary $\partial D(\boldsymbol{\xi})$, s.t. the computational domain D depends smoothly on a parameter vector $\boldsymbol{\xi} \in \Gamma \subset \mathbb{R}^{N_{\boldsymbol{\xi}}}$ which can be used to model deformations in the cavity geometry, for instance. Note that we do not introduce parametric variations for the material parameter ε, μ, as in this work vacuum material properties will be assumed for all cavity models. Following the discretization steps described in the previous sections, taking into account the parameterized domain and employing parameterized basis functions

$$\{\mathbf{w}_j(\mathbf{r};\boldsymbol{\xi})\}_{j=1}^{N_h} \subset H\left(\text{curl}; D(\boldsymbol{\xi})\right), \text{ where } \mathbf{w}_j(\mathbf{r};\boldsymbol{\xi}) : D(\boldsymbol{\xi}) \times \Gamma \to \mathbb{R}^3, \tag{2.52}$$

which will be discussed later, one then obtains the parameterized generalized eigenvalue problem

$$\mathbf{K}(\boldsymbol{\xi})\mathbf{e}(\boldsymbol{\xi}) = \omega_h^2(\boldsymbol{\xi})\mathbf{M}(\boldsymbol{\xi})\mathbf{e}(\boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \in \Gamma. \tag{2.53}$$

2.3.2 Parameterized source problems

For the considered source problems, i.e. the unit cell problem and waveguide problem, a fixed computational domain D but parametric material properties are considered. In particular, it is assumed that the domain D can be decomposed in N_{sub} non-overlapping subdomains D_i such that $\overline{D} = \bigcup_{i=1}^{N_{\text{sub}}} \overline{D_i}$ and that

$$\varepsilon(\mathbf{r},\omega,\boldsymbol{\xi}) = \sum_{i=1}^{N_{\text{sub}}} \varepsilon_i(\omega,\boldsymbol{\xi})\mathbb{1}_i(\mathbf{r},\boldsymbol{\xi}), \quad \mu(\mathbf{r},\omega,\boldsymbol{\xi}) = \sum_{i=1}^{N_{\text{sub}}} \mu_i(\omega,\boldsymbol{\xi})\mathbb{1}_i(\mathbf{r},\boldsymbol{\xi}), \tag{2.54}$$

where

$$\mathbb{1}_i(\mathbf{r},\boldsymbol{\xi}) = \begin{cases} 1\ , & \mathbf{r} \in D_i(\boldsymbol{\xi}), \\ 0\ , & \mathbf{r} \notin D_i(\boldsymbol{\xi}), \end{cases} \tag{2.55}$$

i.e. that the permittivity ε and the permeability μ are spatially piecewise constant on each subdomain D_i and depend on the parameter vector $\boldsymbol{\xi} \in \Gamma \subset \mathbb{R}^{N_{\boldsymbol{\xi}}}$. It is further assumed that the parametric dependencies $\boldsymbol{\xi} \mapsto \varepsilon(\cdot,\cdot,\boldsymbol{\xi}),\ \boldsymbol{\xi} \mapsto \mu(\cdot,\cdot,\boldsymbol{\xi})$ are smooth. We emphasize that the parameter vector $\boldsymbol{\xi}$ can be used to change the material parameter coefficients $\varepsilon_i(\omega,\boldsymbol{\xi}), \mu_i(\omega,\boldsymbol{\xi})$, modeling different permittivities or permeabilities, but also allows to represent geometric variations of the material interfaces inside the unit cell, as the subdomains D_i depend on $\boldsymbol{\xi}$ as well. Inserting (2.54) in (2.40) and (2.46), respectively, one then obtains parameterized weak formulations in the form: find $\mathbf{E}(\boldsymbol{\xi}) \in V$ s.t.

$$a_{\boldsymbol{\xi}}(\mathbf{E}(\boldsymbol{\xi}),\mathbf{E}') = l(\mathbf{E}'),\ \forall \mathbf{E}' \in V. \tag{2.56}$$

In particular, in case of the unit cell problem (2.24), the function space V is given as V^{uc} and the sesquilinear form $a_{\boldsymbol{\xi}}(\cdot,\cdot)$ and the antilinear form $l(\cdot)$ read

$$\begin{aligned} a^{\mathrm{uc}}_{\boldsymbol{\xi}}(\mathbf{E}(\boldsymbol{\xi}),\mathbf{E}') :=& \left(\mu_{\mathrm{r}}^{-1}(\boldsymbol{\xi})\nabla\times\mathbf{E}(\boldsymbol{\xi}),\nabla\times\mathbf{E}'\right)_D - \omega^2\mu_0\left(\varepsilon(\boldsymbol{\xi})\mathbf{E}(\boldsymbol{\xi}),\mathbf{E}'\right)_D \\ &- \frac{\mathrm{i}}{k_z^{\mathrm{inc}}}\left(\mathbf{k}_{\mathrm{t}}^{\mathrm{inc}}\cdot\mathbf{E}_{\mathrm{T}}(\boldsymbol{\xi}),\mathbf{k}_{\mathrm{t}}^{\mathrm{inc}}\cdot\mathbf{E}'_{\mathrm{T}}\right)_{\partial D_{z^+}} - \mathrm{i}k_z^{\mathrm{inc}}\left(\mathbf{E}_{\mathrm{T}}(\boldsymbol{\xi}),\mathbf{E}'_{\mathrm{T}}\right)_{\partial D_{z^+}}, \end{aligned} \tag{2.57}$$

$$l^{\mathrm{uc}}(\mathbf{E}') := 2\mathrm{i}\omega\mu_0 Y^{\mathrm{inc}}\left(\mathbf{E}_{\mathrm{T}}^{\mathrm{inc}},\mathbf{E}'_{\mathrm{T}}\right)_{\partial D_{z^+}}. \tag{2.58}$$

In case of the waveguide problem (2.27), the function space V is given as V^{wg} and

$$a^{\mathrm{wg}}_{\boldsymbol{\xi}}(\mathbf{E}(\boldsymbol{\xi}),\mathbf{E}') := \left(\mu_{\mathrm{r}}^{-1}(\boldsymbol{\xi})\nabla\times\mathbf{E}(\boldsymbol{\xi}),\nabla\times\mathbf{E}'\right)_D - \omega^2\mu_0\left(\varepsilon(\boldsymbol{\xi})\mathbf{E}(\boldsymbol{\xi}),\mathbf{E}'\right)_D + \mathrm{i}k_{z10}\left(\mathbf{E}_{\mathrm{T}}(\boldsymbol{\xi}),\mathbf{E}_{\mathrm{T}}\right)_{\partial D_{\mathrm{P1}}\cup\partial D_{\mathrm{P2}}}, \tag{2.59}$$

$$l^{\mathrm{wg}}(\mathbf{E}') := 2\mathrm{i}k_{z10}\left(\mathbf{E}^{\mathrm{inc,wg}},(\mathbf{w}_i)_{\mathrm{T}}\right)_{\partial D_{\mathrm{P1}}}. \tag{2.60}$$

Finally, the parameterized scattering parameters of the unit cell problem and the waveguide problem read

$$S^{\alpha,mn}(\boldsymbol{\xi}) = \left(\mathbf{E}_{\mathrm{T}}(\boldsymbol{\xi}) - \mathbf{E}_{\mathrm{T}}^{\mathrm{inc}},\mathbf{E}_{\mathrm{T}}^{\alpha,mn}\right)_{\partial D_{z^+}}, \quad \text{where } \alpha \in \{\mathrm{TE},\mathrm{TM}\}, m \in \mathbb{Z}, n \in \mathbb{Z}, \tag{2.61}$$

$$S_{10}^{\mathrm{TE}}(\boldsymbol{\xi}) = \frac{2}{E_0 ab}\left(\mathbf{E}(\boldsymbol{\xi}) - \mathbf{E}^{\mathrm{inc,wg}},\mathbf{E}_{10}^{\mathrm{TE}}\right)_{\partial D_{\mathrm{P1}}}, \tag{2.62}$$

respectively. After discretization, one obtains in both cases linear systems in the form

$$\mathbf{A}(\boldsymbol{\xi})\,\mathbf{e}(\boldsymbol{\xi}) = \mathbf{f}, \tag{2.63a}$$

$$S_h(\boldsymbol{\xi}) = \mathbf{j}\cdot\left(\mathbf{e}(\boldsymbol{\xi}) - \mathbf{e}^{\mathrm{inc}}\right). \tag{2.63b}$$

2.3.3 Uncertainty propagation and quantification

In the context of this thesis, we are not only interested in the electric field solution $\mathbf{e}(\boldsymbol{\xi})$ but particularly interested in specific scalar QoIs $\mathcal{Q}(\boldsymbol{\xi}) \in \mathbb{C}$ which might refer to eigenvalues of Maxwell's eigenproblem or a bounded functional of the electric field $\mathcal{Q}(\boldsymbol{\xi}) = J\big(\mathbf{e}(\boldsymbol{\xi})\big)$, as the scattering parameters of the unit cell or waveguide problem, for instance. We emphasize that the considered models, i.e. the respective maps $\boldsymbol{\xi} \mapsto \mathcal{Q}(\boldsymbol{\xi})$, are deterministic which means that the same input parameter values $\boldsymbol{\xi}$ will always lead to the exact same output $\mathcal{Q}(\boldsymbol{\xi})$. However, these QoIs become stochastic, if we now assume that the input parameters of the parameterized models discussed in previous subsection are given as real RVs Ξ_n, $n = 1,2,\ldots,N_{\boldsymbol{\xi}}$ in order to

model uncertainties due to manufacturing imperfections. Note that the type of uncertainties that originate in the fabrication process, e.g. regarding the final shape of the device, and are assumed to be irreducible or inherently random, is usually referred to as aleatory uncertainty. In contrast, uncertainty about the proper modeling of a physical system due to a lack of knowledge is called epistemic, which could eventually be reduced, e.g. by measurements. However, the distinction between aleatory and epistemic uncertainties is sometimes difficult in practice [172]. In the context of this work, we address both types of uncertainties with a probabilistic description which is referred to as Bayesian probabilistic approach for epistemic uncertainties in [161]. Note that we, however, do not employ the two-level approach suggested in [161] and refer to [162, 172] for further procedures to address epistemic uncertainties. The corresponding random vector $\boldsymbol{\Xi} = [\Xi_1, \dots, \Xi_{N_{\boldsymbol{\xi}}}]^\top$ is defined on a probability space (Θ, Σ, P) where Θ refers to the sample space, i.e. the set of elementary outcomes, Σ denotes the sigma-algebra of events and $P : \Sigma \to [0, 1]$ the probability measure which maps from events to probabilities. The random vector $\boldsymbol{\Xi}$ is a measurable function $\boldsymbol{\Xi} : \Theta \to \Gamma$, where $\Gamma \subset \mathbb{R}^{N_{\boldsymbol{\xi}}}$ denotes the image space, with joint PDF $\rho(\boldsymbol{\xi}) : \Gamma \to \mathbb{R}_0^+$. The parameter vector $\boldsymbol{\xi}$ then represents a realization of $\boldsymbol{\Xi}$, i.e. $\boldsymbol{\xi} = \boldsymbol{\Xi}(\theta) \in \Gamma, \theta \in \Theta$.

We further introduce the notation $\Gamma_i \subset \mathbb{R}$, $i = 1, \dots, N_{\boldsymbol{\xi}}$ for the image space of the single RV Ξ_i with univariate PDF $\rho_i(\xi_i) : \Gamma_i \to \mathbb{R}_0^+$. In the subsequent chapters, we will then sometimes make the assumptions that the RVs Ξ_i, $i = 1, \dots, N_{\boldsymbol{\xi}}$ are mutually independent, i.e. $\rho(\boldsymbol{\xi}) = \prod_{i=1}^{N_{\boldsymbol{\xi}}} \rho_i(\xi_i)$ and $\Gamma = \Gamma_1 \times \ldots \times \Gamma_{N_{\boldsymbol{\xi}}}$. Note that independent RVs could also be obtained by a suitable transformation of dependent inputs, for example, Rosenblatt or Nataf transformations [112, 127]. In particular, the Rosenblatt transformation which maps $\boldsymbol{\Xi}$ to a uniformly distribution random vector on $[0, 1]^{N_{\boldsymbol{\xi}}}$ with independent elements is given as

$$\mathbf{T} : \mathbb{R}^{N_{\boldsymbol{\xi}}} \to \mathbb{R}^{N_{\boldsymbol{\xi}}},\ \boldsymbol{\xi} \mapsto \begin{pmatrix} F_{\boldsymbol{\Xi};1}(\xi_1) \\ F_{\boldsymbol{\Xi};2|1}(\xi_2|\Xi_1 = \xi_1) \\ \vdots \\ F_{\boldsymbol{\Xi};N_{\boldsymbol{\xi}}|1,\dots,N_{\boldsymbol{\xi}}-1}(\xi_{N_{\boldsymbol{\xi}}}|\Xi_1 = \xi_1, \dots, \Xi_{N_{\boldsymbol{\xi}}-1} = \xi_{N_{\boldsymbol{\xi}}-1}) \end{pmatrix}, \tag{2.64}$$

where $F_{\boldsymbol{\Xi};k|1,\dots,k-1}$, $k \in \{1, \dots, N_{\boldsymbol{\xi}}\}$ refers to the cumulative distribution function (CDF) of the conditional RV $\Xi_k|\Xi_1 \dots \Xi_{k-1}$, see [127]. Furthermore, for correlated input parameters the Karhunen-Loève expansion is a powerful tool [62, 184, 211] which will be explained and applied in the numerical examples, see Section 5.2.2.

The random inputs $\boldsymbol{\Xi}$ lead to a random output $\mathcal{Q}(\boldsymbol{\Xi})$. Hence, it can be seen that the uncertainty from the inputs propagates to the outputs, as illustrated in Fig. 1.1. The goal of forward uncertainty propagation, which is the main topic of Chapter 4, then is to quantify the uncertainty in the QoI given a numerical model $\boldsymbol{\xi} \mapsto \mathcal{Q}(\boldsymbol{\xi})$ with uncertain inputs $\boldsymbol{\xi}$. It shall be mentioned that inverse UQ refers to the opposite, i.e. using information about the model output $\mathcal{Q}$, e.g. measurement data, to compute the distribution of the inputs, see [191] and the references therein for details. However, in this thesis we focus on forward UQ.

2.4 Summary

In this chapter, we have derived parametric/stochastic problems from Maxwell's equations as well as their FE discretization. These models comprise, for instance, particle accelerator cavities, nano-optical structures, or electrical waveguides. Afterward, the setting for forward uncertainty propagation was introduced.

Note that Maxwell's source problems (2.41) and (2.48) are parametric systems with respect to the angular frequency ω. This type of parametric problem is addressed in the following chapter, where suitable surrogate modeling techniques for FRFs are discussed. In Chapter 4 we will then address UQ for the parametric problems with uncertain shape or material parameters introduced in Section 2.3.1-2.3.2.

3 Approximation of frequency response functions

The content and structure of this chapter are based on our work [86]. As derived in the previous chapter, the discretization of deterministic Maxwell's source problems in the frequency domain leads to linear systems in the form

$$\left(\mathbf{K} - \omega^2\mathbf{M} + \mathrm{i}\mathbf{D}\right)\mathbf{e} = \mathbf{f}, \tag{3.1}$$

see (2.41) and (2.48). We note that also problems from different engineering fields, e.g. structural mechanics or acoustics, can be brought in the same form. In general, they are usually derived by applying the Fourier transform to a time-dependent linear second-order PDE which is then discretized using a suitable numerical method, e.g. FEM. For simplicity, we focus here on scalar frequency-dependent functionals of the solution $\mathbf{e}$, which are commonly employed to assess engineering designs, e.g. scattering parameters. Such a quantity of interest can be written as

$$Q : \mathbb{C} \to \mathbb{C}, \mathrm{i}\omega \mapsto J(\mathbf{e}(\mathrm{i}\omega)). \tag{3.2}$$

Note that the repeated evaluation of the FRF $\omega \mapsto Q(\mathrm{i}\omega)$ for different frequencies can be computationally demanding as each evaluation requires solving the linear system (3.1). Hence, in this chapter we address the construction of a surrogate model $Q_{N_\omega} \approx Q$ based on a small training set

$$\{\omega_i, Q(\mathrm{i}\omega_i)\}_{i=1}^{N_\omega}, \text{ where } w_i \in \Omega, Q(\mathrm{i}\omega_i) \in \mathbb{C}, \ i = 1, \ldots, N_\omega, \tag{3.3}$$

and where we assume the ω_i, $i = 1, \ldots, N_\omega$ to be distinct. This task has already been considered in the literature using interpolation as well as regression techniques, see for example [99, 125, 147] and the references therein. In particular, we mention the well-established rational approximation techniques vector fitting (VF) [99] and the adaptive Antoulas–Anderson (AAA) algorithm [147]. As both methods will be used as a benchmark, we will recall the key ideas of AAA and VF in the following section. Alternative data-driven techniques related to MOR and rational approximations are the Loewner framework [5] and approaches employing the Heaviside representation [153].

In Section 3.2, we then suggest a kernel-based interpolation approach for FRFs. This topic has so far only received limited attention in the literature, which is eventually caused by the slow convergence of the well-established Matérn kernels [168] or radial basis functions (RBFs) [39] for many dynamical responses. Hence, we suggest a method that is specifically tailored to complex-valued FRFs. We show that this dedicated kernel method, avoiding the individual interpolation of real- and imaginary part, uses the training data more efficiently. It is combined with a low-order rational basis and a tailored model selection scheme, in order to improve the efficiency for FRFs with a few dominant poles. In Section 5.1, we will then compare our RKI method against both AAA and VF and observe an improved or at least comparable performance for a variety of test cases.

3.1 Review of existing methods

For convenience of the reader, we recall the main ideas of AAA and VF in the following but refer to the literature for specific details.

3.1.1 Adaptive Antoulas–Anderson algorithm (AAA)

The AAA method for rational approximation has been proposed in [147]. It employs the barycentric representation $Q_{N_J}^{\mathrm{AAA}}(\omega)$ such that

$$Q(\mathrm{i}\omega) \approx Q_{N_J}^{\mathrm{AAA}}(\omega) = \frac{n(\omega)}{d(\omega)} = \frac{\sum_{j\in J} \frac{w_j Q(\mathrm{i}\omega_j)}{\omega-\omega_j}}{\sum_{j\in J} \frac{w_j}{\omega-\omega_j}}, \tag{3.4}$$

where $J \subseteq \{1,\dots,N_\omega\}$ with cardinality $N_J = \#J$ and $w_j \in \mathbb{C}$ are the weights. The barycentric representation of the rational function $Q_{N_J}^{\mathrm{AAA}}(\omega)$ is well-conditioned, see [147] for details. By multiplying the numeratur $n(\omega)$ and the denominator $d(\omega)$ both by $\prod_{j\in J}(\omega-\omega_j)$, one can see that $Q_{N_J}^{\mathrm{AAA}}(\omega)$ is a rational function of order (N_J-1, N_J-1). Furthermore, it fulfills the interpolating property $Q_{N_J}^{\mathrm{AAA}}(\omega_j) = Q(\mathrm{i}\omega_j)$ for all $j \in J$, see [147, Theorem 2.1]. To avoid instabilities, the support points ω_j are adaptively selected using a greedy scheme:

1. Starting from an ordered subset of given interpolation nodes ω_j, $j \in J = (j_1,\dots,j_{N_J}) \subset (1,\dots,N_\omega)$, the corresponding weights w_j, $j \in J$ are computed by solving the least-square problem

$$\text{minimize} \sum_{i\in(1,\dots,N_\omega)\setminus J} |Q(\mathrm{i}\omega_i)d(\omega_i) - n(\omega_i)|^2, \ \text{ such that } \frac{1}{N_J}\sum_{j\in J}|w_j|^2 = 1. \tag{3.5}$$

2. Then, the next nodal index j_{N_J+1} is selected, which corresponds to the remaining training point, where the nonlinear residual has the largest magnitude, i.e.

$$j_{N_J+1} = \operatorname*{arg\,max}_{i\in(1,\dots,N_\omega)\setminus J} \left| Q(\mathrm{i}\omega_i) - \frac{n(\omega_i)}{d(\omega_i)} \right|. \tag{3.6}$$

For further details on the AAA algorithm, we refer to [147].

3.1.2 Vector Fitting

VF is another method for rational approximation which is specifically tailored to functions in the frequency domain. The name of VF refers to the fact that the method can also be applied to vector-valued QoIs. It was proposed in [99] and is already well established in different engineering fields. VF constructs a rational approximation which is represented in terms of partial fractions as

$$Q(\mathrm{i}\omega) \approx Q_{N_\omega}^{\mathrm{VF}}(\mathrm{i}\omega) = \sum_{n=1}^{N_{\mathrm{pol}}^{\mathrm{VF}}} \frac{r_n}{\mathrm{i}\omega - p_n} + a + \mathrm{i}\omega b, \tag{3.7}$$

with $a, b \in \mathbb{R}$. The implementation ensures that all poles p_n and residuals r_n are either real or come in complex-conjugate pairs.

The VF algorithm consists of two stages. First, a set of initial poles $\{q_n\}$ is iteratively relocated in order to identify the poles of Q. To this end, the linear problem

$$\underbrace{\left(\sum_{n=1}^{N_{\mathrm{pol}}^{\mathrm{VF}}} \frac{\tilde{r}_n}{\mathrm{i}\omega - q_n} + 1\right)}_{\sigma(\mathrm{i}\omega)} Q(\mathrm{i}\omega) = \sum_{n=1}^{N_{\mathrm{pol}}^{\mathrm{VF}}} \frac{\hat{r}_n}{\mathrm{i}\omega - q_n} + \hat{a} + \mathrm{i}\omega\hat{b}, \tag{3.8}$$

with fixed poles q_n, is solved in least-squares sense to obtain $\hat{r}_n, \tilde{r}_n, \hat{a}, \hat{b}$. Then, the zeros of $\sigma(\mathrm{i}\omega)$ are calculated by solving an eigenproblem and employed as new estimates of the poles replacing the initial guesses $\{q_n\}$, see [99] for details. Usually several iterations of this pole relocation procedure are applied. Once the final poles $\{p_n\}$ are obtained, the corresponding values of r_n, a, b are then computed by solving the linear least-squares problem associated to (3.7).

Finally, we give a few concluding remarks. The employed implementation of VF enforces stable poles, i.e. $\mathcal{R}[p_n] < 0$, by eventually flipping them to left half-plane. In [98], a modified version of (3.8) using a relaxed non-triviality constrained was proposed, which is also employed in this work. In [99, Section 3.2] recommendations are given for the choice of suitable starting poles $\{q_n\}$ which might be crucial for the convergence, accuracy or numerical stability of the method. In this work, as suggested, we employ starting poles that are linearly distributed over the frequency range with a small magnitude of the real part. For further details on VF, we refer to [98, 99, 104].

3.2 Rational kernel-based interpolation

In this section, we address FRFs Q, which are elements of the Hardy space

$$H^2(\Upsilon_\alpha) := \{f \in \mathcal{H}(\Upsilon_\alpha) : \|f\|_{H^2(\Upsilon_\alpha)} := \sup_{x>-\alpha} \left(\int_{-\infty}^{\infty} |f(x+\mathrm{i}y)|^2\right)^{\frac{1}{2}} < \infty\}, \tag{3.9}$$

where $\mathcal{H}(\Upsilon_\alpha)$ refers to the space of holomorphic functions on Υ_α, which, in turn, denotes the half-plane domain

$$\Upsilon_\alpha := \{s \in \mathbb{C} \,|\, \mathcal{R}[s] > -\alpha\}. \tag{3.10}$$

Some details on the space $H^2(\Upsilon_\alpha)$ will be given below; however, for a comprehensive background on Hardy spaces, we refer to [121]. We note that any FRF $Q(\mathrm{i}\omega)$, which can be represented as the Laplace transform of a time-signal $f(t) \in L^2(\mathbb{R}^+)$ (supported on the positive axis), has to be an element of $H^2(\Upsilon_0)$ which can be concluded from the Paley-Wiener theorem [206, Theorem 5]. However, in the context of this work, we make a stronger assumption by assuming a sufficient damping component such that $\alpha > 0$.

The assumption, in particular, excludes poles on the frequency axis. This is the case, for example, if the functional J in (3.2) can be represented as a linear functional $J(\mathbf{e}(\mathrm{i}\omega)) = \mathbf{j} \cdot \mathbf{e}(\mathrm{i}\omega)$ with $\mathbf{j} \in \mathbb{R}^{N_h}$ and one of the following two conditions hold. In particular, if the matrices in (3.1) are symmetric positive definite matrices $\mathbf{K}, \mathbf{D}, \mathbf{M} \in \mathbb{R}^{N_h \times N_h}$, then all poles of $Q = \mathbf{j} \cdot \mathbf{e} \in H^2(\Upsilon_\alpha)$ have a negative real part, see [198, Section 3]. Also, if the homogeneous version of the time-domain counterpart of (3.1) is stable, i.e. all solutions decay exponentially for $t \to \infty$, this holds again [198]. The holomorphy of FRFs has also been analyzed in the context of PDEs. In [27], for instance, an acoustic scattering problem is investigated where a suitable damping

component leads to a locally holomorphic frequency response map and it was shown that all eigenvalues have a negative real part[1].

3.2.1 Complex RKHS interpolation

We are interested in kernel-based interpolation of FRFs $Q \in H^2(\Upsilon_\alpha)$. As $H^2(\Upsilon_\alpha)$ is a complex reproducing kernel Hilbert space (RKHS), we recall a few basic facts on RKHSs in the following. Note that we will focus on the key ideas and refer to [163] for a general introduction in RKHSs and to our contribution [86, Section 2] for details on the specific functional analysis background summarized in the following subsections as well as proofs of the statements.

A complex RKHS over a set $\mathbb{X} \neq \emptyset$ is a complex Hilbert space H of functions $f : \mathbb{X} \to \mathbb{C}$, such that for all $x \in \mathbb{X}$ the evaluation functional $\delta_x : H \to \mathbb{C}, f \mapsto f(x), \forall f \in H$ is continuous. The reproducing kernel of H is a unique function $k : \mathbb{X} \times \mathbb{X} \to \mathbb{C}$ such that $k(\cdot, x) \in H$, $\forall x \in \mathbb{X}$ and which fulfills the reproducing property

$$f(x) = \delta_x(f) = \big(f, k(\cdot, x)\big)_H, \quad \forall f \in H, x \in \mathbb{X}, \tag{3.11}$$

where $(\cdot, \cdot)_H$ refers to the Hermitian inner product of H.

The kernel associated with the considered space $H^2(\Upsilon_\alpha)$ is a variant of the Szegö kernel. In particular, in [28, Theorem 2.12] the Szegö kernel is derived for the Hardy space H^2 on the upper half-plane instead of Υ_α. By then employing the mapping $s \mapsto \mathrm{i}(s + \alpha)$ for translation and rotation of the half-plane domain, we obtain the respective Szegö kernel

$$k_{\mathrm{S}}(s_k, s_l; \alpha) = \frac{1}{2\alpha + s_k - \overline{s_l}}, \quad s_k, s_l \in \Upsilon_\alpha, \tag{3.12}$$

for $H^2(\Upsilon_\alpha)$. As we are interested in approximations on the frequency axis, (3.12) can be simplified by evaluating it only on the imaginary axis $s = \mathrm{i}\omega$ which leads to

$$k_{\mathrm{S}}(\mathrm{i}\omega_k, \mathrm{i}\omega_l; \alpha) = \frac{1}{2\alpha + \mathrm{i}(\omega_k - \omega_l)}, \quad \omega_k, \omega_l \in \Omega. \tag{3.13}$$

Employing this kernel and the training data (3.3), we can define a kernel interpolation of an FRF in $H^2(\Upsilon_\alpha)$ as

$$Q_{H^2,N_\omega}(\mathrm{i}\omega) = \sum_{i=1}^{N_\omega} \gamma_i k_{\mathrm{S}}(\mathrm{i}\omega, \mathrm{i}\omega_i; \alpha), \tag{3.14}$$

where the coefficients $\gamma_i \in \mathbb{C}$ can be obtained by solving

$$\begin{bmatrix} k_{\mathrm{S}}(\mathrm{i}\omega_1, \mathrm{i}\omega_1; \alpha) & \dots & k_{\mathrm{S}}(\mathrm{i}\omega_1, \mathrm{i}\omega_{N_\omega}; \alpha) \\ \vdots & \ddots & \vdots \\ k_{\mathrm{S}}(\mathrm{i}\omega_{N_\omega}, \mathrm{i}\omega_1; \alpha) & \dots & k_{\mathrm{S}}(\mathrm{i}\omega_{N_\omega}, \mathrm{i}\omega_{N_\omega}; \alpha) \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_{N_\omega} \end{bmatrix} = \begin{bmatrix} Q(\mathrm{i}\omega_1) \\ \vdots \\ Q(\mathrm{i}\omega_{N_\omega}) \end{bmatrix}. \tag{3.15}$$

It can then be shown that (3.14) is the unique function of minimal norm in $H^2(\Upsilon_\alpha)$ which interpolates the training data (3.3), i.e.

$$Q_{H^2,N_\omega} = \operatorname*{arg\,min}_{\tilde{Q} \in H^2(\Upsilon_\alpha)} \|\tilde{Q}\|_{H^2(\Upsilon_\alpha)}, \text{ such that } \tilde{Q}(\mathrm{i}\omega_i) = Q(\mathrm{i}\omega_i),\ i = 1, \dots, N_\omega, \tag{3.16}$$

see e.g. [163, Theorem 3.4].

[1] In fact, due to another convention, [27, Proposition 5.3] shows that the eigenvalues have a negative imaginay part.

3.2.2 Complex/real RKHS interpolation

The FRF Q can be represented as the Laplace transform of a real-valued time-signal, hence it fulfills the property

$$\overline{Q(s)} = Q(\overline{s}). \tag{3.17}$$

Accordingly, we focus in the following on an approximation in the subspace of functions in

$$H^2_{\mathrm{sym}}(\Upsilon_\alpha) \;:=\; \left\{ f \in H^2(\Upsilon_\alpha) :\; \forall s \in \Upsilon_\alpha,\; \overline{f(s)} = f(\overline{s}) \right\} \subset H^2(\Upsilon_\alpha). \tag{3.18}$$

Although $H^2_{\mathrm{sym}}(\Upsilon_\alpha)$ contains complex-valued functions, it cannot be endowed with the structure of a complex RKHS. This can be easily shown by considering any $f \in H^2_{\mathrm{sym}}(\Upsilon_\alpha)$ and $s \in \Upsilon_\alpha$ such that $f(\overline{s}) \neq 0$, for which we would then obtain the contradiction between $\overline{(\mathrm{i}f)(s)} = -\mathrm{i}\overline{f(s)} = -\mathrm{i}f(\overline{s})$ and $\overline{(\mathrm{i}f)(s)} = (\mathrm{i}f)(\overline{s})$.

As $H^2(\Upsilon_\alpha)$ is a real vector space consisting of complex-valued functions, we define in our work [86] a new type of function space which is called a complex/real RKHS. In particular, a complex/real RKHS refers to a real Hilbert space $\hat{H}$ of complex-valued functions on a set $\mathbb{X} \neq \emptyset$ if the evaluation maps δ_x are continuous, i.e. for all $x \in \mathbb{X}$, the function $\delta_x : \hat{H} \to \mathbb{C}$, $f \mapsto f(x)$, is continuous. Although the complex/real RKHS contains complex-valued functions over $\mathbb{X}$, we can employ the mapping $\mathcal{A} : \mathbb{C}^{\mathbb{X}} \to \mathbb{R}^{\mathbb{X}\times\{0,1\}}$ given by

$$(\mathcal{A}f)(x,a) = \begin{cases} \Re[f(x)] & a = 0, \\ \Im[f(x)] & a = 1, \end{cases} \tag{3.19}$$

to represent these functions as real-valued functions over $\mathbb{X} \times \{0,1\}$. In particular, the complex/real RKHS H is isometrically isomorphic to the real RKHS $\mathcal{A}\hat{H} \subset \mathbb{R}^{\mathbb{X}\times\{0,1\}}$ endowed with the inner product $(f,g)_{\mathcal{A}\hat{H}} = (\mathcal{A}^{-1}f, \mathcal{A}^{-1}g)_{\hat{H}}$, see [86, Section 2]. This observation will be employed on the one side to discuss some properties of complex/real RKHS in the following and on the other side for implementation purposes. In particular, in Section 3.2.3 we will explain how the mapping $\mathcal{A}$ can be employed to implement complex/real RKHS interpolation for complex functions employing existing software for the real-valued case.

In the following, we denote as $k_{\mathcal{A}} : (\mathbb{X}\times\{0,1\}) \times (\mathbb{X}\times\{0,1\}) \to \mathbb{R}$ the symmetric positive definite kernel of the real RKHS $\mathcal{A}\hat{H}$ corresponding to the complex/real RKHS $\hat{H}$. Then, we can define the complex, hermitian, positive definite kernel $\hat{k}$ and the complex, symmetric pseudo-kernel of $\hat{H}$ as

$$\hat{k}(x,x') = (k_{RR} + k_{II}) + \mathrm{i}(k_{IR} - k_{RI}), \tag{3.20}$$
$$\hat{c}(x,x') = (k_{RR} - k_{II}) + \mathrm{i}(k_{IR} + k_{RI}), \tag{3.21}$$

where

$$k_{RR}(x,x') \;=\; k_{\mathcal{A}}\left((x,0),\,(x',0)\right), \qquad k_{RI}(x,x') \;=\; k_{\mathcal{A}}\left((x,0),\,(x',1)\right), \tag{3.22}$$
$$k_{IR}(x,x') \;=\; k_{\mathcal{A}}\left((x,1),\,(x',0)\right), \qquad k_{II}(x,x') \;=\; k_{\mathcal{A}}\left((x,1),\,(x',1)\right), \tag{3.23}$$

[86, Section 2].

It is conjectured [86, Section 2] that the considered function space $H^2_{\mathrm{sym}}(\Upsilon_\alpha)$ defined in (3.18) is a complex/real RKHS on Υ_α with complex kernel k_{S} as defined in (3.12) and corresponding pseudo-kernel c_{S} which is given as

$$c_{\mathrm{S}}(s,s';\alpha) = k_{\mathrm{S}}\big(s,\overline{s'};\alpha\big) \quad \forall s,s' \in \Upsilon_\alpha. \tag{3.24}$$

In particular, the corresponding pseudo-kernel associated to the Szegö kernel (3.13) then reads

$$c_{\mathrm{S}}(\mathrm{i}\omega_k, -\mathrm{i}\omega_l; \alpha) = \frac{1}{2\alpha + \mathrm{i}(\omega_k + \omega_l)}, \tag{3.25}$$

i.e., a non-stationary function, with decreasing magnitude w.r.t. ω. Furthermore, we can define the complex/real kernel interpolation of an FRF in $H^2_{\mathrm{sym}}(\Upsilon_\alpha)$ employing the training data (3.3) as

$$Q_{H^2_{\mathrm{sym}},N_\omega}(\mathrm{i}\omega) = \sum_{i=1}^{N_\omega} \gamma_i k_{\mathrm{S}}(\mathrm{i}\omega, \mathrm{i}\omega_i) + \overline{\gamma_i} c_{\mathrm{S}}(\mathrm{i}\omega, \mathrm{i}\omega_i), \tag{3.26}$$

where the coefficients are a solution of

$$\begin{bmatrix} \Re[\mathbf{K}_{\mathrm{S}} + \mathbf{C}_{\mathrm{S}}] & \Im[-\mathbf{K}_{\mathrm{S}} + \mathbf{C}_{\mathrm{S}}] \\ \Im[\mathbf{K}_{\mathrm{S}} + \mathbf{C}_{\mathrm{S}}] & \Re[\mathbf{K}_{\mathrm{S}} - \mathbf{C}_{\mathrm{S}}] \end{bmatrix} \begin{bmatrix} \Re[\boldsymbol{\gamma}] \\ \Im[\boldsymbol{\gamma}] \end{bmatrix} = \begin{bmatrix} \Re[\mathbf{Q}] \\ \Im[\mathbf{Q}] \end{bmatrix}, \tag{3.27}$$

with $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_{N_\omega}]^\top$, $\mathbf{Q} = [Q(\mathrm{i}\omega_1), \ldots, Q(\mathrm{i}\omega_{N_\omega})]^\top$ and

$$\mathbf{K}_{\mathrm{S}} = \begin{bmatrix} k_{\mathrm{S}}(\mathrm{i}\omega_1, \mathrm{i}\omega_1; \alpha) & \ldots & k_{\mathrm{S}}(\mathrm{i}\omega_1, \mathrm{i}\omega_{N_\omega}; \alpha) \\ \vdots & \ddots & \vdots \\ k_{\mathrm{S}}(\mathrm{i}\omega_{N_\omega}, \mathrm{i}\omega_1; \alpha) & \ldots & k_{\mathrm{S}}(\mathrm{i}\omega_{N_\omega}, \mathrm{i}\omega_{N_\omega}; \alpha) \end{bmatrix}, \mathbf{C}_{\mathrm{S}} = \begin{bmatrix} c_{\mathrm{S}}(\mathrm{i}\omega_1, \mathrm{i}\omega_1; \alpha) & \ldots & c_{\mathrm{S}}(\mathrm{i}\omega_1, \mathrm{i}\omega_{N_\omega}; \alpha) \\ \vdots & \ddots & \vdots \\ c_{\mathrm{S}}(\mathrm{i}\omega_{N_\omega}, \mathrm{i}\omega_1; \alpha) & \ldots & c_{\mathrm{S}}(\mathrm{i}\omega_{N_\omega}, \mathrm{i}\omega_{N_\omega}; \alpha) \end{bmatrix}. \tag{3.28}$$

It shall be noted that the span of basis functions in the form of (3.26) corresponds to the so-called *Widely Linear RKHS*, which has been defined in [26]. It can be shown that this space is a dense subspace of the complex/real RKHS [86, Section 2]. Furthermore, it can be conjectured that (3.26) is the unique function of minimal norm in $H^2_{\mathrm{sym}}(\Upsilon_\alpha)$ which interpolates the training data (3.3), i.e.

$$Q_{H^2_{\mathrm{sym}},N_\omega} = \underset{\tilde{Q} \in H^2_{\mathrm{sym}}(\Upsilon_\alpha)}{\arg\min} \|\tilde{Q}\|_{H^2_{\mathrm{sym}}(\Upsilon_\alpha)}, \text{ such that } \tilde{Q}(\mathrm{i}\omega_i) = Q(\mathrm{i}\omega_i)\ i = 1, \ldots, N_\omega, \tag{3.29}$$

see [86, Section 2].

3.2.3 Relation to complex Gaussian process regression and hyperparameter tuning

RKHS interpolation is closely related to Gaussian process (GP) interpolation, see e.g. [119] for a comprehensive discussion on their connection in the real-valued case. Furthermore, we refer to [168] for a general introduction to GP regression and interpolation. A complex GP is a collection of complex RVs such that any finite subset is jointly Gaussian distributed. In particular, a mean-free complex GP $g(\omega)$ over Ω is completely determined by its covariance function

$$K(\omega, \omega') = \mathbb{E}[g(\omega)\overline{g(\omega')}] \in \mathbb{C},\ \forall \omega, \omega' \in \Omega \tag{3.30}$$

and its pseudo-covariance function

$$C(\omega, \omega') = \mathbb{E}[g(\omega)g(\omega')] \in \mathbb{C},\ \forall \omega, \omega' \in \Omega. \tag{3.31}$$

The GP $g(\omega)$ is here employed to model the FRF $\omega \mapsto Q(\mathrm{i}\omega)$, where it shall be noted that we focus on real-valued inputs for simplicity. To this end, we then employ the Szegö kernel (3.13) as covariance function, i.e.

$$K(\omega, \omega'; [\alpha, \sigma]) = \sigma^2 k_{\mathrm{s}}(\mathrm{i}\omega, \mathrm{i}\omega'; \alpha), \tag{3.32}$$

where $\sigma > 0$ denotes a scaling parameter. Due to the symmetry property (3.17) of the FRF, we obtain the respective pseudo-covariance function

$$C(\omega, \omega'; [\alpha, \sigma]) = \mathbb{E}[g(\omega)g(\omega')] = \mathbb{E}[g(\omega)\overline{g(-\omega')}] = K(\omega, -\omega'; [\alpha, \sigma]) = \sigma^2 c_{\mathrm{s}}(\mathrm{i}\omega, \mathrm{i}\omega'; \alpha), \tag{3.33}$$

which corresponds to (3.25), see also [125, Section 5]. Employing the complex GP $g(\omega)$ as prior and conditioning on the training data (3.3), one can obtain the corresponding posterior distribution, see [25] for details. It can then be conjectured that the respective posterior mean is equivalent to the complex/real RKHS interpolation (3.26) evaluated on the imaginary axis, see [86, Section 2].

However, performing GP regression or interpolation in the complex case requires dedicated routines, as explained in [25, 125]. Hence, we discuss in the following how the mapping $\mathcal{A}$ defined in (3.19) can be employed to realize an implementation based on any available software toolbox for the real-valued case, which allows specifying custom kernel functions. The main idea is to construct a mean-free real GP $g_{\mathcal{A}}([\omega, a])$ on an augmented input space $[\omega, a] \in \Omega \times \{0, 1\}$ with covariance function $K_{\mathcal{A}}$, such that

$$g_{\mathcal{A}}([\omega, 0]) = \Re[g(\omega)], \qquad g_{\mathcal{A}}([\omega, 1]) = \Im[g(\omega)]. \tag{3.34}$$

The required augmented training data $\{\boldsymbol{\omega}_{\mathcal{A},i}, Q_{\mathcal{A},i}\}_{i=1}^{N_{\mathcal{A}}} \subset \Big(\big(\Omega \times \{0,1\}\big) \times \mathbb{R}\Big)^{N_{\mathcal{A}}}$ is obtained through the map

$$\Big\{\big(\omega_i, Q(\mathrm{i}\omega_i)\big)\Big\} \mapsto \begin{cases} \Big\{\big([\omega_i, 0], \Re[Q(\mathrm{i}\omega_i)]\big), \big([\omega_i, 1], \Im[Q(\mathrm{i}\omega_i)]\big)\Big\}, & \omega_i \neq 0, \\ \Big\{\big([\omega_i, 0], \Re[Q(\mathrm{i}\omega_i)]\big)\Big\}, & \omega_i = 0, \end{cases} \tag{3.35}$$

which is applied for each element of the training set (3.3). Note that any FRF in $H^2_{\mathrm{sym}}(\Upsilon_\alpha)$ assumes a real value at $w = 0$ and, hence, the special treatment in (3.35) is required to avoid a singular covariance matrix, see [125] for an extensive discussion on this issue. The covariance function $K_{\mathcal{A}}$ is derived by enforcing (3.34) which leads to

$$K_{\mathcal{A}}([\omega, a], [\omega', a']) = \begin{cases} \mathbb{E}\big[\Re[g(\omega)]\Re[g(\omega')]\big] = \frac{1}{2}\Re[K(\omega, \omega') + C(\omega, \omega')], & a = a' = 0, \\ \mathbb{E}\big[\Im[g(\omega)]\Im[g(\omega')]\big] = \frac{1}{2}\Re[K(\omega, \omega') - C(\omega, \omega')], & a = a' = 1, \\ \mathbb{E}\big[\Re[g(\omega)]\Im[g(\omega')]\big] = \frac{1}{2}\Im[-K(\omega, \omega') + C(\omega, \omega')], & a = 0, a' = 1, \\ \mathbb{E}\big[\Im[g(\omega)]\Re[g(\omega')]\big] = \frac{1}{2}\Im[K(\omega, \omega') + C(\omega, \omega')], & a = 1, a' = 0. \end{cases} \tag{3.36}$$

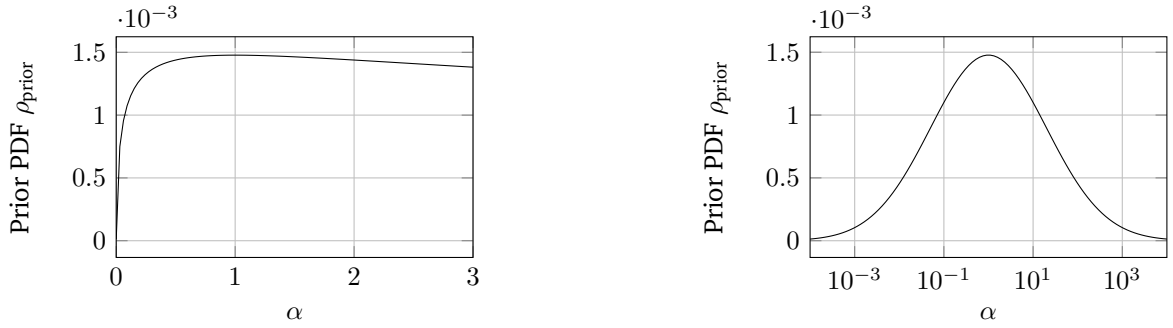
Finally, the posterior distribution of the complex GP $g(\omega)$ can be obtained from the posterior distribution of $g_{\mathcal{A}}([w, a])$ since

$$g(\omega) = g_{\mathcal{A}}([\omega, 0]) + \mathrm{i} g_{\mathcal{A}}([\omega, 1]). \tag{3.37}$$

Note that in the context of this work, we then only employ the posterior mean which corresponds to the complex/real RKHS interpolant (3.26).

Next, we discuss the selection of the hyperparameters α, σ which are present in the covariance function (3.32) and pseudo-covariance function (3.33). As the size of the domain Υ_α of the Hardy space $H^2(\Upsilon_\alpha)$ is, in practice, usually not known a priori, the hyperparameters are estimated based on the likelihood function. Let $\boldsymbol{\omega}_{\mathcal{A}} = [\boldsymbol{\omega}_{\mathcal{A},1}^\top, \ldots, \boldsymbol{\omega}_{\mathcal{A},N_{\mathcal{A}}}^\top]^\top$ and $\mathbf{Q}_{\mathcal{A}} = [Q_{\mathcal{A},1}, \ldots, Q_{\mathcal{A},N_{\mathcal{A}}}]^\top$ collect the augmented input and output training data, respectively. Then, the log marginal likelihood function reads

$$\log p(\mathbf{Q}_{\mathcal{A}} | \boldsymbol{\omega}_{\mathcal{A}}) = -\frac{1}{2}\mathbf{Q}_{\mathcal{A}}^\top (\mathbf{K}_{\mathcal{A}})^{-1} \mathbf{Q}_{\mathcal{A}} - \frac{1}{2}\log|\mathbf{K}_{\mathcal{A}}| - \frac{N_{\mathcal{A}}}{2}\log(2\pi), \tag{3.38}$$



(a) Linear abscissa.

(b) Logarithmic abscissa.

Figure 3.1: Log-normal prior on hyperparameter α for $|\Omega| = 1$.

where $|\mathbf{K}_{\mathcal{A}}|$ denotes the determinant of $\mathbf{K}_{\mathcal{A}}$ and

$$\mathbf{K}_{\mathcal{A}} = \begin{bmatrix} K_{\mathcal{A}}(\boldsymbol{\omega}_{\mathcal{A},1}, \boldsymbol{\omega}_{\mathcal{A},1}) & \dots & K_{\mathcal{A}}(\boldsymbol{\omega}_{\mathcal{A},1}, \boldsymbol{\omega}_{\mathcal{A},N_{\mathcal{A}}}) \\ \vdots & \ddots & \vdots \\ K_{\mathcal{A}}(\boldsymbol{\omega}_{\mathcal{A},N_{\mathcal{A}}}, \boldsymbol{\omega}_{\mathcal{A},1}) & \dots & K_{\mathcal{A}}(\boldsymbol{\omega}_{\mathcal{A},N_{\mathcal{A}}}, \boldsymbol{\omega}_{\mathcal{A},N_{\mathcal{A}}}) \end{bmatrix}, \tag{3.39}$$

see e.g. [168, Section 2.2]. To ensure a positive value of the hyperparameter α, we add a vague lognormal prior on $\alpha \sim \text{Lognormal}(\mu_\alpha, \sigma_\alpha^2)$. In the context of this work, we chose $\sigma_\alpha = 3$ and the mode $\text{Mo} = \omega_{\max} - \omega_{\min}$ as default values which implies $\mu_\alpha = \sigma_\alpha^2 + \log(\text{Mo}) = 9.6931$, leading to the PDF

$$\rho_{\text{prior}}(\alpha) = \frac{1}{\alpha \sigma_\alpha \sqrt{2\pi}} \exp\Big(\frac{-(\log(\alpha) - \mu_\alpha)^2}{2\sigma_\alpha^2}\Big), \tag{3.40}$$

illustrated in Fig. 3.1. It can be seen that the chosen parameters allow for the choice of α within a range of several orders of magnitude. The particular values of the hyperparameters α, σ are then chosen by maximizing the corresponding posterior distribution, i.e. $\arg\max_{\alpha,\sigma} \left(\log p(\mathbf{Q}_{\mathcal{A}} | \boldsymbol{\omega}_{\mathcal{A}}, \alpha, \sigma) + \log \rho_{\text{prior}}(\alpha)\right)$. To this end, we employ gradient-based optimization with a multistart procedure. More details on the optimization will be given at the beginning of Section 5.1.

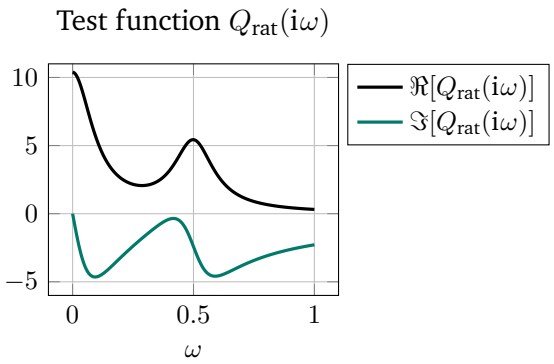
Finally, we briefly discuss the alternative approach of applying kernel interpolation with a standard kernel to real and imaginary part, separately. This approach corresponds to choosing the covariance function

$$K_{\mathcal{A}}^{\text{sep}}([\omega, a], [\omega', a']) = \begin{cases} \mathbb{E}\big[\Re[g(\omega)]\Re[g(\omega')]\big] = K_{\text{real}}\big(\omega, \omega'; \boldsymbol{\theta}_{\text{r}}\big), & a = a' = 0, \\ \mathbb{E}\big[\Im[g(\omega)]\Im[g(\omega')]\big] = K_{\text{real}}\big(\omega, \omega'; \boldsymbol{\theta}_{\text{i}}\big), & a = a' = 1, \\ \mathbb{E}\big[\Re[g(\omega)]\Im[g(\omega')]\big] = 0, & a = 0, a' = 1, \\ \mathbb{E}\big[\Im[g(\omega)]\Re[g(\omega')]\big] = 0, & a = 1, a' = 0. \end{cases} \tag{3.41}$$

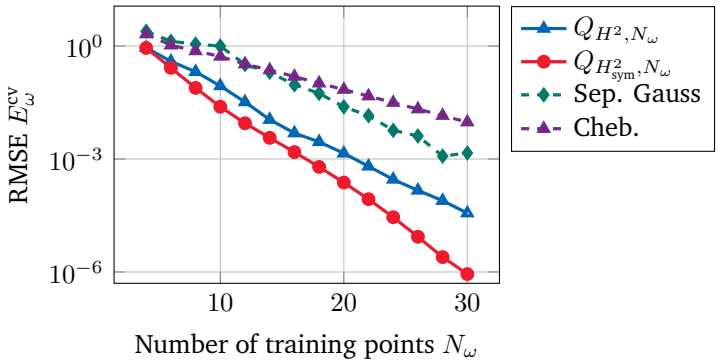
where K_{real} denotes a real-valued covariance kernel and $\boldsymbol{\theta}_{\text{r}}, \boldsymbol{\theta}_{\text{i}}$ the corresponding hyperparameters for real and imaginary part, respectively. It can be seen that in this case no correlation between real and imaginary parts is assumed. In this work, we use this approach with a Gaussian kernel

$$K_{\text{real}}^{\text{Gauss}}\big(\omega, \omega'; [\sigma, l]\big) = \sigma^2 \exp\Big(-\frac{(\omega - \omega')^2}{l^2}\Big), \tag{3.42}$$

which is also referred to as squared exponential kernel, as a reference.



(a) Function $Q_{\text{rat}} \in H^2_{\text{sym}}(\Upsilon_{0.1})$ defined in (3.43) (with real-valued inverse Laplace transform).



(b) Convergence study of the RMSE for different approximations of $Q_{\text{rat}}(\mathrm{i}\omega)$.

Figure 3.2: Numerical example showing the improved convergence of complex/real kernel interpolation with a suitable pseudo-kernel for a test function with real-valued inverse Laplace transform. Additional dashed lines show the inferior convergence of a kernel interpolation with a Gaussian kernel for real and imaginary part separately, as well as polynomial interpolation on Chebyshev nodes.

3.2.4 Numerical example

For illustration, we consider the third-order rational function

$$Q_{\text{rat}}(\mathrm{i}\omega) = \frac{1}{\mathrm{i}\omega - (-0.1)} + \frac{0.5}{\mathrm{i}\omega - (-0.1 - 0.5\mathrm{i})} + \frac{0.5}{\mathrm{i}\omega - (-0.1 + 0.5\mathrm{i})}, \tag{3.43}$$

where $\omega \in \Omega = [\omega_{\min}, \omega_{\max}] = [0, 1]$. This function is the Laplace transform of the real-valued function

$$f(t) = e^{-0.1t}\big(1 + \cos(0.5t)\big), \tag{3.44}$$

and, hence, belongs to the space $Q_{\text{rat}} \in H^2_{\text{sym}}(\Upsilon_{0.1}) \subset H^2(\Upsilon_{0.1})$. We then conduct a convergence study using training data with equidistant frequency sample points

$$\omega_i = \omega_{\min} + (i-1)\frac{\omega_{\max} - \omega_{\min}}{N_\omega - 1},\ i = 1, \ldots, N_\omega, \tag{3.45}$$

for different approximations $\tilde{Q}_{N_\omega}(\mathrm{i}\omega)$ of $Q_{\text{rat}}(\mathrm{i}\omega)$. The accuracy of the approximations $\tilde{Q}_{N_\omega}(\mathrm{i}\omega)$ is quantified in terms of the root-mean-square error (RMSE)

$$E^{\text{cv}}_\omega = \sqrt{\frac{1}{N_{\text{cv}}} \sum_{j=1}^{N_{\text{cv}}} |\tilde{Q}_{N_\omega}(\mathrm{i}\omega_j) - Q_{\text{rat}}(\mathrm{i}\omega_j)|^2}, \tag{3.46}$$

where

$$\omega_j = \omega_{\min} + (j-1)\frac{\omega_{\max} - \omega_{\min}}{N_{\text{cv}} - 1},\ j = 1, \ldots, N_{\text{cv}}, \tag{3.47}$$

is a fine grid and we choose $N_{\text{cv}} = 201$. As explained in the previous section, the hyperparameters are, in all cases, chosen by tuning based on the likelihood function. Further details on the implementation will be given in Section 5.1.

In Fig. 3.2 we demonstrate that employing a suitable pseudo-kernel might significantly impact the convergence properties of kernel interpolation. In particular, for the test function (3.43), the complex/real kernel interpolation (3.26) converges significantly faster than the kernel interpolation (3.14). Note that the test-function is a low order rational function that is here only used to illustrate the impact of the pseudo-kernel. Accordingly, rational interpolation techniques as AAA or VF reach machine accuracy already with ≈ 8 training points and are hence excluded in the convergence plot for clarity. However, it can already be observed that complex/real RKHS interpolation with the Szegö kernel outperforms the alternative approach of separate interpolations for real and imaginary parts with a real Gaussian kernel, as well as polynomial interpolation on Chebyshev nodes. Note that Chebyshev nodes are well-established non-equidistant nodes for polynomial interpolation while we used equidistant nodes for the kernel-based approaches.

3.2.5 Rational basis functions

The convergence of the complex/real RKHS interpolation can be significantly worse compared to rational approximations techniques as AAA or VF, when the FRF has a few dominant poles p_i, i.e. with small attenuation $\mathcal{R}[p_i]$, which strongly limits the domain of the Hardy space $H^2_{\mathrm{sym}}(\Upsilon_\alpha)$. Hence, in this section, we discuss how the complex/real RKHS interpolation can be combined with a low number of rational basis functions in a suitable manner for the approximation of FRFs. To this end, as in VF, we employ rational basis functions with a real-valued inverse Laplace transform. In particular, we employ the additional rational function

$$h(\mathrm{i}\omega;\mathbf{p}) = \sum_{i=1}^{N_\mathbf{p}} \frac{1}{\mathrm{i}\omega - p_i} r_i + \frac{1}{\mathrm{i}\omega - \overline{p_i}} \overline{r_i}, \text{ where } \mathcal{R}[p_i] < 0,\ \mathcal{I}[p_i] > 0,\ i = 1,\dots,N_\mathbf{p}, \tag{3.48}$$

with (stable) complex conjugate poles $p_i, \overline{p_i}$ and residues $r_i, \overline{r_i}$. Note that $h(\mathrm{i}\omega;\mathbf{p})$ is, by construction, element of $H^2_{\mathrm{sym}}(\Gamma_\alpha)$ with $\alpha = \min_{i:1\leq i\leq N_\mathbf{p}} |\mathcal{R}[p_i]| > 0$. For simplicity, we only consider complex poles in (3.48); however, the basis could be extended to additionally include a number of real poles, as in the VF basis (3.7). We emphasize that we do not employ a constant basis function, s.t. the resulting approximation

$$Q_{N_\omega}^{(N_\mathbf{p})}(\mathrm{i}\omega;\alpha,\mathbf{p}) = Q_{H^2_{\mathrm{sym}},N_\omega}(\mathrm{i}\omega;\alpha) + h(\mathrm{i}\omega;\mathbf{p}), \tag{3.49}$$

has the desired properties at infinity, in particular,

$$\lim_{\omega\to\infty} Q(\mathrm{i}\omega) = \lim_{\omega\to\infty} Q_{N_\omega}^{(N_\mathbf{p})}(\mathrm{i}\omega) = 0. \tag{3.50}$$

In the context of the present work, we only consider a small number of pole pairs $N_\mathbf{p}$; in particular, we employ $N_\mathbf{p}^{\max} = 5$. The selection of this discrete parameter will be discussed below. However, we first give a few general remarks on the RKI approximation (3.49). The implementation of the complex/real RKHS interpolation based on the mapping $\mathcal{A}$, which was discussed in Section 3.2.3, can be straightforwardly extended to additionally incorporate the rational function h. In this case, the poles $\mathbf{p}$ are treated as additional hyperparameters and h can then be represented as a linear model with respect to the real and imaginary part of the coefficients r_i, respectively. We refer to [168, Section 2.7] for details on incorporating explicit basis functions.

Next, considering a fixed number of pole pairs $N_\mathbf{p}$, we address the tuning of the continuous hyperparameters α, σ and $\mathbf{p}$ which is, again, based on the likelihood function. In particular, we now choose the hyperparameters such that

$$(\alpha,\sigma,\mathbf{p}) = \operatorname*{arg\,max}_{(\alpha,\sigma,\mathbf{p})} \left(\sup_{\mathbf{r}\in\mathbb{C}^K} \log p\left(\mathbf{Q}_\mathcal{A}|\boldsymbol{\omega}_\mathcal{A},\alpha,\sigma,\mathbf{p},\mathbf{r}\right) + \log \rho_{\mathrm{prior}}(\alpha) \right) = \operatorname*{arg\,max}_{(\alpha,\sigma,\mathbf{p})} \mathcal{C}(\alpha,\sigma,\mathbf{p};\mathbf{Q}_\mathcal{A},\boldsymbol{\omega}_\mathcal{A}), \tag{3.51}$$

where the dependence on the nuisance parameters $\{r_i\}_{i=1}^{N_\mathbf{p}}$ has been eliminated in the criterion $\mathcal{C}$ by employing the profile log-likelihood function, see [145] for details on profiling out nuisance parameters in likelihood functions. Note that choosing a suitable number of pole pairs $N_\mathbf{p}$ is more involved and addressed with another model selection technique below.

We consider complex/real RKHS interpolation in the form of (3.49) with a Szegö kernel as well as up to $N_\mathbf{p}^{\max}$ rational basis functions. For the approximation accuracy, providing suitable initial values for the poles, as well as selecting an appropriate number of poles, is crucial. We tackle these tasks with a two-step procedure. First, we build $N_\mathbf{p}^{\max}+1$ approximations $Q_{N_\omega}^{(N_\mathbf{p})},\ ,N_\mathbf{p}=0,\dots,N_\mathbf{p}^{\max}$, where the superscript $N_\mathbf{p}$ indicates the number of pole pairs. A suitable procedure will be discussed in the following. In a next step, we then perform model selection, which will be discussed in the subsequent section.

In the first iteration, we start with the maximum number of poles $N_\mathbf{p}^{\max}$ and an equidistant pole distribution along the frequency axis (with a small magnitude of the real part)

$$p_i = -10^{-3}|\Omega| + \mathrm{i}\Big(\omega_{\min} + (j+0.5)\frac{\omega_{\max}-\omega_{\min}}{N_\mathbf{p}^{\max}}\Big),\quad j=0,\dots,N_\mathbf{p}^{\max}-1. \tag{3.52}$$

We recall that the basis functions (3.48) include the complex conjugate poles as well. This choice of starting poles is closely related to the standard recommendation for the starting poles of VF for very similar reasons, see e.g. [99]. In particular, the equidistant distribution along the frequency axis shall reduce the probability that the optimal poles are far from the initial poles, while the weak attenuation improves the numerical condition. Alternatively, selecting starting poles by running a few VF iterations beforehand could be another possible option to provide initial values.

In the following iterations, we then reuse the $N_\mathbf{p}$ optimized pole pairs from the previous iterations as starting points. In particular, we remove one pole based on the corresponding change in the profile likelihood-based criterion $\mathcal{C}$, i.e. the least relevant pole, and then restart the tuning with $N_\mathbf{p}-1$ pole pairs. This procedure is described in detail in Algorithm 1 and yields a set of optimized RKI models $\{Q_{N_\omega}^{(N_\mathbf{p})}\}_{N_\mathbf{p}=0}^{N_\mathbf{p}^{\max}}$, where $N_\mathbf{p}$ indicates the number of pole pairs in the rational function h.

3.2.6 Model selection

Model selection is then based on leave-one-out (LOO) cross-validation, i.e. on the error indicators

$$\mathcal{E}_{\mathrm{loo}}^{N_\mathbf{p}} = \frac{1}{N_\omega}\sum_{i=1}^{N_\omega}\Big|Q(\mathrm{i}\omega_i) - \hat{Q}_{N_\omega-1,i}^{(N_\mathbf{p})}(\mathrm{i}\omega_i)\Big|^2,\quad N_\mathbf{p}=0,1,\dots,N_\mathbf{p}^{\max}, \tag{3.53}$$

for the model $Q_{N_\omega}^{(N_\mathbf{p})}$, where $\hat{Q}_{N_\omega-1,i}^{(N_\mathbf{p})}(\mathrm{i}\omega_i)$ denotes a model interpolating the reduced training data

$$\{\omega_j, Q(\mathrm{i}\omega_j)\}_{j\in\{1,\dots,N_\omega\}\setminus\{i\}}, \tag{3.54}$$

evaluated at the removed frequency point ω_i. It is common practice to keep the hyperparameters fixed, which significantly improves the efficiency. However, we numerically observed that this approach, i.e. using the same value of α and the same poles $\mathbf{p}$ as in $Q_{N_\omega}^{(N_\mathbf{p})}$ for each $\hat{Q}_{N_\omega-1,i}^{(N_\mathbf{p})}$, $i=1,\dots,N_\omega$, does not always give satisfactory results for the considered, strongly nonlinear, basis. In particular, it generally tends to favor the models with a larger number of poles. Hence, we instead employ the LOO criterion with re-tuning of the hyperparameters

Data: Training data $\{\omega_i, Q(\mathrm{i}\omega_i)\}_{i=1}^{N_\omega}$; frequency range Ω; maximum number of pole pairs $N_{\mathbf{p}}^{\max}$

Result: Set of optimized models $\{Q_{N_\omega}^{(N_{\mathbf{p}})}\}_{N_{\mathbf{p}}=0}^{N_{\mathbf{p}}^{\max}}$, where $N_{\mathbf{p}}$ indicates the number of pole pairs in the rational function h

Initialize vector of poles $\mathbf{p}_{\text{init}} \in \mathbb{C}^{N_{\mathbf{p}}^{\max}}$ as in (3.52)

Initialize Szegö kernel hyperparameter α_{init} by tuning a model with fixed poles $\mathbf{p}_{\text{init}}$ in $h(\mathrm{i}\omega; \mathbf{p}_{\text{init}})$

Initialize $N_{\mathbf{p}} = N_{\mathbf{p}}^{\max}$

while *number of poles* $N_{\mathbf{p}} \geq 0$ **do**

  Tune model $Q_{N_\omega}^{(N_{\mathbf{p}})}$ with Szegö kernel $k_{\mathrm{S}}(s, s'; \alpha^{(N_{\mathbf{p}})})$ and pseudo-kernel $c_{\mathrm{S}}(s, s'; \alpha^{(N_{\mathbf{p}})})$ in terms of the hyperparameter $\alpha^{(N_{\mathbf{p}})}$ and the poles $\mathbf{p}^{(N_{\mathbf{p}})} \in \mathbb{C}^{N_{\mathbf{p}}}$, where α_{init} and $\boldsymbol{p}_{\text{init}}$ are used as initial values

  Store model $Q_{N_\omega}^{(N_{\mathbf{p}})}(\mathrm{i}\omega; \alpha^{(N_{\mathbf{p}})}, \mathbf{p}^{(N_{\mathbf{p}})})$

  Set $\alpha_{\text{init}} \leftarrow \alpha^{(N_{\mathbf{p}})}$

  `/* Choose initial poles for next iteration by removing one pole pair (s.t. the change in the likelihood is minimal): */`

  for $i \leftarrow 1$ *to* $N_{\mathbf{p}}$ **do**

    Set $\mathbf{p}^{(N_{\mathbf{p}}, \sim i)} \leftarrow \left(p_k^{(N_{\mathbf{p}})}\right)_{k \in \{1,\ldots,N_{\mathbf{p}}\} \setminus \{i\}} \in \mathbb{C}^{N_{\mathbf{p}}-1}$

    Evaluate profile likelihood-based criterion $\mathcal{C}^{N_{\mathbf{p}}, i}$ of a RKI model with hyperparameter $\alpha^{(N_{\mathbf{p}})}$ and fixed poles $\mathbf{p}^{(N_{\mathbf{p}}, \sim i)} \in \mathbb{C}^{N_{\mathbf{p}}-1}$ in h

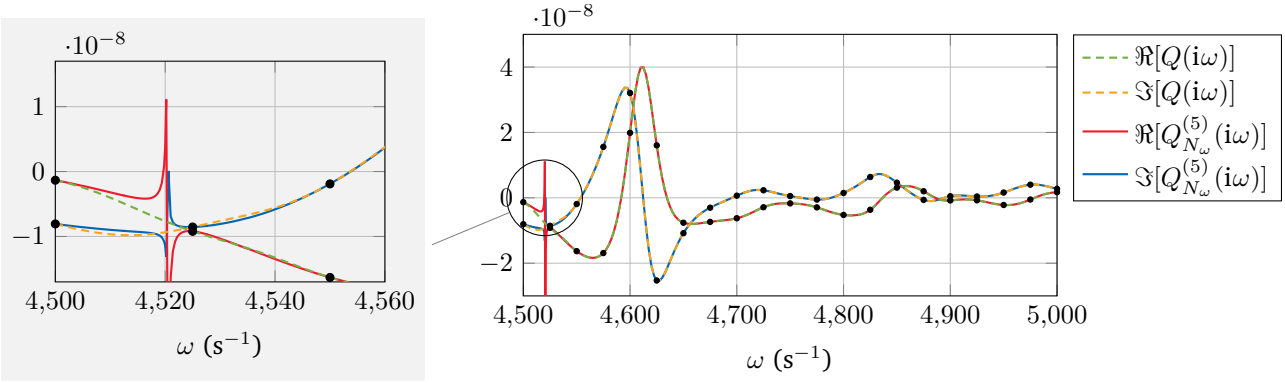
  end

  Set $\mathbf{p}_{\text{init}} \leftarrow \mathbf{p}^{(N_{\mathbf{p}}, \sim i)}$, where $i = \arg\max_i \mathcal{C}^{N_{\mathbf{p}}, i}$

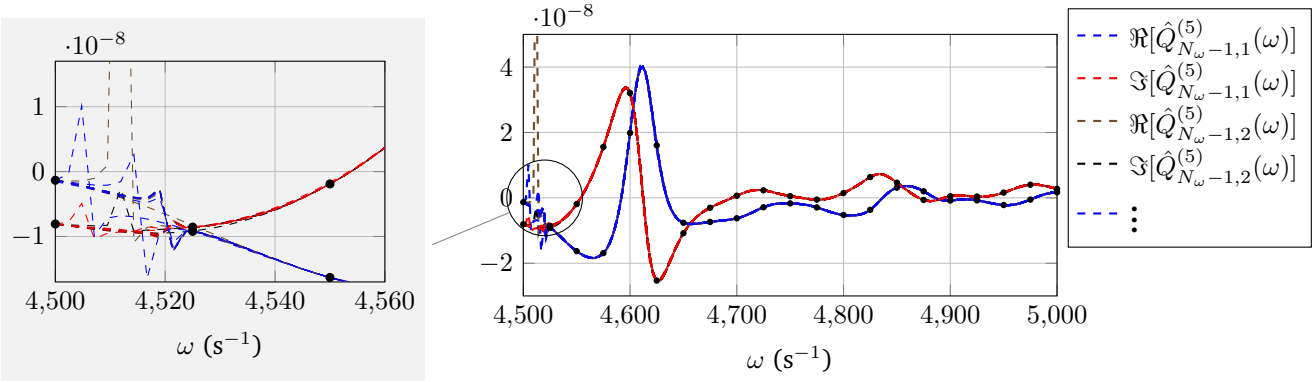
  Set $N_{\mathbf{p}} \leftarrow N_{\mathbf{p}} - 1$

end

Algorithm 1: Heuristic algorithm that employs suitable initial values to optimize RKI models (3.49) with different numbers of poles in the rational function h. In particular, it starts with the maximum number of poles and then subsequently reuses the tuned values of the hyperparameters as starting values after selecting one pole to be removed based on the profile log likelihood function.



(a) Dashed lines show the function to approximate. Black dots indicate the training data. Solid lines represent a *bad* approximation model, which, however, is selected by the LOO-criterion. Zoomed plot (gray background) highlights the influence of a wrongly identified pole.



(b) LOO predictions, which show strong local variations between $4500\,\mathrm{s}^{-1}$ and $4520\,\mathrm{s}^{-1}$. However, these variations do not significantly affect the values at the respective training points.

Figure 3.3: Illustration of model selection issue with the standard LOO criterion as motivating example for the proposed instability penalty.

$\alpha, \mathbf{p}$, i.e. we tune the model $\hat{Q}^{(N_\mathbf{p})}_{N_\omega-1,i}(\mathrm{i}\omega;\alpha,\mathbf{p})$ using the reduced training data (3.54) with gradient-based optimization, where we employ the hyperparameters $\alpha^{(N_\mathbf{p})}$ and poles $\mathbf{p}^{(N_\mathbf{p})}$ of the model $Q^{(N_\mathbf{p})}_{N_\omega}(\mathrm{i}\omega;\alpha^{(N_\mathbf{p})},\mathbf{p}^{(N_\mathbf{p})})$ as starting point (without any multistart procedure).

However, we additionally introduce an additional penalty term, which also takes *global* model variations into account. This approach can be motivated by one particular example, illustrated in Fig. 3.3 (top). The corresponding vibro-acoustic benchmark model will be described in Section 5.1.2; however, here we simply consider the approximation of the dashed function, based on interpolation of the training points (black dots), as a general example. At the top, it can be observed that the LOO-criterion (3.53) leads to the selection of a model (solid lines) $Q^{(5)}_{N_\omega}$ which wrongly identifies a pole at $\approx 4520\,\mathrm{s}^{-1}$. However, this effect is rather local, it mainly takes place between two training points (illustrated by black dots). At the bottom, we show the models $\hat{Q}^{(N_\mathbf{p})}_{N_\omega-1,i},\ \ i=1,\dots,N_\omega$, which show strong local variations close to $\approx 4510\,\mathrm{s}^{-1}$ but rather small errors at the training points ω_i. To take this into account, we introduce an additional instability penalty term, which leads to the criterion

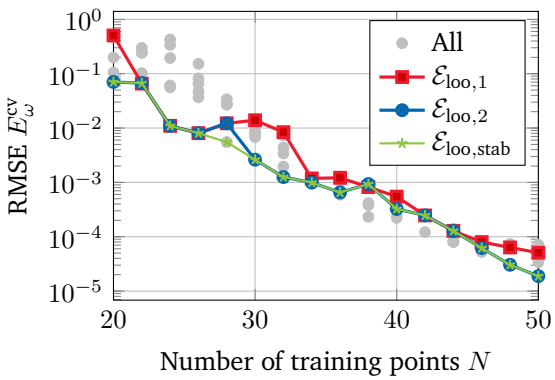
$$\mathcal{E}^{N_\mathbf{p}}_{\mathrm{loo,stab}} = \mathcal{E}^{N_\mathbf{p}}_{\mathrm{loo}} + \lambda \frac{1}{N_\omega}\frac{1}{N^{\mathrm{fine}}_\omega}\sum_{i=1}^{N_\omega}\sum_{j=1}^{N^{\mathrm{fine}}_\omega}\left|Q^{(N_\mathbf{p})}_{N_\omega}(\mathrm{i}\hat{\omega}_j) - \hat{Q}^{(N_\mathbf{p})}_{N_\omega-1,i}(\mathrm{i}\hat{\omega}_j)\right|^2, \tag{3.55}$$

where $\{\hat{\omega}_j\}_{j=1}^{N^{\mathrm{fine}}_\omega}$ denotes a fine grid of sample points. In this work, we employ an equidistant grid on Ω with $N^{\mathrm{fine}}_\omega = 10N_\omega + 1$ sample points. The weighting factor λ is chosen as

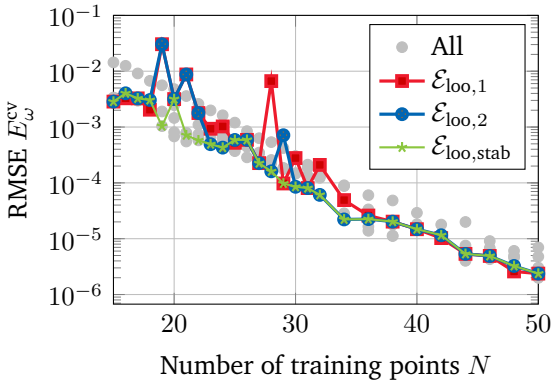
$$\lambda = 0.2\frac{\mathcal{E}^0_{\mathrm{loo}}}{\frac{1}{N_\omega}\frac{1}{N^{\mathrm{fine}}_\omega}\sum_{i=1}^{N_\omega}\sum_{j=1}^{N^{\mathrm{fine}}_\omega}\left|Q^{(0)}_{N_\omega}(\mathrm{i}\hat{\omega}_j) - \hat{Q}^{(0)}_{N_\omega-1,i}(\mathrm{i}\hat{\omega}_j)\right|^2}, \tag{3.56}$$

i.e., 0.2 after normalizing both terms with respect to the respective values of the kernel interpolation model. To our knowledge, this approach for model selection proposed in our contribution [86] has not been considered before, although it is related to the continuously-defined LOO error [80, 116, 120]. The latter LOO error has been employed for sequential sampling, while we propose to use the (numerical) integral of the squared quantity as instability penalty for model selection. Stability selection [134, 142, 176] is another loosely related approach, which is also based on data resampling, but is usually employed for variable selection. In particular, stability selection chooses the variables that are included sufficiently many times when a particular variable selection technique, e.g. the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) [41], is employed repeatedly for different replicas of the training data. Note that AIC and BIC are different and well-established model selection criteria that are not considered here since preliminary numerical tests indicated an inferior performance (usually leading to over-fitting phenomena) compared to LOO for the considered strongly nonlinear function h.

Employing the stabilized criterion (3.55) for model selection gives satisfactory results for the benchmark examples considered in this work. For illustration, we consider the convergence studies for two models, which will be described in Section 5.1.2. Fig. 3.4 shows the RMSEs E^{cv}_ω of the available models with gray dots and the accuracy of the selected models by the different criteria. It can be observed that the stabilized criterion (3.55) gives the best results, while the LOO residual with retuning is superior to the approach without retuning.



(a) Spiral Antenna Model.



(b) Vibro-Acoustics Model.

Figure 3.4: Comparison of different model selection criteria for two benchmark problems which are described in Section 5.1.2. The gray dots illustrate the RMSEs of all available models, which are obtained using Algorithm 1. $\mathcal{E}_{\mathrm{loo},1}$ and $\mathcal{E}_{\mathrm{loo},2}$ denote the LOO residual without and with retuning of hyperparameters, respectively. The stabilized criterion $\mathcal{E}_{\mathrm{loo,stab}}$ (with retuning) defined in (3.55) gives the best results.

3.3 Summary

In this chapter, we have addressed the approximation of FRFs with a dedicated RKI method. To this end, complex/real kernel interpolation using a suitable pseudo-kernel has been introduced. For an academic example, this method has shown superior convergence with respect to the alternative approach of using separate Gauss kernel interpolation for real and imaginary parts. It has also outperformed polynomial interpolation which could be expected as the kernel method employed the Szegö kernel and is, hence, specifically tailored to the Hardy function space of the considered FRFs. However, in the next chapter, we will address different parameters, i.e. uncertain shape or material parameters. In this case, the associated map from the parameters to the QoI usually has significantly different properties. In particular, in Section 4.2 we will address the case where the domain of holomorphy is given by an ϵ-neighborhood instead of a half-plane. In addition, special emphasis will be put on the multivariate case considering a moderately large number of parameters where sparse approximations are crucial, which is a well-established concept for polynomial interpolation. Hence, in the next chapter, we will then mainly focus on polynomial-based surrogate modeling techniques.

In the last part of this chapter, the kernel interpolation method has been combined with a few rational basis functions, as well as a dedicated hyper-parameter tuning and model selection scheme. The application of this algorithm, which also addresses FRFs with a few dominant poles, will then be presented in Section 5.1. In particular, we will numerically investigate its convergence for various benchmark problems from different fields and compare its performance with the established rational approximation techniques AAA and VF.

4 Uncertainty quantification

In this chapter, we address the problem of forward uncertainty propagation which was introduced in Section 2.3.3. In particular, we address a deterministic computational model $\boldsymbol{\xi} \mapsto \mathcal{Q}(\boldsymbol{\xi})$ which maps from known uncertain inputs $\boldsymbol{\xi}$ to an uncertain output $\mathcal{Q}(\boldsymbol{\xi})$. To this end, in the first section we review a number of existing UQ methods, namely (multifidelity) MC, GPC, SC and global sensitivity analysis. We then suggest improved methods in the subsequent sections. First, we propose transformed basis functions using conformal maps for both GPC and SC, in order to accelerate the convergence by suitably transforming the region of holomorphy. The main ideas as well as the enhanced convergence rates are illustrated numerically by means of an academic model problem, i.e. a stochastic RLC circuit. We then address moderately high-dimensional approximations with a dimension-adaptive scheme based on mapped Leja nodes. To further accelerate the convergence, an adjoint-error indicator is incorporated to steer the dimension-adaptivity and for error correction. In this context, we also derive the adjoint formulations for Maxwell's source problems introduced in Chapter 2. Next, efficient yield estimation using the adjoint-based mapped SC approximation is discussed. To obtain reliable results efficiently, we suggest a hybrid decision process for the classification of sample points which utilizes not only the SC approximation but also FE models of different fidelity as well as adjoint-error estimation. Finally, we address UQ for Maxwell's eigenproblem where standard spectral UQ methods cannot be straightforwardly applied using standard eigenvalue solvers due to possible eigenvalue crossings with respect to parameter changes. Hence, we suggest an eigenvalue tracking method based on homotopies between collocation points.

We note that our contributions are motivated by the high frequency electromagnetic field problems, introduced in Chapter 2, however, the majority can be applied in a more general setting also. The content and structure of this chapter follows our works [60, 81, 88–90].

4.1 Review of existing methods

4.1.1 Monte Carlo

In the context of UQ, MC simulation [47, 103, 130] employs a random sample $\{\boldsymbol{\Xi}_i, \mathcal{Q}(\boldsymbol{\Xi}_i)\}_{i=1}^{N_{\mathrm{MC}}}$, where $\boldsymbol{\Xi}_i$ are independently drawn according to the underlying PDF ρ, to estimate statistical measures of the QoI, for example the mean value, variance, confidence intervals or failure probabilities. The MC estimate of the mean is given as

$$\mathbb{E}\left[\mathcal{Q}(\boldsymbol{\Xi})\right] = \int_{\Gamma} \mathcal{Q}(\boldsymbol{\xi})\rho(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \approx E_{N_{\mathrm{MC}}}\left[\mathcal{Q}(\boldsymbol{\Xi})\right] \coloneqq \frac{1}{N_{\mathrm{MC}}} \sum_{i=1}^{N_{\mathrm{MC}}} \mathcal{Q}(\boldsymbol{\Xi}_i), \tag{4.1}$$

with RMSE

$$\left(\mathbb{E}\left[\left|\mathbb{E}\left[\mathcal{Q}(\boldsymbol{\Xi})\right] - E_{N_{\mathrm{MC}}}\left[\mathcal{Q}(\boldsymbol{\Xi})\right]\right|^2\right]\right)^{\frac{1}{2}} = \frac{\mathrm{Std}\left[\mathcal{Q}(\boldsymbol{\Xi})\right]}{\sqrt{N_{\mathrm{MC}}}}, \tag{4.2}$$

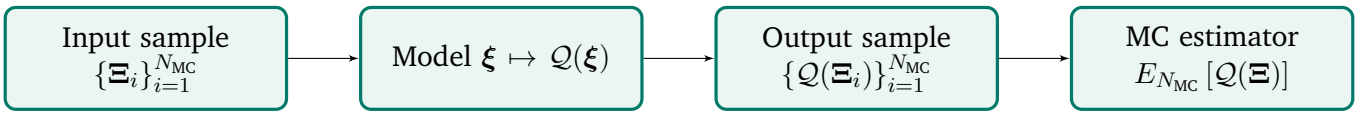


Figure 4.1: Illustration of MC method: the computational model can be treated as a black-box, hence, the method is considered as non-intrusive.

where $\mathrm{Std}[\mathcal{Q}(\boldsymbol{\Xi})] = \sqrt{\mathbb{V}[\mathcal{Q}(\boldsymbol{\Xi})]}$. Note that the MC estimate is unbiased, i.e.

$$\mathbb{E}\left[E_{N_{\mathrm{MC}}}[\mathcal{Q}(\boldsymbol{\Xi})]\right] = \mathbb{E}[\mathcal{Q}(\boldsymbol{\Xi})], \quad \forall N_{\mathrm{MC}} \geq 1. \tag{4.3}$$

Eq. (4.2) shows that the accuracy of the MC estimator only mildly depends on the number of RVs $N_{\boldsymbol{\xi}}$, as the output variance might in turn depend on $N_{\boldsymbol{\xi}}$. This is a main advantage of the MC estimate as the computational cost of other methods can be very strongly affected by $N_{\boldsymbol{\xi}}$ which will be discussed later. Furthermore, the MC estimator is very robust and easily applicable, as it can be applied for all square-integrable RVs, without any smoothness requirements, and only requires an output sample which can be often obtained by treating the computational model as a parameterized blackbox, see Fig. 4.1 for an illustration. The main drawback of the MC method is the rather slow convergence, i.e. the RMSE only converges as $\mathcal{O}(\frac{1}{\sqrt{N_{\mathrm{MC}}}})$ which can lead to high computational cost for certain accuracy requirements.

We also introduce the unbiased MC estimate for the variance

$$\mathbb{V}[\mathcal{Q}(\boldsymbol{\Xi})] \approx V_{N_{\mathrm{MC}}}[\mathcal{Q}(\boldsymbol{\Xi})] \coloneqq \frac{1}{N_{\mathrm{MC}}-1}\sum_{i=1}^{N_{\mathrm{MC}}}\left(\mathcal{Q}(\boldsymbol{\Xi}_i) - E_{N_{\mathrm{MC}}}[\mathcal{Q}(\boldsymbol{\Xi})]\right)^2. \tag{4.4}$$

We note that the MC sample $\{\mathcal{Q}(\boldsymbol{\Xi}_i)\}_{i=1}^{N_{\mathrm{MC}}}$ can further be employed to infer the PDF of $\mathcal{Q}(\boldsymbol{\Xi})$, for instance using kernel density estimation (KDE) [72]. KDE is a non-parametric method to estimate a continuous PDF as

$$\rho_{\mathcal{Q}}(q) \approx \rho_{\mathcal{Q},N_{\mathrm{MC}}}(q) \coloneqq \frac{1}{hN_{\mathrm{MC}}}\sum_{i=1}^{N_{\mathrm{MC}}} K\left(\frac{q-\mathcal{Q}(\boldsymbol{\Xi}_i)}{h}\right), \tag{4.5}$$

with a non-negative kernel function K and bandwidth parameter $h > 0$. In this work we employ the Epanechnikov kernel [72]

$$K(q) \coloneqq \begin{cases} \frac{3}{4}\left(1-q^2\right), & q \in [-1,1], \\ 0, & \text{else.} \end{cases} \tag{4.6}$$

Finally, we define a performance feature specification [97] as

$$\mathcal{Q}(\boldsymbol{\xi}) \leq c, \tag{4.7}$$

i.e. the QoI is required to be below a constant c, which is considered, without loss of generality, as an upper bound. The indicator function

$$\mathcal{I}_{\Gamma_{\mathrm{s}}}(\boldsymbol{\xi}) = \begin{cases} 1 & \boldsymbol{\xi} \in \Gamma_{\mathrm{s}}, \\ 0 & \text{else,} \end{cases} \tag{4.8}$$

then indicates whether a parameter vector belongs to the safe domain

$$\Gamma_{\mathrm{s}} \coloneqq \{\boldsymbol{\xi} : \mathcal{Q}(\boldsymbol{\xi}) \leq c\}, \tag{4.9}$$

s.t. the yield $\mathcal{Y}$ can be expressed as

$$\mathcal{Y} \coloneqq \mathbb{E}\left[\mathcal{I}_{\Gamma_s}(\boldsymbol{\Xi})\right] = \int_{\Gamma} \mathcal{I}_{\Gamma_s}(\boldsymbol{\xi})\rho(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}. \tag{4.10}$$

The yield $\mathcal{Y}$ corresponds to the percentage of realizations in a manufacturing process which fulfill the predefined performance feature specification (4.7), see [97]. Note that the yield $\mathcal{Y}$ is related to the failure probability $\mathcal{F} \coloneqq P(\mathcal{Q}(\boldsymbol{\Xi}) > c)$ as $\mathcal{Y} = 1 - \mathcal{F}$. Using the MC method, the yield (4.10) can be estimated as

$$\mathcal{Y} \approx \mathcal{Y}^{N_{\mathrm{MC}}} \coloneqq E_{N_{\mathrm{MC}}}\left[\mathcal{I}_{\Gamma_s}(\boldsymbol{\Xi})\right] = \frac{1}{N_{\mathrm{MC}}} \sum_{i=1}^{N_{\mathrm{MC}}} \mathcal{I}_{\Gamma_s}(\boldsymbol{\Xi}_i). \tag{4.11}$$

Following [94],[81, Section 3.1] and exploiting that all observations are independent, the variance of $\mathcal{Y}^{N_{\mathrm{MC}}}$ is obtained as

$$\mathbb{V}\left[\mathcal{Y}^{N_{\mathrm{MC}}}\right] = \frac{1}{N_{\mathrm{MC}}^2}\mathbb{V}\left[\sum_{i=1}^{N_{\mathrm{MC}}} \mathcal{I}_{\Gamma_s}(\boldsymbol{\Xi}_i)\right] = \frac{1}{N_{\mathrm{MC}}^2}\sum_{i=1}^{N_{\mathrm{MC}}} \mathbb{V}\left[\mathcal{I}_{\Gamma_s}(\boldsymbol{\Xi}_i)\right] = \frac{1}{N_{\mathrm{MC}}^2} N_{\mathrm{MC}}\mathcal{Y}(1-\mathcal{Y}) = \frac{\mathcal{Y}(1-\mathcal{Y})}{N_{\mathrm{MC}}}. \tag{4.12}$$

Accordingly, as the MC estimate $\mathcal{Y}^{N_{\mathrm{MC}}}$ is unbiased, the RMSE of $\mathcal{Y}^{N_{\mathrm{MC}}}$ is obtained as

$$\left(\mathbb{E}\left[|\mathcal{Y} - \mathcal{Y}^{N_{\mathrm{MC}}}|^2\right]\right)^{\frac{1}{2}} = \sqrt{\frac{\mathcal{Y}(1-\mathcal{Y})}{N_{\mathrm{MC}}}} \leq \frac{0.5}{\sqrt{N_{\mathrm{MC}}}}, \tag{4.13}$$

and depends on the value of the yield $0 \leq \mathcal{Y} \leq 1$, attaining its maximum for a yield of 0.5. Again it can be observed that the MC approach converges rather slowly with $\mathcal{O}(\frac{1}{\sqrt{N_{\mathrm{MC}}}})$ which can lead to high computational cost if the model evaluations are expensive and require the solution of a PDE, for instance.

4.1.2 Multifidelity Monte Carlo

There exist a number of approaches that aim for improved convergence of the MC method under certain assumptions. One approach is quasi-MC [47, 130] where the random sample $\{\boldsymbol{\Xi}_i\}_{i=1}^{N_{\mathrm{MC}}}$ is replaced by a low-discrepancy sequence. Another approach is the multilevel MC method [94] where a hierarchy of different computational models with increasing cost and increasing accuracy is combined into a single estimator. In particular, this model hierarchy might correspond to a set of FE models as introduced in Chapter 2 with different grid resolutions h. The multilevel MC approach then generates predictions with the cheap coarse grid models which are corrected using a few evaluations of the reference model on the finest grid. Note that multilevel MC was recently used for a high-frequency application [133]. However, multilevel MC relies on a number of assumptions, e.g. certain convergence rates, see, e.g. [94, Theorem 1]. Hence, we focus in the following on the MFMC method [164, 165] which generalizes the multilevel approach. In particular, MFMC allows combining low-fidelity models of different kinds, without any quantification of the model errors, into an efficient sampling-based estimator, which is illustrated in Fig. 4.2. We recall the key aspects of the method in the following and refer to [164, 165] for details. Note that [151] recently proposed a related method, referred to as Bayesian MFMC, which employs a Bayesian approach to estimate the output PDF $\rho_{\mathcal{Q}}$ of the high-fidelity model $\mathcal{Q}$.

MFMC relies on a set of models $\{\mathcal{Q}^{(i)}\}_{i=1}^{N_{\mathrm{mod}}}$, where $\mathcal{Q}^{(1)} \coloneqq \mathcal{Q}$ is the high-fidelity model while $\mathcal{Q}^{(i)}$, $i = 2, \ldots, N_{\mathrm{mod}}$ denote the low-fidelity models. The low-fidelity models might refer not only to coarse-grid FE

models but also different simplified models, for example 2D instead of 3D models or equivalent network/circuit models. One particular way to obtain a low-fidelity model which is very well suited for MFMC for periodic optical structures will be presented in Section 5.3.3. The MFMC estimate for the expectation is then obtained as

$$\mathbb{E}\left[\mathcal{Q}(\boldsymbol{\Xi})\right] \approx E_{\mathbf{N}_{\mathrm{MC}}}\left[\mathcal{Q}(\boldsymbol{\Xi})\right] := E_{N_{\mathrm{MC}}^{(1)}}\left[\mathcal{Q}^{(1)}(\boldsymbol{\Xi})\right] + \sum_{i=2}^{N_{\mathrm{mod}}} \alpha_i \left(E_{N_{\mathrm{MC}}^{(i)}}\left[\mathcal{Q}^{(i)}(\boldsymbol{\Xi})\right] - E_{N_{\mathrm{MC}}^{(i-1)}}\left[\mathcal{Q}^{(i)}(\boldsymbol{\Xi})\right]\right), \tag{4.14}$$

where $E_{N_{\mathrm{MC}}^{(i)}}\left[\mathcal{Q}^{(i)}\right]$ denotes the standard MC estimator based on the sample $\{\boldsymbol{\Xi}_j, \mathcal{Q}^{(i)}(\boldsymbol{\Xi}_j)\}_{j=1}^{N_{\mathrm{MC}}^{(i)}}$ and

$$0 < N_{\mathrm{MC}}^{(1)} \leq N_{\mathrm{MC}}^{(2)} \leq \ldots \leq N_{\mathrm{MC}}^{(N_{\mathrm{mod}})}. \tag{4.15}$$

Note that Eq. (4.14) is an unbiased estimator of $\mathbb{E}\left[\mathcal{Q}^{(1)}\right]$, as the high-fidelity model $\mathcal{Q}^{(1)}$ is sampled at least one time. The coefficients α_i and the sample sizes $N_{\mathrm{MC}}^{(i)}$, $i = 1, \ldots, N_{\mathrm{mod}}$ are chosen by an optimal model management strategy which minimizes the variance, and, hence, the RMSE, of the estimator for a fixed computational budget $\mathcal{B}$. The model management is based on the Pearson correlation coefficients

$$\rho_{1,i} = \frac{\mathrm{Cov}\left[\mathcal{Q}^{(1)}(\boldsymbol{\Xi}), \mathcal{Q}^{(i)}(\boldsymbol{\Xi})\right]}{\mathrm{Std}\left[\mathcal{Q}^{(1)}(\boldsymbol{\Xi})\right]\mathrm{Std}\left[\mathcal{Q}^{(i)}(\boldsymbol{\Xi})\right]}, \quad i = 2, \ldots, N_{\mathrm{mod}}, \tag{4.16}$$

and the computational cost $C^{(i)}$, $i = 1, \ldots, N_{\mathrm{mod}}$ per model evaluation. In particular, low-fidelity models are usually sampled extensively if they have a high correlation $\rho_{1,i}$ and low computational cost $C^{(i)}$. If one chooses the optimal values of $\alpha_i, N_{\mathrm{MC}}^{(i)}$ which can be computed analytically, see [165], the MFMC estimator (4.14) achieves a RMSE of

$$\mathbb{E}\left[\left|E_{\mathbf{N}_{\mathrm{MC}}}\left[\mathcal{Q}(\boldsymbol{\Xi})\right] - \mathbb{E}\left[\mathcal{Q}(\boldsymbol{\Xi})\right]\right|^2\right]^{\frac{1}{2}} = \left(\frac{\sigma_1^2}{N_{\mathrm{MC}}^{(1)}} + \sum_{i=2}^{N_{\mathrm{mod}}} \left(\frac{1}{N_{\mathrm{MC}}^{(i-1)}} - \frac{1}{N_{\mathrm{MC}}^{(i)}}\right)\left(\alpha_i^2\sigma_i^2 - 2\alpha_i\rho_{1,i}\sigma_1\sigma_i\right)\right)^{\frac{1}{2}}, \tag{4.17}$$

where $\sigma_i = \mathrm{Std}\left[\mathcal{Q}^{(i)}(\boldsymbol{\Xi})\right]$. Finally, it shall be noted that, if the values of $\rho_{1,i}, C^{(i)}$, $i = 1, \ldots, N_{\mathrm{mod}}$ are not known a priori, they can in practice be estimated based on a pilot run with a small sample, see [164, 165].

4.1.3 Spectral polynomial methods

If the map $\boldsymbol{\xi} \mapsto \mathcal{Q}(\boldsymbol{\xi})$ is smooth, the use of spectral UQ methods [126, 211] is appealing. In particular, we denote as $\mathcal{Q}_M$ the surrogate approximation

$$\mathcal{Q}\left(\boldsymbol{\xi}\right) \approx \mathcal{Q}_M\left(\boldsymbol{\xi}\right) := \sum_{m=0}^{M} s_m \Psi_m\left(\boldsymbol{\xi}\right), \tag{4.18}$$

where $\Psi_m : \Gamma \to \mathbb{R}$ are global polynomials and $s_m \in \mathbb{C}$ the corresponding coefficients. Depending on the regularity of the map $\boldsymbol{\xi} \mapsto \mathcal{Q}(\boldsymbol{\xi})$, fast convergence can be expected, e.g. assuming that the map $\boldsymbol{\xi} \mapsto \mathcal{Q}(\boldsymbol{\xi})$ is analytic, one can even expect exponential convergence, as will be discussed in the following. Once an accurate approximation $\mathcal{Q}_M\left(\boldsymbol{\xi}\right)$ is available, statistical measures about the QoI can then be

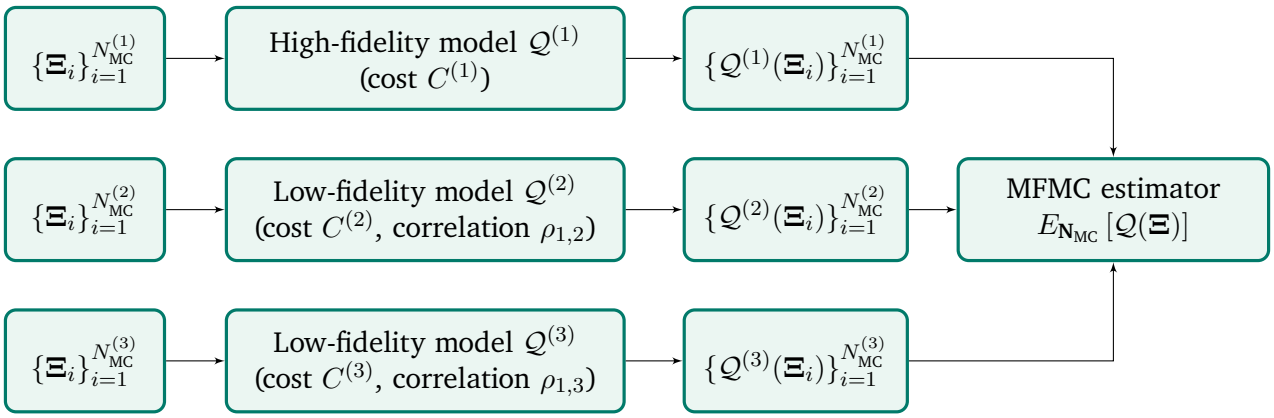


Figure 4.2: Illustration of MFMC method: the approach relies on a family of models with different computational costs $C^{(i)}$. Each model is evaluated separately with a portion of the input sample and the corresponding output samples are then combined for the multifidelity estimator.

computed in post-processing. In particular, stochastic moments or sensitivity indices can be derived directly from the coefficients, while other quantities may also be computed by sampling the inexpensive surrogate model.

In the following, we assume that the RVs Ξ_i, $i = 1, \dots, N_{\boldsymbol{\xi}}$ are mutually independent, as explained in Section 2.3.3. In this thesis, we will address polynomial approximations (4.18) based either on GPC [23, 66, 143, 213] or on sparse grid interpolation [10, 14, 40, 53, 149, 155, 179, 212], as explained in the following.

4.1.3.1 Generalized polynomial chaos

We start by recalling the standard polynomial chaos, which goes back to Wiener [205]. If one considers $\boldsymbol{\Xi}$ as Gaussian RVs, any $\mathcal{Q}(\boldsymbol{\Xi})$, where $\mathbb{V}[\mathcal{Q}(\boldsymbol{\Xi})] < \infty$, can be accurately represented in the form of (4.18) for $M \to \infty$ by using Hermite polynomials as basis functions Ψ_m, see [49]. GPC expansions are then obtained using the Askey-Scheme [213], which yields for different PDFs $\rho(\boldsymbol{\xi})$, suitable basis functions $\Psi_m : \Gamma \to \mathbb{R}$. In particular, these GPC basis function are orthogonal w.r.t. $\rho(\boldsymbol{\xi})$, i.e.

$$\mathbb{E}[\Psi_i(\boldsymbol{\Xi})\Psi_j(\boldsymbol{\Xi})] := \int_\Gamma \Psi_i(\boldsymbol{\xi})\Psi_j(\boldsymbol{\xi})\rho(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} = \mathbb{E}\left[\Psi_i^2\right]\delta_{ij}. \tag{4.19}$$

In the following, we further assume that the GPC basis function are normalized, i.e.

$$\mathbb{E}\left[\left(\Psi_m(\boldsymbol{\Xi})\right)^2\right] = 1,\ m = 0, \dots, M. \tag{4.20}$$

Note that GPC expansions can also be computed for arbitrary densities $\rho(\boldsymbol{\xi})$, see [190]. It was shown, that GPC converges for any square-integrable function in the norm

$$||\mathcal{Q}(\boldsymbol{\Xi}) - \mathcal{Q}_M(\boldsymbol{\Xi})||_{L^2_\rho}^2 := \mathbb{E}\left[\left(\mathcal{Q}(\boldsymbol{\Xi}) - \mathcal{Q}_M(\boldsymbol{\Xi})\right)^2\right] = \int_\Gamma (\mathcal{Q}(\boldsymbol{\xi}) - \mathcal{Q}_M(\boldsymbol{\xi})^2\rho(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}, \tag{4.21}$$

if any of the conditions specified in [74, Theorem 3.4] is fulfilled. In particular, the GPC basis is dense in L^2_ρ, for instance if normal RVs are considered or if the density $\rho(\boldsymbol{\xi})$ has compact support.

For practical computations, the sum in (4.18) is truncated to $M < \infty$ and limited polynomial degrees are considered. To this end, we introduce the multi-index $\mathbf{m} = [m_1, \ldots, m_{N_{\boldsymbol{\xi}}}] \in \mathbb{N}_0^{N_{\boldsymbol{\xi}}}$ which holds the polynomial degree per parameter and rewrite the GPC basis functions as

$$\Psi_{\mathbf{m}}(\boldsymbol{\xi}) = \prod_{n=1}^{N_{\boldsymbol{\xi}}} \psi_{m_n}(\xi_n), \tag{4.22}$$

where $\psi_{m_n}(\xi_n)$ denote univariate GPC basis functions w.r.t. $\rho_n(\xi_n)$. Note that, although not explicitly specified, there is a one-to-one map between the scalar indices m and the multi-indices $\mathbf{m}$. Typically, a total-degree polynomial basis is used, consisting of all polynomials $\Psi_{\mathbf{m}} : \mathbf{m} \in \Lambda^{\mathrm{TD}}_{m_{\max}}$, where

$$\Lambda^{\mathrm{TD}}_{m_{\max}} \coloneqq \{\mathbf{m} \,:\, |\mathbf{m}| = m_1 + \ldots + m_{N_{\boldsymbol{\xi}}} \leq m_{\max}\}. \tag{4.23}$$

In this case, the number of terms in (4.18) is $M + 1 = \frac{(N_{\boldsymbol{\xi}} + m_{\max})!}{N_{\boldsymbol{\xi}}!\, m_{\max}!} + 1$ and the GPC approximation can be represented as

$$\mathcal{Q}(\boldsymbol{\xi}) \approx \mathcal{Q}_M(\boldsymbol{\xi}) \coloneqq \sum_{\mathbf{m} \in \Lambda^{\mathrm{TD}}_{m_{\max}}} s_{\mathbf{m}} \Psi_{\mathbf{m}}(\boldsymbol{\xi}). \tag{4.24}$$

The coefficients $s_{\mathbf{m}}$ can be determined non-intrusively, e.g. by regression, collocation or projection, see [211] for details. In this work, we rely on projection to compute the GPC coefficients $s_{\mathbf{m}}$, i.e.

$$s_{\mathbf{m}} = \mathbb{E}\left[\mathcal{Q}(\boldsymbol{\Xi}) \Psi_{\mathbf{m}}(\boldsymbol{\Xi})\right] = \int_{\Gamma} \mathcal{Q}(\boldsymbol{\xi}) \Psi_{\mathbf{m}}(\boldsymbol{\xi}) \rho(\boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi}. \tag{4.25}$$

In particular, we employ pseudo-spectral projection, i.e. we approximate the integral (4.25), which is usually not readily computable, by numerical quadrature.

Stochastic moments can then be directly calculated from the coefficients $s_{\mathbf{m}}$, due to the orthogonality of the basis. For the expectation, one obtains

$$\mathbb{E}\left[\mathcal{Q}_M(\boldsymbol{\Xi})\right] = \int_{\Gamma} \sum_{\mathbf{m} \in \Lambda^{\mathrm{TD}}_{m_{\max}}} s_{\mathbf{m}} \Psi_{\mathbf{m}}(\boldsymbol{\xi}) \rho(\boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi} = \sum_{\mathbf{m} \in \Lambda^{\mathrm{TD}}_{m_{\max}}} s_{\mathbf{m}} \int_{\Gamma} \Psi_{\mathbf{m}}(\boldsymbol{\xi}) \underbrace{\Psi_{\mathbf{0}}(\boldsymbol{\xi})}_{1} \rho(\boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi} = s_{\mathbf{0}}, \tag{4.26}$$

by exploiting that $\Psi_{\mathbf{0}}(\boldsymbol{\xi}) = 1$, due to (4.20), and the orthogonality condition (4.19). Accordingly, the variance is obtained as

$$\mathbb{V}\left[\mathcal{Q}_M(\boldsymbol{\Xi})\right] = \mathbb{E}\left[\mathcal{Q}_M^2(\boldsymbol{\Xi})\right] - \mathbb{E}\left[\mathcal{Q}_M(\boldsymbol{\Xi})\right]^2 = \sum_{\mathbf{m} \in \Lambda^{\mathrm{TD}}_{m_{\max}} \setminus \{\mathbf{0}\}} s_{\mathbf{m}}^2. \tag{4.27}$$

4.1.3.2 Sparse-Grid Stochastic Collocation

Next, we address approximations based on sparse grid interpolation which are often called sparse grid SC methods [10, 155, 212]. Sparse grid SC is based on combinations of sets of $m_{\mathrm{l2n}}(\ell_n)$ univariate interpolation nodes

$$Z_{\ell_n} = \left\{ \xi_n^{(i_n)} \right\}_{i_n=0}^{m_{\mathrm{l2n}}(\ell_n)-1} \subset \Gamma_n, \; n = 1, \ldots, N_{\boldsymbol{\xi}}, \tag{4.28}$$

where $\ell_n \in \mathbb{N}_0$ denotes the interpolation level and $m_{\mathrm{l2n}} : \mathbb{N}_0 \to \mathbb{N}$ denotes a monotonically increasing "level-to-nodes" function. Introducing the tensor grid $Z^{\mathrm{TP}}_{\boldsymbol{\ell}} = Z_{\ell_1} \times \cdots \times Z_{\ell_N}$ of multivariate interpolation nodes

$\boldsymbol{\xi}^{(\mathbf{i})} = \left(\xi_1^{(i_1)}, \cdots, \xi_{N_{\boldsymbol{\xi}}}^{(i_{N_{\boldsymbol{\xi}}})}\right) \in Z_{\boldsymbol{\ell}}$, which can be uniquely identified by the multi-index $\mathbf{i} = \left(i_1, i_2, \ldots, i_{N_{\boldsymbol{\xi}}}\right) \in \mathbb{N}_0^{N_{\boldsymbol{\xi}}}$, the corresponding tensor-product multivariate approximation reads

$$\mathcal{Q}\left(\boldsymbol{\xi}\right) \approx \mathcal{Q}_{Z_{\boldsymbol{\ell}}^{\mathrm{TP}}}\left(\boldsymbol{\xi}\right) \coloneqq \sum_{\mathbf{i}:\boldsymbol{\xi}^{(\mathbf{i})} \in Z_{\boldsymbol{\ell}}^{\mathrm{TP}}} \mathcal{Q}(\boldsymbol{\xi}^{(\mathbf{i})}) L_{\boldsymbol{\ell},\mathbf{i}}\left(\boldsymbol{\xi}\right), \tag{4.29}$$

where $L_{\boldsymbol{\ell},\mathbf{i}} : \Gamma \to \mathbb{R}$ are multivariate Lagrange polynomials

$$L_{\boldsymbol{\ell},\mathbf{i}}\left(\boldsymbol{\xi}\right) = \prod_{n=1}^{N_{\boldsymbol{\xi}}} l_{\ell_n,i_n}\left(\xi_n\right), \quad \text{where } l_{\ell_n,i_n}(y_n) \coloneqq \begin{cases} \prod_{k=0,k\neq i_n}^{m_{\mathrm{l2n}}(\ell_n)-1} \frac{\xi_n - \xi_n^{(k)}}{\xi_n^{(i_n)} - \xi_n^{(k)}}, & \ell_n \neq 0, \\ 1, & \ell_n = 0. \end{cases} \tag{4.30}$$

Note that (4.30) is used for readability, while the barycentric representation should be used in the actual implementation [21]. The computational model has to be evaluated for all $\#Z_{\boldsymbol{\ell}}^{\mathrm{TP}} = \prod_{i=1}^{N_{\boldsymbol{\xi}}} m_{\mathrm{l2n}}(\ell_i)$ nodes $\boldsymbol{\xi}^{(\mathbf{i})} \in Z_{\boldsymbol{\ell}}^{\mathrm{TP}}$, resulting in a computational complexity of $\mathcal{O}\left(m_{\max}^{N_{\boldsymbol{\xi}}}\right)$, where $m_{\max} = \max_{1\leq n\leq N_{\boldsymbol{\xi}}} m_{\mathrm{l2n}}(\ell_n)$. Hence, the tensor-product approach is often intractable, even if the number of parameters $N_{\boldsymbol{\xi}}$ is relatively small. This is sometimes referred to as *curse-of-dimensionality* [18].

Smolyak sparse grids offer the possibility to reduce the complexity to $\mathcal{O}\left(m_{\max}\left(\log m_{\max}\right)^{N_{\boldsymbol{\xi}}-1}\right)$ while only mildly harming the accuracy of the approximation under certain regularity assumptions [40, 188]. The Smolyak sparse grid Z_l^{Smolyak} of multivariate interpolation nodes w.r.t. the approximation level $l \in \mathbb{N}_0$, can be constructed as a combination of (anisotropic) tensor grids $Z_{\boldsymbol{\ell}}^{\mathrm{TP}}$

$$Z_l^{\mathrm{Smolyak}} = \bigcup_{l-N_{\boldsymbol{\xi}}+1\leq|\boldsymbol{\ell}|\leq l} Z_{\boldsymbol{\ell}}^{\mathrm{TP}}. \tag{4.31}$$

Accordingly, the corresponding Smolyak approximation formula can be viewed as a linear combination of different tensor grid interpolation formulas (4.29), see [40] for details. Note that the Smolyak grid (4.31) is isotropic, however, the approach can be generalized to construct anisotropic sparse grids, see e.g. [152]. Furthermore, Smolyak grids are particularly efficient and become interpolatory [14] if they are based on nested sequences of univariate interpolation nodes, i.e. $Z_{\ell_n-1} \subset Z_{\ell_n}$. Hence, a common choice are Clenshaw Curtis nodes [56, 157] with the (rather rapidly growing) "level-to-node" function

$$m_{\mathrm{l2n}}^{\mathrm{CC}}(i) = 2^i + 1, \tag{4.32}$$

for instance. An alternative choice, weighted Leja nodes, was proposed in [149] where the (monotonic) level-to-node function m_{l2n} can be arbitrarily chosen and which can be tailored to any given PDF $\rho_n(\xi_n)$. In particular, following [149], the univariate nodes $\xi_n^{(k)} \in \Gamma_n$, $k = 0, 1, 2, \ldots$ are constructed as

$$\xi_n^{(K)} = \underset{\xi_n\in\Gamma_n}{\arg\min} \sqrt{\rho_n\left(\xi_n\right)} \prod_{k=0}^{K-1} \left|\xi_n - \xi_n^{(k)}\right|, \tag{4.33}$$

with arbitrary starting node $\xi_n^{(0)} \in \Gamma_n$. Note that the nodes (4.33) are fully nested by construction. Finally, we outline that the anisotropic sparse-grids can also be constructed adaptively in order to detect sensitive parameters ξ_n automatically, see e.g. [52, 53, 91, 148, 179]. More details about the dimension-adaptive scheme will be given in the subsequent sections, where we propose generalizations/extensions of the algorithm.

Finally, we note that, although the stochastic moments of a polynomial SC approximation can usually not be obtained directly from the coefficients as in the GPC case, they can still be calculated exactly, for instance by a Gaussian quadrature rule of suitable order.

4.1.4 Sensitivity analysis

Sensitivity analysis studies how the individual uncertain input parameters $\Xi_i, i = 1, \dots, N_{\boldsymbol{\xi}}$ affect the uncertain output quantity $\mathcal{Q}(\boldsymbol{\Xi})$ which gives important insights into the model behavior [31]. Note that sensitivities are often estimated by computing the gradient of the QoI at the nominal parameter values. However, this approach can lead to high errors, particularly for strongly non-linear models, as it only provides local information. Therefore, we rely in this work on methods for global sensitivity analysis [174] which consider the whole input space Γ. Sobol indices [175] are global sensitivity indices that are well established and based on the assumption of independent inputs $\boldsymbol{\Xi}$. Borgonovo indices [30] provide an alternative without this requirement, however, they have received less attention in engineering so far. In the following, we recall the key ideas of both approaches.

Sobol indices are based on an analysis of variances (ANOVA) [189], i.e. on the idea of decomposing the variance $\mathbb{V}\left[\mathcal{Q}(\boldsymbol{\Xi})\right]$. Here, the ANOVA decomposition can be obtained from the GPC approximation (4.24). The Sobol indices of the GPC approximation can be directly obtained without further approximations. In particular, we will first address the so-called main-effect (first-order) indices which indicate the individual contribution of each input parameter Ξ_i to the total variance. To this end, we introduce the multi-index sets

$$\Lambda_i^{\text{main}} \coloneqq \{\mathbf{m} \in \Lambda_{m_{\max}}^{\text{TD}} \,:\, m_i \neq 0 \,\wedge\, m_j = 0,\, j \neq i\}, \quad i = 1, \dots, N_{\boldsymbol{\xi}}, \tag{4.34}$$

and the corresponding partial variances

$$\mathbb{V}_i^{\text{main}}[\mathcal{Q}_M(\boldsymbol{\Xi})] \coloneqq \sum_{\mathbf{m} \in \Lambda_i^{\text{main}}} s_{\mathbf{m}}^2, \quad i = 1, \dots, N_{\boldsymbol{\xi}}. \tag{4.35}$$

The main-effect Sobol indices are then given as

$$S_i^{\text{main}}[\mathcal{Q}_M(\boldsymbol{\Xi})] \coloneqq \frac{\mathbb{V}_i^{\text{main}}[\mathcal{Q}_M(\boldsymbol{\Xi})]}{\mathbb{V}\left[\mathcal{Q}_M(\boldsymbol{\Xi})\right]}, \quad i = 1, \dots, N_{\boldsymbol{\xi}}. \tag{4.36}$$

In addition, one can define the total-effect (total-order) Sobol indices [106] which take additional higher order effects, i.e. interactions with other parameters, into account. The corresponding multi-index sets read

$$\Lambda_i^{\text{total}} \coloneqq \{\mathbf{m} \in \Lambda_{m_{\max}}^{\text{TD}} \,:\, m_i \neq 0\}, \quad i = 1, \dots, N_{\boldsymbol{\xi}}, \tag{4.37}$$

such that the total-effect indices are then given as

$$S_i^{\text{total}}[\mathcal{Q}_M(\boldsymbol{\Xi})] \coloneqq \frac{\mathbb{V}_i^{\text{total}}[\mathcal{Q}_M(\boldsymbol{\Xi})]}{\mathbb{V}\left[\mathcal{Q}_M(\boldsymbol{\Xi})\right]}, \text{ where } \mathbb{V}_i^{\text{total}}[\mathcal{Q}_M(\boldsymbol{\Xi})] \coloneqq \sum_{\mathbf{m} \in \Lambda_i^{\text{total}}} s_{\mathbf{m}}^2, \quad i = 1, \dots, N_{\boldsymbol{\xi}}. \tag{4.38}$$

We note that, by construction, the Sobol indices fulfill the relation $0 \leq \sum_{i=1}^{N_{\boldsymbol{\xi}}} S_i^{\text{main}} \leq 1 \leq \sum_{i=1}^{N_{\boldsymbol{\xi}}} S_i^{\text{total}}$. We further note that, although computationally efficient, the GPC approximation is by no means necessary to compute Sobol indices. In particular, Saltelli's algorithm provides an MC based sampling approach as an alternative, see [174] for details.

Sobol indices, as any variance based method, are based on the implicit assumption that considering only the variance, i.e. one particular stochastic moment, adequately captures the output uncertainty [173]. This issue was addressed by Borgonovo [30] who introduced a moment-independent importance measure for the input parameters based on the individual effects on the output PDF. Furthermore, these Borgonovo indices do not require that the inputs $\boldsymbol{\Xi}$ are statistically independent.

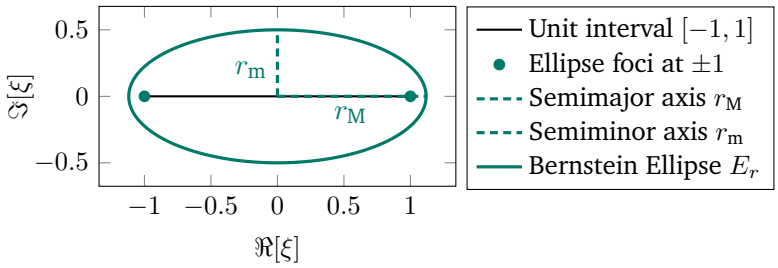


Figure 4.3: Illustration of Bernstein ellipse E_r where the subscript r indicates the size $r = r_\mathrm{M} + r_\mathrm{m}$, based on [89, Figure 1].

Let the unconditional PDF of the output distribution, i.e. all inputs are considered as RVs, be denoted as $\rho_{\mathcal{Q}}$. We then introduce the conditional PDFs $\rho_{\mathcal{Q}|\Xi_i}$, $i = 1, \dots, N_{\boldsymbol{\xi}}$ which are obtained by removing the uncertainty in Ξ_i and setting the input to a certain fixed value. The difference between the PDFs of the output distribution is quantified by the shift functions

$$s_i(\Xi_i) = \int_{\Gamma_\mathcal{Q}} |\rho_\mathcal{Q}(q) - \rho_{\mathcal{Q}|\Xi_i}(q)| \, \mathrm{d}q, \quad i = 1, \dots, N_{\boldsymbol{\xi}}, \tag{4.39}$$

where $\Gamma_\mathcal{Q}$ denotes the image set of the uncertain QoI $\mathcal{Q}(\boldsymbol{\Xi})$. Borgonovo sensitivity indices are then obtained as the normalized expected shift which is caused by Ξ_i as

$$\delta_i = \frac{1}{2}\mathbb{E}_{\Xi_i}[s_i(\Xi_i)] = \frac{1}{2}\int_{\Gamma_i} \rho_i(\xi_i) \left[\int_{\Gamma_\mathcal{Q}} |\rho_\mathcal{Q}(q) - \rho_{\mathcal{Q}|\Xi_i=\xi_i}(q)| \, \mathrm{d}q \right] \mathrm{d}\xi_i, \quad i = 1, \dots, N_{\boldsymbol{\xi}}. \tag{4.40}$$

Among others, the following properties are proven in [30]: The Borgonovo indices δ_i fulfill $0 \leq \delta_i \leq 1$. Furthermore, a Borgonovo index assumes zero, $\delta_i = 0$, if the QoI is independent of Ξ_i. For further details on dependence-measure based sensitivity indices, we refer to [31, 63].

4.2 Conformal mappings

The spectral methods introduced in Section 4.1.3 are already well established for surrogate modeling and UQ in many application fields, as the associated computational effort is often reduced in comparison to Monte Carlo techniques, for instance. However, further convergence acceleration might be required to address problems with large parameter uncertainties and high parametric sensitivities. Hence, in order to enhance the convergence order, we propose to construct transformed basis functions for the spectral UQ methods based on conformal maps. This section is structured as follows: We first recall some basic facts about the convergence of polynomial approximations for analytic functions before presenting the basic idea of using conformal maps to improve polynomial-based methods, which was first introduced in [102] in the context of quadrature methods. We then propose transformed basis functions for GPC [90] and SC [89] which preserve the advantages of the respective methods and discuss the implementation of the corresponding mapped spectral UQ approaches. Finally, the methods are applied to an academic model problem, i.e. an RLC circuit, with available closed-form solution s.t. the main ideas can be illustrated in detail. The content and structure of this section are based on our works [89, 90].

In this section, we assume mutually independent inputs $\boldsymbol{\Xi}$ and that the image set Γ is given as the hypercube $[-1, 1]^{N_{\boldsymbol{\xi}}}$, for simplicity. Note that this assumption would be immediately fulfilled if one employs, for instance,

the Rosenblatt transformation

$$\boldsymbol{\xi} \mapsto -[1,1,\dots,1]^\top + 2\mathbf{T}(\boldsymbol{\xi}), \tag{4.41}$$

where $\mathbf{T}$ is defined in (2.64). We first address the univariate case, i.e. $N_{\boldsymbol{\xi}} = 1$, in some detail before discussing the generalizations to the multivariate case $N_{\boldsymbol{\xi}} > 1$. In particular, we make in the following the assumption that the univariate function $\mathcal{Q}_{1\mathrm{D}} : \Gamma_{1\mathrm{D}} \coloneqq [-1,1] \to \mathbb{C}$ has an analytic continuation onto an open Bernstein ellipse $E_r \subset \mathbb{C}$ with foci at ± 1. The Bernstein ellipse E_r is illustrated in Fig. 4.3, where the subscript r refers to the size of the ellipse. The size r is defined as $r = r_\mathrm{m} + r_\mathrm{M}$ where $r_\mathrm{m}, r_\mathrm{M}$ denote the length of semi-minor and semi-major axis, respectively. Then, the polynomial best approximation $\mathcal{Q}_M^*$ of degree M converges exponentially, as

$$\|\mathcal{Q}_{1\mathrm{D}} - \mathcal{Q}_M^*\|_\infty \leq \frac{C_\mathrm{B} e^{-\log(r)M}}{r-1}, \tag{4.42}$$

where $\|\cdot\|_\infty$ refers to the supremum norm

$$\|u\|_\infty = \sup_{\xi \in \Gamma_{1\mathrm{D}}} |u(\xi)|, \tag{4.43}$$

and the constant C_B depends on the uniform bound of the analytic extension of $\mathcal{Q}_{1\mathrm{D}}$ onto E_r , see, e.g. [199, Theorem 8.2]. We emphasize that (4.42) also guarantees convergence in the $||\cdot||_{L^2_\rho}$ norm for any bounded PDF $\rho_{1\mathrm{D}}$, as

$$\begin{aligned} ||\mathcal{Q}_{1\mathrm{D}}(\Xi) - \mathcal{Q}_M^*(\Xi)||_{L^2_\rho} &= \left(\int_{[-1,1]} \big(\mathcal{Q}_{1\mathrm{D}}(\xi) - \mathcal{Q}_M^*(\xi)\big)^2 \rho_{1\mathrm{D}}(\xi)\,\mathrm{d}\xi \right)^{\frac{1}{2}} \\ &\leq \|\sqrt{\rho_{1\mathrm{D}}}\|_\infty \, \|\mathcal{Q}_{1\mathrm{D}} - \mathcal{Q}_M^*\|_\infty \left(\int_{[-1,1]} 1\,\mathrm{d}\xi \right)^{\frac{1}{2}} \\ &= \sqrt{2}\, \|\sqrt{\rho_{1\mathrm{D}}}\|_\infty \, \|\mathcal{Q}_{1\mathrm{D}} - \mathcal{Q}_M^*\|_\infty. \end{aligned} \tag{4.44}$$

The convergence estimate (4.42) is derived for the polynomial best approximation $\mathcal{Q}_M^*$, however, it is closely related to the convergence of the spectral UQ methods introduced in Section 4.1.3. In particular, regarding the GPC approximation, it is stated for instance in [213, Chapter 3.6] that the additional aliasing error caused by the pseudo-spectral projection is not harming the convergence order for well-resolved smooth functions. Regarding the SC method, we consider a polynomial interpolant $\mathcal{Q}_{Z_{1\mathrm{D}}}$ of order $M = \#Z_{1\mathrm{D}} - 1$ in the form of (4.29) where $Z_{1\mathrm{D}} = \{\xi^{(i)}\}_{i=0}^M$ is a set of $\#Z_{1\mathrm{D}}$ univariate distinct interpolation nodes. Then, there holds

$$\|\mathcal{Q}_{1\mathrm{D}} - \mathcal{Q}_{Z_{1\mathrm{D}}}\|_\infty \leq (1 + \Delta_M)\|\mathcal{Q}_{1\mathrm{D}} - \mathcal{Q}_M^*\|_\infty \leq (1+\Delta_M)\frac{C_\mathrm{B} e^{-\log(r)M}}{r-1}, \tag{4.45}$$

where

$$\Delta_M \coloneqq \max_{\xi \in [-1,1]} \sum_{i=0}^M |l_{M,i}(\xi)|, \tag{4.46}$$

denotes the Lebesgue constant. In the above expression $l_{M,i}(\xi)$ refers to the univariate Lagrange polynomial defined in (4.30) with $m_{\mathrm{l2n}} : n \mapsto n$. If the Lebesgue constant Δ_M grows sub-exponentially, it can be seen that the polynomial interpolant $\mathcal{Q}_{Z_{1\mathrm{D}}}$ converges uniformly if $r > 1$, i.e. for analytic functions.

In summary, (4.42) shows that the asymptotic rate of geometric convergence [33, Definition 6] for the polynomial best approximation is given as $\log r$ and hence, the convergence of spectral UQ methods as GPC

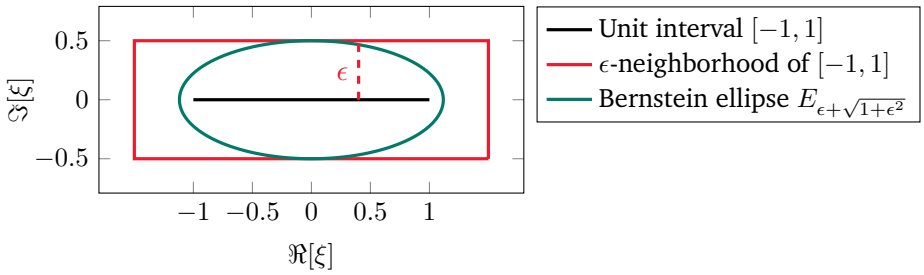


Figure 4.4: Illustration of ϵ-neighborhood and largest Bernstein ellipse which is fully contained, based on [89, Figure 3].

and SC depends on the region of analyticity of the function $\mathcal{Q}_{\mathrm{1D}}$. In particular, it depends on the size $r_{\max}$ of the largest Bernstein ellipse which does not contain any poles of the extension of $\mathcal{Q}_{\mathrm{1D}}$. However, established procedures [10] which infer the regularity based on derivatives of the parametric problem prove analyticity in an ϵ-neighborhood of Γ_{1D} instead of elliptical regions. Such an $\epsilon-$neighborhood is depicted in Fig. 4.4 as well as the largest Bernstein ellipse which it fully contains. To address this mismatch and enlarge the region of analyticity, one can employ a conformal mapping g to map Bernstein ellipses E_r to *straighter* domains $g(E_r)$, as shown in Fig. 4.5. This approach was originally introduced in [102] to derive new quadrature formulas. The acceleration of quadrature methods using conformal maps was further studied in [101, 113, 199]. In this work, we show how it can also be employed to improve spectral surrogate modeling and UQ methods, in particular GPC [90] and (adaptive) SC [89]. As in [102], we only consider conformal mappings which fulfill

$$g(\Gamma_{\mathrm{1D}}) = \Gamma_{\mathrm{1D}}, \tag{4.47a}$$
$$g(\pm 1) = \pm 1. \tag{4.47b}$$

The condition (4.47a) guarantees that the resulting methods will only require model evaluations for input values which are elements of the image set Γ_{1D}. This property is highly relevant, as, for instance, evaluating the analytic continuation is often not possible in practice. There are various mappings that have been proposed and fulfill the intended properties, see [101, 102] for instance. In this thesis, we mainly rely on the so-called sausage mapping which was proposed in [102]. It is derived by normalization of the Taylor expansion of the inverse sine function such that (4.47b) is fulfilled. The d-th order sausage mapping then reads

$$g_{\mathrm{S}}(\xi; d) = \left(\sum_{i=0}^{\lfloor (d-1)/2 \rfloor} \frac{(2i)!}{4^i (2i+1)(i!)^2} \right)^{-1} \sum_{i=0}^{\lfloor (d-1)/2 \rfloor} \frac{(2i)!}{4^i (2i+1)(i!)^2} \xi^{2i+1}. \tag{4.48}$$

An alternative choice is given by the mapping

$$g_{\mathrm{KTE}}(\xi; \alpha) = \frac{\arcsin(\alpha \xi)}{\arcsin \alpha}, \ \alpha \in (0, 1) \tag{4.49}$$

from Kosloff and Tal-Ezer [122]. In the following, we first discuss how the conformal maps g can be combined with GPC before addressing the mapped SC approach.

4.2.1 Conformally mapped generalized polynomial chaos

In this subsection, we propose a transformed basis by combining GPC and conformal maps which still features the orthogonality property (4.19) such that the advantages of GPC are preserved. For instance, it still

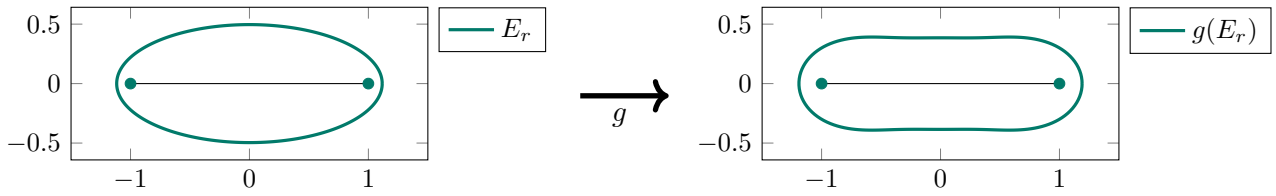


Figure 4.5: A conformal map is applied to a Bernstein ellipse E_r which is mapped to a *straighter* region $g(E_r)$. The illustration is based on [89, Figure 4].

facilitates the straightforward computation of stochastic moments and Sobol indices. Furthermore, we employ pseudo-spectral projection based on mapped quadrature rules to compute the respective surrogate models with an improved cost accuracy ratio. Note that alternative approaches which also rely on mapped GPC, i.e. GPC expansions with basis rotation, were recently studied, see [161, 200]. However, these works only consider affine-linear transformations and address high-dimensional approximations instead of convergence acceleration.

As introduced in the last subsection, starting from the assumption of analyticity of $\mathcal{Q}_{1\mathrm{D}}$ in an ϵ-neighborhood of $\Gamma_{1\mathrm{D}}$, we then assume that the composition

$$h := \mathcal{Q}_{1\mathrm{D}} \circ g : \Gamma_{1\mathrm{D}} \to \mathbb{C}, \tag{4.50}$$

has an enlarged Bernstein ellipse of size $\hat{r} > r$ which is advantageous for polynomial approximations, see (4.42). Hence, we suggest the basis functions

$$\hat{\Psi}_m := \tilde{\Psi}_m \circ g^{-1}, \quad m = 0, \dots, M, \tag{4.51}$$

where $\tilde{\Psi}_m$ denote normalized GPC basis functions w.r.t. the transformed density

$$\tilde{\rho}_{1\mathrm{D}}(s) := g'(s)\rho_{1\mathrm{D}}(g(s)). \tag{4.52}$$

Note that the suggested basis $\{\hat{\Psi}_m\}_{m=0}^M$ is then orthonormal w.r.t. the original input PDF $\rho_{1\mathrm{D}}$ which can be shown using a change of variables $\xi = g(s)$, as

$$\mathbb{E}\left[\hat{\Psi}_i(\Xi)\hat{\Psi}_j(\Xi)\right] = \int_{\Gamma_{1\mathrm{D}}} (\tilde{\Psi}_i \circ g^{-1})(\xi)(\tilde{\Psi}_j \circ g^{-1})(\xi)\rho_{1\mathrm{D}}(\xi)\,\mathrm{d}\xi, \tag{4.53}$$

$$= \int_{\Gamma_{1\mathrm{D}}} \tilde{\Psi}_i(s)\tilde{\Psi}_j(s) \underbrace{\rho_{1\mathrm{D}}(g(s))g'(s)}_{\tilde{\rho}_{1\mathrm{D}}(s)}\,\mathrm{d}s = \delta_{ij}, \tag{4.54}$$

where the last line is fulfilled by construction of the GPC basis $\{\tilde{\Psi}_m\}_{m=0}^M$. Exploiting the orthogonality, the coefficients $\hat{s}_m$ of the conformally mapped GPC approximation

$$\hat{Q}_M(\xi) = \sum_{m=0}^{M} \hat{s}_m \hat{\Psi}_m(\xi), \tag{4.55}$$

can be computed by projection as

$$\hat{s}_m = \mathbb{E}\left[\hat{\Psi}_m(\Xi)Q_{1\mathrm{D}}(\Xi)\right] = \int_{\Gamma_{1\mathrm{D}}} \hat{\Psi}_m(\xi)Q_{1\mathrm{D}}(\xi)\rho_{1\mathrm{D}}(\xi)\,\mathrm{d}\xi. \tag{4.56}$$

Introducing the symbol $\hat{\mathcal{Q}}_M^*$ for the best approximation in terms of the mapped polynomials $\{\hat{\Psi}_m\}_{m=0}^M$, its error can then be estimated as

$$\|\mathcal{Q}_{1\mathrm{D}} - \hat{\mathcal{Q}}_M^*\|_\infty = \|\mathcal{Q}_{1\mathrm{D}} \circ g - \hat{\mathcal{Q}}_M^* \circ g\|_\infty = \|h - h_M^*\|_\infty \leq \frac{\hat{C}_B e^{-\log(\hat{r})M}}{\hat{r} - 1}, \tag{4.57}$$

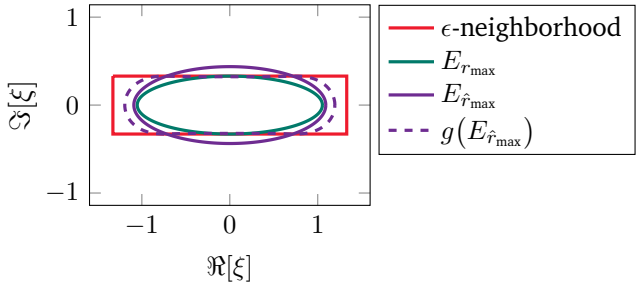
where h_M^* refers to the polynomial best approximation of the function h, defined in (4.50), which admits an analytic continuation to a Bernstein ellipse $E_{\hat{r}}$ where it is uniformly bounded. It can be seen that the mapped approximation has an asymptotic rate of geometric convergence of $\log \hat{r}$ where $\hat{r}$ refers to the size of the largest Bernstein ellipse $E_{\hat{r}_{\max}}$ such that $g(E_{\hat{r}_{\max}})$ is still fully contained in the region of analyticity of $\mathcal{Q}_{1\mathrm{D}}(\xi)$. Hence, if the analyticity region is known, the convergence improvement, which can be attributed to the use of conformal maps, can be estimated a priori which is illustrated in Fig. 4.6a. In particular, we denote the largest Bernstein ellipse in the interior by $E_{r_{\max}}$. Then, the convergence of the polynomial best approximation is given by $\mathcal{O}\big(\exp\left(-\log(r_{\max})M\right)\big)$, see (4.42), while (4.57) shows that the best approximation using mapped polynomials converges as $\mathcal{O}\big(\exp\left(-\log(\hat{r}_{\max})M\right)\big)$. Accordingly, we can define

$$G = \frac{\log \hat{r}_{\max}}{\log r_{\max}} - 1, \tag{4.58}$$

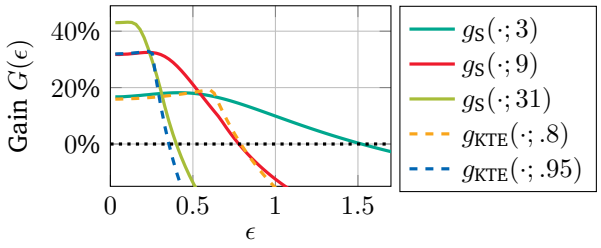
as the relative improvement in the asymptotic rate of geometric convergence [33, Definition 6] due to the use of conformal maps. To study the effect of different mappings g, we compute the gain G by considering different ϵ-neighborhoods as analyticity regions, where we numerically compute $\hat{r}_{\max}$ for the sausage mapping $g_{\mathrm{S}}(\cdot, d)$ of different orders d and the KTE mapping which are defined in (4.48) and (4.49), respectively. The results are shown in Fig. 4.6b. Note that even higher improvements could be expected if one would employ mappings which are specifically constructed based on the exact locations of the poles of $\mathcal{Q}_{1\mathrm{D}}$ in the complex plane, see [101]. However, as, in practice, the poles are usually not known a priori, this approach is not considered in this thesis. In the following, we only employ the 9-th order sausage mapping $g_{\mathrm{S}}(\xi; 9)$ which leads to a substantial convergence improvement for a significant range of ϵ-neighborhoods, as shown in Fig. 4.6b. In particular, it can be seen that $\hat{r}_{\max} > r_{\max}$ for any positive $\epsilon < 0.75$. Furthermore, it was already established in [101, 102] and does not introduce any artificial singularities, in contrast to the KTE map (4.49). Since the inverse mapping $g_{\mathrm{S}}(\xi; 9)^{-1}$ which is required for the mapped basis functions (4.51) is not known analytically, it is approximated numerically up to machine precision using a Chebyshev approximation of degree 100. However, it should be noted that the chosen mapping $g_{\mathrm{S}}(\xi; 9)$ does not necessarily guarantee an improved convergence. In particular, one could also imagine a function where the poles in the complex plane are located such that a Bernstein ellipse better fits the region of analyticity than a strip/sausage-like shape. However, we also note that for the numerical examples investigated in this thesis an improved convergence rate was indeed observed.

We now proceed with the numerical computation of the mapped GPC approximation (4.56). To this end, mapped quadrature rules are derived, cf. [101, 102]. Note that standard Gaussian quadrature is constructed based on polynomial approximations and, hence, its convergence properties also depend on the regularity of the integrand. For instance, in [102, Theorem 1] it is shown that the geometric convergence order of Gaussian quadrature for analytic functions also depends on the size of the Bernstein ellipse associated to the integrand. Hence, we apply a change of variables $\xi = g(s)$ in (4.56)

$$\hat{s}_m = \mathbb{E}\left[\hat{\Psi}_m(\Xi)\mathcal{Q}_{1\mathrm{D}}(\Xi)\right] = \int_{\Gamma_{1\mathrm{D}}} \hat{\Psi}_m(\xi)\mathcal{Q}_{1\mathrm{D}}(\xi)\rho_{1\mathrm{D}}(\xi)\,\mathrm{d}\xi = \int_{\Gamma_{1\mathrm{D}}} \hat{\Psi}_m(g(s))\mathcal{Q}_{1\mathrm{D}}(g(s))\underbrace{\rho_{1\mathrm{D}}(g(s))g'(s)}_{\tilde{\rho}_{1\mathrm{D}}}\,\mathrm{d}s, \tag{4.59}$$



(a) Illustration of geometric estimation of convergence gain G for one particular value of ϵ.



(b) Gain in convergence for different mappings, i.e. the sausage mapping $g_S(\cdot, d)$ of different orders d and the KTE map g_{KTE} defined in (4.49).

Figure 4.6: Investigation of the relative improvement in the asymptotic rate of geometric convergence G by employing mapped approximations for functions analytic in an ϵ-neighborhood, based on [89, Figure 7]

as we assume again that $h = \mathcal{Q}_{1D} \circ g$ has a larger Bernstein ellipse than $\mathcal{Q}_{1D}$. Let $\{\tilde{\xi}_q^{(i)}\}_{i=1}^{N_{quad}}$ denote the quadrature nodes and $\{\tilde{w}_q^{(i)}\}_{i=1}^{N_{quad}}$ the correspondings weights of Gaussian quadrature w.r.t. the transformed density $\tilde{\rho}_{1D}$. Then, application of Gaussian quadrature to the transformed integrand in (4.59) yields

$$\hat{s}_m \approx \sum_{i=1}^{N_{quad}} \hat{\Psi}_m\left(g\left(\tilde{\xi}_q^{(i)}\right)\right) \mathcal{Q}_{1D}\left(g\left(\tilde{\xi}_q^{(i)}\right)\right) \tilde{w}_q^{(i)} = \sum_{i=1}^{N_{quad}} \hat{\Psi}_m\left(\hat{\xi}_q^{(i)}\right) \mathcal{Q}_{1D}\left(\hat{\xi}_q^{(i)}\right) \hat{w}_q^{(i)}, \tag{4.60}$$

where $\hat{\xi}_q^{(i)} := g\left(\tilde{\xi}_q^{(i)}\right)$ are the mapped quadrature nodes corresponding to the mapped weights $\hat{w}_q^{(i)} := \tilde{w}_q^{(i)}$. We then expect an improved convergence of the mapped quadrature scheme, as we, again, assume that the transformed integrand in (4.59) is associated to a larger Bernstein ellipse, see [102] for details on mapped quadrature methods. However, it should be mentioned that [102] only addresses unweighted Gaussian quadrature and, thereby, considers $g'(s)$ as part of the integrand instead of the weight while we employ Gaussian quadrature w.r.t. the transformed density $\tilde{\rho}_{1D}$.

In the following, we address the generalization to the multivariate case $N_{\boldsymbol{\xi}} > 1$ based on a multivariate mapping $\mathbf{g}(\mathbf{s}) = [g_1(s_1), \ldots, g_N(s_N)]$ which is conformal in each parameter s_i. For simplicity, in this thesis we usually employ the same mapping for all coordinates, that is $g_1 = \ldots = g_N = g_S(\cdot, 9)$, however, this is not required. We also note that standard GPC would be recovered if one employs the identity map $\mathbf{g}_{id} : \mathbf{s} \mapsto \mathbf{s}$. Defining the univariate transformed PDFs

$$\tilde{\rho}_i(\xi_i) := \rho_i(g_i(\xi_i)) g_i'(\xi_i), \quad i = 1, \ldots, N_{\boldsymbol{\xi}}, \tag{4.61}$$

the associated transformed joint PDF reads

$$\tilde{\rho}(\boldsymbol{\xi}) = \prod_{i=1}^{N_{\boldsymbol{\xi}}} \tilde{\rho}_i(\xi_i). \tag{4.62}$$

Then, we consider a (normalized) multivariate GPC basis $\{\tilde{\Psi}_{\mathbf{m}}\}_{\mathbf{m}}$ w.r.t. $\tilde{\rho}$, where we use the same multi-index notation as in Section 4.1.3, that is, $\tilde{\Psi}_{\mathbf{m}}$ is a tensor-product polynomial of order m_j in dimension $j = 1, \ldots, N_{\boldsymbol{\xi}}$. The corresponding mapped polynomials are then given by

$$\hat{\Psi}_{\mathbf{m}}(\boldsymbol{\xi}) := (\tilde{\Psi}_{\mathbf{m}} \circ \mathbf{g}^{-1})(\boldsymbol{\xi}), \tag{4.63}$$

and the multivariate mapped approximation reads

$$\hat{\mathcal{Q}}_M(\boldsymbol{\xi}) := \sum_{\mathbf{m}\in\Lambda^{\mathrm{TD}}_{m_{\max}}} \hat{s}_{\mathbf{m}}\hat{\Psi}_{\mathbf{m}}(\boldsymbol{\xi}). \tag{4.64}$$

The coefficients $\hat{s}_{\mathbf{m}}$ can again be computed by projection

$$\hat{s}_{\mathbf{m}} = \mathbb{E}\left[\hat{\Psi}_{\mathbf{m}}(\Xi)\mathcal{Q}(\Xi)\right] = \int_{\Xi} \hat{\Psi}_{\mathbf{m}}(\boldsymbol{\xi})\mathcal{Q}(\boldsymbol{\xi})\rho(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}. \tag{4.65}$$

The multi-dimensional integral in (4.65) is numerically evaluated using a mapped Gaussian quadrature scheme with mapped nodes $\{\hat{\boldsymbol{\xi}}_{\mathrm{q}}^{(i)}\}_{i=1}^{N_{\mathrm{quad}}} := \{\mathbf{g}(\tilde{\boldsymbol{\xi}}_{\mathrm{q}}^{(i)})\}_{i=1}^{N_{\mathrm{quad}}}$ and associated weights $\{\hat{w}_{\mathrm{q}}^{(i)}\}_{i=1}^{N_{\mathrm{quad}}} := \{\tilde{w}_{\mathrm{q}}^{(i)}\}_{i=1}^{N_{\mathrm{quad}}}$, where, in turn, $\tilde{\boldsymbol{\xi}}_{\mathrm{q}}^{(i)}$ denotes the nodes and $\tilde{w}_{\mathrm{q}}^{(i)}$ denotes the weights of a standard Gaussian quadrature w.r.t. the transformed PDF $\tilde{\rho}$.

Note that the mapped GPC approximation (4.64) uses an orthogonal basis, and, hence, stochastic moments and Sobol indices can still be directly computed from the coefficients $\hat{s}_{\mathbf{m}}$. In particular, the mean value and variance are obtained as

$$\mathbb{E}\left[\hat{\mathcal{Q}}_M(\boldsymbol{\Xi})\right] = \hat{s}_{\mathbf{0}}, \qquad \mathbb{V}\left[\hat{\mathcal{Q}}_M(\boldsymbol{\Xi})\right] = \sum_{\mathbf{m}\in\Lambda^{\mathrm{TD}}_{m_{\max}}\setminus\{\mathbf{0}\}} \hat{s}_{\mathbf{m}}^2, \tag{4.66}$$

where we also employed that $\hat{\Psi}_{\mathbf{0}}(\boldsymbol{\xi}) = 1$. Then, the main-effect and total-effect Sobol indices, introduced in Section 4.1.4, can be computed as

$$S_i^{\mathrm{main}}[\hat{\mathcal{Q}}_M(\boldsymbol{\Xi})] := \frac{\mathbb{V}_i^{\mathrm{main}}[\hat{\mathcal{Q}}_M(\boldsymbol{\Xi})]}{\mathbb{V}\left[\hat{\mathcal{Q}}_M(\boldsymbol{\Xi})\right]}, \text{ where } \mathbb{V}_i^{\mathrm{main}}[\hat{\mathcal{Q}}_M(\boldsymbol{\Xi})] := \sum_{\mathbf{m}\in\Lambda_i^{\mathrm{main}}} \hat{s}_{\mathbf{m}}^2, \quad i = 1,\dots,N_{\boldsymbol{\xi}}, \tag{4.67}$$

and

$$S_i^{\mathrm{total}}[\hat{\mathcal{Q}}_M(\boldsymbol{\Xi})] := \frac{\mathbb{V}_i^{\mathrm{total}}[\hat{\mathcal{Q}}_M(\boldsymbol{\Xi})]}{\mathbb{V}\left[\hat{\mathcal{Q}}_M(\boldsymbol{\Xi})\right]}, \text{ where } \mathbb{V}_i^{\mathrm{total}}[\hat{\mathcal{Q}}_M(\boldsymbol{\Xi})] := \sum_{\mathbf{m}\in\Lambda_i^{\mathrm{total}}} \hat{s}_{\mathbf{m}}^2, \quad i = 1,\dots,N_{\boldsymbol{\xi}}, \tag{4.68}$$

respectively.

4.2.2 Conformally mapped stochastic collocation

Next, we discuss how the conformal map g can be employed in the context of SC methods. Once again, we first consider the univariate case $N_{\boldsymbol{\xi}} = 1$, where we now address the polynomial interpolant

$$\mathcal{Q}_{Z_{1\mathrm{D}}}(\xi) := \sum_{i=0}^{M} \mathcal{Q}_{1\mathrm{D}}\big(\xi^{(i)}\big) l_{M,i}(\xi), \tag{4.69}$$

of $\mathcal{Q}_{1\mathrm{D}} : \Gamma_{1\mathrm{D}} \to \mathbb{C}$ on a set $Z_{1\mathrm{D}} = \{\xi^{(i)}\}_{i=0}^{M} \subset \Gamma_{1\mathrm{D}}$ of univariate interpolation nodes, represented with univariate Lagrange polynomials $l_{M,i}(\xi)$. Note that $\mathcal{Q}_{Z_{1\mathrm{D}}}$ converges as estimated in (4.45). To accelerate the convergence, we compute the mapped interpolation nodes

$$\hat{Z}_{1\mathrm{D}} = \{\hat{\xi}^{(i)}\}_{i=0}^{M} = \{g\big(\xi^{(i)}\big)\}_{i=0}^{M} \subset \Gamma_{1\mathrm{D}}, \tag{4.70}$$

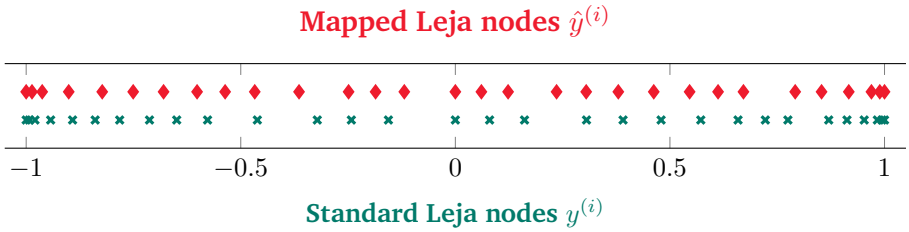


Figure 4.7: Comparison of $M + 1 = 30$ standard Leja and mapped Leja interpolation nodes, based on [89, Figure 5].

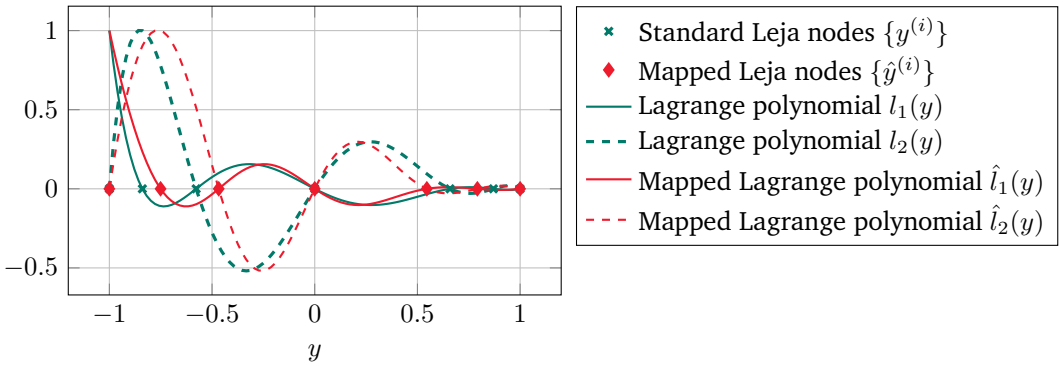


Figure 4.8: Comparison of standard and mapped Lagrange polynomials (degree $M = 6$), based on [89, Figure 6a].

which are more evenly distributed, see Fig. 4.7 for an illustration of (mapped) Leja nodes defined in (4.33). The mapped nodes $\{\hat{\xi}^{(i)}\}_{i=0}^{M}$ are then interpolated using mapped Lagrange polynomials

$$\hat{l}_{M,i} = l_{M,i} \circ g^{-1}, \tag{4.71}$$

which are illustrated in Fig. 4.8. Note that the mapped Lagrange polynomials $\hat{l}_{M,i}$ also fulfill

$$\hat{l}_{M,j}(\hat{\xi}^{(i)}) = l_{M,j} \circ g^{-1}(\hat{\xi}^{(i)}) = l_{M,j}(\xi^{(i)}) = \delta_{ij}, \tag{4.72}$$

and, hence, the mapped interpolant can be represented as

$$\hat{\mathcal{Q}}_{\hat{Z}_{1\mathrm{D}}}(\xi) := \sum_{i=0}^{M} \mathcal{Q}_{1\mathrm{D}}\big(\hat{\xi}^{(i)}\big)\hat{l}_{M,i}(\xi). \tag{4.73}$$

In order to derive an error estimate for the mapped approximation $\hat{\mathcal{Q}}_{\hat{Z}_{1\mathrm{D}}}$, we introduce the notation $h_{Z_{1\mathrm{D}}}$ for the M-th order polynomial interpolant of the function h defined in (4.50) on the nodes $Z_{1\mathrm{D}}$. Note that $\hat{\mathcal{Q}}_{\hat{Z}_{1\mathrm{D}}}$ equals $h_M \circ g^{-1}$ as

$$\hat{\mathcal{Q}}_{\hat{Z}_{1\mathrm{D}}} = \sum_{i=0}^{M} \mathcal{Q}_{1\mathrm{D}}\Big(g\big(\xi^{(i)}\big)\Big) l_{M,i} \circ g^{-1} = \left(\sum_{i=0}^{M} h\big(\xi^{(i)}\big) l_{M,i}\right) \circ g^{-1} = h_{Z_{1\mathrm{D}}} \circ g^{-1}. \tag{4.74}$$

We then obtain

$$||\mathcal{Q}_{1\mathrm{D}} - \hat{\mathcal{Q}}_{\hat{Z}_{1\mathrm{D}}}||_\infty = ||\big(\mathcal{Q}_{1\mathrm{D}} - \hat{\mathcal{Q}}_{\hat{Z}_{1\mathrm{D}}}\big) \circ g \circ g^{-1}||_\infty \tag{4.75a}$$

$$= ||(h - h_{Z_{1\mathrm{D}}}) \circ g^{-1}||_\infty \tag{4.75b}$$

$$= ||h - h_{Z_{1\mathrm{D}}}||_\infty \tag{4.75c}$$

$$\leq (1 + \Delta_M)||h - h_M^*||_\infty \tag{4.75d}$$

$$\leq (1 + \Delta_M)\frac{\hat{C}_{\mathrm{B}} e^{-\log(\hat{r})M}}{\hat{r} - 1}. \tag{4.75e}$$

In summary, standard polynomial interpolation converges as $\mathcal{O}\big((1 + \Delta_M)\exp\left(-\log(r_{\max})M\right)\big)$, see (4.45), while the convergence of the mapped interpolation is $\mathcal{O}\big((1 + \Delta_M)\exp\left(-\log(\hat{r}_{\max})M\right)\big)$. Hence, assuming that the Lebesgue constant Δ_M grows sufficiently slow, the quantity G defined in (4.58) also represents the relative improvement in the asymptotic rate of geometric convergence for mapped interpolation.

Next, we address the multivariate case $N_{\boldsymbol{\xi}} > 1$. To this end, we first replace (4.28) with its mapped counter part, i.e.

$$\hat{Z}_{\ell_n} \coloneqq \big\{g_n\big(\xi_n^{(i_n)}\big) \,:\, \xi_n^{(i_n)} \in Z_{\ell_n}\big\} \subset \Gamma_n,\ n = 1, \ldots, N_{\boldsymbol{\xi}}. \tag{4.76}$$

Then, the we introduce the tensor-grid of mapped interpolation nodes

$$\hat{Z}_{\boldsymbol{\ell}}^{\mathrm{TP}} \coloneqq \{\mathbf{g}(\boldsymbol{\xi}^{(\mathbf{i})}) \,:\, \boldsymbol{\xi}^{(\mathbf{i})} \in Z_{\boldsymbol{\ell}}^{\mathrm{TP}}\} = \hat{Z}_{\ell_1} \times \ldots \times \hat{Z}_{\ell_{N_{\boldsymbol{\xi}}}}, \tag{4.77}$$

such that the respective tensor-product multivariate mapped approximation reads

$$\hat{\mathcal{Q}}_{\hat{Z}_{\boldsymbol{\ell}}^{\mathrm{TP}}}(\boldsymbol{\xi}) \coloneqq \sum_{\mathbf{i}:\hat{\boldsymbol{\xi}}^{(i)} \in \hat{Z}_{\boldsymbol{\ell}}^{\mathrm{TP}}} \mathcal{Q}\Big(\hat{\boldsymbol{\xi}}^{(\mathbf{i})}\Big)\hat{L}_{\boldsymbol{\ell},\mathbf{i}}(\boldsymbol{\xi}), \tag{4.78}$$

where $\hat{L}_{\boldsymbol{\ell},\mathbf{i}} \coloneqq L_{\boldsymbol{\ell},\mathbf{i}} \circ \mathbf{g}^{-1} : \Gamma \to \mathbb{R}$ are mapped multivariate Lagrange polynomials. Note that for the identity mapping $\mathbf{g} : \mathbf{s} \mapsto \mathbf{s}$ the polynomial approximation (4.29) would be recovered.

However, as already discussed in Section 4.1.3, the computational complexity of tensor-product approaches does not scale well with the number of random parameters $N_{\boldsymbol{\xi}}$. Hence, in the following, we discuss the adaptive construction of an anisotropic sparse grid of mapped interpolation nodes. To this end we employ weighted Leja nodes (4.33) with the level-to-nodes function $m_{\mathrm{l2n}} : \ell_n \mapsto \ell_n + 1, \mathbb{N}_0 \to \mathbb{N}$ such that single extra node of an interpolation level ℓ_n is denoted as $\xi_n^{(\ell_n)} = Z_{\ell_n} \setminus Z_{\ell_n - 1}$. The corresponding mapped univariate Leja node is denoted by $\hat{\xi}_n^{(\ell_n)} = g_n(\xi_n^{(\ell_n)}) = \hat{Z}_{\ell_n} \setminus \hat{Z}_{\ell_n - 1}$. As outlined in Section 4.1.3 nested grids are particularly appealing for the construction of sparse grids, resulting in polynomial approximations of increasing accuracy [53]. In the multivariate case they can be obtained by enforcing monotone multi-index sets [53, 91]. To this end, we consider a multi-index set Λ and define

$$\Lambda_+ \coloneqq \{\boldsymbol{\ell} + \mathbf{e}_n, \forall \boldsymbol{\ell} \in \Lambda, \forall n = 1, \ldots, N_{\boldsymbol{\xi}}\}, \tag{4.79a}$$

$$\Lambda_- \coloneqq \{\boldsymbol{\ell} - \mathbf{e}_n, \forall \boldsymbol{\ell} \in \Lambda, \forall n = 1, \ldots, N_{\boldsymbol{\xi}} \,:\, \ell_n > 0\}, \tag{4.79b}$$

as its forward and backward neighbor multi-index sets, respectively, where $\mathbf{e}_n$ denotes the n-th unit vector. Then, the multi-index set Λ is monotone if and only if $\Lambda_- \subset \Lambda$.

A multi-index $\boldsymbol{\ell} \notin \Lambda$ such that $\Lambda \cup \boldsymbol{\ell}$ is monotone, corresponds to the single extra node

$$\hat{\boldsymbol{\xi}}^{(\boldsymbol{\ell})} = \hat{Z}_{\Lambda \cup \boldsymbol{\ell}} \setminus \hat{Z}_\Lambda, \quad \text{where } \hat{Z}_\Lambda = \bigcup_{\boldsymbol{m} \in \Lambda} \{\hat{\boldsymbol{\xi}}^{(\boldsymbol{m})}\}. \tag{4.80}$$

Then, (4.78) can be replaced by the mapped hierarchical interpolation

$$\hat{\mathcal{Q}}_{\hat{Z}_{\Lambda\cup\boldsymbol{\ell}}}(\boldsymbol{\xi}) = \hat{\mathcal{Q}}_{\hat{Z}_{\Lambda}}(\boldsymbol{\xi}) + \mathfrak{s}_{\boldsymbol{\ell}}\,\hat{H}_{\boldsymbol{\ell}}(\boldsymbol{\xi}), \tag{4.81}$$

where $\hat{H}_{\boldsymbol{\ell}}$ are multivariate mapped hierarchical Lagrange polynomials

$$\hat{H}_{\boldsymbol{\ell}}(\boldsymbol{\xi}) = \prod_{n=1}^{N} \hat{h}_{\ell_n}(\xi_n), \quad \text{where } \hat{h}_{\ell_n}(\xi_n) \coloneqq \begin{cases} \prod_{k=0}^{\ell_n-1} \frac{g_n^{-1}(\xi_n)-\xi_n^{(k)}}{\xi_n^{(\ell_n)}-\xi_n^{(k)}}, & \ell_n \neq 0, \\ 1, & \ell_n = 0, \end{cases} \tag{4.82}$$

and $\mathfrak{s}_{\boldsymbol{\ell}} \in \mathbb{C}$ denotes the corresponding coefficients, referred to as hierarchical surpluses, and given by

$$\mathfrak{s}_{\boldsymbol{\ell}} = \mathcal{Q}\left(\hat{\boldsymbol{\xi}}^{(\boldsymbol{\ell})}\right) - \hat{\mathcal{Q}}_{\hat{Z}_{\Lambda}}\left(\hat{\boldsymbol{\xi}}^{(\boldsymbol{\ell})}\right). \tag{4.83}$$

Note that the mapped hierarchical polynomials do not change when additional nodes are added. Furthermore, the corresponding coefficients $\mathfrak{s}_{\boldsymbol{\ell}} \in \mathbb{C}$ can be employed as a posteriori error indicators, which quantify the individual contribution of an additional node $\hat{\boldsymbol{\xi}}^{(\boldsymbol{\ell})}$ to the current approximation $\hat{\mathcal{Q}}_{\hat{Z}_{\Lambda}}(\boldsymbol{\xi})$. Hence, we now discuss the dimension-adaptive construction of the mapped sparse grid interpolation. Note that, by choosing $\mathbf{g}$ as the identity map, the presented approach becomes similar to the polynomial-based schemes considered in [53, 91, 148, 179], however, minor changes are required to address the complex-valued QoI.

For a given monotone multi-index set Λ, corresponding to the mapped grid $\hat{Z}_{\Lambda}$ and the mapped interpolation $\hat{\mathcal{Q}}_{\hat{Z}_{\Lambda}}$, we define the set of admissible multi-indices

$$\Lambda_{+}^{\mathrm{adm}} \coloneqq \{\boldsymbol{\ell} \in \Lambda_{+} \,:\, \boldsymbol{\ell} \notin \Lambda \text{ and } \{\boldsymbol{\ell}\}_{-} \subset \Lambda\}. \tag{4.84}$$

for refinement which ensures that we obtain a sequence of nested monotone sets [53]. We then employ the modulus of the complex hierarchical surpluses $|\mathfrak{s}_{\boldsymbol{\ell}}|$ for each multi-index $\boldsymbol{\ell} \in \Lambda_{+}^{\mathrm{adm}}$ as error indicator and expand Λ with the multi-index with the largest error indicator, i.e. $\arg\max_{\boldsymbol{\ell}\in\Lambda_{+}^{\mathrm{adm}}} |\mathfrak{s}_{\boldsymbol{\ell}}|$. Note that, depending on the application, alternative choices for the error indicator are equally valid, e.g. $\max\left(|\Re\{\mathfrak{s}_{\boldsymbol{\ell}}\}|, |\Im\{\mathfrak{s}_{\boldsymbol{\ell}}\}|\right)$. After updating the grid $\hat{Z}_{\Lambda}$ and the approximation $\hat{\mathcal{Q}}_{\hat{Z}_{\Lambda}}$, accordingly, the procedure is then repeated until a computational budget which specifies the maximum number of model evaluation $\mathcal{B}$ is reached, i.e. the algorithm is terminated if

$$\#\hat{Z}_{\Lambda\cup\Lambda_{+}^{\mathrm{adm}}} \geq \mathcal{B}. \tag{4.85}$$

Then, the final approximation is given as $\hat{\mathcal{Q}}_{\hat{Z}_{\Lambda\cup\Lambda_{+}^{\mathrm{adm}}}}$, i.e. the readily available model evaluations for the admissible neighbor set $\Lambda_{+}^{\mathrm{adm}}$ are employed as well. We note that the algorithm is usually initialized with the multi-index set $\Lambda = \{(0,\ldots,0)\}$. The presented scheme is summarized in Algorithm 2.

Finally, we discuss the minor modifications which are required in order to generalize the method to multiple QoIs, i.e. $\mathcal{Q} : \Gamma \to \mathbb{C}^{N_{\mathcal{Q}}}$. To this end, we replace the scalar hierarchical surplus (4.83) by its vector-valued counterpart $\mathfrak{s}_{\boldsymbol{\ell}} \in \mathbb{C}^{N_{\mathcal{Q}}}$. The adaptivity is then steered by expanding with the multi-index

$$\boldsymbol{\ell} = \underset{\tilde{\boldsymbol{\ell}}\in\Lambda_{+}^{\mathrm{adm}}}{\arg\max} \left(\max\left\{ \left|w_1\mathfrak{s}_{\tilde{\boldsymbol{\ell}},1}\right|, \left|w_2\mathfrak{s}_{\tilde{\boldsymbol{\ell}},2}\right|, \ldots, \left|w_{N_{\mathcal{Q}}}\mathfrak{s}_{\tilde{\boldsymbol{\ell}},N_{\mathcal{Q}}}\right|, \right\}\right), \tag{4.86}$$

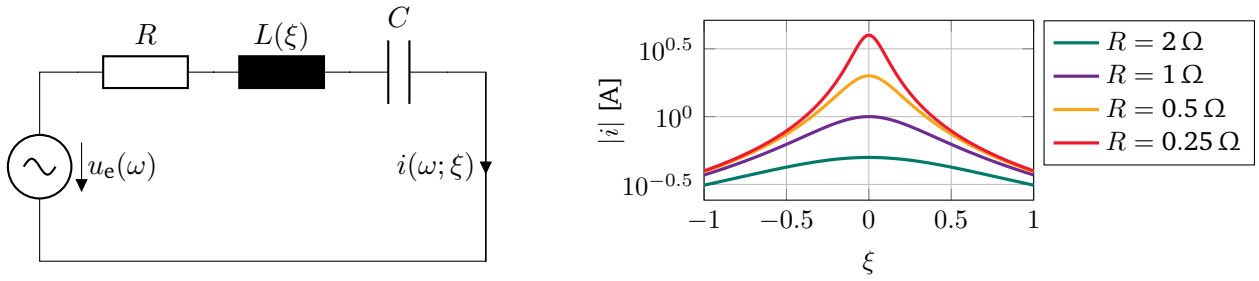
where $\mathbf{w} = [w_1, w_2, \ldots, w_{N_{\mathcal{Q}}}] \in \mathbb{R}^{N_{\mathcal{Q}}}$ denotes a vector of weights.

Data: QoI $\mathcal{Q}(\boldsymbol{\xi})$, coordinate-wise conformal map $\mathbf{g}$, multi-index set Λ, computational budget $\mathcal{B}$
Result: sparse grid $\hat{Z}_{\Lambda\cup\Lambda_+^{\mathrm{adm}}}$ of mapped interpolation nodes, mapped approximation $\hat{\mathcal{Q}}_{\hat{Z}_{\Lambda\cup\Lambda_+^{\mathrm{adm}}}}$

repeat
- Compute the admissible set Λ_+^{adm}, as in (4.84).
- Compute the hierarchical surpluses $\mathfrak{s}_{\boldsymbol{\ell}}$, $\forall \boldsymbol{\ell} \in \Lambda_+^{\mathrm{adm}}$, as in (4.83).
- Select the multi-index $\boldsymbol{\ell} \in \Lambda_+^{\mathrm{adm}}$ associated to the maximum error indicator $|\mathfrak{s}_{\boldsymbol{\ell}}|$.
- Compute the refined approximation $\hat{\mathcal{Q}}_{\hat{Z}_{\Lambda\cup\boldsymbol{\ell}}}$, as in (4.81).
- Enlarge the multi-index set $\Lambda = \Lambda \cup \boldsymbol{\ell}$.

until *simulation budget $\mathcal{B}$ is reached, as in* (4.85);

Algorithm 2: Dimension-adaptive conformally mapped SC method, based on [89, Algorithm 1].



(a) Diagram of considered RLC circuit.

(b) Amplitude of electric current w.r.t. input parameter ξ.

Figure 4.9: An RLC circuit with uncertain inductance L and variable resistance R is considered as benchmark problem. The illustration is based on [90, Figure 3].

4.2.3 Numerical example: RLC circuit

We now consider an academic model problem, i.e. an RLC circuit, with an available closed-form solution, in order to illustrate the main ideas of the presented methods. The application to a real-world example as well as numerical tests with a larger number of parameters can then be found in Section 5.3.1.

The considered RLC circuit is shown in Fig. 4.9a where we consider the current amplitude as QoI, i.e. $\mathcal{Q} := |i|$. In turn, the electric current is obtained as

$$\left(-L\omega^2 + j\omega R + \frac{1}{C}\right) i = \mathrm{i}\omega u_{\mathrm{e}}, \tag{4.87}$$

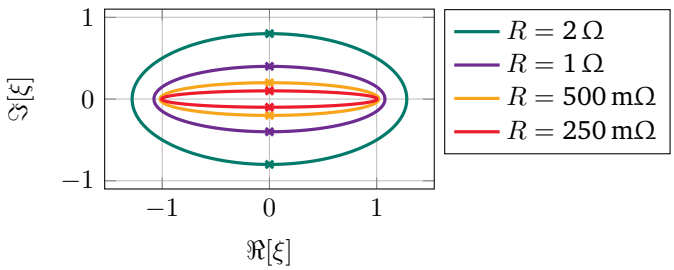
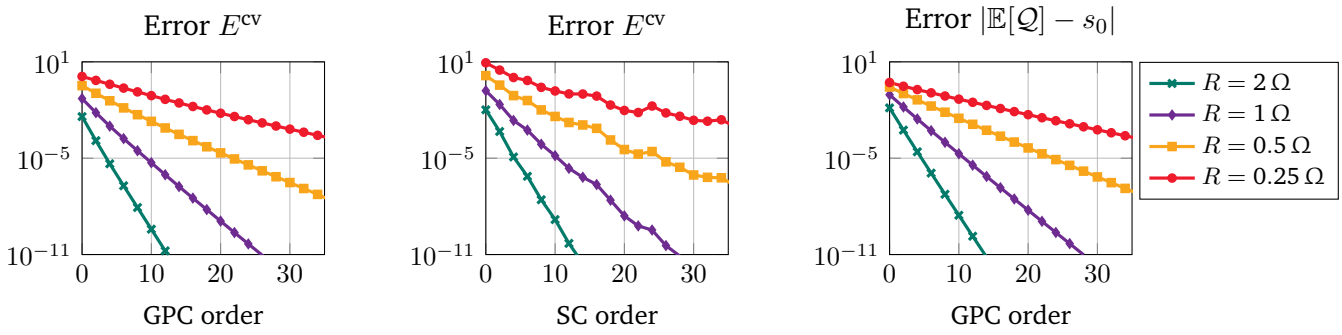


Figure 4.10: Associated Bernstein ellipses of $\xi \mapsto \mathcal{Q}(\xi)$ for different values of R. The respective poles are illustrated by crosses. The illustration is based on [90, Fig. 4a].



(a) Convergence of empirical L^2_ρ error using GPC.

(b) Convergence of empirical L^2_ρ error using SC.

(c) Convergence of estimated mean using GPC.

Figure 4.11: Convergence study using spectral polynomial methods for stochastic RLC circuits with different values of R and, hence, different locations of the poles in the complex plane. The plots are based on [89, Figure 4].

where we assumed a harmonic time-dependency. The respective circuit parameters are chosen as follows: We assume an angular frequency of $\omega = 10^4\,\mathrm{s}^{-1}$, a capacitance $C = 10\,\mu\mathrm{F}$, an exciting voltage $u_\mathrm{e} = 1\,\mathrm{V}$ and we will consider different (rather small) values for the resistance, i.e. $R \in \{2\,\Omega, 1\,\Omega, 0.5\,\Omega, 0.25\,\Omega\}$. Furthermore, the model becomes a stochastic model, as we assume a variable inductance such that $L(\xi) = 1\,\mathrm{mH} + 0.25\,\mathrm{mH}\cdot\xi$, where, in turn, the parameter ξ is then modeled as a realization of a uniformly distributed RV $\Xi \sim \mathcal{U}(-1,1)$ with PDF $\rho_{1\mathrm{D}} = 0.5$. The parametric mapping $\xi \mapsto \mathcal{Q}(\xi) = |i(\xi)|$ is depicted in Fig. 4.9b for different values of R. Note that this mapping is analytic for $\xi \in \Gamma_{1\mathrm{D}} = [-1,1]$, however, its extension to complex inputs has poles at

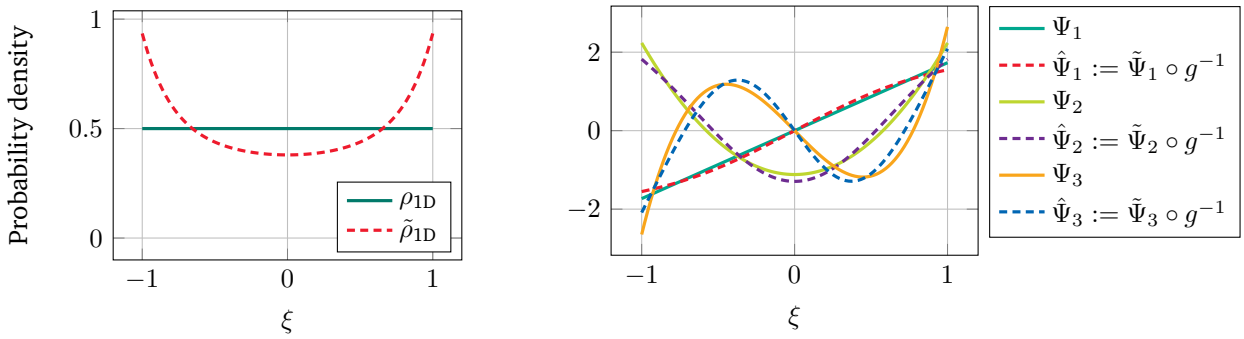
$$\xi = \pm i \frac{R}{\omega \cdot 0.25\,\mathrm{mH}}. \tag{4.88}$$

In Fig. 4.10 the associated Bernstein ellipses of $\xi \mapsto \mathcal{Q}(\xi)$ are shown for different values of R. It can be observed how the size of the Bernstein ellipses is restricted by the complex conjugate pole pair (4.88).

In each case, we then apply the spectral methods introduced in Section 4.1.3 and assess the accuracy of the respective surrogate models in terms of the RMSE, i.e. the square root of the error in the empirical $L^2_{\rho_{1\mathrm{D}}}$ norm. In particular, we draw $N_\mathrm{cv} = 1000$ random realizations $\{\xi_\mathrm{cv}^{(i)}\}_{i=1}^{N_\mathrm{cv}}$ of Ξ and evaluate the error

$$E^{\mathrm{cv}} = \sqrt{\frac{1}{N_\mathrm{cv}} \sum_{i=1}^{N_\mathrm{cv}} \left| \tilde{\mathcal{Q}}\left(\xi_\mathrm{cv}^{(i)}\right) - \mathcal{Q}\left(\xi_\mathrm{cv}^{(i)}\right) \right|^2}, \tag{4.89}$$

where $\tilde{\mathcal{Q}}$ refers to the surrogate model, e.g. computed by GPC or SC. In particular, on the one hand we employ the `Chaospy` toolbox [78] to construct GPC approximations of (increasing) degree M where we use Gaussian quadrature of order $M+1$ to compute the GPC coefficients by pseudo-spectral projection. On the other hand, we employ SC on (weighted) Leja nodes. For both methods, we then conduct convergence studies w.r.t. an increasing polynomial degree M and evaluate the error (4.89). The results are shown in Fig. 4.11a for GPC and in Fig. 4.11b for SC. It can be seen that (4.42) is confirmed numerically, as, for both methods, an increasing convergence order can indeed be observed for increasing values of R which are associated to an increased size of the respective Bernstein ellipses. It is worth noting that a similar behavior is to be expected for decreasing amplitudes of the input variation, according to (4.88). Next, we also evaluate the error in the mean value of the GPC approximation, which is, according to (4.26), given by the GPC coefficient s_0. To this end, we compute the reference solution for $\mathbb{E}[\mathcal{Q}(\Xi)]$ with Gaussian quadrature of degree 200 up to machine



(a) Uniform input distribution ρ_{1D} and respective transformed density $\tilde{\rho}_{1D}$, as in (4.52).

(b) Illustration of some exemplary GPC basis functions Ψ_m and the respective mapped GPC basis functions $\hat{\Psi}_m$.

Figure 4.12: Illustration of (mapped) GPC basis for a uniformly distributed input RV, based on [89, Figure 5].

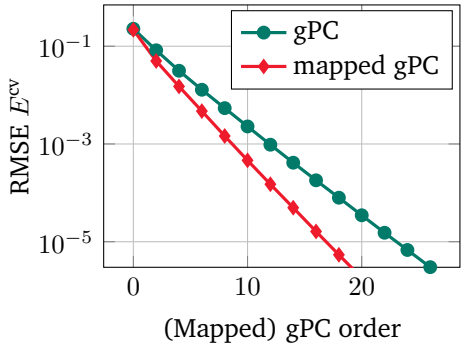
accuracy. The resulting convergence study is presented in Fig. 4.11c and shows that a comparable convergence behavior can be observed for the stochastic moments.

In the following, we employ the conformally mapped approaches proposed in the previous subsections which are, again, implemented in `Python` based on `Chaospy` [78]. The PDF ρ_{1D} along with the respective transformed density ρ_{1D}, defined in (4.52), are shown in Fig. 4.12a. The respective basis functions for GPC and conformally mapped GPC are illustrated in Fig. 4.12b. It is worth noting that the particular GPC basis is given by Legendre polynomials while the mapped basis functions are no polynomials, see (4.51). We recall that the (mapped) basis functions for (mapped) SC have already been presented in Fig. 4.8. We then conduct a convergence study with mapped GPC and mapped SC in terms of the RMSE (4.89), where we employ in the former case mapped Gaussian quadrature of order $M+1$ to compute the mapped coefficients $\hat{s}_m$ for a mapped GPC expansion of degree M. The results are presented in Fig. 4.13a and Fig. 4.13b and clearly show the improved convergence rate of the mapped methods. Finally, it is demonstrated in Fig. 4.13c and Fig. 4.13d that the mapped GPC approach also leads to an improved convergence of the stochastic moments, i.e. the computed mean value and the computed standard deviation.

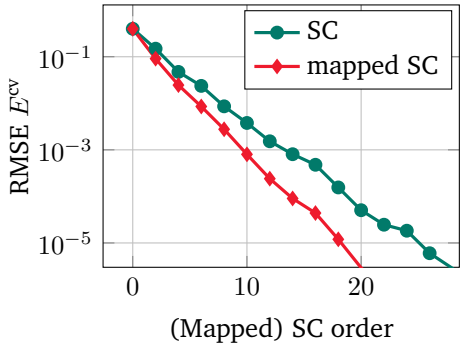
In summary, it was observed that the suggested mapped spectral methods show an improved convergence rate for the considered benchmark problem of a stochastic RLC circuit, leading to significant improvements in either computational cost or accuracy. They are non-intrusive methods and, hence, do not require any modifications of an existing computer code of a parameterized model. Conformally mapped GPC features an orthogonal basis and, hence, the straightforward computation of stochastic moments and Sobol indices while the proposed dimension-adaptive mapped SC scheme based on mapped Leja nodes addresses moderately high-dimensional problems. As will also be explained in the following, the mapped spectral methods can be combined with further techniques for convergence acceleration, e.g. adjoint-error correction.

4.3 Adjoint-based adaptive mapped stochastic collocation

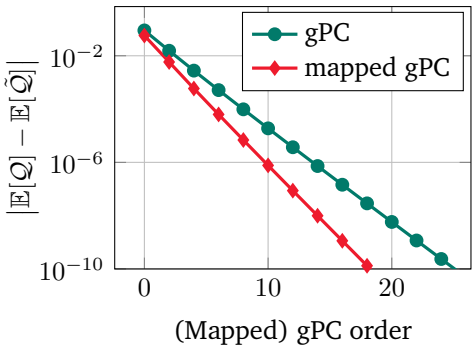
In the following, we discuss how an adjoint error indicator can be employed to improve the dimension-adaptive conformally mapped SC Algorithm 2 proposed in the last section. Note that duality-based a posteriori error estimation is already well established for the FEM, see for instance [16]. More recently, it was also addressed in a parametric setting [43, 44, 171]. In particular, in [111, 179] adaptive SC methods based on Clenshaw



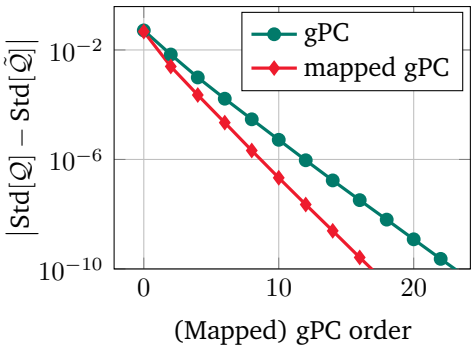
(a) Convergence of RMSE using GPC and mapped GPC.



(b) Convergence of RMSE using SC and mapped SC.



(c) Convergence of estimated mean using GPC and mapped GPC.



(d) Convergence of estimated standard deviation using GPC and mapped GPC.

Figure 4.13: Convergence study using (mapped) spectral methods for stochastic RLC circuit with $R = 1\,\Omega$. The plots are based on [90, Figure 5].

Curtis nodes were proposed where adjoint techniques lead to significant improvements in computational efficiency which can be partially attributed to the exponentially growing level-to-node function (4.32). In this thesis, it is shown that adjoint-based adaptivity and error correction are also beneficial for the proposed SC method based on the more granular Leja nodes and mapped polynomials. The content and structure of this section are based on our work [89].

In this section, we add the additional assumption that the deterministic computational model $\boldsymbol{\xi} \mapsto \mathcal{Q}(\boldsymbol{\xi})$ is represented by a parametric model in the form of (2.56) which is repeated here, for convenience of the reader: find $\mathbf{u}(\boldsymbol{\xi}) \in V$ s.t.

$$a_{\boldsymbol{\xi}}(\mathbf{u}(\boldsymbol{\xi}), \mathbf{v}) = l(\boldsymbol{\xi}) \quad \forall \mathbf{v} \in V, \tag{4.90}$$

where the QoI is a bounded linear functional $\mathcal{Q}(\boldsymbol{\xi}) = J(\mathbf{u}(\boldsymbol{\xi}))$. Note that in this section (4.90) could refer to the models (2.57)-(2.60) but also other parametric PDEs where $a_{\boldsymbol{\xi}}(\cdot,\cdot)$ denotes a continuous sesquilinearform, $l(\cdot)$ a continuous antilinear form and V a suitable Hilbert space. In any case, it is assumed that $\mathbf{u} : \Gamma \to V$ is well-defined and smooth. Note that this assumption is often fulfilled for parameterized PDEs, see for instance [10] where elliptic PDEs are considered and [52] for different types of problems. We further note that the presented method could be generalized to non-linear functionals, see [196, Chapter 3.2].

Introducing the primal operator $A_{\boldsymbol{\xi}} : V \to V^*$, where V^* denotes the dual space to V, the primal problem (4.90) can be stated as an operator equation, i.e. $\forall \boldsymbol{\xi} \in \Gamma$, find $\mathbf{u}(\boldsymbol{\xi}) \in V$ s.t.

$$\langle A_{\boldsymbol{\xi}} \mathbf{u}(\boldsymbol{\xi}), \mathbf{v} \rangle = a_{\boldsymbol{\xi}}(\mathbf{u}(\boldsymbol{\xi}), \mathbf{v}) = l(\mathbf{v}) \;\; \forall \mathbf{v} \in V. \tag{4.91}$$

The adjoint problem then reads: $\forall \boldsymbol{\xi} \in \Gamma$, find $\mathbf{z}(\boldsymbol{\xi}) \in V$, s.t.

$$\langle \mathbf{w}, A^*_{\boldsymbol{\xi}} \mathbf{z}(\boldsymbol{\xi}) \rangle = a_{\boldsymbol{\xi}}(\mathbf{w}, \mathbf{z}(\boldsymbol{\xi})) = J(\mathbf{w}) \;\; \forall \mathbf{w} \in V, \tag{4.92}$$

where the adjoint operator $A^*_{\boldsymbol{\xi}} : V \to V^*$ is defined s.t.

$$\langle A_{\boldsymbol{\xi}} \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, A^*_{\boldsymbol{\xi}} \mathbf{v} \rangle \;\; \forall \mathbf{u}, \mathbf{v} \in V, \; \forall \boldsymbol{\xi} \in \Gamma. \tag{4.93}$$

Note that these definitions directly imply the so-called primal-dual equivalence

$$J(\mathbf{u}(\boldsymbol{\xi})) = \langle \mathbf{u}(\boldsymbol{\xi}), A^*_{\boldsymbol{\xi}} \mathbf{z}(\boldsymbol{\xi}) \rangle = \langle A_{\boldsymbol{\xi}} \mathbf{u}(\boldsymbol{\xi}), \mathbf{z}(\boldsymbol{\xi}) \rangle = l(\mathbf{z}(\boldsymbol{\xi})). \tag{4.94}$$

We are now interested in the parametric error in the QoI

$$\mathcal{E}^{\mathrm{SC}}(\boldsymbol{\xi}) = J\big(\mathbf{u}(\boldsymbol{\xi}) - \mathbf{u}_Z(\boldsymbol{\xi})\big) = a_{\boldsymbol{\xi}}\big(\mathbf{u}(\boldsymbol{\xi}) - \mathbf{u}_Z(\boldsymbol{\xi}), \mathbf{z}(\boldsymbol{\xi})\big) = l\big(\mathbf{z}(\boldsymbol{\xi})\big) - a_{\boldsymbol{\xi}}\big(\mathbf{u}_Z(\boldsymbol{\xi}), \mathbf{z}(\boldsymbol{\xi})\big), \tag{4.95}$$

for a given (mapped) SC approximation $\mathbf{u}_Z$ of the mapping $\mathbf{u} : \Gamma \to V$ on a grid Z. Note that none of the expressions in (4.95) can be computed efficiently, as they require either the solution of the primal problem (4.91) or the dual problem (4.92), for all $\boldsymbol{\xi} \in \Gamma$. Hence, following [43, 44], we instead employ the error indicator

$$\tilde{\mathcal{E}}^{\mathrm{SC}}(\boldsymbol{\xi}) = l\big(\mathbf{z}_Z(\boldsymbol{\xi})\big) - a_{\boldsymbol{\xi}}\big(\mathbf{u}_Z(\boldsymbol{\xi}), \mathbf{z}_Z(\boldsymbol{\xi})\big), \tag{4.96}$$

where $\mathbf{z}_Z$ denotes a (mapped) SC approximation of the mapping $\mathbf{z} : \Gamma \to V$ on the grid Z.

Data: coordinate-wise conformal map $\mathbf{g}$, multi-index set Λ, budget $\mathcal{B}$ and $a_{\boldsymbol{\xi}}$, l, $\mathcal{Q}$ as defined in (4.90)
Result: sparse grid $\hat{Z}_{\Lambda\cup\Lambda_+^{\mathrm{adm}}}$ of mapped interpolation nodes, mapped approximation $\hat{\mathcal{Q}}_{\hat{Z}_{\Lambda\cup\Lambda_+^{\mathrm{adm}}}}$

repeat
- Compute the admissible set Λ_+^{adm}, as in (4.84).
- Compute the adjoint-based error indicators $|\tilde{\mathfrak{s}}_{\boldsymbol{\ell}}|$, where $\tilde{\mathfrak{s}}_{\boldsymbol{\ell}} = \tilde{\mathcal{E}}^{\mathrm{SC}}\left(\boldsymbol{\xi}^{(\boldsymbol{\ell})}\right)$, $\forall \boldsymbol{\ell} \in \Lambda_+^{\mathrm{adm}}$.
- Select the multi-index $\boldsymbol{\ell} \in \Lambda_+^{\mathrm{adm}}$ associated to the maximum error indicator $|\tilde{\mathfrak{s}}_{\boldsymbol{\ell}}|$.
- Compute the hierarchical surpluses $s_{\boldsymbol{\ell}}$, $\mathbf{u}_{\boldsymbol{\ell}}$, $\mathbf{z}_{\boldsymbol{\ell}}$ as in (4.83), by solving (4.91) and (4.92).
- Compute the refined approximation $\hat{\mathcal{Q}}_{\hat{Z}_{\Lambda\cup\boldsymbol{\ell}}}$, as in (4.81), and the corresponding approximations of primal and dual solution.
- Enlarge the multi-index set $\Lambda = \Lambda \cup \boldsymbol{\ell}$.

until *simulation budget $\mathcal{B}$ is reached*;
Algorithm 3: Adjoint error-based, dimension-adaptive mapped SC method, based on [89, Algorithm 2].

Note that, due to continuity of the sesquilinearform $a_{\boldsymbol{\xi}}(\cdot,\cdot)$, we obtain

$$|\mathcal{E}^{\mathrm{SC}}(\boldsymbol{\xi}) - \tilde{\mathcal{E}}^{\mathrm{SC}}(\boldsymbol{\xi})| = |a_{\boldsymbol{\xi}}\big(\mathbf{u}(\boldsymbol{\xi}) - \mathbf{u}_Z(\boldsymbol{\xi}), \mathbf{z}(\boldsymbol{\xi})\big) - a_{\boldsymbol{\xi}}\big(\mathbf{u}(\boldsymbol{\xi}) - \mathbf{u}_Z(\boldsymbol{\xi}), \mathbf{z}_Z(\boldsymbol{\xi})\big)| \tag{4.97a}$$
$$= |a_{\boldsymbol{\xi}}\big(\mathbf{u}(\boldsymbol{\xi}) - \mathbf{u}_Z(\boldsymbol{\xi}), \mathbf{z}(\boldsymbol{\xi}) - \mathbf{z}_Z(\boldsymbol{\xi})\big)| \tag{4.97b}$$
$$\leq C\|\mathbf{u}(\boldsymbol{\xi}) - \mathbf{u}_Z(\boldsymbol{\xi})\|_V \|\mathbf{z}(\boldsymbol{\xi}) - \mathbf{z}_Z(\boldsymbol{\xi})\|_V, \tag{4.97c}$$

i.e. the error indicator $\tilde{\mathcal{E}}^{\mathrm{SC}}$ shows faster convergence than the (mapped) SC approximations $\mathbf{u}_Z, \mathbf{z}_Z$. In particular, if one assumes the same convergence order for both, primal and dual solution, the error indicator (4.96) exhibits a doubled convergence rate.

Next, we address the adaptions which are required to incorporate the error indicator $\tilde{\mathcal{E}}^{\mathrm{SC}}(\boldsymbol{\xi})$ into the dimension-adaptive (mapped) SC Algorithm 2. In particular, one needs to construct the additional (mapped) SC approximations $\mathbf{u}_Z, \mathbf{z}_Z$ which are here computed using the same multi-index set Λ, corresponding grid Z and respective mapped hierarchical polynomials $\hat{H}_{\boldsymbol{\ell}}(\boldsymbol{\xi})$ as for the scalar QoI $\mathcal{Q}$. The approximations are then given in the form of (4.81) where the hierarchical surpluses $\mathfrak{s}_{\boldsymbol{\ell}} \in \mathbb{C}$ are replaced by vector-valued coefficients $\mathbf{u}_{\boldsymbol{\ell}}, \mathbf{z}_{\boldsymbol{\ell}} \in \mathbb{C}^{N_h}$. Next, following [111], one can employ the adjoint error indicator (4.96) to replace the error indicators $|\mathfrak{s}_{\boldsymbol{\ell}}|$, $\forall \boldsymbol{\ell} \in \Lambda_+^{\mathrm{adm}}$ which are employed in Algorithm 2 and require the solution of (4.91) for all $\hat{\boldsymbol{\xi}}^{(\boldsymbol{\ell})} \in \hat{Z}_{\Lambda_+^{\mathrm{adm}}}$. In particular, we employ the adjoint error indicators $|\tilde{\mathfrak{s}}_{\boldsymbol{\ell}}|$, where $\tilde{\mathfrak{s}}_{\boldsymbol{\ell}} \coloneqq \tilde{\mathcal{E}}^{\mathrm{SC}}\big(\hat{\boldsymbol{\xi}}^{(\boldsymbol{\ell})}\big)$ to steer the adaptivity, i.e. we then select the multi-index with the largest error indicator $\arg\max_{\boldsymbol{\ell}\in\Lambda_+^{\mathrm{adm}}} |\tilde{\mathfrak{s}}_{\boldsymbol{\ell}}|$ and solve the corresponding primal and dual problem in order to update the mapped SC approximations $\mathcal{Q}_Z, \mathbf{u}_Z, \mathbf{z}_Z$. Furthermore, after the termination of the algorithm, the adjoint-error indicator is also employed for error correction. To this end, one can employ the multi-index set $\Lambda \cup \Lambda_+^{\mathrm{adm}}$ to construct the final approximation, such that all readily available evaluations of (4.96) are used as coefficients $\mathfrak{s}_{\boldsymbol{\ell}}$ by setting $\mathfrak{s}_{\boldsymbol{\ell}} = \tilde{\mathfrak{s}}_{\boldsymbol{\ell}}, \forall \boldsymbol{\ell} \in \Lambda_+^{\mathrm{adm}}$. This procedure is outlined in Algorithm 3. However, we also note that based on (4.97c) we expect an even higher accuracy of the adjoint-error corrected approximation

$$\tilde{\mathcal{Q}}_{\hat{Z}_\Lambda}(\boldsymbol{\xi}) = \mathcal{Q}_{\hat{Z}_\Lambda}(\boldsymbol{\xi}) + \tilde{\mathcal{E}}^{\mathrm{SC}}(\boldsymbol{\xi}). \tag{4.98}$$

The computational cost of evaluating (4.98), which requires the evaluation of a residual of the primal problem, see (4.96), is usually significantly lower than the cost for solving the primal problem (4.91) to evaluate the QoI but higher than the cost for evaluating the (mapped) SC approximation $\mathcal{Q}_{\hat{Z}_\Lambda}(\boldsymbol{\xi})$. Hence, to obtain a (mapped) SC approximation with higher accuracy, one can continue with further refinements of $\mathcal{Q}_{\hat{Z}_\Lambda}(\boldsymbol{\xi})$ by

then employing Algorithm 2 on the adjoint-error corrected approximation (4.98). The application of the proposed method is investigated in detail in Section 5.3.1.

4.3.1 Adjoint formulation for Maxwell's source problems

In the following, we relate the methodology from the previous subsection to the model problems introduced in Chapter 2. In particular, the strong formulation of the dual problem (4.92) is derived for the case of the unit cell problem (2.24) and the waveguide model (2.27) by following the procedure outlined in [75]. To this end, we define the differential operator L by rewriting (2.3) as

$$L\mathbf{E} \coloneqq \nabla \times \left(\mu_{\mathrm{r}}^{-1}\nabla \times \mathbf{E}\right) - \omega^2 \varepsilon \mu_0 \mathbf{E} = 0 \qquad \text{in } D. \tag{4.99}$$

Then, the formal adjoint L^* refers to the differential operator which fulfills

$$(L\mathbf{E}, \mathbf{E}')_D = (\mathbf{E}, L^*\mathbf{E}')_D, \tag{4.100}$$

for all compactly supported and sufficiently smooth functions $\mathbf{E}, \mathbf{E}'$ on D. Hence, the formal adjoint L^* can be obtained by applying integration by parts to (2.33) and neglecting boundary terms

$$\left(\mathbf{E}, \nabla \times \left(\overline{\mu_{\mathrm{r}}^{-1}}\nabla \times \mathbf{E}'\right)\right)_D - \omega^2 \mu_0 \left(\mathbf{E}, \bar{\varepsilon}\mathbf{E}'\right)_D = 0, \tag{4.101}$$

which then yields

$$L^*\mathbf{E}' := \nabla \times \left(\overline{\mu_{\mathrm{r}}^{-1}}\nabla \times \mathbf{E}'\right) - \omega^2 \bar{\varepsilon} \mu_0 \mathbf{E}' = 0 \qquad \text{in } D. \tag{4.102}$$

To derive the adjoint boundary conditions, we then drop the assumptions that $\mathbf{E}$ and $\mathbf{E}'$ are compactly supported on D and require that

$$(L\mathbf{E}, \mathbf{E}')_D - (\mathbf{E}, L^*\mathbf{E}')_D = \left((\mu_{\mathrm{r}}^{-1}\nabla \times \mathbf{E})_{\mathrm{t}}, \mathbf{E}'_{\mathrm{T}}\right)_{\partial D} - (\mathbf{E}_{\mathrm{T}}, (\mu_{\mathrm{r}}^{-1}\nabla \times \mathbf{E}')_{\mathrm{t}})_{\partial D} = 0, \tag{4.103}$$

holds for general smooth functions $\mathbf{E}, \mathbf{E}'$. In the following, we first address the adjoint boundary conditions of the unit cell problem (2.24) and afterward the waveguide model (2.27) is considered. We note that in both cases, the QoI is given by affine-linear functionals, i.e. (2.25) and (2.28), while we consider linear functionals in this section, however, the extension to address the constant offset is straightforward and, hence, is omitted here.

4.3.1.1 Unit cell problem

For the unit cell model we consider the linear functional

$$J^{\alpha,mn}(\mathbf{E}) \coloneqq \left(\mathbf{E}_{\mathrm{T}}, \mathbf{E}_{\mathrm{T}}^{\alpha,mn}\right)_{\partial D_{z^+}}, \; V^{\mathrm{uc}} \to \mathbb{C}, \quad \text{where } \alpha \in \{\mathrm{TE}, \mathrm{TM}\}, m \in \mathbb{Z}, n \in \mathbb{Z}, \tag{4.104}$$

see (2.25). We now discuss the adjoint boundary conditions which are implied by (4.103) for the different parts of the boundary ∂D of the unit cell separately. On ∂D_{z^-}, the second integral vanishes due to the PEC boundary condition (2.24d) and, hence, we also need to enforce $\mathbf{E}'_{\mathrm{T}} = 0$ on ∂D_{z^-}. Due to the quasi-periodic

boundary conditions (2.24b)-(2.24c), we obtain the following conditions for the Dirichlet $\mathbf{E}'_{\mathrm{T}}$ and the Neumann trace $\left(\mu_{\mathrm{r}}^{-1}\nabla\times\mathbf{E}'\right)_{\mathrm{t}}$

$$\mathbf{E}'_{\mathrm{T}}|_{\partial D_{x^+}}e^{\mathrm{i}k_x^{\mathrm{inc}}d_x}=\mathbf{E}'_{\mathrm{T}}|_{\partial D_{x^-}},\qquad (\mu_{\mathrm{r}}^{-1}\nabla\times\mathbf{E}')_{\mathrm{t}}|_{\partial D_{x^+}}e^{\mathrm{i}k_x^{\mathrm{inc}}d_y}=(\mu_{\mathrm{r}}^{-1}\nabla\times\mathbf{E}')_{\mathrm{t}}|_{\partial D_{x^-}}\qquad \text{on } \partial D_{x^+}\cup\partial D_{x^-}, \tag{4.105a}$$

$$\mathbf{E}'_{\mathrm{T}}|_{\partial D_{y^+}}e^{\mathrm{i}k_y^{\mathrm{inc}}d_y}=\mathbf{E}'_{\mathrm{T}}|_{\partial D_{y^-}},\qquad (\mu_{\mathrm{r}}^{-1}\nabla\times\mathbf{E}')_{\mathrm{t}}|_{\partial D_{y^+}}e^{\mathrm{i}k_y^{\mathrm{inc}}d_y}=(\mu_{\mathrm{r}}^{-1}\nabla\times\mathbf{E}')_{\mathrm{t}}|_{\partial D_{y^-}}\qquad \text{on } \partial D_{y^+}\cup\partial D_{y^-}, \tag{4.105b}$$

such that the respective parts of the integrals on ∂D_{x^-}, ∂D_{y^-} and ∂D_{x^+}, ∂D_{y^+} cancel each other, respectively. Next, we address ∂D_{z^+} where we employ the Floquet boundary condition (2.24e) in (4.103) considering a homogeneous system, i.e. $\mathbf{E}_{\mathrm{T}}^{\mathrm{inc}}=0$, which yields

$$-\left(\mathbf{E}_{\mathrm{T}},\left(\nabla\times\mathbf{E}'\right)_{\mathrm{t}}\right)_{\partial D_{z^+}}-\frac{\mathrm{i}}{k_z^{\mathrm{inc}}}\left(\mathbf{k}_{\mathrm{t}}^{\mathrm{inc}}\cdot\mathbf{E}_{\mathrm{T}},\mathbf{k}_{\mathrm{t}}^{\mathrm{inc}}\cdot\mathbf{E}'_{\mathrm{T}}\right)_{\partial D_{z^+}}-\mathrm{i}k_z^{\mathrm{inc}}\left(\mathbf{E}_{\mathrm{T}},\mathbf{E}'_{\mathrm{T}}\right)_{\partial D_{z^+}}=0, \tag{4.106}$$

$$-\left(\mathbf{E}_{\mathrm{T}},\left(\nabla\times\mathbf{E}'\right)_{\mathrm{t}}-\frac{\mathrm{i}\mathbf{k}_{\mathrm{t}}^{\mathrm{inc}}}{k_z^{\mathrm{inc}}}\mathbf{k}_{\mathrm{t}}^{\mathrm{inc}}\cdot\mathbf{E}'_{\mathrm{T}}-\mathrm{i}k_z^{\mathrm{inc}}\mathbf{E}'_{\mathrm{T}}\right)_{\partial D_{z^+}}=0. \tag{4.107}$$

Following the procedure outlined in [75], the adjoint boundary condition on ∂D_{z^+} is then obtained by substituting $\mathbf{E}'$ with the dual solution $\mathbf{z}$, $\mathbf{E}$ with a general test function $\mathbf{w}$ and the RHS with the linear functional $J^{\alpha,mn}(\cdot)$ which then yields

$$\left(\frac{\mathrm{i}}{\omega\mu_0}\nabla\times\mathbf{z}\right)_{\mathrm{t}}+\frac{\mathbf{k}_{\mathrm{t}}^{\mathrm{inc}}}{\omega\mu_0 k_z^{\mathrm{inc}}}\left(\mathbf{k}_{\mathrm{t}}^{\mathrm{inc}}\cdot\mathbf{z}_{\mathrm{T}}\right)+\frac{k_z^{\mathrm{inc}}}{\omega\mu_0}\mathbf{z}_{\mathrm{T}}=-\frac{\mathrm{i}}{\omega\mu_0}\mathbf{E}_{\mathrm{T}}^{\alpha,mn}\qquad \text{on } \partial D_{z^+}. \tag{4.108}$$

In summary, the strong formulation of the adjoint problem associated to (2.24) reads

$$\nabla\times\left(\overline{\mu_{\mathrm{r}}^{-1}}\nabla\times\mathbf{z}\right)-\omega^2\bar{\varepsilon}\mu_0\mathbf{z}=0\qquad \text{in } D, \tag{4.109a}$$

$$\mathbf{z}_{\mathrm{T}}|_{\partial D_{x^+}}e^{\mathrm{i}k_x^{\mathrm{inc}}d_x}=\mathbf{z}_{\mathrm{T}}|_{\partial D_{x^-}}\qquad \text{on } \partial D_{x^+}\cup\partial D_{x^-}, \tag{4.109b}$$

$$\mathbf{z}_{\mathrm{T}}|_{\partial D_{y^+}}e^{\mathrm{i}k_y^{\mathrm{inc}}d_y}=\mathbf{z}_{\mathrm{T}}|_{\partial D_{y^-}}\qquad \text{on } \partial D_{y^+}\cup\partial D_{y^-}, \tag{4.109c}$$

$$\mathbf{z}_{\mathrm{t}}=0\qquad \text{on } \partial D_{z^-}, \tag{4.109d}$$

$$\left(\frac{\mathrm{i}}{\omega\mu_0}\nabla\times\mathbf{z}\right)_{\mathrm{t}}+\frac{\mathbf{k}_{\mathrm{t}}^{\mathrm{inc}}}{\omega\mu_0 k_z^{\mathrm{inc}}}\left(\mathbf{k}_{\mathrm{t}}^{\mathrm{inc}}\cdot\mathbf{z}_{\mathrm{T}}\right)+\frac{k_z^{\mathrm{inc}}}{\omega\mu_0}\mathbf{z}_{\mathrm{T}}=-\frac{\mathrm{i}}{\omega\mu_0}\mathbf{E}_{\mathrm{T}}^{\alpha,mn}\ \text{on } \partial D_{z^+}, \tag{4.109e}$$

where $\alpha\in\{\mathrm{TE},\mathrm{TM}\},m\in\mathbb{Z},n\in\mathbb{Z}$, as in (4.104).

4.3.1.2 Waveguide problem

For the waveguide model (2.27) we consider the linear functional

$$J^{\mathrm{wg}}(\mathbf{E}):=\frac{2}{E_0 ab}\left(\mathbf{E},\mathbf{E}_{10}^{\mathrm{TE}}\right)_{\partial D_{\mathrm{P1}}}, \tag{4.110}$$

see (2.28). Again, the PEC boundary conditions (2.27d), imply the respective adjoint boundary conditions $\mathbf{z}_{\mathrm{t}}=0$ on $\partial D_{\mathrm{PEC}}$. Next, we employ the port boundary conditions (2.27b)-(2.27c) in (4.103), where we consider again a homogeneous system with $\mathbf{E}^{\mathrm{inc}}=0$ which yields

$$\left(\mathrm{i}k_{z10}\mathbf{E}_{\mathrm{T}},\mathbf{E}'_{\mathrm{T}}\right)_{\partial D_{\mathrm{P1}}\cup\partial D_{\mathrm{P2}}}-\left(\mathbf{E}_{\mathrm{T}},\left(\nabla\times\mathbf{E}'\right)_{\mathrm{t}}\right)_{\partial D_{\mathrm{P1}}\cup\partial D_{\mathrm{P2}}}=0, \tag{4.111}$$

$$-\left(\mathbf{E}_{\mathrm{T}},\mathrm{i}k_{z10}\mathbf{E}'_{\mathrm{T}}\right)_{\partial D_{\mathrm{P1}}\cup\partial D_{\mathrm{P2}}}-\left(\mathbf{E}_{\mathrm{T}},\left(\nabla\times\mathbf{E}'\right)_{\mathrm{t}}\right)_{\partial D_{\mathrm{P1}}\cup\partial D_{\mathrm{P2}}}=0. \tag{4.112}$$

As before, we replace $\mathbf{E}'$ by the adjoint solution $\mathbf{z}$, $\mathbf{E}$ with a general test function and the RHS with the linear functional $J^{\mathrm{wg}}(\cdot)$ leading to the adjoint boundary conditions (4.113b)+(4.113c) of the adjoint BVP

$$\nabla \times \left(\overline{\mu_{\mathrm{r}}^{-1}} \nabla \times \mathbf{z}\right) - \omega^2 \overline{\varepsilon} \mu_0 \mathbf{z} = 0 \qquad \text{in } D, \tag{4.113a}$$

$$\left(\nabla \times \mathbf{z}'\right)_{\mathrm{t}} + \mathrm{i} k_{z10} \mathbf{z}_{\mathrm{T}} = -\frac{2}{abE_0} \mathbf{E}_{10}^{\mathrm{TE}} \qquad \text{on } \partial D_{\mathrm{P1}}, \tag{4.113b}$$

$$\left(\nabla \times \mathbf{z}'\right)_{\mathrm{t}} + \mathrm{i} k_{z10} \mathbf{z}_{\mathrm{T}} = 0 \qquad \text{on } \partial D_{\mathrm{P2}}, \tag{4.113c}$$

$$\mathbf{z}_{\mathrm{t}} = 0 \qquad \text{on } \partial D_{\mathrm{PEC}}, \tag{4.113d}$$

which is associated to (2.27).

4.3.1.3 Discretization

Note that the discretization of Maxwell's source problem with the FEM yields a linear system

$$\mathbf{A}\mathbf{e} = \mathbf{f}, \tag{4.114}$$

where $\mathbf{A} = \mathbf{A}^{\mathrm{uc}}, \mathbf{f} = \mathbf{f}^{\mathrm{uc}}$ in case of the unit cell problem, see (2.41), and $\mathbf{A} = \mathbf{A}^{\mathrm{wg}}, \mathbf{f} = \mathbf{f}^{\mathrm{wg}}$ in case of the waveguide problem, see (2.48). Discretization of the respective adjoint problems (4.109) and (4.113) then yields linear systems in the form of

$$\mathbf{A}^{\mathrm{H}} \mathbf{z}_h = \mathbf{j}, \tag{4.115}$$

where $\mathbf{z}_h, \mathbf{j} \in \mathbb{C}^{N_h}$. Hence, the adjoint solution $\mathbf{z}_h$ can often be computed with negligible computational cost, for instance if one solves the primal problem (4.114) with a sparse LU decomposition $\mathbf{A} = \mathbf{L}\mathbf{U}$. In this case, $\mathbf{L} \in \mathbb{C}^{N_h \times N_h}$ and $\mathbf{U} \in \mathbb{C}^{N_h \times N_h}$ are a lower triangular and an upper triangular matrix, respectively, which immediately also yield the required LU decomposition for the dual problem (4.115) as $\mathbf{A}^{\mathrm{H}} = (\mathbf{L}\mathbf{U})^{\mathrm{H}} = \mathbf{U}^{\mathrm{H}}\mathbf{L}^{\mathrm{H}}$. Note that one can then employ forward and backward substitution to solve the linear systems

$$\mathbf{L}\mathbf{b}_{\text{primal}} = \mathbf{f}, \qquad \mathbf{U}^{\mathrm{H}}\mathbf{b}_{\text{dual}} = \mathbf{j}, \tag{4.116a}$$

$$\mathbf{U}\mathbf{e} = \mathbf{b}_{\text{primal}}, \qquad \mathbf{L}^{\mathrm{H}}\mathbf{z}_h = \mathbf{b}_{\text{dual}}, \tag{4.116b}$$

in order to obtain the primal solution $\mathbf{e}$ and the dual solution $\mathbf{z}_h$.

4.4 Adjoint-based yield estimation

In this section, we address the estimation of the yield (4.10) and suggest an adjoint-based hybrid method that comprises accuracy and computational efficiency. We note that many algorithms for yield estimation have been proposed where the MC method (4.11) might be the most established one [103]. However, it often requires many evaluations of the QoI [94] which can become computationally prohibitive if the underlying model is given by a highly resolved FE discretization of a PDE, as introduced in Chapter 2 for instance. Hence, the proposed hybrid method evaluates the majority of an MC sample with an adjoint-based SC surrogate model, as introduced in the previous subsection, and only employs the high fidelity model for certain critical sample points where the surrogate accuracy is insufficient to ensure a correct classification. The content and structure of this section are based on our work [81].

Before we describe the proposed scheme in detail, we recall a few alternative approaches for yield estimation suggested in the literature which also aim for a reduced number of high fidelity evaluations. The first-order reliability method (FORM) or second-order reliability method (SORM) are sampling-free methods which search for the most probable point of failure and then construct an approximation of the limit state function at this point [36, 54]. Importance sampling [85] and subset simulation [7, 17] are sampling-based methods which lead to a reduction in sample size compared to standard MC. Further approaches rely on first constructing a response surface or surrogate model and then conducting a MC analysis on the approximation, see e.g. [131] and the references therein. However, [131, Example 3.1] shows that even arbitrary small approximation errors can lead to completely wrong estimates of the yield (or the failure probability). Hence, a hybrid approach is suggested which evaluates the high fidelity model instead of the surrogate whenever a point is *close* to the limit state [131]. A related method that employs pointwise adjoint-based error estimates instead of a fixed interval size to classify the *critical* sample points was proposed in [45].

In this work, we present a hybrid approach for efficient yield estimation based on the adaptive adjoint-based SC Algorithm 3. We note that in contrast to [131], we identify the *critical* sample points based on adjoint error indicators as (4.96). Contrary to [45], we employ SC and incorporate an additional adjoint-based estimate of the FE error in the hybrid decision criterion. Furthermore, if required, we then iteratively refine the FE model for certain sample points which is in the spirit of the multilevel and multifidelity methods discussed in Section 4.1.2. Note that we neglected the FE error in the previous subsections as it can usually be controlled using established techniques, see, e.g. [2], to be sufficiently small. However, as we now address the estimation of (failure) probabilities instead of moments, any error source has to be incorporated as it could lead to erroneous results, see the aforementioned example [131, Example 3.1].

To estimate the parametric error of the SC approximation, we can employ $\tilde{\mathcal{E}}^{\mathrm{SC}}$ introduced in (4.96). To additionally estimate the FE error, we also rely on adjoint techniques, following [16, 73]:

$$\mathcal{E}^{\mathrm{FE},h} = J(\mathbf{u}-\mathbf{u}_h) = \left(\mathbf{u}-\mathbf{u}_h, A^*_{\boldsymbol{\xi}}\mathbf{z}\right)_D = \left(A_{\boldsymbol{\xi}}(\mathbf{u}-\mathbf{u}_h), \mathbf{z}\right)_D = l(\mathbf{z}) - a_{\boldsymbol{\xi}}(\mathbf{u}_h, \mathbf{z}). \tag{4.117}$$

This expression only becomes computable if one employs a FE approximation of $\mathbf{z}$. However, if one would simply choose $\mathbf{z}_h \in V_h$ the respective error estimate is zero as $\mathbf{z}_h$ has to be orthogonal to the residual, see (4.90). Hence, we need an adjoint solution of higher accuracy to estimate $\mathcal{E}^{\mathrm{FE},h}$, contrary to $\mathcal{E}^{\mathrm{SC}}$, see [43]. In this work, we compute the adjoint on a refined FE mesh, however, there are equally valid alternatives, e.g. employing higher order basis functions or recovery techniques [218]. This leads to the error indicator

$$\tilde{\mathcal{E}}^{\mathrm{FE},h} = l(\mathbf{z}_{h/2}) - a_{\boldsymbol{\xi}}(\mathbf{u}_h, \mathbf{z}_{h/2}), \tag{4.118}$$

where $\mathbf{z}_{h/2} \in V_{h/2}$. Eq. (4.118) represents a computable expression, however, it requires the computation of $\mathbf{z}_{h/2}$ for all $\boldsymbol{\xi} \in \Gamma$. Hence, we construct the corresponding SC approximation $\tilde{\mathcal{E}}^{\mathrm{FE},h}{}_Z$. Finally, the total error can be decomposed into the SC and the FE contribution as

$$\mathcal{E} \coloneqq |\mathcal{Q} - \mathcal{Q}_{h,Z}| \leq \left|\mathcal{Q} - \mathcal{Q}_h\right| + |\mathcal{Q}_h - \mathcal{Q}_{h,Z}| \approx \left|\tilde{\mathcal{E}}_Z^{\mathrm{FE},h}\right| + \left|\tilde{\mathcal{E}}_h^{\mathrm{SC}}\right| =: \tilde{\mathcal{E}}, \tag{4.119}$$

where $\tilde{\mathcal{E}}_h^{\mathrm{SC}}$ is obtained by replacing $\mathbf{u}_Z, \mathbf{z}_Z$ in (4.96) with $\mathbf{u}_{h,Z}, \mathbf{z}_{h,Z}$, respectively. Note that the adjoint-based estimation of the combined error resulting from the spatial and the parametric discretization has also been considered in different applications, for instance in [44], where the stochastic Galerkin method is employed for time-dependent forward and inverse problems.

Next, we discuss how (4.119) can be employed in a hybrid approach for yield estimation. The goal is to reach the same accuracy as MC on the high-fidelity model while reducing the computational cost by instead

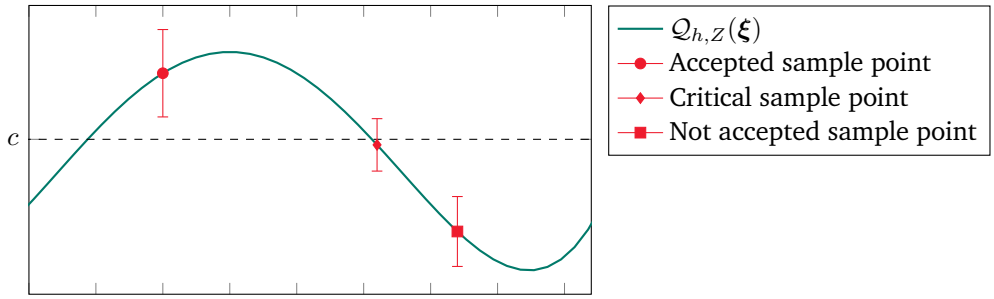


Figure 4.14: Illustration of the error estimate based classification of sample points as accepted, critical or not accepted, according to the performance feature specification (4.7).

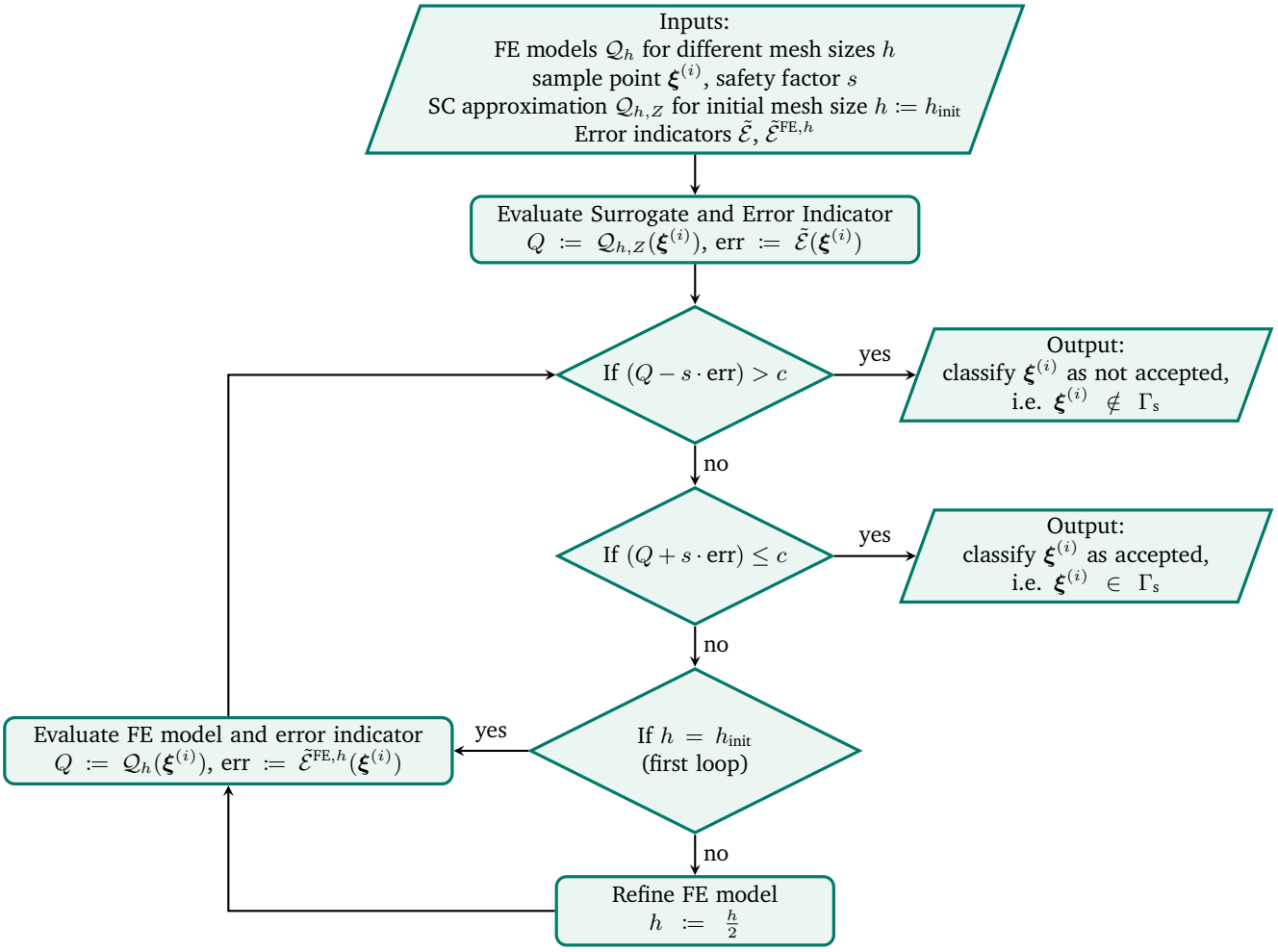


Figure 4.15: Flowchart describing the hybrid decision process. The content is based on [81, Algorithm 1].

evaluating only the available surrogate models whenever possible. To this end, we employ the MC estimator (4.11) but replace the indicator function (4.8) with a hybrid decision process. In particular, we introduce the interval

$$\mathcal{I}^1_{\tilde{\mathcal{E}}}(\boldsymbol{\xi}^{(i)}) = \left[\mathcal{Q}_{h,Z}(\boldsymbol{\xi}^{(i)}) - s\tilde{\mathcal{E}}(\boldsymbol{\xi}^{(i)}), \mathcal{Q}_{h,Z}(\boldsymbol{\xi}^{(i)}) + s\tilde{\mathcal{E}}(\boldsymbol{\xi}^{(i)})\right], \tag{4.120}$$

for each sample point $\boldsymbol{\xi}^{(i)},\ i = 1, \dots, N_{\mathrm{MC}}$, where $s \geq 1$ denotes a safety factor which will be discussed later. If each element of the interval $\mathcal{I}^1_{\tilde{\mathcal{E}}}(\boldsymbol{\xi}^{(i)})$ fulfills the performance feature specifications (4.7), that is $Q_j \leq c\ \forall Q_j \in \mathcal{I}^1_{\tilde{\mathcal{E}}}(\boldsymbol{\xi}^{(i)})$ the sample point can be reliably classified as accepted. Similarly, if $Q_j > c\ \forall Q_j \in \mathcal{I}^1_{\tilde{\mathcal{E}}}(\boldsymbol{\xi}^{(i)})$ the sample point can be reliably classified as not accepted. Finally, if only a subset of the interval fulfills the performance feature specifications, the sample point is classified as *critical* and further treatment is necessary. Fig. 4.14 illustrates this distinction of sample points. Note that, except for the critical sample points, the classification can be done purely based on the (cheap) surrogate models. Next, the FE model $\mathcal{Q}_h$ is evaluated for all critical sample points which then leads to a new interval

$$\mathcal{I}^2_{\tilde{\mathcal{E}}}(\boldsymbol{\xi}^{(i)}) = \left[\mathcal{Q}_h(\boldsymbol{\xi}^{(i)}) - s\left|\tilde{\mathcal{E}}^{\mathrm{FE},h}(\boldsymbol{\xi}^{(i)})\right|, \mathcal{Q}_h(\boldsymbol{\xi}^{(i)}) + s\left|\tilde{\mathcal{E}}^{\mathrm{FE},h}(\boldsymbol{\xi}^{(i)})\right|\right]. \tag{4.121}$$

Following the same procedure as before, the sample points are, in turn, classified as accepted, not accepted, or critical. For the remaining critical sample points, we then iteratively refine the mesh until the respective sample points can be reliably classified. The full decision procedure is presented in Fig. 4.15. Note that, in practice, one might restrict the maximum number of refinement steps. However, the presented approach should lead to the same accuracy as MC on the most-refined FE model unless errors at certain sample points would be greatly underestimated and, hence, the respective sample points wrongly classified. To this end, the safety factor $s \geq 1$ is chosen rather conservatively, as the adjoint error indicators are not strict upper bounds. In particular, the safety factor might be determined by evaluating the error of $\mathcal{Q}_{h,Z}$ w.r.t. the FE model with the most refined mesh on a small random sample. These errors could then be compared to the respective values of the error indicator $\tilde{\mathcal{E}}$, e.g. by considering the maximum ratio. In this work, we then set the safety factor to $s = 2$ for the considered numerical example. Note that a conservative choice of the safety factor increases the computational effort but may lead to a higher accuracy as it avoids the misclassification of sample points. We further note that the method could possibly be extended by exploiting the (expensive) FE model evaluations to train an additional model discrepancy term of the surrogate model, see [45, 82], which is excluded here, for simplicity.

Finally, we emphasize that the presented approach for yield estimation allows to take all relevant error sources into account. The hybrid decision process, outlined in Fig. 4.15, considers the FE error and the SC error of the surrogate model. Additionally, the MC error, associated to the finite size of the random sample, can be controlled based on (4.13). In particular, in this work we specify a maximum standard deviation $\sigma_{\mathcal{Y}} \coloneqq \mathrm{Std}\left[\mathcal{Y}\right]$ to derive the corresponding sample size N_{MC} using (4.13). In Section 5.3.4, the proposed yield estimation procedure is then applied to the benchmark problem of an electric waveguide.

4.5 Eigenvalue tracking

In this section, we address UQ for generalized eigenvalue problems in the form of (2.53), where we now consider the eigenfrequency of a specific eigenmode as QoI $\mathcal{Q}$. The content and structure of this section are based on our work [88]. To quantify the uncertainty, we are interested in employing an SC method as described in the previous sections. To this end, we make the assumption that the matrices $\mathbf{K}(\boldsymbol{\xi})$ and $\mathbf{M}(\boldsymbol{\xi})$ as well as the corresponding eigenpairs $\left(\omega_{h,n}(\boldsymbol{\xi}), \mathbf{e}_n(\boldsymbol{\xi})\right)$ depend smoothly on the parameters $\boldsymbol{\xi}$. Note that

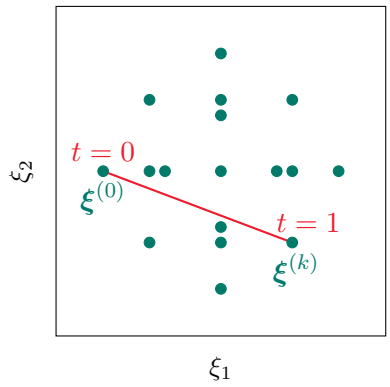


Figure 4.16: Illustration of physical homotopy for two points of a Smolyak sparse grid, based on [88, Figure 4c].

this condition might not be fulfilled if a new FE mesh is constructed for each $\boldsymbol{\xi} \in \Gamma$ and, hence, a mesh transformation is required. In this work, we either rely on a design element approach [35] for the FE mesh, as explained in Section 5.3.1, or realize the deformation by moving control points of an isogeometric analysis (IGA) discretization [38], see Section 5.2 or [88]. However, applying SC still requires a dedicated procedure to correctly match the respective eigenfrequency at each collocation point $\boldsymbol{\xi}^{(k)}$, as the natural ordering based on the eigenvalue is in general not sufficient to guarantee consistency. In particular, even small shape variations, for instance, can lead to eigenvalue crossings or eigenmode separation/degeneration, see Section 5.2.1 for an illustrative example. Hence, we employ a tracking procedure that is adapted from [138]. In particular, we start from one collocation point $\boldsymbol{\xi}^{(0)}$, e.g. a parameter configuration which describes a nominal geometry and then track the respective eigenpair corresponding to the QoI to the different collocation points $\boldsymbol{\xi}^{(k)},\ k \geq 1$. To this end, we introduce the algebraic homotopy between the respective stiffness and mass matrices defined in (2.38) as

$$\mathbf{K}_{k,t} := (1-t)\mathbf{K}\big(\boldsymbol{\xi}^{(0)}\big) + t\mathbf{K}\big(\boldsymbol{\xi}^{(k)}\big), \tag{4.122a}$$

$$\mathbf{M}_{k,t} := (1-t)\mathbf{M}\big(\boldsymbol{\xi}^{(0)}\big) + t\mathbf{M}\big(\boldsymbol{\xi}^{(k)}\big), \tag{4.122b}$$

where $t \in [0,1]$. In this section, we assume for simplicity that there are no eigenvalue bifurcations for $t \in [0,1]$ and that the eigenmodes are not degenerated at $t = 0$. It shall be noted that other homotopies would be equally possible, for instance

$$\mathbf{K}_{k,t} := \mathbf{K}\big((1-t)\boldsymbol{\xi}_0 + t\boldsymbol{\xi}^{(k)}\big), \tag{4.123a}$$

$$\mathbf{M}_{k,t} := \mathbf{M}\big((1-t)\boldsymbol{\xi}_0 + t\boldsymbol{\xi}^{(k)}\big), \tag{4.123b}$$

where the eigenvalue problems for $t \in (0,1)$ would refer to actual shapes. The physical homotopy (4.123) is illustrated in Fig. 4.16 for two points in a Smolyak sparse grid. However, (4.122) allows obtaining the derivatives w.r.t. t as

$$\mathbf{K}'_{k,t} := \frac{\partial}{\partial t}\mathbf{K}_{k,t} = \mathbf{K}\big(\boldsymbol{\xi}^{(k)}\big) - \mathbf{K}\big(\boldsymbol{\xi}^{(0)}\big), \tag{4.124a}$$

$$\mathbf{M}'_{k,t} := \frac{\partial}{\partial t}\mathbf{M}_{k,t} = \mathbf{M}\big(\boldsymbol{\xi}^{(k)}\big) - \mathbf{M}\big(\boldsymbol{\xi}^{(0)}\big). \tag{4.124b}$$

The homotopy enables a one-dimensional analysis for each collocation point $\boldsymbol{\xi}^{(k)} \in \Gamma \subset \mathbb{R}^{N_{\boldsymbol{\xi}}}$ by rewriting the

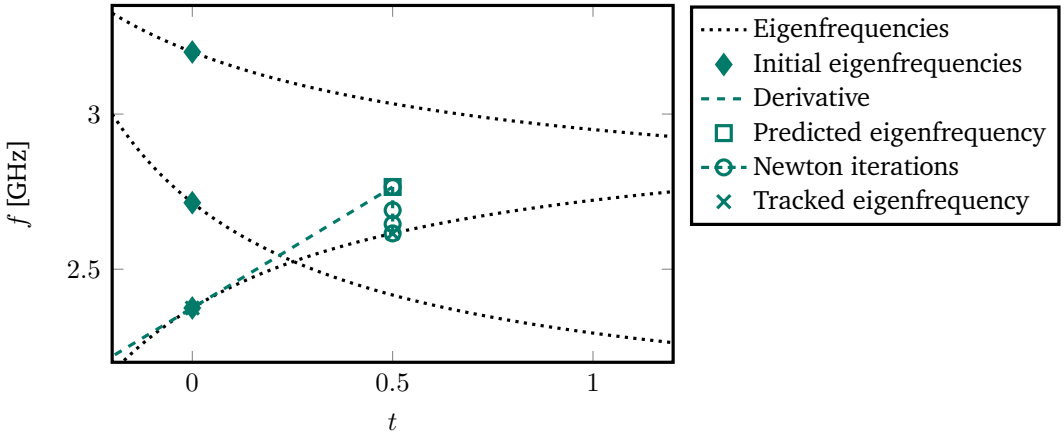


Figure 4.17: Illustration of one step of tracking procedure with stepsize $h_\mathrm{t} = 0.5$ starting at $t^* = 0$.

eigenvalue problem (2.53) depending on the scalar parameter t

$$\begin{bmatrix} \mathbf{K}_t \mathbf{e}_t - \lambda_t \mathbf{M}_t \mathbf{e}_t \\ \mathbf{c}^\top \mathbf{e}_t - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{4.125}$$

where $\lambda_t = \omega_{h,t}^2$ and the last line represents a normalization constraint based on a suitable vector $\mathbf{c} \in \mathbb{C}^{N_h}$. Note that the quantities in (4.125) depend on the collocation point but we suppress the index k for brevity of notation. In particular, we only address the case of one eigenpair $(\mathbf{e}_t, \lambda_t)$ for one single collocation point $\boldsymbol{\xi}^{(k)}$ in the following.

One step of the proposed tracking procedure is illustrated in Fig. 4.17 and described in the following. We start with a first-order Taylor expansion at $t^* \in [0, 1]$ w.r.t. t

$$\begin{bmatrix} \tilde{\mathbf{e}}_t \\ \tilde{\lambda}_t \end{bmatrix} = \begin{bmatrix} \mathbf{e}_{t^*} \\ \lambda_{t^*} \end{bmatrix} + (t - t^*) \begin{bmatrix} \mathbf{e}'_{t^*} \\ \lambda'_{t^*} \end{bmatrix}, \tag{4.126}$$

to obtain an approximation of the eigenpair for $t \neq t^*$, i.e. $[\tilde{\mathbf{e}}_t^\top, \tilde{\lambda}_t]^\top \approx [\mathbf{e}_t^\top, \lambda_t]^\top$. Note that the derivatives $[\mathbf{e}'^\top_{t^*}, \lambda'_{t^*}]^\top$ can be obtained by solving the linear system

$$\begin{bmatrix} \mathbf{K}_{t^*} - \lambda_{t^*} \mathbf{M}_{t^*} & -\mathbf{M}_{t^*} \mathbf{e}_{t^*} \\ \mathbf{c}^\top & 0 \end{bmatrix} \begin{bmatrix} \mathbf{e}'_{t^*} \\ \lambda'_{t^*} \end{bmatrix} = \begin{bmatrix} -\mathbf{K}'_{t^*} \mathbf{e}_{t^*} + \lambda_{t^*} \mathbf{M}'_{t^*} \mathbf{e}_{t^*} \\ 0 \end{bmatrix}, \tag{4.127}$$

which is, in turn, obtained by differentiating (4.125) w.r.t. the parameter t. Then, the Newton-Raphson method can be employed to compute the exact eigenpair $[\mathbf{e}_t^\top, \lambda_t]^\top$ by employing the predicted eigenpair $[\tilde{\mathbf{e}}_t^\top, \tilde{\lambda}_t]^\top$ as an initial guess

$$\begin{bmatrix} \mathbf{e}_t^{(0)} \\ \lambda_t^{(0)} \end{bmatrix} := \begin{bmatrix} \tilde{\mathbf{e}}_t \\ \tilde{\lambda}_t \end{bmatrix}, \tag{4.128}$$

for the first iteration. In particular, the $i-$th iteration of the Newton Raphson method requires to solve the linear system

$$\begin{bmatrix} \mathbf{K}_t - \lambda_t^{(i-1)} \mathbf{M}_t & -\mathbf{M}_t \mathbf{e}_t^{(i-1)} \\ \mathbf{c}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{e}_t^{(i)} \\ \Delta \lambda_t^{(i)} \end{bmatrix} = - \begin{bmatrix} \mathbf{K}_t \mathbf{e}_t^{(i-1)} - \lambda_t^{(i-1)} \mathbf{M}_t \mathbf{e}_t^{(i-1)} \\ \mathbf{c}^\top \mathbf{e}_t^{(i-1)} - 1 \end{bmatrix}, \tag{4.129}$$

and to then improve the prediction as

$$\begin{bmatrix} \mathbf{e}_t^{(i)} \\ \lambda_t^{(i)} \end{bmatrix} := \begin{bmatrix} \mathbf{e}_t^{(i-1)} \\ \lambda_t^{(i-1)} \end{bmatrix} + \begin{bmatrix} \Delta\mathbf{e}_t^{(i)} \\ \Delta\lambda_t^{(i)} \end{bmatrix}. \tag{4.130}$$

This procedure is repeated until the step $|\Delta\lambda_t^{(i)}|$ and the norm of the residual of (4.125) are both sufficiently small. In particular, we terminate the Newton-Raphson method if

$$\max\left\{|\Delta\lambda_t^{(i)}|, \left\|\begin{bmatrix} \mathbf{K}_t\mathbf{e}_t - \lambda_t\mathbf{M}_t\mathbf{e}_t \\ \mathbf{c}^\top\mathbf{e}_t - 1 \end{bmatrix}\right\|\right\} < \mathrm{tol}, \tag{4.131}$$

where tol denotes a prescribed tolerance. Then, the required number of iterations until (4.131) is fulfilled, is further used for a stepsize control which shall improve the robustness and efficiency of the procedure. The complete tracking algorithm is presented in a flowchart in Fig. 4.18.

Finally, we give a few concluding remarks. The case of degenerate eigenmodes at $t = 0$, i.e. an intersection of eigenvalues which was excluded for simplicity only, could be taken into account by employing Ojalvo's method [64, 158] to compute the derivatives instead of (4.127). However, simply choosing a starting configuration $\boldsymbol{\xi}^{(0)}$ which is far from the nominal configuration and, hence, corresponds to a geometry with fewer symmetries, is also often sufficient to meet the assumption. We further emphasize that the proposed procedure for UQ only requires to solve one eigenvalue problem for the initial configuration $\boldsymbol{\xi}^{(0)}$ and, hence, on the one hand, avoids the computational cost of solving eigenvalue problems at the other collocation points. On the other hand, several linear systems have to be solved during the tracking. It shall be noted that the computational efficiency of the suggested tracking algorithm could also be further improved, for instance, by a higher order Taylor approximation in (4.126), by replacing (4.129) with a simplified Newton method [68], or by tracking a subspace of eigenpairs where the eigenpairs could then be matched using a correlation coefficient [118, 214].

4.6 Summary

In this chapter, we have suggested various improved UQ methods. First, conformally mapped GPC and conformally mapped SC have been developed. Their enhanced convergence rate was then shown for the academic model problem of a stochastic RLC circuit with one uncertain parameter. Afterward, extensions based on an adjoint-error indicator have been suggested. In particular, it can be employed for steering the adaptivity in an efficient dimension-adaptive SC scheme and for error correction. Then, a multifidelity yield estimation procedure has been proposed based on the mapped SC approximation and adjoint-error estimation. Finally, a homotopy-based eigenvalue tracking method has been suggested to enable the application of the SC methods for Maxwell's eigenproblem with uncertain input data and ensure a correct classification of eigenmodes.

The application of these methods is investigated numerically in the following chapter. In particular, the different (adjoint-based) conformally mapped spectral UQ methods are studied in detail in Section 5.3.1, where an optical grating coupler is employed as a benchmark model. The adjoint-based yield estimation procedure is employed in Section 5.3.4, where the benchmark problem of an electromagnetic waveguide is considered. The application of the proposed tracking method for UQ for superconducting cavities is presented in Section 5.2.

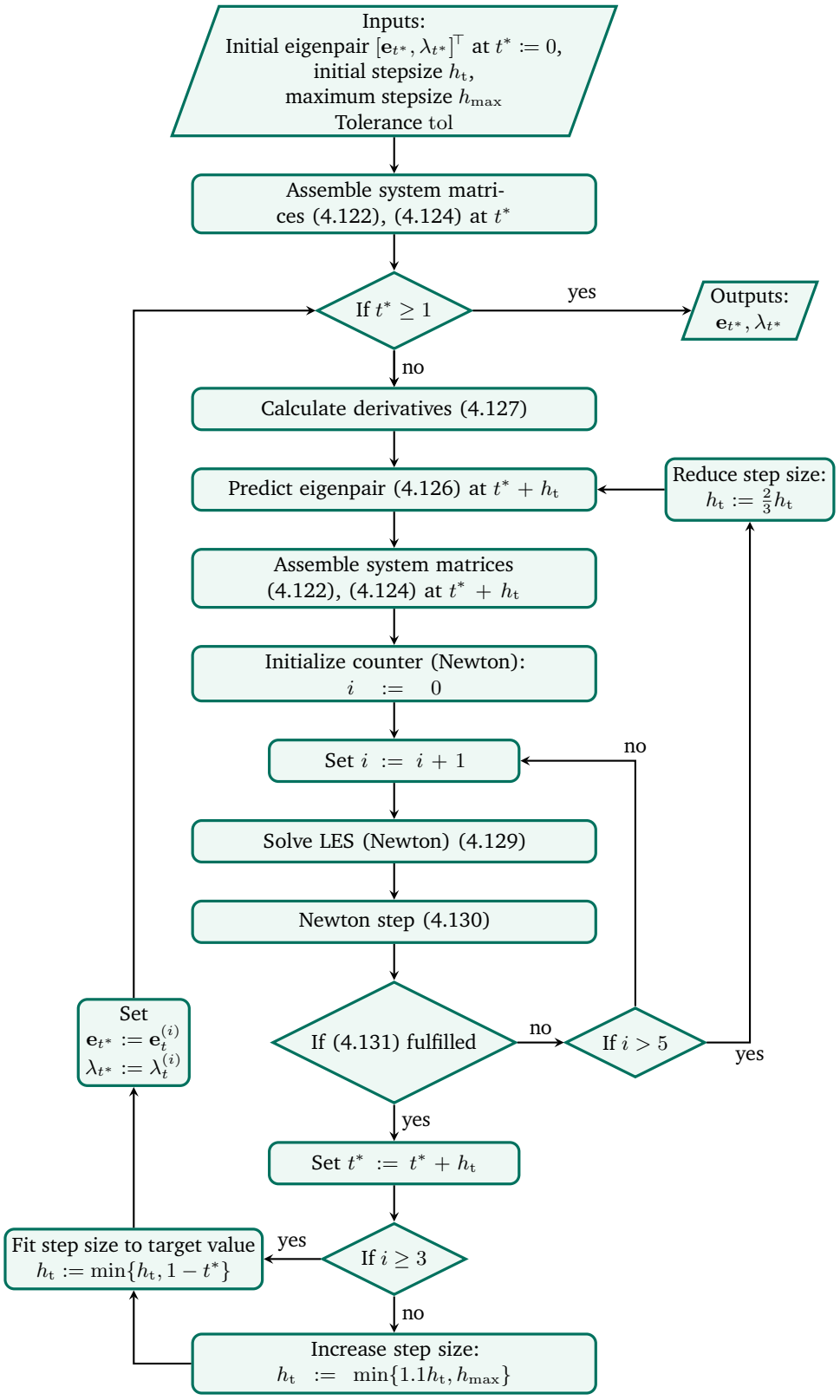


Figure 4.18: Flowchart describing the proposed eigenmode tracking procedure, based on [88, Fig. 5].

5 Applications

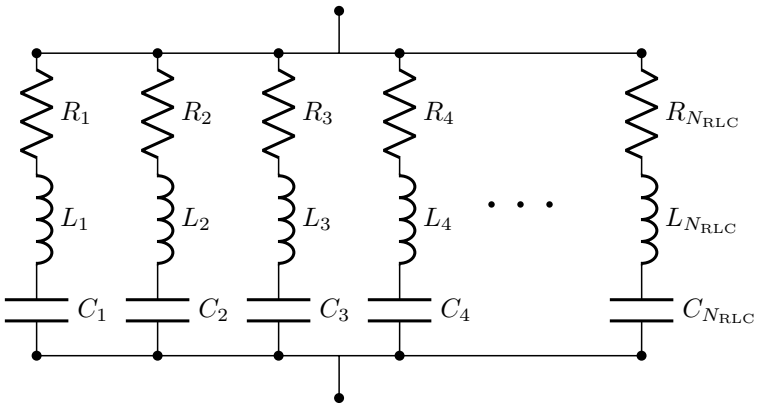
In this chapter, the application of the proposed surrogate modeling and UQ techniques is discussed. First, the approximation of FRFs employing the RKI scheme proposed in Chapter 3 is investigated for a number of benchmark problems from different fields. Then, the UQ methods suggested in Chapter 4 are applied to the parameterized Maxwell's eigen- and source problems introduced in Chapter 2. For a proper modeling and treatment of the involved uncertainties, dedicated UQ workflows are set up for the different applications. In particular, in Section 5.2 comprehensive UQ studies for particle accelerator cavities are presented where we mainly focus on the 9-cell TESLA cavities of the EXFEL at Deutsches Elektronen-Synchrotron (DESY). To this end, SC based on the suggested eigenvalue tracking technique is applied and investigated numerically. In Section 5.3 we then address two different optical gratings that can be modeled as infinitely periodic structures as well as a split ring resonator (SRR) array of finite periodicity with $N_{\text{cells}} = 7$ elements. This section starts with a detailed investigation and comparison of different surrogate modeling techniques for UQ, specifically considering the suggested approaches based on conformal maps and adjoint error indicators. Next, different studies regarding UQ for (quasi-)periodic structures are presented. In particular, a decoupled uncertainty propagation based on unit cell surrogates is proposed for the SRR array. Finally, the efficiency of the adjoint-based yield estimation procedure is demonstrated for the academic benchmark problem of an electric waveguide.

5.1 Approximation of frequency response functions

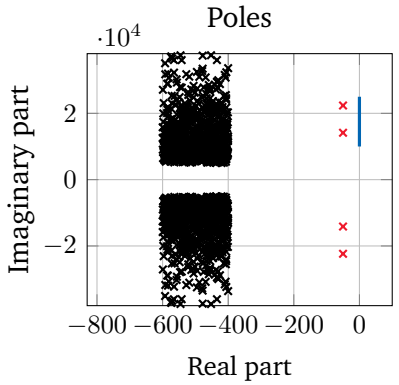
The content and structure of this section follow our work [86]. In the following, we apply the approximation techniques presented in Chapter 3 to a number of benchmark FRFs from different fields. In particular, we employ N_ω training points (3.3), where we use the equidistant frequency sample points (3.45), for simplicity. The accuracy of different approximations of the FRF is then quantified in terms of the RMSE E_ω^{cv}, which is in this section computed as in (3.46) with $N_{\text{cv}} = 201$ for all numerical examples.

In the following, we give a few details on the implementation. For AAA [147], we rely on the implementation of the chebfun toolbox [70]. For VF, we employ the VectFit3 toolbox [67, 98, 99], where we use complex equidistant starting poles, distributed according to the general recommendation and always run 30 iterations. We apply the relaxed non-triviality constraint, include the constant but not the linear term, and enforce stable poles. The number of complex starting pole pairs is set to the maximum number of $2\lfloor \frac{N_\omega - 1}{2} \rfloor$, which generally seems to lead to the best results for the very smooth test functions considered. The implementation of the (complex/real) RKHS interpolation is done in Matlab as well, based on STK[1] (Small Matlab/Octave Toolbox for Kriging) using the non-intrusive approach presented in Section 3.2.3. The tuning of the hyper-parameters and poles based on the profile likelihood function is carried out using `fmincon` in Matlab, i.e. gradient-based optimization, which we combine with a multistart procedure using $N_{\text{ms}} \approx 20$ starting points for α_0, which are

[1] `https://github.com/stk-kriging/stk`



(a) Circuit diagram.



(b) Distribution of poles.

Figure 5.1: a) Parallel connection of (underdamped) series RLC circuits. b) Black crosses indicate the distribution of $2N_{\mathrm{RLC}}^{\{1\}} = 2000$ poles of the circuit admittance $Y^{\{1\}}$ in the complex plane. Red crosses indicate the two additional poles considered for the circuit admittance $Y^{\{2\}}$ with $2N_{\mathrm{RLC}}^{\{2\}} = 2004$ poles. Blue line indicates the considered frequency range.

chosen within the optimization bounds given as $10^{-6}|\Omega| \leq \alpha \leq |\Omega|$. The corresponding scaling parameter is simultaneously optimized, where we use the optimal value, obtained by a generalized least squares (GLS) estimate, corresponding to α_0 as initial value σ_0. By investigating the shape of the likelihood function for a number of benchmark problems, we conclude that the logarithmic reparameterization, discussed in [15] for instance, is only beneficial for the scaling parameter σ but disadvantageous for the parameter α; hence, it is only applied for the scaling parameter. For the poles, which are simultaneously tuned with the kernel hyper-parameters, we employ the optimization bounds

$$-|\Omega| \leq \mathcal{R}[p_i] \leq -10^{-6}|\Omega|, \tag{5.1}$$

$$\max\left\{10^{-6}|\Omega|,\ \omega_{\min} - \frac{|\Omega|}{3}\right\} \leq \Im[p_i] \leq \omega_{\max} + \frac{|\Omega|}{3}, \tag{5.2}$$

where we set the maximum number of poles pairs to $N_{\mathbf{p}}^{\max} = \min\{5, \lfloor \frac{N_\omega}{4} \rfloor\}$.

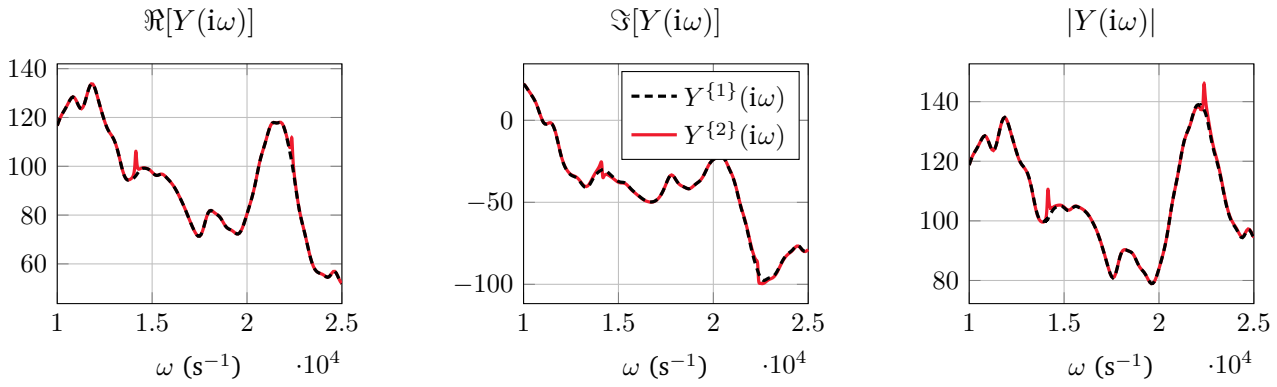
5.1.1 Electric circuit

We consider in the following a parallel connection of N_{RLC} underdamped series RLC circuits, as illustrated in Fig. 5.1a. The admittance of this electric circuit is given as

$$Y(\mathrm{i}\omega) = \sum_{i=1}^{N_{\mathrm{RLC}}} \frac{\mathrm{i}\omega}{(\mathrm{i}\omega)^2 L_i + \mathrm{i}\omega R_i + C_i^{-1}} = \sum_{i=1}^{N_{\mathrm{RLC}}} \frac{c_i}{\mathrm{i}\omega - a_i} + \frac{c_i^*}{\mathrm{i}\omega - a_i^*}. \tag{5.3}$$

The corresponding residues c_i, c_i^* and poles a_i, a_i^* of the partial fraction representation in (5.3) are here obtained as

$$c_i = \frac{\sqrt{\frac{1}{L_i C_i} - \left(\frac{R_i}{2L_i}\right)^2} + \frac{R_i}{2L_i}\mathrm{i}}{2L_i\sqrt{\frac{1}{L_i C_i} - \left(\frac{R_i}{2L_i}\right)^2}}, \quad a_i = \frac{-R_i}{2L_i} + \mathrm{i}\sqrt{\frac{1}{L_i C_i} - \left(\frac{R_i}{2L_i}\right)^2}, \tag{5.4}$$



(a) Real part. **(b)** Imaginary part. **(c)** Magnitude.

Figure 5.2: Complex admittances $Y^{\{1\}}$ and $Y^{\{2\}}$ of the electric circuits w.r.t. frequency for a particular random parameter realization and $N_{\mathrm{RLC}}^{\{1\}} = 1000$ and $N_{\mathrm{RLC}}^{\{2\}} = 1002$, respectively.

since we assume underdamped RLC elements, i.e.

$$\frac{R_i}{2}\sqrt{\frac{C_i}{L_i}} < 1, \quad i = 1, \dots, N_{\mathrm{RLC}}, \tag{5.5}$$

which implies that the argument of the square roots is positive. Note that we consider the frequency range $\Omega = [10\,\mathrm{kHz}, 25\,\mathrm{kHz}]$. From (5.3) and (5.4) it can be concluded that Y belongs to the Hardy space $H^2(\Upsilon_\alpha)$ defined in (3.9) where $\alpha = \min_{1\le i\le N_{\mathrm{RLC}}} \frac{R_i}{2L_i}$.

First, we assume $N_{\mathrm{RLC}}^{\{1\}} = 1000$ random series RLC elements, where

$$C_i \sim \mathcal{U}(1, 20)\,\mu\mathrm{F}, \quad L_i \sim \mathcal{U}(0.1, 2)\,\mathrm{mH}, \tag{5.6}$$

and we assume the resistance R_i to be roughly proportional to the inductance except for random variations of $\pm 20\%$, i.e.

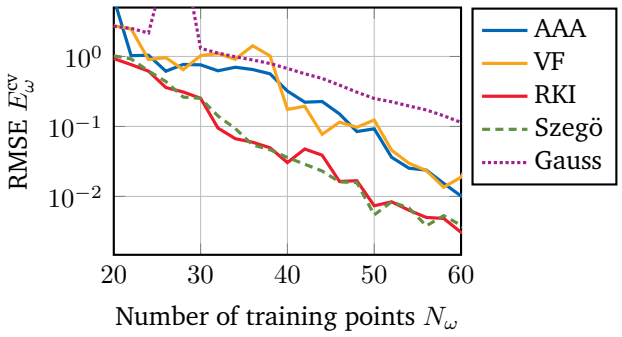
$$R_i = L_i(1 + \Delta)\frac{\Omega}{\mathrm{mH}}, \text{ where } \Delta \sim \mathcal{U}(-0.2, 0.2). \tag{5.7}$$

Note that for any combination of those parameters, the corresponding series RLC circuits are underdamped. For one particular realization, the distribution of the $2N_{\mathrm{RLC}}^{\{1\}} = 2000$ poles is illustrated in Fig. 5.1. The corresponding admittance $Y^{\{1\}}$ is shown in Fig. 5.2 in black color. We then conduct a convergence study for the particular realization of the electric circuit, which is shown in Fig. 5.3a. To obtain a smoother convergence behavior, we additionally repeat the convergence study for 100 random realizations and show the median RMSE at each point, see Fig. 5.3b. It can be observed that for the considered range of the number of training points, where $N_\omega \ll N_{\mathrm{RLC}}$, the complex/real Szegö kernel interpolation outperforms AAA and VF. Employing the adaptive RKI algorithm does not yield an improvement but leads to similar good results. Note that employing real Gauss kernel interpolation for real and imaginary part, separately, leads to the worst results out of the considered approaches, as shown by the dotted line in Fig. 5.3b.

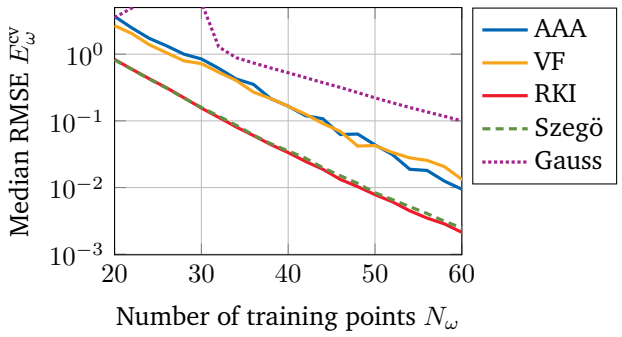
In a next step, we add two additional circuit elements with a very small damping, i.e. we now consider $N_{\mathrm{RLC}}^{\{2\}} = 1002$ and

$$C_{1001} = 5\,\mathrm{pF}, \qquad L_{1001} = 1\,\mathrm{mH}, \qquad R_{1001} = 0.1\,\Omega, \tag{5.8}$$

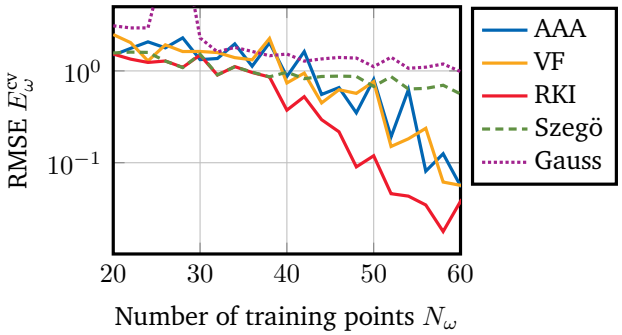
$$C_{1002} = 2\,\mathrm{pF}, \qquad L_{1002} = 1\,\mathrm{mH}, \qquad R_{1002} = 0.1\,\Omega. \tag{5.9}$$



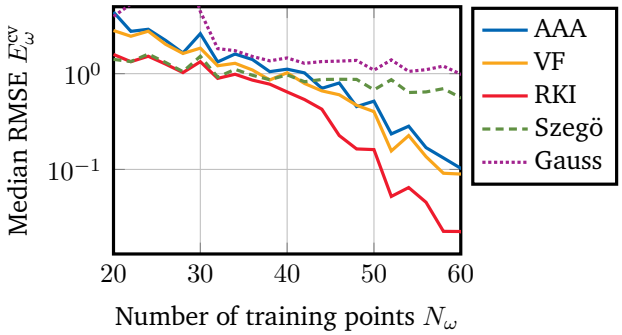
(a) Admittance $Y^{\{1\}}$ (one particular realization).



(b) Admittance $Y^{\{1\}}$ (100 random realization).

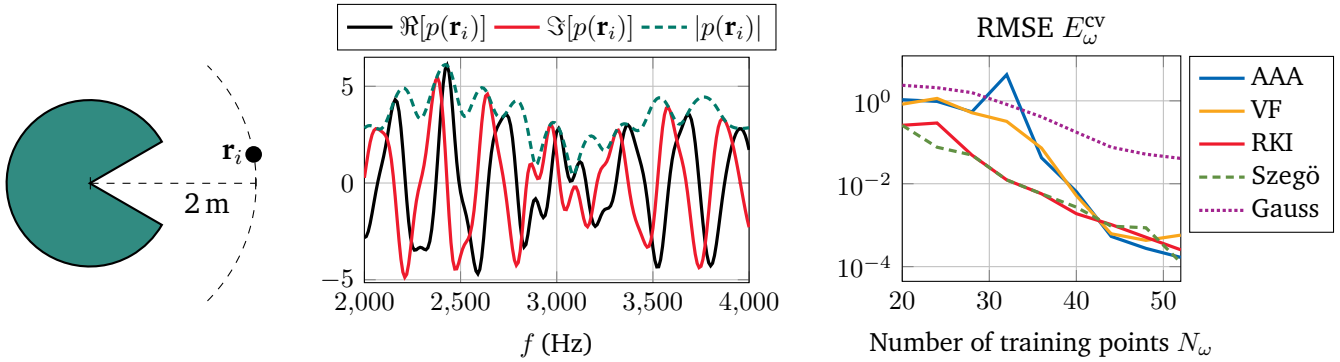


(c) Admittance $Y^{\{2\}}$ (one particular realization).



(d) Admittance $Y^{\{2\}}$ (100 random realizations).

Figure 5.3: Convergence studies for admittances $Y^{\{1\}}$ (top) and $Y^{\{2\}}$ (bottom). Left: RMSE E_ω^{cv} for one particular circuit realization. Right: Median RMSE E_ω^{cv} for 100 random realizations.



(a) PAC-MAN model. **(b)** Complex FRF. **(c)** Convergence study.

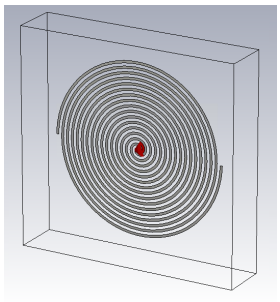
Figure 5.4: a) We consider a surface vibration of the PAC-MAN model and evaluate the radiated acoustic field $p(\mathbf{r}_i)$ at a point (black dot) in $2\,\mathrm{m}$ distance to the center. b) Complex FRF. c) Convergence study w.r.t. the number of training points.

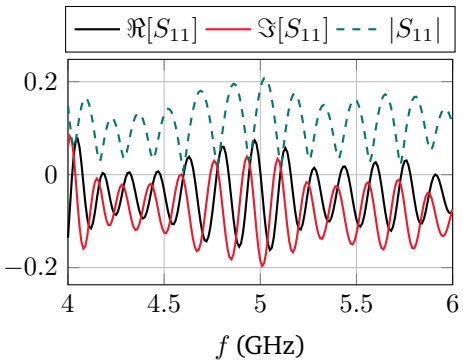
This leads to two additional poles which are closer to the input domain, as illustrated by the red crosses in Fig. 5.1b. The corresponding admittance $Y^{\{2\}}(\mathrm{i}\omega)$ only minor differs from $Y^{\{1\}}(\mathrm{i}\omega)$, except for two sharp peaks, as can be seen in Fig. 5.2. However, as $Y^{\{2\}} \in H^2_{\mathrm{sym}}(\Upsilon_{50})$ belongs to a larger Hardy space than $Y^{\{1\}} \in H^2_{\mathrm{sym}}(\Upsilon_{400})$, the accuracy of the respective RKHS interpolation is significantly affected. In particular, Figs. 5.3c-5.3d show that the convergence order of Szegö kernel interpolation is significantly reduced. However, the adaptive RKI algorithm is able to *cancel* the impact of the *dominant* poles limiting the region of analyticity. Accordingly, it exhibits fast convergence and an improvement w.r.t. AAA and VF can again be observed.

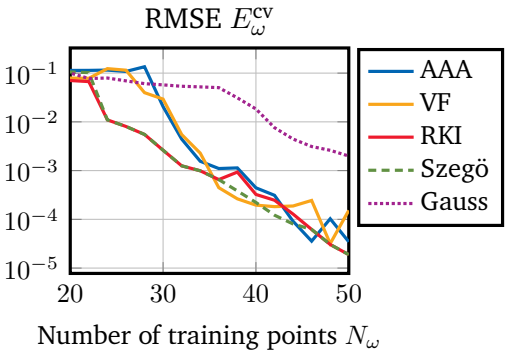
5.1.2 Wave propagation models

In the following, we investigate a number of PDE-based examples. As these models are primarily employed as benchmarks here, we mainly focus on the numerical results in this section and refer to the literature for specific details about the considered models. We start with the acoustic Helmholtz equation, particularly the PAC-MAN benchmark example, introduced in [217] which is also included in the platform for benchmark cases in computational acoustics from the European acoustics associations [107]. The model, shown in Fig. 5.4a, has the PAC-MAN shape with an opening angle of $30°$ and a radius of $1\,\mathrm{m}$. As in [217, Section 6.1], we consider as excitation a vibration of the surface of the PAC-MAN with cylindrical modes and observe the radiated field p at a point $\mathbf{r}_i$ in $2\,\mathrm{m}$ distance at an angle of $10°$. As in [107], the computation was done based on the implementation of the analytical solution provided in [217] by replacing the python module scipy with mpmath for the computation of higher order Bessel functions. In particular, we set the truncation order to 300. The complex acoustic pressure field phasor p of the total sound-field w.r.t. the frequency $f \in [2000\,\mathrm{Hz}, 4000\,\mathrm{Hz}]$ is shown in Fig. 5.4b. We then conduct a convergence study w.r.t. the number of training points, which is depicted in Fig. 5.4c. It can be observed that the Szegö kernel-based interpolation approaches outperform the alternative approaches in the range up to about 40 training points. However, adding the rational basis functions does not further improve the accuracy but does not significantly harm the accuracy either.

Next, we consider two electromagnetic models problems which are demo examples of CST Microwave Studio [65], solving the full Maxwell's equations in frequency domain. The first model is a spiral antenna, depicted in

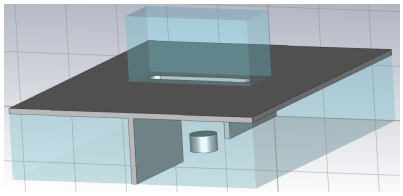


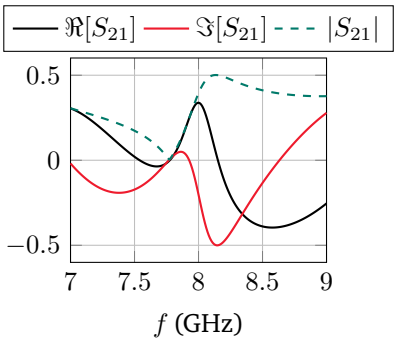


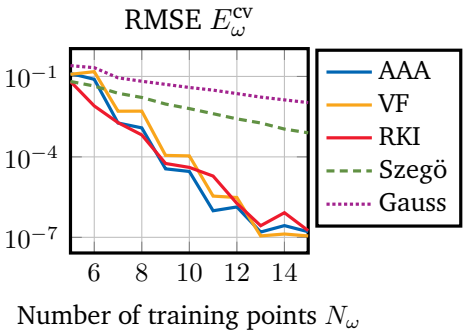


(a) Antenna model. **(b)** Complex FRF. **(c)** Convergence study.

Figure 5.5: a) Spiral antenna model, taken from CST Microwave Studio [65]. b) Complex FRF S_{11}. c) Convergence study w.r.t. the number of training points.

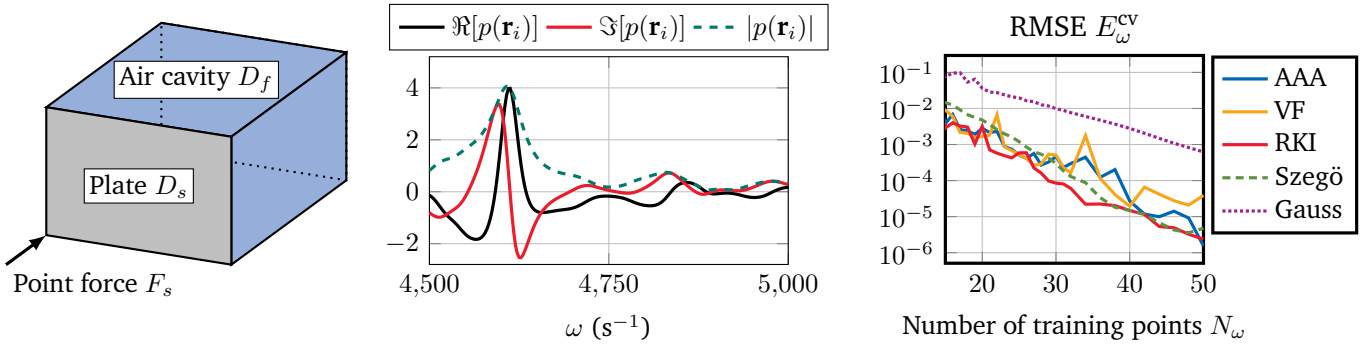






(a) Waveguide junction model. **(b)** Complex FRF. **(c)** Convergency study.

Figure 5.6: a) Waveguide junction model, taken from CST Microwave Studio [65]. b) Complex FRF S_{21}. c) Convergence studies w.r.t. the number of training points.



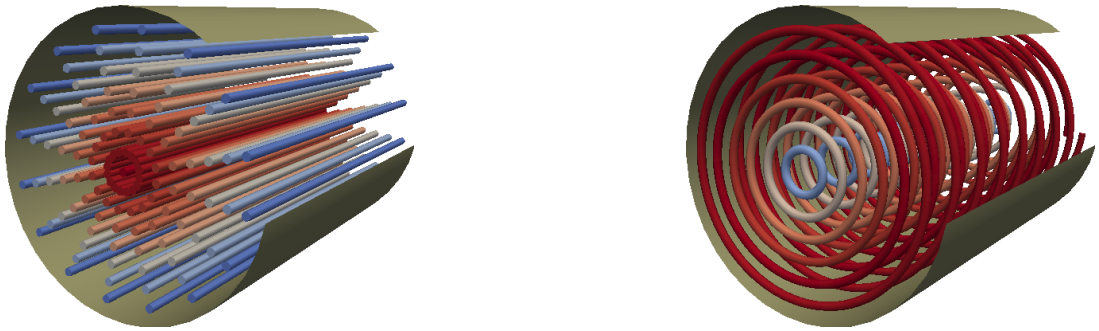
(a) Vibro-acoustic model. **(b)** Complex FRF. **(c)** Convergence study.

Figure 5.7: a) Vibro-acoustic benchmark problem, based on [170]. b) Complex FRF. c) Convergence study w.r.t. the number of training points.

Fig. 5.5a, where we consider the reflection coefficient S_{11} on a frequency range of $[4\,\text{GHz}, 6\,\text{GHz}]$ as QoI, which is shown in Fig. 5.5b. The data sets are obtained using the BEM in CST Microwave Studio [65]. The results are qualitatively the same as for the PAC-MAN model, see Fig. 5.5c. The second electromagnetic problem is a waveguide junction model with 4 ports, which contains a small metallic disk and is connected to an external cavity resonator, see Fig. 5.6a. The structure is excited at the first port and simulated using the FEM in the frequency domain. In particular, we set the solver accuracy of the 3rd order solver to 10^{-6} and use a curved mesh with standard settings. We employ an initial adaptive mesh refinement at $9\,\text{GHz}$, where we set the S-parameter criterion threshold with 2 subsequent checks to 10^{-4}. As QoI we consider the scattering parameters on a frequency range of $[7\,\text{GHz}, 9\,\text{GHz}]$ using equidistant sample points, where we restrict ourselves to S_{21} for brevity; however, the results are qualitatively similar for all 4 scattering parameters. It can be seen in Fig. 5.6b that the QoIs are *simpler* than in the previous examples but have a dominant pole at around $8\,\text{GHz}$. This causes the purely kernel-based interpolations to be inferior compared to the rational approximations. However, the proposed combination of kernel-based interpolation and rational approximations leads to satisfactory results with comparable accuracy as AAA and VF, see Fig. 5.6c.

The final test case is a vibroacoustic FE model, taken from [170] and depicted in Fig. 5.7a. A 2D Mindlin plate (vibrating structure D_s) is excited by a point force and strongly coupled to a 3D acoustic domain (air cavity D_f). Then, the response p at a point $\mathbf{r}_i$ in the fluid is evaluated. For details on the model, we refer to [170]. We consider the frequency response on a frequency interval $\omega \in [4500\,\text{s}^{-1}, 5000\,\text{s}^{-1}]$, shown in Fig. 5.7b. The convergence study, given in Fig. 5.7c, indicates that the proposed RKI approach usually achieves at least a comparable accuracy as AAA and VF while at certain points an improvement by about an order of magnitude can be observed. It can also be seen that the additional rational basis functions in the suggested RKI algorithm improve the approximation accuracy at the majority of points compared to the pure Szegö kernel-based interpolation.

In summary, it was observed that the suggested RKI algorithm yields satisfactory results for all considered benchmark models, leading to, at least, a comparable accuracy as the alternative approaches. In all cases, the *naive* approach of employing real Gauss-kernel interpolation for real and imaginary part separately, was outperformed by the rational approximation techniques as well as the complex/real Szegö kernel interpolation. Hence, it can be concluded that the data is used more efficiently if the individual interpolation of real and imaginary part is avoided. It could also be observed that adding a few rational basis functions for the RKI method yields strong improvements with respect to the complex/real Szegö kernel interpolation for the subset of benchmark functions with a few *dominant* poles. We emphasize that the resulting RKI method



(a) Electric field lines.

(b) Magnetic field lines.

Figure 5.8: Electromagnetic field of the TM_{010} mode in a pillbox cavity, taken from [88, Figure 6].

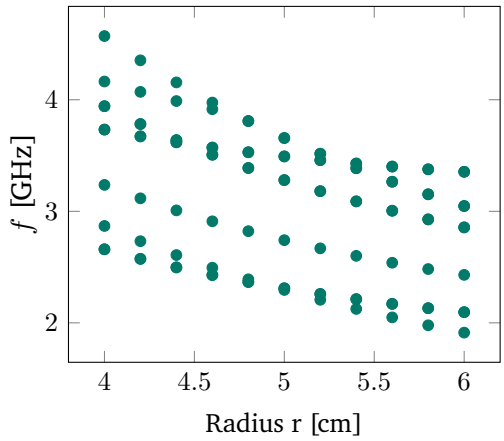
even shows significant accuracy improvements for the majority of benchmark models with respect to the established state-of-the-art techniques AAA and VF. Furthermore, the suggested method should allow for a straightforward incorporation of adaptive sequential sampling strategies, which is still subject to further research. In addition, further work could also address the efficient inclusion of derivative information, which can often be obtained with reduced computational cost, e.g. using adjoint techniques as discussed in Chapter 4.

5.2 Uncertainty quantification for Maxwell's eigenproblem

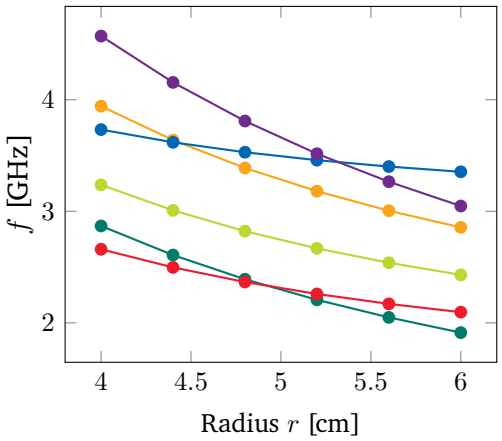
In this section, we apply the presented UQ methods from Chapter 4 to accelerator cavities with uncertain shapes. First, we consider the academic example of a cylindrical (pillbox) cavity with an uncertain radius to investigate the tracking method discussed in Section 4.5 and show that its application is clearly necessary in certain cases. Afterward, we address UQ for a 9-cell TESLA type cavity [8]. We will first study the impact of eccentric deformations [88] before we address variations of the iris and equatorial radii [89] in the subsequent section. In all cases, the model problem is a parameterized eigenvalue problem in the form of (2.53) where we treat the eigenfrequencies as QoIs and then employ SC for UQ as described in Chapter 4. The content and structure of this section are based on our works [60, 88]. Note that UQ for accelerator cavities has also been addressed in other references, see [1, 58, 180, 209]. However, in this section, we demonstrate the need for eigenvalue tracking techniques and employ a more sophisticated SC method in order to take a large number of uncertain parameters into account, i.e. up to 19 parameters. Furthermore, we address correlated input data and the numerical studies are substantiated by the available measurement data for the EXFEL cavities [208].

5.2.1 Pillbox cavity

A pillbox cavity, i.e. a cavity with a cylindrical shape, is considered as a numerical benchmark problem with an available closed-form reference solution [110, Chapter 8.7]. The electromagnetic field of the fundamental TM mode in a pillbox cavity is illustrated in Fig. 5.8. The cylindrical shape is assumed to have a length $l = 10\,\mathrm{cm}$ and an uncertain radius modeled as a uniformly distributed RV $r \sim \mathcal{U}(4\,\mathrm{cm}, 6\,\mathrm{cm})$. The discrete parameterized eigenvalue problem (2.53) is, in this case, obtained by employing a curl-conforming IGA discretization [38]. IGA [108] aims for an exact representation of the shapes defined by a computer-aided design (CAD) software



(a) Separate computation of eigenfrequencies for different pillbox radii.



(b) Tracked eigenfrequencies (markers indicate the sample points).

Figure 5.9: Illustration of the 10 lowest eigenfrequencies (including degenerate modes) of a pillbox cavity for different values of the radius r. Note that the tracking procedure only follows the 10 eigenmodes with the smallest eigenvalue at $r = 6\,\mathrm{cm}$, which do not necessarily coincide with the first 10 eigenmodes for other radii. Hence, the difference between (a) and (b) regarding the computed eigenfrequencies f_i, for instance at $r = 4\,\mathrm{cm}$, is expected. The illustrations are based on [88, Figure 7].

by employing B-splines or non-uniform rational B-splines (NURBS) [29, 166] as basis functions (2.52). It was shown in [58, 59] that using IGA for cavity simulations can lead to an improved accuracy or reduced computational cost. In addition, the isogeometric basis functions often allow for a straightforward and smooth domain deformation, i.e. by moving control points, see [88] for details. Furthermore, the accurate representation of the simulation domain is also particularly appealing in the context of shape UQ. Hence, we employ the GeoPDEs and NURBS Octave packages [76] for an IGA discretization with second order basis functions, leading to $N_h = 21692$ degrees of freedom (DoFs).

To show the relevance of mode tracking, we illustrate the eigenfrequencies of (2.53) at discrete points $\{r_i\}_{i=0}^{10} \subset [4\,\mathrm{cm}, 6\,\mathrm{cm}]$ in Fig. 5.9a. The eigenvalue problems are solved using the command EIGS in Matlab R2017a, i.e. based on Arnoldi's method [129]. It can be seen that a tracking method is required in order to assign the eigenvalues to specific modes without relying on cumbersome post-processing heuristics [34]. Next, we then employ the tracking procedure introduced in Section 4.5 where we use, for this first example, the physical homotopy (4.123) instead of the algebraic homotopy (4.122). This ensures that all computed eigenvalues are associated to physical problems and enables us to compare with the available reference solution [110] for all intermediate points during the tracking. However, as we thus cannot employ (4.124), the derivatives $\mathbf{K}'(r), \mathbf{M}'(r)$ are here approximated with finite differences. The result of the tracking for the 10 lowest eigenfrequencies at $r = 6\,\mathrm{cm}$ is shown in Fig. 5.9b. It can be observed that the transition at $r \approx 4.92\,\mathrm{cm}$ of the fundamental mode from TM to TE which is predicted by theory, is correctly captured by the tracking method. It should be mentioned that in Section 4.5 we made for simplicity the assumption of non-degenerate eigenmodes at $t = 0$, which corresponds to $r = 6\,\mathrm{cm}$ in this test case. Although this assumption is violated here, the tracking algorithm proposed in Fig. 4.18 can nevertheless be employed, as the symmetry of the geometry which causes the degeneracy is preserved for all values of r, or equivalently t.

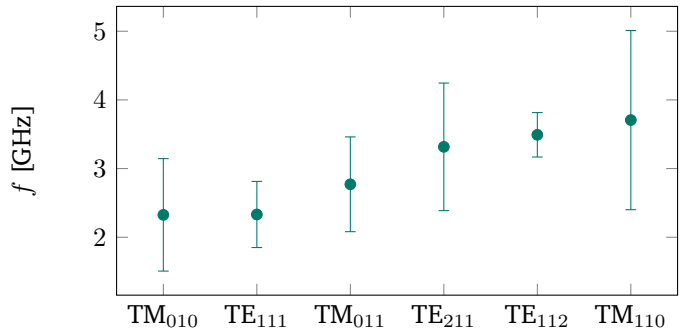


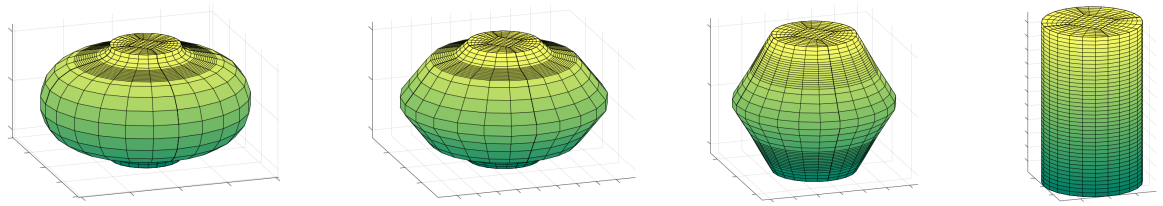
Figure 5.10: Mean value and 3σ-interval for the eigenfrequencies of the pillbox cavity with uncertain radius, based on [88, Figure 8].

Next, we address UQ for the 6 different eigenmodes with the lowest eigenfrequencies at $r = 5\,\mathrm{cm}$. In particular, we employ SC based on 5 univariate Clenshaw Curtis collocation nodes. Only the initial eigenvalue problem for the first collocation point at $r = 5\,\mathrm{cm}$ is solved with EIGS, for the remaining collocation points the tracking method is applied. We note that we always select a fitting initial stepsize such that the respective collocation point is reached within one step. We further note that the stepsize is never reduced as the Newton iteration (4.129) converges in all cases using less than 5 iterations. In particular, the Newton scheme takes on average only 2.2 iterations to converge. Finally, we compute the mean and the standard deviation of the respective eigenfrequencies, which are depicted in Fig. 5.10. For all modes, the relative errors in the numerical estimates of both, expectation and standard deviation, are below $3.5 \cdot 10^{-4}$ (with respect to the closed-form reference solution).

We recall that Fig. 5.9b clearly shows that the tracking procedure is indeed necessary to obtain the correct UQ results. However, as will be demonstrated in the following, it is also computationally more efficient than the alternative approach of solving the eigenvalue problems at each collocation point independently. We note that all respective computations are executed on a workstation with an Intel(R) Xeon(R) CPU E5-2687W $3.1\,\mathrm{GHz}$ processor and 256 GB RAM. To allow for a fair comparison, we employ the Matlab *backslash* operator, i.e. sparse Gaussian elimination, to solve the occurring linear equation systems in all eigenvalue solvers which are considered. All computations are then repeated 10 times and the respective time measurements averaged.

To evaluate the computational cost of the alternative approach of solving independent eigenproblems at each collocation point, we employ on the one hand the Matlab command EIGS which, in turn, internally relies on the ARPACK library [128] and on the other hand a Matlab implementation of the Jacobi-Davidson algorithm [79, 187]. In both cases, we calculate 20 eigenvalues at each collocation point, as we are interested in UQ for 10 eigenmodes and want to ensure that the respective eigenvalues are always included. This leads to an averaged computation time of $192.6\,\mathrm{s}$ for solving 80.6 linear systems in the case of EIGS and $931.4\,\mathrm{s}$ for, on average, 42.0 iterations of the Jacobi-Davidson algorithm. Next, we discuss the computational cost of the eigenvalue tracking method, which consists of the time to compute the derivatives by solving (4.127) and the cost associated to the Newton iterations (4.129). Note that we exclude the cost for assembling the system matrices because for UQ one could employ the algebraic homotopy (4.122) where no additional matrix assemblies are needed. This leads to a computation time of $16.0\,\mathrm{s}$ for solving 3.2 linear equations systems per collocation point and eigenpair.

Finally, we further validate the tracking method by considering a more sophisticated shape deformation. To this end, we deform a 1-cell TESLA cavity shape [8] to a cylindrical pillbox cavity with length $l = 11.54\,\mathrm{cm}$



(a) $t = 0.0$ **(b)** $t = 0.5$ **(c)** $t = 0.75$ **(d)** $t = 1$

Figure 5.11: Deformation of a 1-cell TESLA cavity ($t = 0$) to a pillbox cavity ($t = 1$), taken from [88, Figure 6].

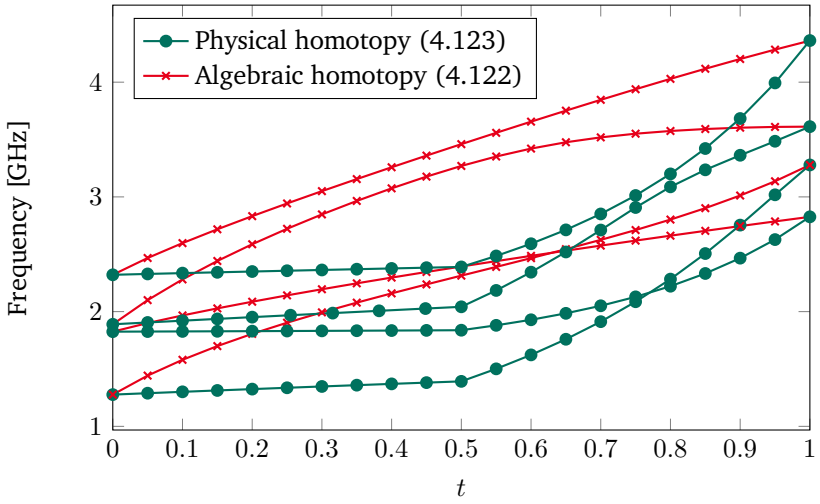


Figure 5.12: Application of the eigenvalue tracking method during the shape deformation of a 1-cell TESLA cavity ($t = 0$) to a pillbox cavity ($t = 1$), based on [88, Figure 10].

and radius $r = 3.5\,\mathrm{cm}$, as depicted in Fig. 5.11. For the discretization we use $N_h = 20\,712$ DoFs which are, again, associated to second order basis functions. We then employ both homotopies, i.e. the algebraic homotopy (4.122) and the physical homotopy (4.122) which is associated to the illustrated shapes. Fig. 5.12 shows the result of the tracking algorithm applied for both homotopies and for the lowest eigenfrequencies of the TESLA cavity. We emphasize that both approaches, as expected, lead to the same pillbox eigenmodes.

5.2.2 TESLA cavity with eccentric uncertainty

Next, we consider a real-world example, i.e. the TESLA cavity, illustrated in Fig. 2.1. A large number of these superconducting RF cavities are, for instance, installed at DESY in Hamburg. Each cavity is a 9-cell structure with a length of about $1\,\mathrm{m}$. A detailed description of the production process will be given in the subsequent section; here we only note that single half-cells are produced by deep drawing of niobium sheets which are then welded together using electron-beam welding (EBM) [71]. In this section, we study the impact of possible misalignment of the cavity shape w.r.t. to the ideal axis due to manufacturing imperfections. To this end, we collect measurement data of the individual cell center coordinates for approximately $N_{\mathrm{meas}} \approx 700$ cavities from the DESY database [84, 208], as illustrated for one particular cavity in Fig. 5.13. In particular,

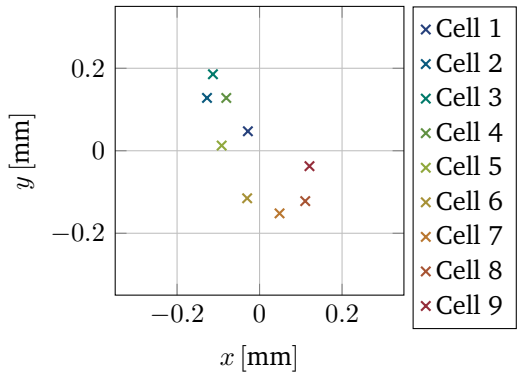


Figure 5.13: Measurements of the cell center coordinates for a particular TESLA cavity, based on [88, Figure 5.7].

this data is then arranged in a matrix

$$\mathbf{T} = \begin{bmatrix} x_{\text{cell},1}^{(1)} & y_{\text{cell},1}^{(1)} & x_{\text{cell},2}^{(1)} & y_{\text{cell},2}^{(1)} & \cdots & x_{\text{cell},9}^{(1)} & y_{\text{cell},9}^{(1)} \\ x_{\text{cell},1}^{(2)} & y_{\text{cell},1}^{(2)} & x_{\text{cell},2}^{(2)} & y_{\text{cell},2}^{(2)} & \cdots & x_{\text{cell},9}^{(2)} & y_{\text{cell},9}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{\text{cell},1}^{(N_{\text{meas}})} & y_{\text{cell},1}^{(N_{\text{meas}})} & x_{\text{cell},2}^{(N_{\text{meas}})} & y_{\text{cell},2}^{(N_{\text{meas}})} & \cdots & x_{\text{cell},9}^{(N_{\text{meas}})} & y_{\text{cell},9}^{(N_{\text{meas}})} \end{bmatrix} \in \mathbb{R}^{N_{\text{meas}} \times 18}, \tag{5.10}$$

where $(x_{\text{cell},i}^{(j)}, y_{\text{cell},i}^{(j)})$ denote the center coordinates of the i-th cell of the j-th cavity w.r.t. the ideal cavity axis. We note that each of the 18 columns approximately corresponds to realizations of a normally distributed RV, however, there is a significant correlation between the coordinates of neighboring cells which can already be seen in Fig. 5.13. Furthermore, considering 18 RVs with a spectral UQ method is challenging due to the curse-of-dimensionality as discussed in Section 4.1.3. Hence, we apply a (truncated) discrete Karhunen–Loève expansion (or equivalently principal component analysis [211]) in order to obtain (a reduced number of) uncorrelated variables, in the following.

To this end, we introduce the symmetric and positive semi-definite covariance matrix $\mathbf{C} \in \mathbb{R}^{18\times 18}$ associated to (5.10) where

$$C_{ij} := \frac{1}{N_{\text{meas}} - 1} \sum_{m=1}^{N_{\text{meas}}} (T_{mi} - \mu_i)(T_{mj} - \mu_j), \tag{5.11}$$

and $\boldsymbol{\mu} \in \mathbb{R}^{18}$ denotes the sample mean

$$\mu_i = \frac{1}{N_{\text{meas}}} \sum_{m=1}^{N_{\text{meas}}} T_{mi}. \tag{5.12}$$

We then compute an eigendecomposition of $\mathbf{C}$ as

$$\mathbf{C} = \mathbf{V}\boldsymbol{\Sigma}\mathbf{V}^\top, \tag{5.13}$$

where $\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_1, \ldots, \sigma_{18})$ denotes a diagonal matrix containing the eigenvalues $\{\sigma_i\}_{i=1}^{18}$ of $\mathbf{C}$ and $\mathbf{V} \in \mathbb{R}^{18\times 18}$ contains the corresponding orthonormal eigenvectors. In the following, we assume without loss of generality that $\boldsymbol{\Sigma}, \mathbf{V}$ are ordered s.t. $\sigma_1 \geq \ldots \geq \sigma_{18}$. Then, the least relevant contributions can be neglected by approximating the covariance matrix as

$$\mathbf{C} \approx \mathbf{C}_{\text{t}} := \mathbf{V}_{\text{t}}\boldsymbol{\Sigma}_{\text{t}}\mathbf{V}_{\text{t}}^\top, \tag{5.14}$$

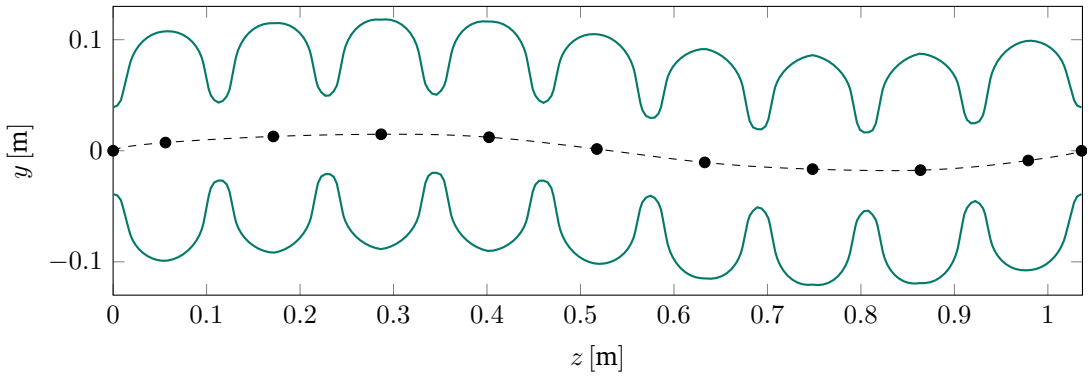


Figure 5.14: Illustration of a 9-cell TESLA cavity with eccentric deformation where the cell centers are illustrated by black dots. For visualization purposes the deformation is scaled by a factor of 500. The illustration is based on [88, Figure 13].

Refinement level	N_h	$f_{\mathrm{TM}_{010}}$ [GHz]	Δ [kHz]
1	7772	1.313 948 189	-
2	24 960	1.301 478 100	12 470.089
4	111 140	1.299 939 735	1538.365
6	299 992	1.299 971 920	32.185
8	631 836	1.299 969 453	2.467

Table 5.1: The eigenfrequency $f_{\mathrm{TM}_{010}}$ of the accelerating mode in the 9-cell TESLA cavity is numerically evaluated for different levels of mesh refinement. In the last column, the difference w.r.t. the eigenfrequency of the previous refinement level is shown. The table is taken from [88, Table 1].

where $\mathbf{\Sigma}_\mathrm{t} = \mathrm{diag}(\sigma_1, \dots, \sigma_{N_\mathrm{t}})$ with $N_\mathrm{t} < 18$ and $\mathbf{V}_\mathrm{t} \in \mathbb{R}^{18\times N_\mathrm{t}}$ contains the respective eigenvectors. Accordingly, the 18 RVs associated to the cell center coordinates can be approximately represented as

$$\begin{bmatrix} x_{\mathrm{cell},1} \\ y_{\mathrm{cell},1} \\ \vdots \\ x_{\mathrm{cell},9} \\ y_{\mathrm{cell},9} \end{bmatrix} \approx \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_{17} \\ \mu_{18} \end{bmatrix} + \mathbf{V}_\mathrm{t}\mathbf{\Sigma}_\mathrm{t}^{1/2} \begin{bmatrix} \Xi_1 \\ \vdots \\ \Xi_{N_\mathrm{t}} \end{bmatrix}, \tag{5.15}$$

where $\mathbf{\Xi}$ is then given as standard normally distributed random vector with N_t mutually independent elements [62]. In particular, in this case we choose $N_\mathrm{t} = 7$ as it fulfills the following condition for the explained variance

$$\frac{\sum_{i=1}^{N_\mathrm{t}=7} \sigma_i}{\sum_{i=1}^{N=18} \sigma_i} \geq 0.95. \tag{5.16}$$

We note that $\mathbf{\Xi}$ may not have a clear physical meaning in contrast to the cell center coordinates, however, the reduced number of RVs, i.e. 7 instead of 18, strongly improves the computational efficiency of an SC method.

We then employ SC on a Smolyak sparse grid as introduced in Section 4.1.3 of level 2 based on Gaussian points, which has in total 127 collocation points [10, 155]. For each collocation point, we can construct the respective

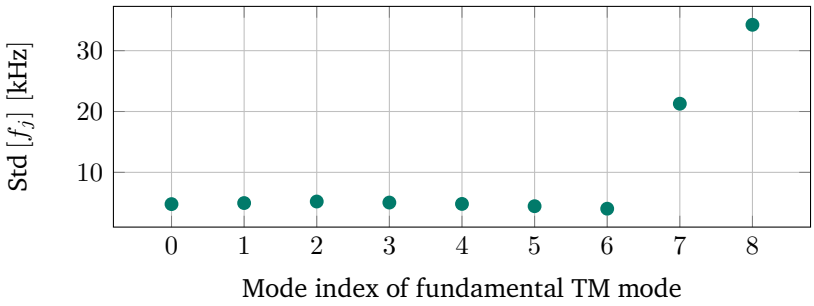


Figure 5.15: Standard deviations of the eigenfrequencies in the first passband for a TESLA cavity with eccentric deformations which are 50 times bigger than the measured ones, based on [88, Figure 12].

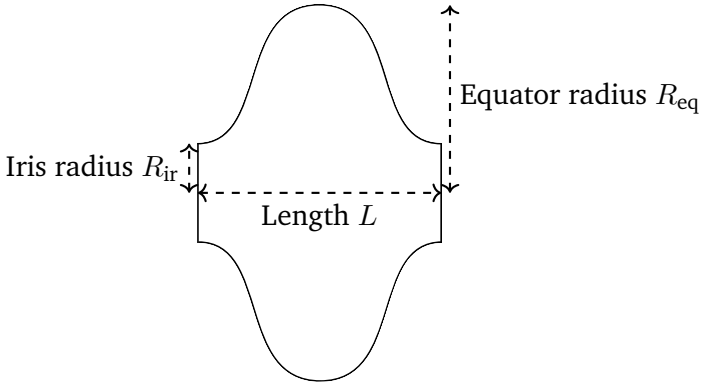


Figure 5.16: Shape parameters $L, R_{\text{ir}}, R_{\text{eq}}$ of a TESLA cavity cell.

deformed geometry based on the corresponding cell center coordinates as illustrated in Fig. 5.14, see [88] for details. Then, the eigenvalue tracking algorithm can be applied for each collocation point based on the algebraic homotopy (4.122). However, the measured eccentric deformations, which are for instance shown in Fig. 5.13, only mildly affect the eigenfrequencies of the first monopole passband, i.e. the magnitude of the variation is below $1\,\text{kHz}$ and can be compensated by the initial compulsory tuning procedure [123]. Hence, for the numerical study, we artificially scale the deformations by 50 to show that the proposed UQ workflow can be employed for arbitrary variations in the geometry where for instance linearization approaches might fail. The frequency shifts for the fundamental modes caused by these increased deformations have a magnitude of about $\sim 10\,\text{kHz}$ and, hence, an IGA discretization with second order basis functions and $N_h \approx 600000$ provides a suitable accuracy, see Table 5.1. Finally, the calculated standard deviations of the eigenfrequencies of the fundamental monopole modes are shown in Fig. 5.15. It can be observed that the first 7 modes are significantly less sensitive than the last two modes. However, in practice, before installation of the cavity, the eigenfrequency and field flatness of the accelerating mode is tuned by further mechanical deformations with an automatic cavity tuning machine which was not yet considered in this study. It can be expected that this procedure then reduces the uncertainty in the last modes, as will be shown in the next section, where we address a UQ analysis that takes the full cavity production chain into account.

5.2.3 TESLA cavity with uncertain radii

We now report on a comprehensive UQ study for the EXFEL cavities taking another source of uncertainty into account. In particular, Fig. 5.16 depicts the shape of a TESLA cavity cell where we will consider random

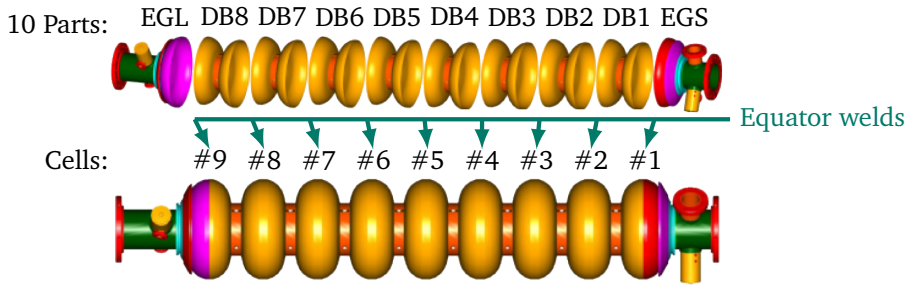


Figure 5.17: Illustration of the welding of 8 dumb-bells (DB) and end-groups (EGS, EGL) to one EXFEL cavity, taken from [60, Figure 2].

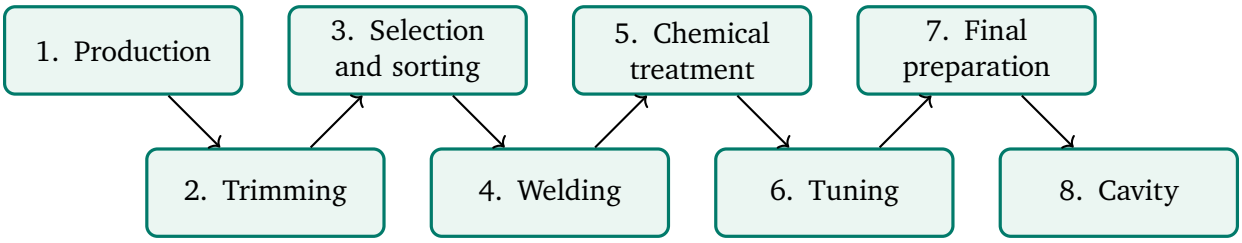


Figure 5.18: Flowchart describing the TESLA cavity manufacturing chain, based on [60, Figure 3c].

perturbations of the equatorial radii and iris radii of each cavity cell. For this study, we will discuss all steps of the manufacturing chain and propose to mimic most of them in a simulation workflow. Surrogate modeling based on dimension-adaptive SC as introduced in Section 4.2.2 is employed to keep the computational workload manageable and allows to quantify uncertainties and compute sensitivity indices, in particular Sobol indices and Borgonovo indices which were introduced in Section 4.1.4.

We address, again, the parametrized eigenvalue problem (2.53), where the domain D corresponds to the shape of a TESLA cavity with $N_\mathrm{c} = 9$ cells. We now assume that the shape is parameterized in terms of the equatorial radii $R_\mathrm{eq}^{(i)}$, $i = 1, \dots, N_\mathrm{c}$ and the iris radii $R_\mathrm{ir}^{(i)}$, $i = 1, \dots, N_\mathrm{c}+1$ of the different cells. The perturbations of these values with respect to their nominal values are then denoted by $\Delta R_\mathrm{eq}^{(i)}$ and $\Delta R_\mathrm{ir}^{(i)}$ which are considered as parameters in the following, i.e.

$$\boldsymbol{\xi} = [\Delta R_\mathrm{eq}^{(1)}, \dots, \Delta R_\mathrm{eq}^{(9)}, \Delta R_\mathrm{ir}^{(1)}, \dots, R_\mathrm{ir}^{(10)}]. \tag{5.17}$$

The parameterized simulation model is, in this case, set up based on Matlab and the Superlans code [146], i.e. a 2D-axisymmetric FEM model, which can now be employed as the parameters (5.17) do not lead to eccentricity and, hence, preserve the azimuthal symmetry of the cavity.

The EXFEL [4] accelerates electrons to a particle velocity close to the speed of light with an energy of up to $17.5\,\mathrm{GeV}$ by employing 808 TESLA cavities. Each of the $N_\mathrm{cav} > 808$ produced cavities consists of 8 dumb-bells and 2 end-groups, which are welded together, as illustrated in Fig. 5.17. To compensate for manufacturing imperfections and obtain acceptable tolerances, e.g. to ensure a field flatness $\mathrm{FF} > 90\%$ and a deviation of the π-mode frequency from its nominal value $|f_8 - 1.3\,\mathrm{GHz}| < 100\,\mathrm{kHz}$, dedicated procedures are employed. These different procedures are summarized in Fig. 5.18. In the following, we discuss the different steps and how they can be represented in the simulation workflow.

1. **Step:** Production
   In this step, the main subcomponents of N_cav cavities, as shown in Fig. 5.17, are produced. To mimic

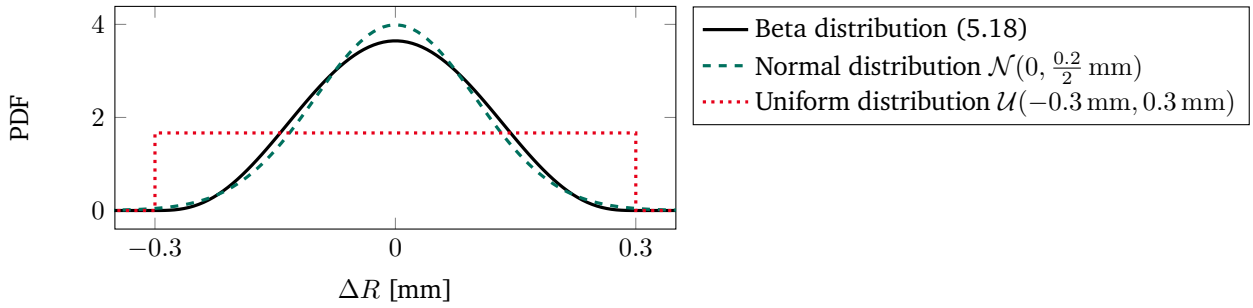


Figure 5.19: Comparison of different PDFs. The beta distribution and the uniform distribution have bounded support in $[-0.3\,\mathrm{mm}, 0.3\,\mathrm{mm}]$ while the normal distribution has unbounded support. The illustration is based on [60, Figure 7].

the imperfect manufacturing process on the simulation side, we create $\tilde{N}_{\mathrm{cav}} = 10^6$ random ("virtual") cavities. To this end, we generate for each virtual cavity seven independent random mid cells and two end cells by drawing random realizations of the random vector $\boldsymbol{\Xi}_{\mathrm{prod}}$ associated to the parameters given in (5.17). We assume that the elements of $\boldsymbol{\Xi}_{\mathrm{prod}}$ are mutually independent which is often justified for mass-production processes where the individual parts are produced independently of each other. We further assume that each RV $\Xi_{\mathrm{prod},i}$ is beta-distributed with PDF

$$\rho(\xi) = \frac{140}{(u-l)^7} \begin{cases} (\xi - l)^3 (u - \xi)^3, & l < \xi < u, \\ 0, & \text{else}, \end{cases} \tag{5.18}$$

where $l = -0.3\,\mathrm{mm}$ is the lower bound and $u = 0.3\,\mathrm{mm}$ the upper bound. As can be seen in Fig. 5.19, the particular beta distribution (5.18) represents an approximation of a normal distribution with a 2σ interval of $\pm 0.2\,\mathrm{mm}$. However, contrary to the respective normal distribution, (5.18) has bounded support [211, Appendix B] which avoids non-physical parameter values in numerical studies.

2. **Step:** Trimming
   The dumb-bells and end groups are trimmed in order to compensate for shape deviations and achieve a suitable cavity length. However, as we generate the virtual cavities based on elementary cells instead of dumb-bells, an explicit modeling of the trimming is not possible but it will be indirectly considered due to the length constraint introduced in the following step.

3. **Step:** Selection and sorting
   Two end-groups and 8 dumb-bells are selected and the dumb-bells then sorted such that the impact of the manufacturing imperfections on the desired field distribution of the HOMs is minimized. To this end, a dumb-bell with average frequency is placed at position 8 (last position) while the remaining dumb-bells are sorted w.r.t. a decreasing fundamental eigenfrequency. The corresponding virtual 9-cell cavity is constructed from the random virtual components of the first step by applying a respective sorting procedure: For all middle cells the fundamental eigenfrequency is computed by solving a one-cell-eigenvalue problem. Then, the cell with the eigenfrequency which is the closest to the average frequency is placed at position 8 while the remaining cells are ordered according to a decreasing frequency on positions 2 to 7. Furthermore, a constraint on the total cavity length

$$\sum_{i=1}^{9} \Delta L^{(i)} < 3\,\mathrm{mm}, \tag{5.19}$$

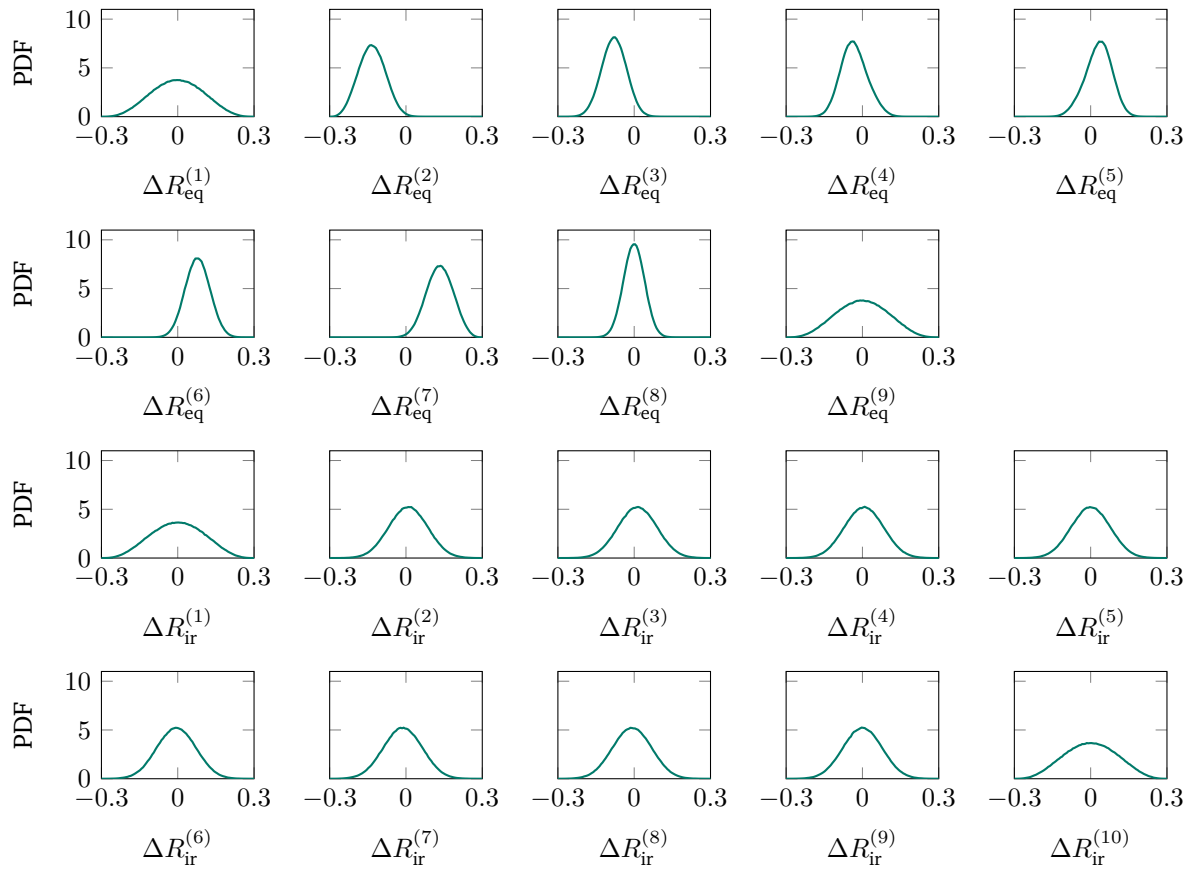


Figure 5.20: Estimated PDFs of the 19 correlated input RVs, based on [60, Figure 8].

will be imposed. If (5.19) is not fulfilled after the tuning in the 6-th step, we disregard the respective virtual cavity. We note that this constraint is in real manufacturing already taken into account by trimming and compensation effects.

4. **Step:** Welding
   All components are welded to each other, see [192] for details. In the simulation workflow, given that we simulate cells instead of dumb-bells, we model this step by averaging the respective iris radii of adjacent cells. In the previous section, the effect of an eccentric miss-alignment of cells during the welding was investigated. As its impact on the fundamental eigenfrequencies was negligible, we neglect it here. We note that the procedures for selection, sorting and welding affect the PDFs of the uncertain radii. The corresponding random vector is denoted by $\boldsymbol{\Xi}_{\text{sort}}$, which now contains stochastic dependencies between its elements. The virtual cavities at this stage (which also comply with the length constraint) represent a sample $\{\boldsymbol{\Xi}_{\text{sort}}^{(m)}\}_{m=1}^{\tilde{N}_{\text{sel}}}$ of size $\tilde{N}_{\text{sel}} = 809641$. This sample is then employed to infer the univariate PDFs of $\boldsymbol{\Xi}_{\text{sort}}$ with KDE as explained in Section 4.1.1. The resulting PDFs are shown in Fig. 5.20.

5. **Step:** Chemical treatment
   The chemical treatment of the cavities reduces spikes and impurities, see [186, 194] for details. However, it cannot be modeled appropriately in the present setting as a complex parameterization and multi-scale analyses or at least very fine resolutions would be necessary.

6. **Step:** Tuning

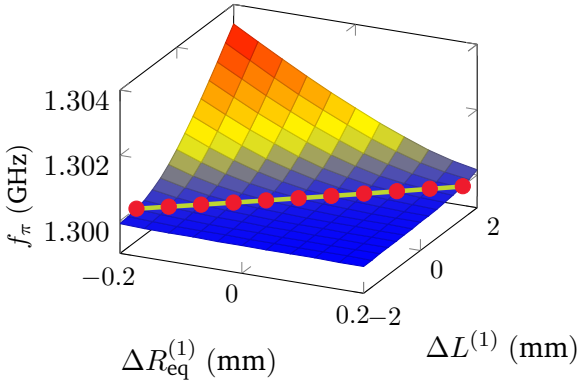


Figure 5.21: The surface shows the π-mode frequency of the 9-cell cavity w.r.t. the equatorial radius $\Delta R_{\text{eq}}^{(1)}$ and length $\Delta L^{(1)}$ of the first cell. For given values of $\Delta R_{\text{eq}}^{(1)}$, the red points indicate the result for $\Delta L^{(1)}$ obtained by the virtual tuning procedure. It can be seen that the tuning values lie on the $1.3\,\text{GHz}$ contour line shown in green. The figure is based on [60, Figure 9].

In this step, the cavity is mechanically compressed or stretched by an automatic tuning machine [123] such that the π-mode frequency and the field flatness are improved. We apply a virtual tuning scheme which tunes each individual cell to $1.3\,\text{GHz}$ by adapting its length $L^{(i)}$. To this end, a non-linear root-finding problem

$$f_0^{(i)}\left(L^{(i)}\right) - 1.3\,\text{GHz} = 0, \tag{5.20}$$

is solved where $f_0^{(i)}$ refers to the fundamental eigenfrequency of a one-cell eigenvalue problem in the i-th cell. In this case, we employ the Matlab function FZERO to solve (5.20); however, other methods, as bisection or Newton's method, would be equally possible. Fig. 5.21 shows the π-mode frequency of the 9-cell cavity with respect to the equator radius and the length of the first cell. It can be seen that the tuning procedure, as expected, chooses suitable cell lengths such that the nominal value of $1.3\,\text{GHz}$ is achieved. Furthermore, in 500 test cases, which we considered, all resulting cavities had an acceptable field flatness of $\text{FF} > 96\%$.

7. **Step:** Final preparation (for operation)
   The final preparation procedures include, for instance, a final buffered chemical polishing etching and the cooldown to $2\,\text{K}$. They are described in detail in [186, Figure 2] and differ depending on the manufacturer. This step cannot be considered in the present simulation setting because of involved numerical modeling and insufficient data.

8. **Step:** Cavity
   At this step, the fundamental mode frequencies of the N_{cav} cavities are measured under operation temperature. Accordingly, the eigenproblem (2.53) is solved for each virtual cavity in order to compute statistics based on the first nine eigenfrequencies. We note that eigenvalue tracking as described in Section 4.5 should be generally used in this step; however, we observed that for the considered particular parameter variations no crossings of the fundamental eigenfrequencies occur, see Fig. 5.22.

Next, we discuss an efficient implementation of the described simulation and UQ workflow. The tremendous computational effort of solving Maxwell's eigenproblem in steps 3 and 8 for a large number of virtual cavities can be avoided by constructing surrogate models. In particular, we employ the dimension-adaptive SC Algorithm (2) based on unweighted Leja nodes in order to reach a high uniform accuracy. For step 3, we

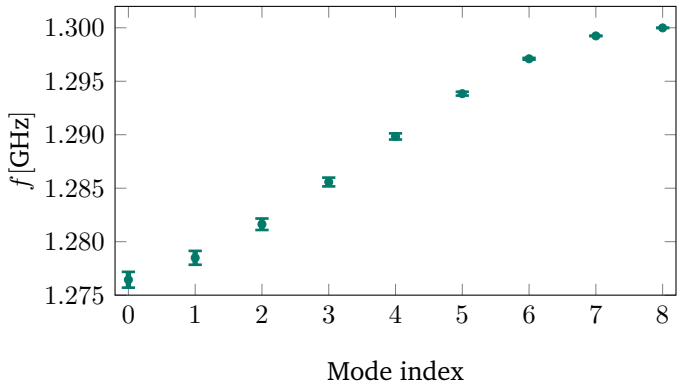


Figure 5.22: Expected eigenfrequencies of the fundamental modes in a 9-cell TESLA with uncertain radii. The bars show the 5σ intervals. The figure is based on [60, Figure 11].

Table 5.2: Expected eigenfrequencies of the fundamental modes and their standard deviations, taken from [60, Table 2].

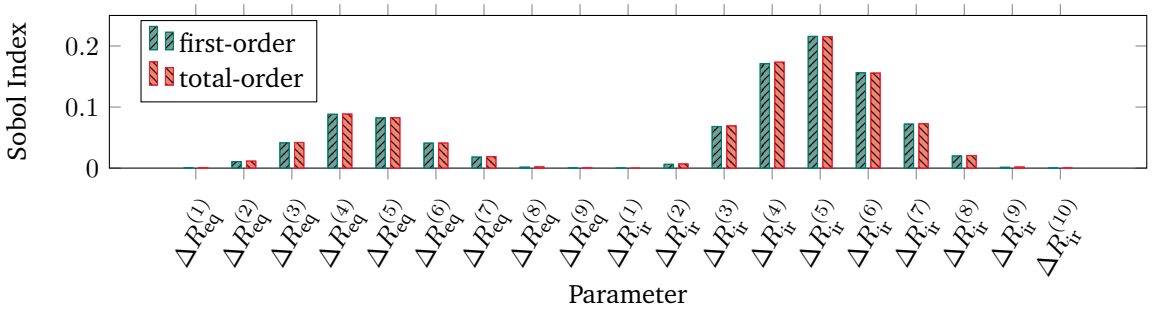
Mode	Mean [MHz]	Std. dev. [MHz]
0	1,276.45	0.15
1	1,278.50	0.13
2	1,281.64	0.11
3	1,285.60	0.08
4	1,289.85	0.06
5	1,293.84	0.04
6	1,297.11	0.02
7	1,299.25	0.00
8	1,300.00	0.00

construct 4-variate SC approximations with a computational budget $\mathcal{B}$ such that 50 one-cell eigenproblems are solved. For the 19-variate surrogate model required for step 8, the budget $\mathcal{B}$ is chosen such that 500 9-cell eigenproblems are solved. Cross-validation at random sample points indicates for all polynomial surrogates models an error in the fundamental eigenfrequencies below $10\,\mathrm{kHz}$ and, hence, also below the expected deviations.

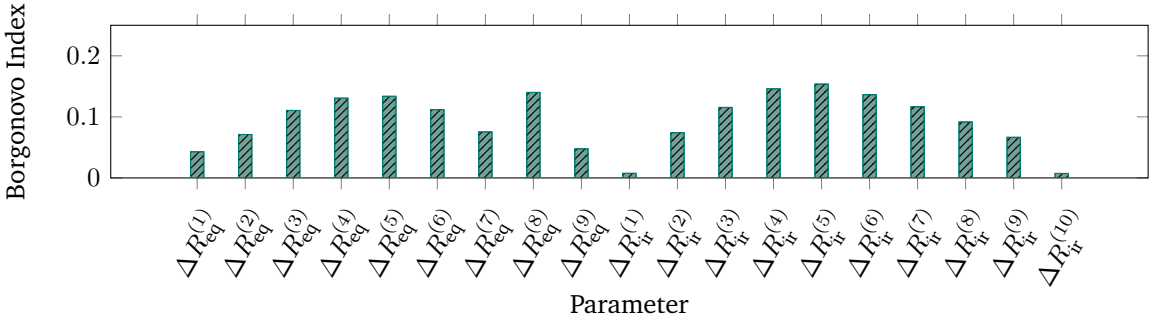
Next, all statistical measures can be computed based on the sample of eigenfrequencies evaluated in step 8. In particular, we use the MC estimates (4.1) and (4.4) to calculate the mean and the standard deviation of the eigenfrequencies, which are shown in Fig. 5.22 and Table 5.2. Note that the results agree well with the measurement data reported in [60, Figure 5 and Table 1] where an increasing standard deviation w.r.t. a decreasing mode number can be observed as well. The mean and standard deviation of the cell-to-cell coefficient k_{cc} defined in (2.7) are then computed in the same way

$$\mathbb{E}[k_{\mathrm{cc}}] \approx 1.82802, \quad \mathrm{Std}[k_{\mathrm{cc}}] \approx 0.01897. \tag{5.21}$$

In the following, we conduct a sensitivity analysis as introduced in Section 4.1.4 for the cell-to-cell coupling coefficient k_{cc}. First, we estimate Sobol indices, although they are based on the assumption of independent parameters, using the OpenTurns implementation of Saltelli's algorithm based on the kernel density estimates illustrated in Fig. 5.20. The results are computed by evaluating the surrogate model $4 \cdot 10^7$ times and shown



(a) Sobol indices.

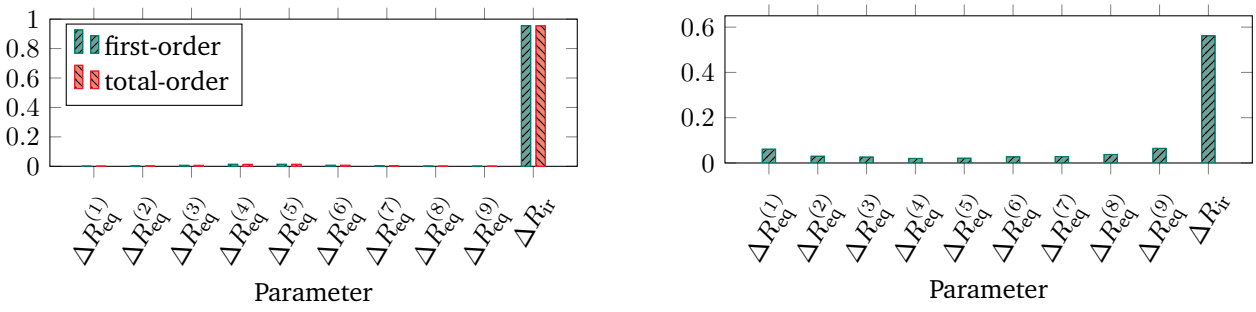


(b) Borgonovo indices.

Figure 5.23: Estimated sensitivity indices for the coupling coefficient k_{cc}, taken from [60, Figure 12].

in Fig. 5.23a. It can be seen that the variations of the cells in the center have a stronger influence on k_{cc}. Next, we estimate Borgonovo indices, which are suited for the correlated input data, using an approach based on KDE [63]. The results are shown in Fig. 5.23b. As explained in Section 4.1.4, Borgonovo indices are based on another sensitivity measure than Sobol indices, and, hence, different magnitudes are expected. However, we note that the results are still similar as they identify, again, a larger influence of the center cell variations. Contrary to the Sobol indices results, the Borgonovo index of $\Delta R_{\mathrm{eq}}^{(8)}$ is surprisingly large, which, in our opinion, can be attributed to the special treatment of the respective cell in step 3 during the sorting procedure. In particular, a large magnitude of $\Delta R_{\mathrm{eq}}^{(8)}$ can implicitly also imply larger deviations of the other middle cells, as we place the cell which has the closest eigenfrequency to the average eigenfrequency of the cavity cells on this position. We note that we also repeated the whole study where we omitted the special treatment of cell 8 in the sorting procedure, in which case the Borgonovo index of $\Delta R_{\mathrm{eq}}^{(8)}$ is then strongly reduced, confirming our interpretation.

To conduct a preliminary inverse analysis, we proceed, as in [193], for the remainder of this section with the simplifying assumption that all iris radii deviations $\Delta R_{\mathrm{ir}}^{(i)}$, $i = 1, \ldots, N_{\mathrm{c}} + 1$ of a cavity are equal and, thus, also drop the index i. This allows to estimate geometric deviations based on measurements of the cell-to-cell coupling coefficient k_{cc}. To this end, we start by rerunning the full UQ workflow where the overall iris radii deviation ΔR_{ir} is now modeled as a single beta distributed RV. The resulting sensitivity indices are presented in Fig. 5.24. It can be seen that the coupling coefficient k_{cc} is strongly impacted by ΔR_{ir} while the other parameters only have a minor influence. Note that both Sobol indices attribute more than 95% of the output variance to ΔR_{ir}. Hence, we neglect the deformations in the equatorial radii in the following, since considering all input parameters in an inverse problem would require significantly more sophisticated numerical methods, see e.g. [191]. Here, we then focus on estimating the parameter ΔR_{ir} based on the available measurement



(a) Sobol indices

(b) Borgonovo indices.

Figure 5.24: Sensitivity indices for the coupling coefficient k_{cc} when the same iris radius deviation ΔR_{ir} is considered for all cells, taken from [60, Figure 13].

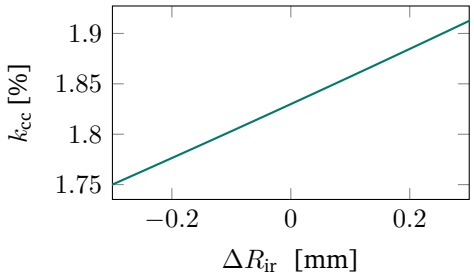


Figure 5.25: Cell-to-cell coupling k_{cc} w.r.t. variations of the iris radius, taken from [60, Figure 14].

data of the coupling coefficient k_{cc} for the EXFEL cavities. The map $\Delta R_{\text{ir}} \mapsto k_{\text{cc}}$ is illustrated in Fig. 5.25. It can be seen that, in the considered range, the function is monotonic and even almost linear. Next, we collect the measured fundamental eigenfrequencies from the DESY Database [208] for $N_{\text{cav}} = 826$ TESLA cavities. Then, by numerically inverting

$$\Delta R_{\text{ir},i} \mapsto k_{\text{cc},i},\ i = 1, \dots, N_{\text{cav}}, \tag{5.22}$$

one can estimate the parameter $\Delta R_{\text{ir},i}$ for the i-th cavity. Note that the *subscript* i, $1 \leq i \leq N_{\text{cav}}$ denotes the cavity number while previously the *superscript* i denoted the location of an individual iris radius deviation. In this case, $\Delta R_{\text{ir},i}$ is computed using the scipy implementation of the L-BFGS-B algorithm [46, 215], i.e. a quasi-Newton method. However, again, other methods, as bisection or Newton's method, could be employed as well. Finally, statistical measures of the estimated iris radius variations are given in Table 5.3. Note that according to the EXFEL specification, the iris radii before welding should be within $\pm 0.2\,\text{mm}$ around their nominal value. This distribution of iris radii is expected to vary during certain steps of the manufacturing chain, for instance the chemical treatment can have a significant impact. This shift is estimated by the presented method as $\mathbb{E}[\Delta R_{\text{ir},i}] = 0.243\,\text{mm}$, which is shown in Table 5.3 (last row). The corresponding standard deviation $\text{Std}[R_{\text{ir},i}]$, where the cavities from both vendors are considered together after the full manufacturing chain, is still within the specification limit as $0.17\,\text{mm} < 0.2\,\text{mm}$. When the manufacturers are considered separately, it can be observed that both achieve approximately a 3σ level as the respective standard deviations are only $0.057\,\text{mm}$ and $0.073\,\text{mm}$.

In summary, in this section UQ for superconducting cavities was addressed. It was demonstrated that, in general, eigenvalue tracking techniques, as proposed in Section 4.5, are necessary to ensure a consistent matching of eigenmodes and allow for an efficient SC approximation. First, the impact of eccentric deformations was studied based on existing measurement data for the EXFEL TESLA cavities where the Karhunen-Loève decomposition

Table 5.3: Sample mean and standard deviation of k_{cc} for in total $N_{\mathrm{cav}} = 826$ cavities as well as statistics about the respective estimated iris deformations, taken from [60, Table 3].

Manufacturer	$\mathbb{E}[k_{\mathrm{cc},i}]$	$\mathrm{Std}[k_{\mathrm{cc},i}]$	$\mathbb{E}[\Delta R_{\mathrm{ir},i}]$	$\mathrm{Std}[\Delta R_{\mathrm{ir},i}]$
Research Instruments GmbH (RI)	1.854	0.016	$0.087\,\mathrm{mm}$	$0.057\,\mathrm{mm}$
Ettore Zanon S.p.A. (ZA)	1.941	0.021	$0.400\,\mathrm{mm}$	$0.073\,\mathrm{mm}$
RI+ZA	1.897	0.047	$0.243\,\mathrm{mm}$	$0.170\,\mathrm{mm}$

could be applied to address correlations in the input data and reduce the dimension of the parametric space. The eccentric deformations were found to have a small impact on the fundamental mode spectra. Next, the uncertainties in the highly relevant shape parameters for the equatorial and iris radii were considered. For a proper consideration, the production chain of the EXFEL cavities was modeled by a comprehensive simulation procedure. To keep the computational effort manageable, the dimension-adaptive scheme suggested in Section 4.2.2 was employed at various steps. The estimated sensitivity indices give valuable insights, e.g. the expert knowledge that the cell-to-cell coupling coefficient is highly sensitive to the iris radius is confirmed systematically. Finally, the SC approximation was employed to estimate statistics about the actual iris radius variations based on measurements of the fundamental mode spectra at cryogenic temperatures. The estimated standard deviations are for both vendors within the specification.

5.3 Uncertainty quantification for Maxwell's source problem

In this section, we apply the proposed UQ methods to Maxwell's source problems with uncertain input data, as introduced in Chapter 2. In particular, we first numerically investigate the suggested conformally mapped (adjoint-based) spectral methods where we consider an optical grating coupler [167] as a benchmark application. Next, we present another UQ study for a periodic optical structure, i.e. a gradient index meta surface [181], where we study and compare the impact of periodic and non-periodic shape deformations, respectively. Then, an SRR array is addressed as a benchmark problem for periodic optical structures of finite size subject to shape uncertainty. For this type of models, we propose a dedicated UQ workflow based on a decoupled uncertainty propagation. In particular, we combine SC on the unit cell level with an SMA as well as the MFMC method introduced in Section 4.1.2. Finally, we demonstrate the effectiveness of the suggested adjoint-based yield estimation method from Section 4.4, considering an established academic benchmark problem of an electric waveguide.

5.3.1 Optical grating coupler

In this subsection, we investigate the spectral UQ methods proposed in Section 4.2.1, 4.2.2, and 4.3, in detail. To this end, an optical grating coupler [61, 167] is employed as a non-trivial benchmark model in the form of the parametrized unit cell problem (2.57)-(2.58) introduced in Section 2.3.2. We first describe the numerical model as well as the considered uncertainties in material and geometry. Then, conformally mapped GPC and mapped SC as well as the extension in terms of adjoint error estimates are applied in subsequent order. The content and structure of this section are based on our works [88, 90]. Note that the considered model is a periodic structure where one can distinguish between two different classes of uncertainties referred to as *global* and *local* in the following. In this section, we only address *global* uncertainties, that model a systematic offset in the fabrication, e.g. variations of the material properties for the full structure or global variations of

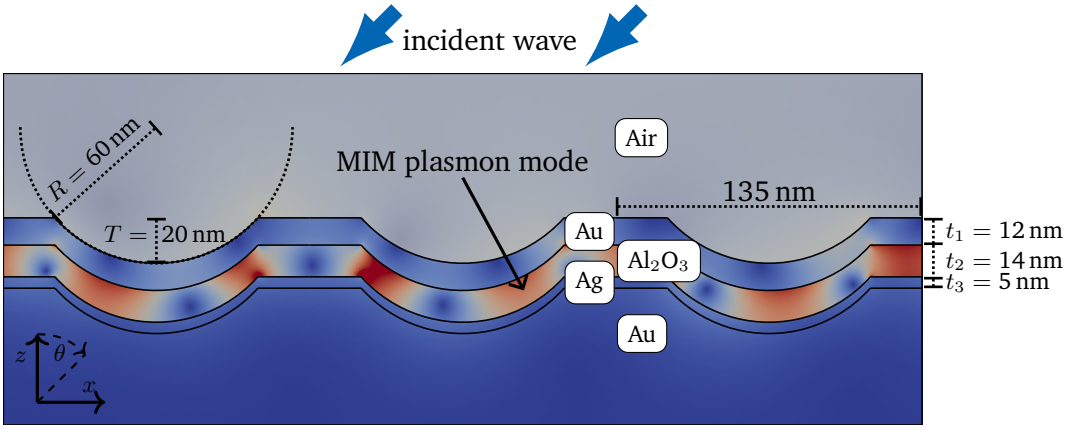


Figure 5.26: Three unit cells of an optical grating coupler [167] which is periodically extended in x-direction: an incident plane wave in free space directly excites a MIM plasmon mode which then propagates in horizontal direction. The coloring indicates the magnitude of the electric field (for a specific, arbitrary chosen, point in time). The geometry parameters representing the radius R and depth T of the grating as well as t_1, t_2, t_3 denoting the width of the three upper layers are illustrated and their nominal values are given. The illustration is based on [89, Figure 9].

geometry parameters, and affect all unit cells in the same way. However, in the next section, we then present another study where also *local* uncertainties are taken into account, which break the periodicity of the unit cells.

5.3.1.1 Numerical model

We consider an optical grating [61, 167] which is illustrated in Fig. 5.26 and couples power from a plane wave in free-space at optical frequency directly into a metal-insulator-metal (MIM) plasmon mode, propagating along the metallic surfaces. In [167], it was shown that the MIM resonance strongly depends on the depth of the grating, and, hence, evaluating the impact of the nano-scale manufacturing tolerances is highly relevant. The grating coupler can be represented by the model introduced in Section 2.1.2, as we assume that the structure is periodic in x-direction with $d_x = 135\,\text{nm}$ and infinitely extended in y-direction such that an arbitrary unit cell width d_y can be chosen. It is excited from the top (at ∂D_{z^+}) by a TM plane wave, such that

$$\pi_\text{T}\left[\mathbf{E}^\text{inc}\right] = \pi_\text{T}\left[\mathbf{E}^\text{TM,00}\right] \quad \text{at } \partial D_{z^+}, \tag{5.23}$$

with oblique propagation direction $\theta^\text{inc} = 53°$, $\phi^\text{inc} = 0°$ and angular frequency $\omega = 2\pi\,(414\,\text{THz})$. As QoIs we consider the reflection coefficients (2.25) as they indicate the coupling efficiency of the coupler. In particular, here we only address the fundamental scattering parameter s.t. $\mathcal{Q} := S^\text{TM,00}$ due to the negligible magnitude of all other scattering parameters for this particular model.

The dispersive material properties of the noble metals are modeled based on the measurement data of the reflectivity studies from Johnson and Christy [117]. These are shown in Fig. 5.27 along with their given error estimates. Note that from the refractive indices n and the extinction coefficients κ, the complex permittivity is obtained as

$$\varepsilon = \left(n^2 - \kappa^2 - \mathrm{i}(2n\kappa)\right)\varepsilon_0, \tag{5.24}$$

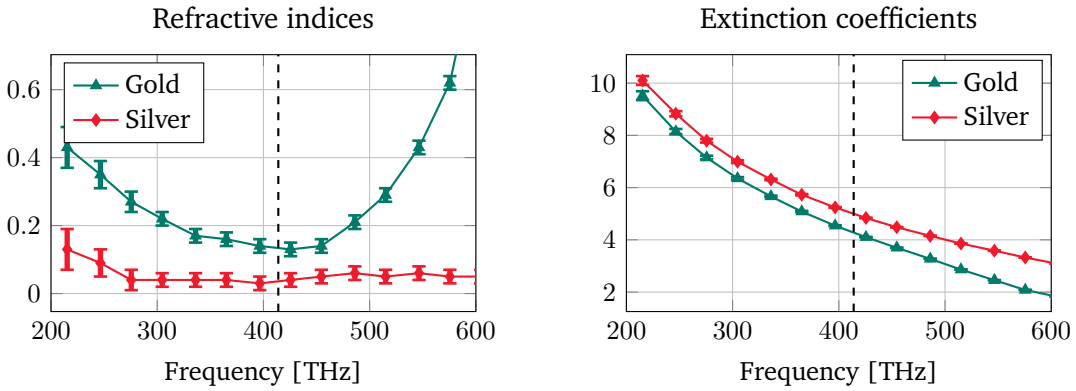


Figure 5.27: Dispersive material properties of gold and silver [117]. The bars indicate the respective error estimates and the dashed line the considered frequency in the numerical model.

see e.g. [139, Chapter 1.1]. However, the data is only provided for discrete frequency sample points. Therefore, we need to interpolate between those sample points in order to obtain the material properties at the considered frequency of $f = 414\,\mathrm{THz}$. To this end, we employ the three available frequency sample points which are the closest to $f = 414\,\mathrm{THz}$, i.e. $[f_0, f_1, f_2] = [396.55, 425.57, 454.58]\,\mathrm{THz}$ where we denote the respective material parameters as

$$\beta_i^\alpha, \text{ where } \alpha \in \{\mathrm{Au}, \mathrm{Ag}\},\ \beta \in \{n, \kappa\},\ i \in 0, 1, 2. \tag{5.25}$$

These data values are then interpolated using 2nd order Lagrange polynomials as

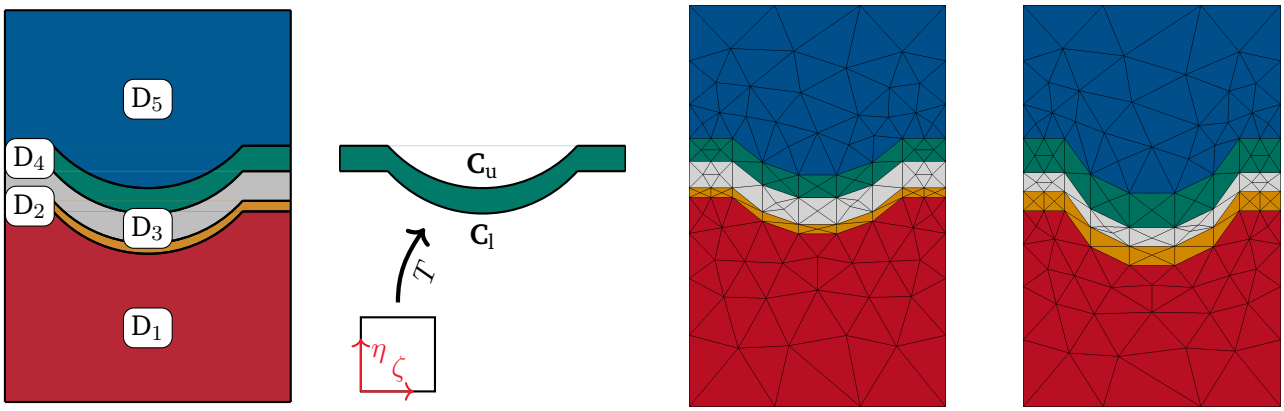
$$\beta^\alpha(f) = \sum_{i=0}^{2} \beta_i^\alpha l_i(f), \qquad l_i(f) = \prod_{j=0, j\neq i}^{2} \frac{f - f_j}{f_i - f_j}, \tag{5.26}$$

to obtain the desired material parameter values at $f = 414\,\mathrm{THz}$.

Next, we discuss the implementation of the deterministic numerical model and its parameterization in terms of the geometry parameters depicted in Fig. 5.26. The strong formulation is given by (2.24) where the first order Floquet Boundary condition (2.24e) is indeed a suitable choice as the fundamental Floquet modes are the only propagating modes for the considered unit cell dimensions. GMSH [92] is used to create the required periodic mesh for the nominal design. The discrete system (2.41) is obtained based on the FE library FENICS [3]. We note that, as FENICS 2017.2.0 does not work with complex numbers, the real and imaginary parts of $\mathbf{A}^{\mathrm{uc}}, \mathbf{f}^{\mathrm{uc}}, \mathbf{q}^{\mathrm{TM},00}$ are assembled separately, employing 2nd order Nédélec elements of the 1st kind. Then, using SCIPY and NUMPY, the quasi-periodic boundary conditions (2.10a), (2.10b) are imposed and a sparse LU decomposition is employed to solve the linear system (2.41) with $N_h = 56200$ DoFs. Note that, according to (4.116), the dual solution $\mathbf{z}_h$ can then be obtained with negligible computational cost.

The numerical model is parameterized in terms of the geometry parameters, using a design element approach [35] which avoids the need to re-mesh for different parameter configurations. To this end, the unit cell domain is split into 5 design elements D_m, $m = 1, \ldots, 5$ as illustrated in Fig. 5.28a. The upper and lower boundary of each design element are described using NURBS [166] curves

$$\mathbf{C}_i(\zeta; \boldsymbol{\xi}) = \sum_{j=0}^{n} R_j(\zeta) \mathbf{P}_{j,i}(\boldsymbol{\xi}), \quad \zeta \in [0, 1],\ i = 1, \ldots, 6, \tag{5.27}$$



(a) Design elements. **(b)** Mapping (5.28). **(c)** Initial mesh. **(d)** Deformed mesh.

Figure 5.28: Illustration of mesh deformation approach employing design elements, based on [89, Figure 11]. a) The unit cell domain D is decomposed in 5 design elements $D_1, \dots, D_5$. b) A mapping from the unit square is defined for each design element, see (5.28). c) Initial mesh for the nominal design parameter $\mathbf{y}^{\text{nom}}$. The mesh is chosen inadequately coarse for illustration purposes. d) By evaluating the mapping (5.28) for each mesh node of the original mesh, the respective deformed mesh for the geometry parameters $R = 50\,\text{nm},\ T = 25\,\text{nm},\ t_1 = 18\,\text{nm},\ t_2 = 10\,\text{nm},\ t_3 = 10\,\text{nm}$ is obtained.

where $R_j : [0,1] \to \mathbb{R}$ denote the rational basis functions and $\mathbf{P}_{j,i}(\boldsymbol{\xi}) \in \mathbb{R}^3$ the control points, see [166] for details. Next, one can define mappings from the unit square $[0,1] \times [0,1]$ to each design element $D_i(\boldsymbol{\xi})$ such that

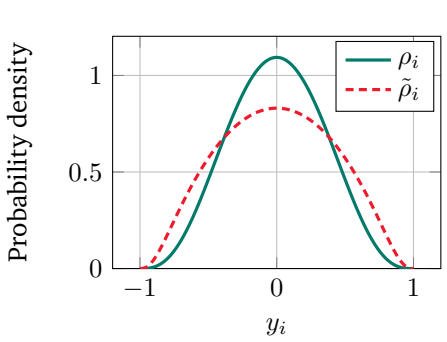
$$\mathbf{T}_m(\zeta, \eta; \boldsymbol{\xi}) = \eta \mathbf{C}_{m,\text{u}}(\zeta; \boldsymbol{\xi}) + (1-\eta)\mathbf{C}_{m,\text{l}}(\zeta; \boldsymbol{\xi}), \quad \zeta, \eta \in [0,1],\ m = 1, \dots, 5, \tag{5.28}$$

where $\mathbf{C}_{m,\text{l}}$ and $\mathbf{C}_{m,\text{u}}$ denote the lower and upper interface curve of the design element D_m, as illustrated in Fig. 5.28b. Then, given the initial mesh for the nominal design parameters $\boldsymbol{\xi}^{\text{nom}}$, for each mesh node with coordinates $\mathbf{x}_j \in D_m$, the corresponding coordinates ζ_j, η_j on the unit square are found by solving the non-linear root-finding problem: find $\zeta_j, \eta_j \in [0,1],$ s.t.

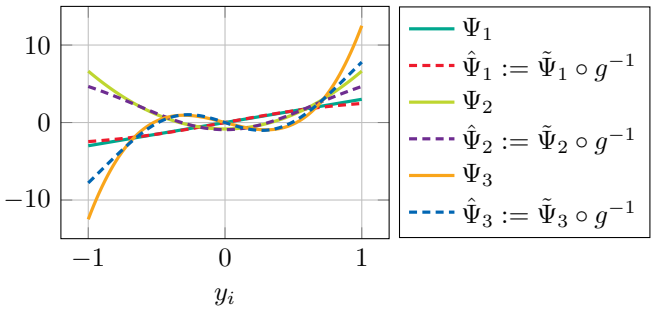
$$\mathbf{T}_m(\zeta_j, \eta_j; \boldsymbol{\xi}^{\text{nom}}) - \mathbf{x}_j = \mathbf{0},\ \mathbf{x}_j \in D_m(\boldsymbol{\xi}^{\text{nom}}). \tag{5.29}$$

Problem (5.29) can be reformulated as an optimization problem and is solved here using sequential quadratic programming (SQP) [156, Chapter 18]. Note that choosing adequate starting values for SQP might be required to achieve convergence. Although they depend on the particular parameterization of the NURBS curves $\mathbf{C}_{m,\text{l}}, \mathbf{C}_{m,\text{u}}$, it is usually straightforward to choose suitable starting values. Finally, one can obtain for different geometry parameters $\boldsymbol{\xi}$ the respective deformed meshes by evaluating (5.28), which yields the new coordinates for each mesh node. This is illustrated in Fig. 5.28c and Fig. 5.28d for one particular parameter configuration.

In the following, we investigate the mapped spectral UQ methods proposed in Chapter 4. To this end, we first only consider $N_{\boldsymbol{\xi}} = 3$ sensitive geometry parameters for the conformally mapped GPC approximation and postpone the high-dimensional UQ study to the subsequent subsection, which is then addressed with the dimension-adaptive mapped SC method.



(a) Univariate beta input PDF ρ_i and respective transformed density $\tilde{\rho}_i$ as in (4.52).



(b) Illustration of some exemplary GPC basis functions Ψ_m and the respective mapped basis functions $\hat{\Psi}_m$, based on [90, Figure 7].

Figure 5.29: Conformally mapped GPC for an optical grating coupler with $N_{\boldsymbol{\xi}} = 3$ beta distributed input parameters.

5.3.1.2 Conformally mapped GPC approximation

In this subsection, three uncertain parameters $\boldsymbol{\xi} \in [-1, 1]^3$ are considered. They are used to model variations around the nominal geometry parameters illustrated in Fig. 5.26 as follows: The thicknesses of the two upper layers are represented as $t_1 = 12\,\mathrm{nm} + \Delta\xi_1,\ t_2 = 14\,\mathrm{nm} + \Delta\xi_2$, respectively, and the grating depth as $T = 20\,\mathrm{nm} + \Delta\xi_3$, where $\Delta = 2\,\mathrm{nm}$. These parameters $\xi_i,\ i = 1, \ldots, 3$ are then considered as realizations of independent beta distributed RVs $\Xi_i,\ i = 1, \ldots, 3$ with PDF ρ_i in the form of (5.18) where $l = -1$ and $u = 1$. The corresponding transformed density (4.52) is illustrated in Fig. 5.29a. Furthermore, Fig. 5.29b presents a few GPC basis functions, i.e. Jacobi polynomials in this case, and the corresponding mapped basis functions.

In a first step, we investigate the decay of the GPC coefficients in order to study the smoothness of the map $\boldsymbol{\xi} \mapsto \mathcal{Q}$, illustrated in Fig. 5.30, numerically. If this mapping is analytic, the Fourier coefficients $s_{\mathbf{m}}$ of the GPC approximation (4.24) are expected to decay exponentially s.t.

$$|s_{\mathbf{m}}|^2 \leq C e^{-\sum_{n=1}^{N_{\boldsymbol{\xi}}} c_n m_n}, \tag{5.30}$$

where $C, c_1, c_2 \ldots, c_{N_{\boldsymbol{\xi}}}$ are positive constants, see, e.g. [154] where Legendre polynomials are addressed. Considering the maximum of the magnitude of the GPC coefficients $s_{\mathbf{m}}$ with the same maximum degree w, we obtain

$$\max_{||\mathbf{m}||_\infty = w} |s_{\mathbf{m}}|^2 \leq \max_{||\mathbf{m}||_\infty = w} C e^{-\sum_{n=1}^{N} c_n m_n} = C e^{-\min_{||\mathbf{m}||_\infty = w} \sum_{n=1}^{N} c_n m_n} \leq C e^{-(\min_n c_n) w}, \tag{5.31}$$

i.e. an exponential convergence w.r.t. the maximum-degree w is expected.

We then compute a GPC approximation in the form of (4.24), however, in this subsection we replace the total-degree basis by a tensor-product construction of maximum degree p, i.e. the multi-index set $\Lambda^{\mathrm{TD}}_{m_{\max}}$ is replaced by

$$\Lambda^{\mathrm{TP}}_{m_{\max}} := \left\{ \mathbf{m} \,:\, 0 \leq \|\mathbf{m}\|_\infty = \left(\max_{1 \leq n \leq N_{\boldsymbol{\xi}}} m_n \right) \leq m_{\max} \right\}, \tag{5.32}$$

in all definitions of the sections 4.1.3.1-4.2.1. The tensor-product construction is here employed in order to study the convergence of the coefficients in the form of (5.31) and it is also beneficial for the particular

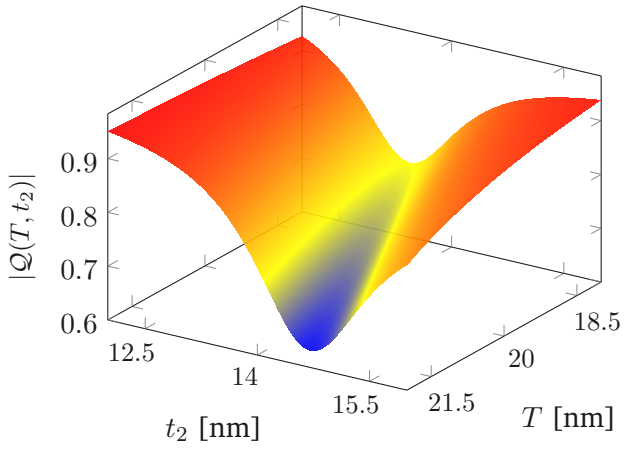


Figure 5.30: Fundamental scattering parameter of optical grating coupler w.r.t. two geometry parameters.

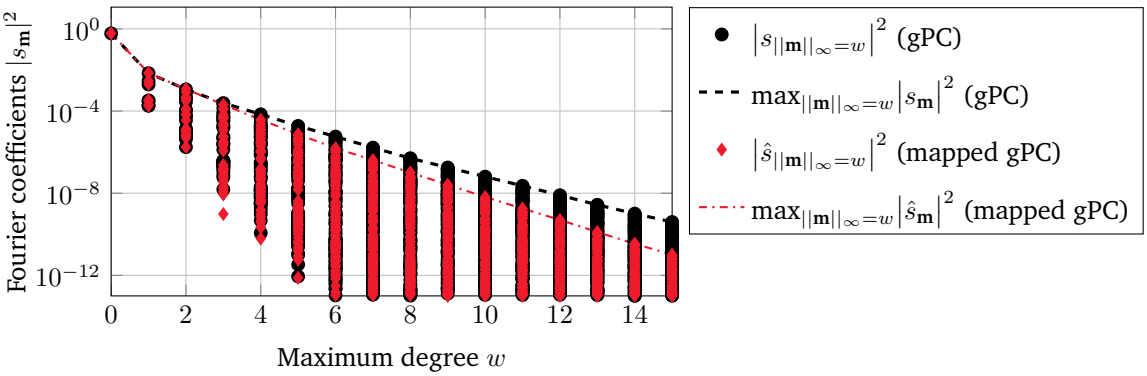


Figure 5.31: Fourier coefficients of a multivariate (mapped) GPC expansion, based on [90, Figure 8].

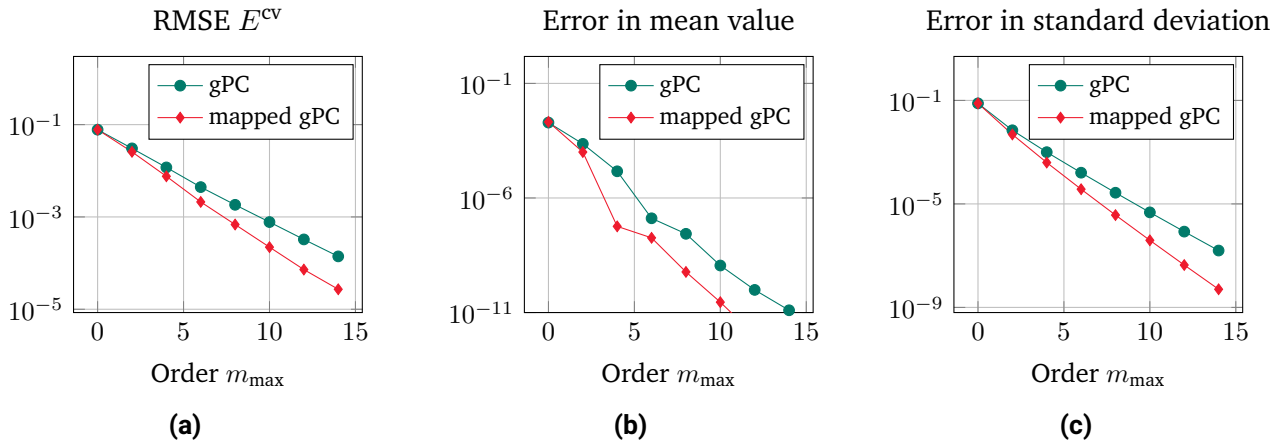


Figure 5.32: Convergence study employing the (mapped) GPC approximations for the optical grating coupler with $N_{\boldsymbol{\xi}} = 3$ uncertain parameters, based on [90, Figure 9]. a) Convergence of the empirical RMSE (4.89). b) Convergence of the mean value. c) Convergence of the standard deviation.

numerical model due to the strong interaction effects of the parameters, which will be discussed later. We employ a tensor-product basis of degree $m_{\max} = 15$ and use a Gauss quadrature of order 17 for the pseudo-spectral projection. All GPC coefficients are then depicted in Fig. 5.31 which indeed show the exponential decay (5.31). This is considered as a numerical indication that the approximated mapping from the inputs to the complex-valued scattering parameter $\boldsymbol{\xi} \mapsto \mathcal{Q}$ is smooth and justifies the use of (mapped) spectral polynomial methods. In addition, we compute a mapped GPC approximation (4.64) of the same degree where the respective mapped coefficients $\hat{s}_{\mathbf{m}}$ are shown in Fig. 5.31 as well. A faster convergence of the mapped coefficients can be seen and, hence, the mapped GPC approach is expected to lead to an enhanced convergence which is confirmed in the following.

We construct (mapped) tensor-product GPC expansions of the magnitude of the S-parameter $|\mathcal{Q}(\boldsymbol{\xi})|$ with increasing degree $m_{\max}$. The corresponding (mapped) coefficients are computed using pseudo-spectral projection of degree $m_{\max} + 1$. To access the accuracy of the respective approximations, we employ the empirical RMSE (4.89) with $N_{\text{cv}} = 1000$. The resulting convergence plots are presented in Fig. 5.32a where the mapped approach exhibits an approximately 30% faster convergence than standard GPC. We emphasize that the corresponding speed-up factor (for a fixed accuracy) scales exponentially w.r.t. $N_{\boldsymbol{\xi}}$ and, hence, in this case the computational cost, i.e. the number of model evaluations, can be reduced by a factor of approximately 2. Fig. 5.32b and Fig. 5.32c show very similar results for the computed mean values and standard deviations where the respective reference solutions are here computed using Gaussian quadrature of degree 30. Finally, the mapped GPC approximation of degree 14, i.e. the most accurate of the readily available surrogate models, is then employed to estimate the Sobol sensitivity indices, which show a significant discrepancy between the first-order and total-order Sobol indices, see Fig. 5.33. We note that the main-effect indices only add up to $\approx 34\%$ while the other $\approx 66\%$ correspond to the strong interaction effects of the inputs.

5.3.1.3 Conformally mapped adaptive SC approximation

Next, we apply the (mapped) dimension-adaptive SC Algorithm 2 and its adjoint-based counterpart Algorithm 3. First, for comparison we address the three-dimensional setting studied in the previous section before

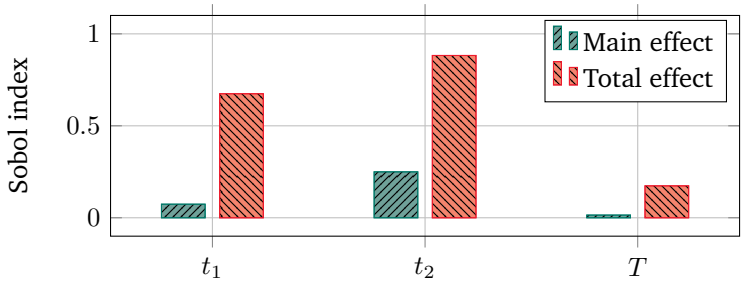


Figure 5.33: Sobol sensitivity indices for the optical grating coupler with 3 sensitive uncertain parameters, based on [90, Figure 10].

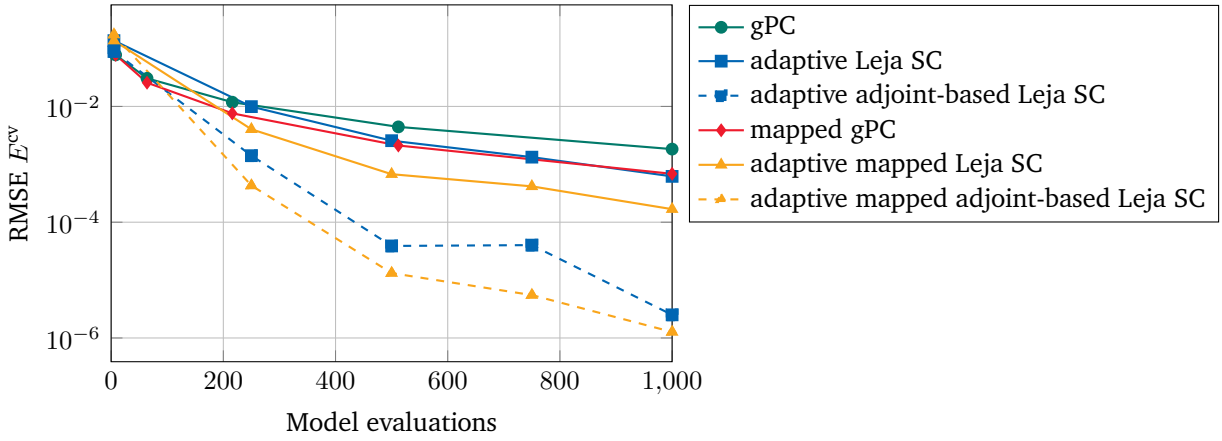


Figure 5.34: Convergence study in terms of the empirical RMSE (4.89) for an optical grating coupler with 3 sensitive uncertain geometry parameters t_1, t_2, T. Here, for the adjoint-based methods, the error of the adjoint-error corrected surrogate models (4.98) are shown by the dashed lines. Note that these models require a residual evaluation at each cross-validation point.

proceeding with a moderately high-dimensional UQ study. The resulting convergence plot is presented in Fig. 5.34, where the abscissa now represents the number of model evaluations. It can be seen that, although the number of parameters $N_{\boldsymbol{\xi}} = 3$ is rather small, the dimension-adaptive sparse SC algorithms are already beneficial compared to the isotropic (mapped) GPC approach. We emphasize that the improved (doubled) convergence order of the adjoint error indicator (4.97c) is confirmed numerically. Furthermore, it shall be noted that the conformal 9-th order sausage map (4.48) leads to a significant convergence improvement for all spectral UQ methods considered in the present setting.

In the following, we address a moderately high dimensional setting with $N_{\boldsymbol{\xi}} = 17$ random inputs $\boldsymbol{\Xi}$, i.e. the 12 material parameters (5.25) and the 5 geometrical parameters R, T, t_1, t_2, t_3 introduced in Fig. 5.26. Note that all uncertain parameters are taken into account by the uncertain complex permittivity in the form of (2.54). The RVs Ξ_i, $i = 1, \dots, 17$ are assumed to be independent and beta-distributed in the form of (5.18). The range parameters l and u of the beta distribution associated to the geometrical parameters are now given by the respective nominal values $\pm 1.5\,\mathrm{nm}$, i.e. modeling a small 2σ interval of $\pm 1\,\mathrm{nm}$. For the material parameters (2.54) the ranges are chosen according to the error estimate provided by Johnson and Christy for the *instrumental accuracy of the reflection and transmission measurements* [117]. In particular, it is assumed that the given error estimate refers to a 2σ interval as no further information on the measurement uncertainties are specified.

Table 5.4: Computational cost and accuracy of different spectral approximations of the optical grating coupler with 17 uncertain input parameters, based on [89, Table 3]. The column *Solves* specifies the computational cost in terms of the number of assemblies and sparse LU decompositions of different system matrices. The column *Residuals* specifies the number of residual evaluations where the associated computational cost can almost be neglected to the solver cost.

Method	Solves	Residuals	Max. Error (5.33)
Total-degree GPC	613	0	7.61×10^{-1}
Smolyak sparse-grid (level 2)	613	0	1.73×10^{-1}
Adaptive Leja SC	613	0	8.56×10^{-2}
Adaptive mapped Leja SC (Iso-Map)	613	0	3.59×10^{-2}
Adaptive mapped Leja SC (Aniso-Map)	613	0	3.85×10^{-2}
Adaptive adjoint-based Leja SC	558	613	8.46×10^{-2}
Adaptive adjoint-based mapped Leja SC (Iso-Map)	563	613	3.57×10^{-2}
Adaptive adjoint-based mapped Leja SC (Aniso-Map)	563	613	3.82×10^{-2}
Adaptive adjoint-based mapped Leja (Aniso-Map) with adjoint-error correction	3000	30000	1.25×10^{-3}

Next, we consider different spectral approximations. The proposed (mapped) dimension-adaptive schemes Algorithm 2 and Algorithm 3 are compared with two sparse non-adaptive approximations, i.e. total-degree GPC and isotropic Smolyak sparse-grid SC based on Clenshaw Curtis nodes. As before, we use CHAOSPY [78] for GPC, the SPARSE-GRID-MATLAB-KIT [12] for the Smolyak sparse-grid SC and an in-house code for the (adjoint-based) mapped adaptive Leja approximations based on DALI[2] [135, 137]. The different approximation techniques are then compared in terms of computational cost and accuracy. For GPC, the Smolyak sparse-grid and the adaptive Algorithm 2, the computational costs can be straightforwardly quantified by the number of model evaluations that are required to construct the surrogate. However, for the adjoint-based Algorithm 3 the estimation of costs is more involved. First, as already mentioned in Section 4.3, the additional computational costs for the computation of the adjoint solution $\mathbf{z}_h(\boldsymbol{\xi})$ can here be neglected as the respective primal solutions are computed with a sparse LU decomposition. Second, for the evaluation of the adjoint-based error indicator $\tilde{\mathcal{E}}^{\mathrm{SC}}(\boldsymbol{\xi})$, it is only required to evaluate a residual of (2.56) for which the associated computational cost is almost negligible compared to the cost of the assembly and sparse LU decomposition of the system matrix $\mathbf{A}^{\mathrm{uc}}(\boldsymbol{\xi})$. Hence, we distinguish between solver calls and residual evaluations in the following.

In addition to the empirical RMSE E^{cv} defined in (4.89), we also evaluate the maximum error over all N_{cv} sample points $\{\boldsymbol{\xi}^{(i)}\}_{i=1}^{N_{\mathrm{cv}}}$

$$E^{\max} := \max_{i=1,\dots,N_{\mathrm{cv}}} \left| \mathcal{Q}(\boldsymbol{\xi}^{(i)}) - \tilde{\mathcal{Q}}(\boldsymbol{\xi}^{(i)}) \right|. \tag{5.33}$$

The respective maximum error of different spectral approximations is presented in Table 5.4, along with the associated computational cost. The respective 2nd order total-degree GPC approximation based on 171 Jacobi polynomials is computed with a sparse 2nd order Gauss quadrature formula consisting of 613 quadrature nodes, and, hence, model evaluations. The considered Smolyak sparse-grid SC approximation is also based on 613 model evaluations. Hence, we employ the same budget $\mathcal{B} = 613$ for the adaptive Algorithm 2. Thereby

[2] https://github.com/dlouk/DALI3

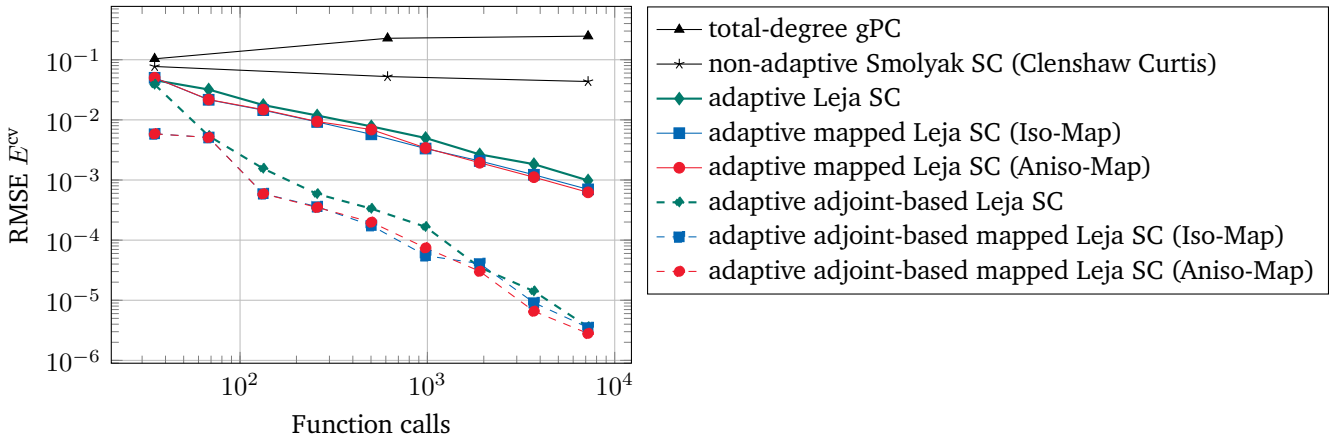


Figure 5.35: Convergence study employing different spectral approximations of the optical grating coupler with $N_{\boldsymbol{\xi}} = 17$ uncertain input parameters, based on [89, Figure 13]. All dimension-adaptive approaches clearly outperform the isotropic methods. For the adjoint-based methods, the errors of the respective adjoint-error corrected surrogate models (4.98) are shown by the dashed lines. Note that these models require a residual evaluation at each cross-validation point.

we compare three different mappings. First, we employ the identity mapping $\mathbf{g} = \mathbf{g}_{\mathrm{id}} : \mathbf{s} \mapsto \mathbf{s}$ which recovers a standard Leja algorithm. Second, we refer with *Iso-Map* to the multivariate mapping where the conformal 9-th order sausage map is applied to all coordinates, i.e. $g_1 = \ldots = g_{N_{\boldsymbol{\xi}}} = g_{\mathrm{S}}(\cdot, 9)$. Third, *Aniso-Map* refers to a multivariate mapping where the 9-th order sausage map $g_{\mathrm{S}}(\cdot, 9)$ is only applied to the sensitive parameters t_1, t_2, T, i.e. the parameters where we numerically observed the slowest univariate convergence rates, while the identity map is employed for the other parameters. It can be observed that all adaptive approximations are significantly more accurate than the non-adaptive approaches. Furthermore, it can be observed that both, the *Iso-Map* and the *Aniso-Map* approach, lead to a similar improvement in terms of the maximum error w.r.t. the standard adaptive Leja SC. It shall be noted that both approaches do not require any (relevant) extra computational cost. Next, the Algorithm 3, which employs the adjoint error indicator to steer the adaptivity, is employed. We compute, again, approximations with 613 (mapped) polynomials which leads to very similar errors as in the case of Algorithm 2. However, the number of required LU decompositions is reduced by approximately 50 and, hence, the computational cost. We note that even larger relative improvements, e.g. $> 50\%$, are obtained in different settings, see [89] for details.

The convergence of the different approximation techniques is further illustrated in Fig. 5.35. In particular, the empirical RMSE (4.89) is plotted w.r.t. the number of sparse LU decompositions. Note that the error decay of the isotropic GPC approximation is still pre-asymptotic for the considered degrees (only up to degree 3), see [89, Appendix C] for details. However, it can be seen that both isotropic reference approaches only reach a very poor accuracy within the considered computational budget and are greatly outperformed by the dimension-adaptive approaches. Again, an improvement due to the use of conformal maps can be observed where the difference between the *Iso-Map* and the *Aniso-Map* approach is rather negligible. The respective adjoint-error corrected surrogate models (4.98), illustrated by dashed lines, show a doubled convergence order. However, it shall be emphasized that they do not correspond to the mapped (polynomial) surrogate models shown in Table 5.4 but require a residual evaluation at each cross-validation point $\{\boldsymbol{\xi}^{(i)}\}_{i=1}^{N_{\mathrm{cv}}}$.

Finally, we construct a very accurate surrogate following the procedure outlined at the end of Section 4.3. To this end, we first employ the adjoint-based mapped SC Algorithm 3 with a larger computational budget of $\mathcal{B} = 3000$ sparse LU decomposition. Then, the resulting mapped polynomial approximation is further

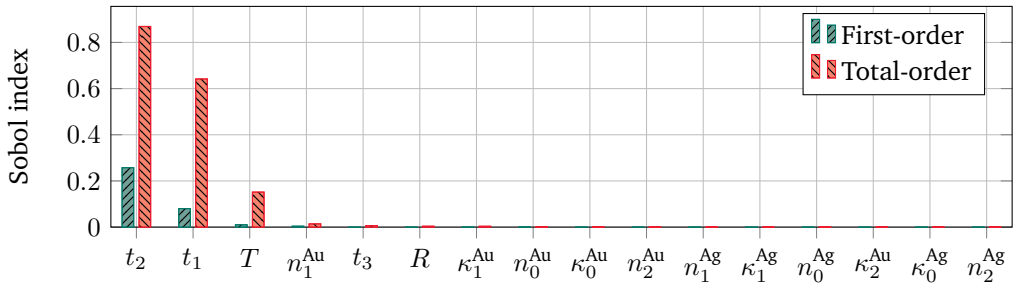


Figure 5.36: Sobol sensitivity indices for scattering parameter magnitude $|\mathcal{Q}|$ of the optical grating coupler with $N_{\boldsymbol{\xi}} = 17$ uncertain parameters, based on [89, Figure 14b]. The input parameters are ordered w.r.t. their respective total-effect indices.

refined by employing Algorithm 2 on the adjoint-error corrected approximation (4.98) until 30000 mapped polynomials are employed, which further reduces the associated error by more than an order of magnitude and leads to an RMSE E^{cv} of $1.2 \cdot 10^{-4}$. We emphasize that in this way, the adjoint approach leads to tremendous computational saving with respect to the classical Leja Algorithm 2, as 27000 sparse LU decomposition are avoided in the considered setting. The accuracy of the resulting mapped polynomial approximation is shown in the last row in Table 5.4. We note that the SC error has a similar magnitude as the FE discretization error and, hence, the mapped SC approximation is in the following employed in post-processing for UQ of the magnitude of the scattering parameter $|\mathcal{Q}|$.

In particular, as the mapped polynomial SC approximation $\mathcal{Q}_{\hat{Z}_\Lambda}(\boldsymbol{\xi})$ can be evaluated efficiently, we employ the sampling based approaches introduced in Section 4.1 with a large number of surrogate model evaluations to estimate statistical moments and sensitivity indices. Using $N_{\mathrm{MC}} = 10^7$ surrogate model evaluations in (4.1) and (4.4), we obtain for the mean and standard deviation of the scattering parameter magnitude

$$\mathbb{E}\left[|\mathcal{Q}|\right] \approx E_{N_{\mathrm{MC}}}\left[|\mathcal{Q}_{\hat{Z}_\Lambda}(\boldsymbol{\xi})|\right] \approx 0.7607, \qquad \mathrm{Std}\left[|\mathcal{Q}|\right] \approx \sqrt{V_{N_{\mathrm{MC}}}\left[|\mathcal{Q}_{\hat{Z}_\Lambda}(\boldsymbol{\xi})|\right]} \approx 0.0660, \tag{5.34}$$

respectively. Next, we employ Saltelli's algorithm using 3.6×10^6 evaluations of the surrogate model, in order to estimate the Sobol indices of $|\mathcal{Q}|$. The results are presented in Fig. 5.36 where it can be observed that the alumina layer thickness t_2 is the most sensitive parameter. We emphasize that the Sobol indices show that the particular model is very sensitive to the small variations of the geometry. In particular, we only considered variations within a range of $\pm 1.5\,\mathrm{nm}$ and yet their influence is significantly stronger than the partial variances which can be attributed to the uncertainty in the material data modeled based on the specified measurement error in [117]. Furthermore, as in the previous subsection, we observe again strong interaction effects of the inputs, as the sum of all main-effect indices is only 33%.

In summary, in this subsection we studied the application of the proposed enhanced surrogate modeling techniques to the challenging benchmark problem of an optical grating coupler. In particular, the proposed combinations of conformal maps, adjoint-based adaptivity and error correction with spectral UQ methods was confirmed to delay the curse-of-dimensionality, allowing us to consider up to $N_{\boldsymbol{\xi}} = 17$ uncertain (and sensitive) parameters. Significant gains in terms of either accuracy or computational costs could be observed due to the improved convergence rates.

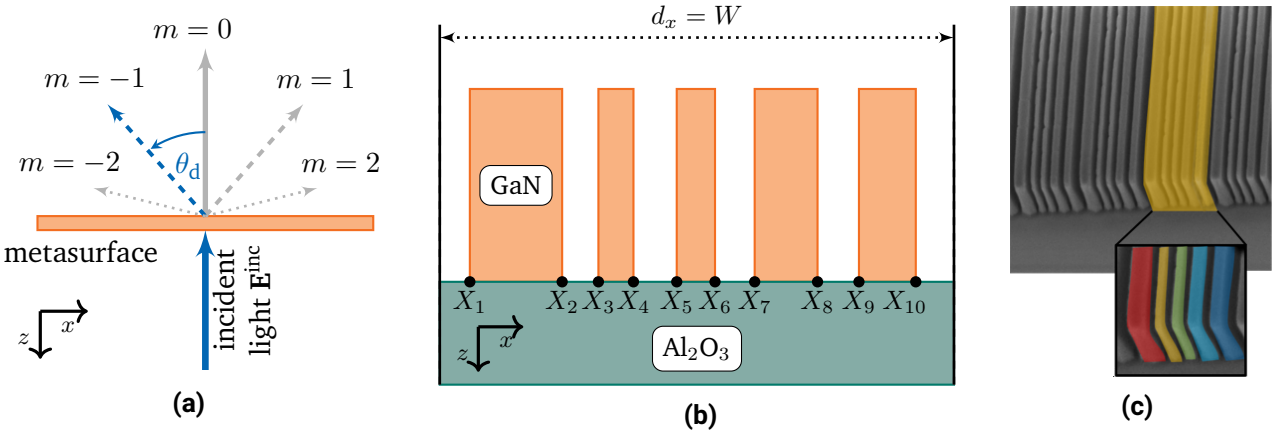
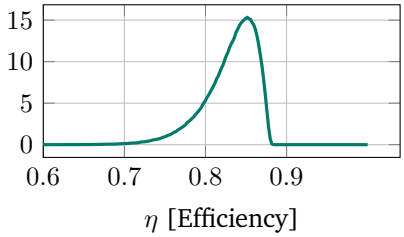


Figure 5.37: Illustration of gradient index metasurface, based on [181, Figure 1, 6, 10]. a) Discrete diffracted Floquet modes of different orders m for a periodic microscopic structure with periodicity $W = 1322\,\mathrm{nm} > \lambda$. The particular structure is optimized such that the transmission efficiency for an incident TM polarized plane wave with $\lambda = 600\,\mathrm{nm}$ in the Floquet mode $\mathbf{E}^{\mathrm{TM},-1\,0}$ with propagation angle $\theta_\mathrm{d} = 27°$ is maximized. b) The particular metasurface consists of 5 GaN nanoridges with different locations and widths to achieve the desired transmission behavior. The periodic boundaries are illustrated by black lines. The parameters $X_1, \dots, X_{10}$ are considered as uncertain. c) Electron micrograph of a fabricated structure with a zoom at one unit cell highlighted in yellow.

5.3.2 Gradient index metasurface

Next, we address UQ for a gradient index metasurface, that is, again, a periodic microscopic structure. The particular metasurface [181] is shown in Fig. 5.37 and consists of Gallium Nitride (GaN) nanoridges of subwavelength dimensions. It is designed to deflect an incident TM polarized plane wave with wavelength $\lambda = 600\,\mathrm{nm}$ and normal propagation direction $\theta^\mathrm{inc} = 0°, \phi^\mathrm{inc} = 0°$ into a particular Floquet mode with a desired angle θ_d. In particular, the widths and relative positions of these nanoridges are optimized such that the transmission efficiency η in the first diffraction order with $m = -1$, i.e. the Floquet mode $\mathbf{E}^{\mathrm{TM},-1\,0}$ is maximized. The content and structure of this section are based on our work [181], where an ensemble of similar structures with different blaze angles θ_d, different feature sizes, and different numbers of nanoridges per unit cell was designed and studied. Here, for brevity, we only address the particular grating, illustrated in Fig. 5.37 which employs 5 GaN nanoridges of height $1\,\mu\mathrm{m}$ per unit cell and is designed for a deflection angle of $\theta_\mathrm{d} = 27°$ which implies a grating period of $d_x = 1322\,\mathrm{nm}$. For further details on diffraction gratings, metasurfaces and the particular structure, we refer to [181] and the references therein.

Shape variabilities in the manufacturing process are taken into account by considering variations in the x-coordinates of the bounds of the ridges $X_1, \dots, X_{10}$, which are illustrated in Fig. 5.37 by black dots. In particular, the random x-coordinates are obtained as $X_i = X_i^\mathrm{nom} + Y_i$, $i = 1, \dots, 10$, where X_i^nom, $i = 1, \dots, 10$, refer to the nominal geometry parameters and $\boldsymbol{\Xi}$ to the corresponding RVs modeling the uncertainty. In the following, we assume that RVs $\boldsymbol{\Xi}$ are independent and identically uniformly distributed in the range of $[-5\,\mathrm{nm},\ 5\,\mathrm{nm}]$. In [181], the nominal design of the considered grating as well as UQ studies with different distributions are reported. As QoI $\mathcal{Q}$ we then consider the transmission efficiency $\eta \in [0, 1]$ in the first Floquet mode $\mathbf{E}^{\mathrm{TM},-1\,0}$. The transmission efficiency is here numerically evaluated by employing the rigorous coupled wave analysis (RCWA) implementation RETICOLO, which allows very efficient monochromatic simulations of



(a) PDF of the deflection efficiency η.

Mean	0.831
Standard deviation	0.032
Minimum value	0.541
Maximum value	0.891
Measurement	0.860

(b) UQ and measurement results for the efficency η.

Figure 5.38: UQ results for the gradient index metasurface and periodic deformations.

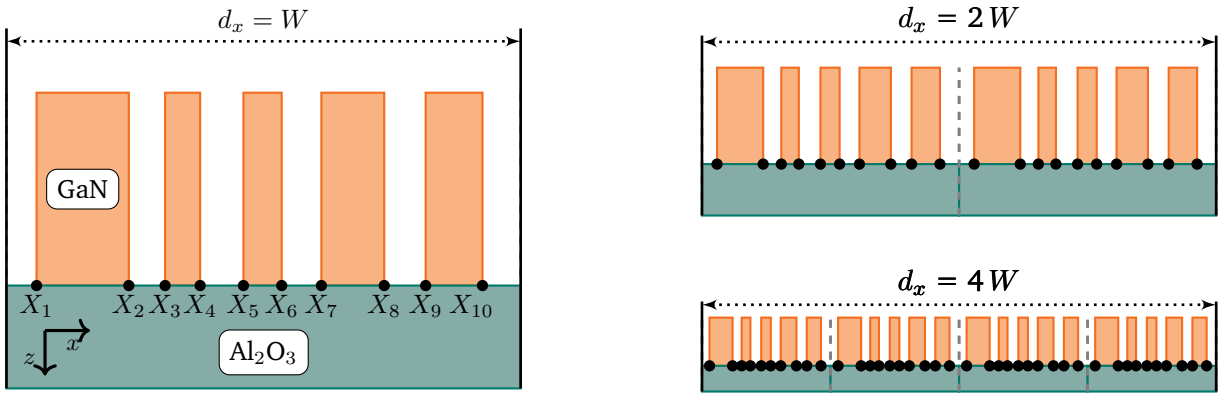


Figure 5.39: Sketches of different numerical models where an increasing number of uncertain grating elements is considered as a unit cell, based on [181, Figure 10]. The solid black lines indicate the quasi-periodic boundary conditions and the black dots illustrate the (increasing number of) uncertain parameters. Left: unit cell consisting of $n = 1$ grating period subject to periodic deformations. Right: The impact of the periodic boundary conditions is systematically minimized by considering $n = 2$ (top) and $n = 4$ (bottom) grating periods per unit cell.

the considered grating, see [109, 181] for details.

Preliminary numerical tests indicate that the spectral UQ methods discussed in Chapter 4 are not readily applicable for the particular structure since the map $\boldsymbol{\xi} \mapsto \mathcal{Q}$ is not sufficiently smooth. Hence, the MC method is employed here, which is appealing as a single model evaluation using the RCWA solver takes less than $1\,\mathrm{s}$ of computation time. Furthermore, as the MC method is embarrassingly parallelizable, even millions of model evaluations are feasible. In particular, we employ the MC estimates (4.1) and (4.4) with $N_{\mathrm{MC}} = 10^6$ random realizations to estimate the mean and standard deviation of the random deflection efficiency $\mathcal{Q}(\boldsymbol{\Xi})$ with a negligible sampling error (4.2). Furthermore, we evaluate the maximum and minimum values of $\mathcal{Q}$ within the MC output sample $\{\mathcal{Q}(\boldsymbol{\xi}^{(i)})\}_{i=1}^{N_{\mathrm{MC}}}$ and employ KDE to estimate the output PDF. These results are presented in Fig. 5.38. In addition, the measured deflection efficiency of a fabricated structure, shown in Fig. 5.37c, is given in the last row of Table 5.38b, which deviates less than one standard deviation of the estimated mean deflection efficiency. Note that for this UQ study, the computational unit cell domain D was confined to a single grating as illustrated in Fig. 5.37b with quasi-periodic boundary conditions that imply periodic geometry variations as well.

For comparison, we proceed with another study, where we minimize the impact of the quasi-periodic boundary conditions systematically. To this end, we consider an increasing number of uncertain grating periods n per unit

Data:
function handle `get_sample(n)` for unit cell with n grating periods,
tolerance tol,
minimum number of MC samples $N_{\mathrm{MC}}^{\mathrm{min}}$
Result:
Expectation E_n and standard deviation σ_n w.r.t. number of gratings n

Initialize grating period number $n = 1$
repeat
  Initialize number of samples $N_{\mathrm{MC}} = 0$
  repeat
    $N_{\mathrm{MC}} = N_{\mathrm{MC}} + 1$
    Draw MC sample $\eta^{(N_{\mathrm{MC}})} =$ `get_sample(`n`)`
    Estimate expectation E_n as in (4.1)
    Estimate variance σ_n^2 as in (4.4)
    Estimate error ϵ_{RMS} as in (4.2)
  until $N_{\mathrm{MC}} > N_{\mathrm{MC}}^{\mathrm{min}}$ and $\epsilon_{\mathrm{RMS}} < \frac{1}{6}\mathrm{tol}$;
  Increase number of gratings $n = 2n$
until $n > 4$ and $|E_{n/2} - E_{n/4}| < \mathrm{tol}$ and $|E_{n/4} - E_{n/8}| < \mathrm{tol}$;

Algorithm 4: MC-based UQ method for optical gratings with independent non-periodic geometry uncertainties, based on [181, Figure 12].

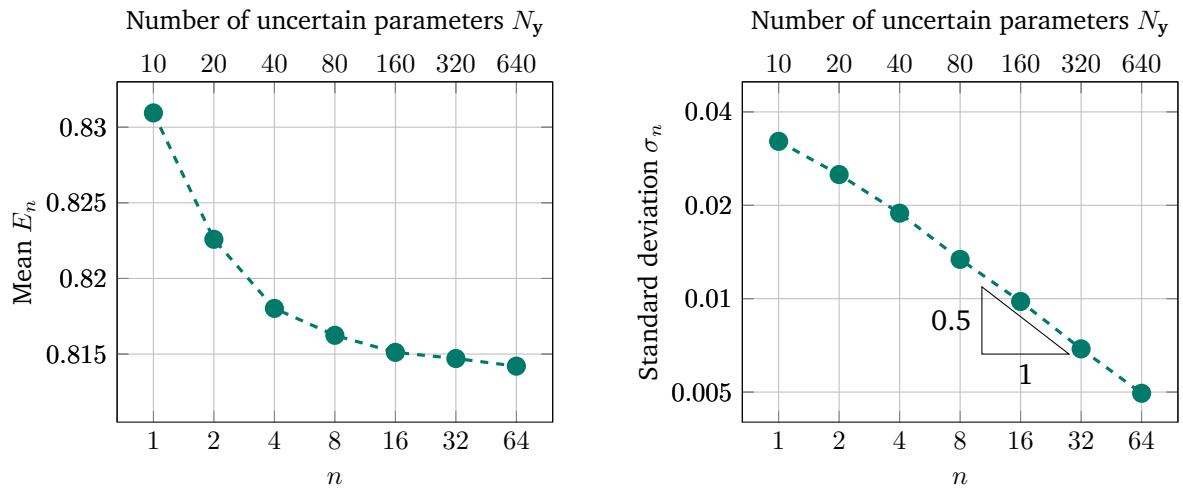


Figure 5.40: Mean and standard deviation of the deflection efficiency η w.r.t. the number of uncertain grating periods n within a periodic unit cell of width $d_x = nW$, based on [181, Figure 11].

Table 5.5: Mean E_n and standard deviation σ_n w.r.t. the number of uncertain grating periods n within a periodic unit cell, based on [181, Table 5]. N_{MC} refers to the required number of MC samples for the chosen tolerance tol.

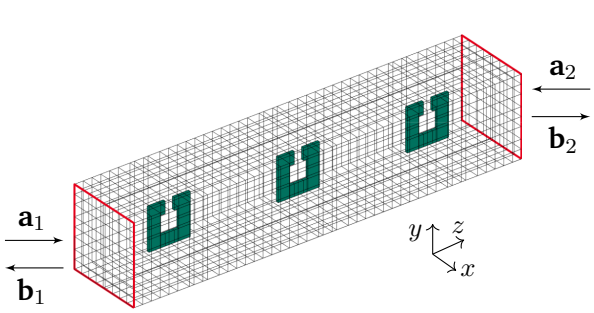
n	$N_{\boldsymbol{\xi}}$	N_{MC}	Mean E_n	Std. dev. σ_n
1	10	103,339	0.831	0.032
2	20	63,123	0.823	0.025
4	40	35,579	0.818	0.019
8	80	17,923	0.816	0.013
16	160	9,594	0.815	0.010
32	320	4,746	0.815	0.007
64	640	2,457	0.814	0.005

cell, as illustrated in Fig. 5.39, where the respective grating elements are assumed to be subject to independent geometric uncertainties. The number of grating periods n and hence the number of uncertain parameters, is then increased until changes in the estimated mean become negligible. The procedure is described in detail in Algorithm 4. We employ Algorithm 4 with $M_{\mathrm{min}} = 100$ and $\epsilon = 6 \cdot 10^{-4}$ and present the results in Fig. 4. It can be seen that the mean deflection efficiency decreases slightly before saturating after $n \approx 16$. In contrast, the standard deviation is continuously decreasing with a slope of -0.5 which, as shown in Table 5.5, also leads to a decreasing number of required MC samples with respect to n. This is computationally beneficial as it partially mitigates the increased computational cost associated to the larger unit cell domains. The decreasing standard deviation shows that in absence of systematic, i.e. stochastic dependent, manufacturing variabilities, the considered grating becomes a more deterministic behavior with an increased number of grating periods. This can be physically explained by averaging interference effects between the neighboring structures with different uncertain grating elements. In particular, the overall efficiency η is composed of the individual contributions from the different grating periods and, hence, in accordance with the central limit theorem, it is approximately normally distributed with a standard deviation $\propto 1/\sqrt{n}$, as can be seen in Fig. 5.40.

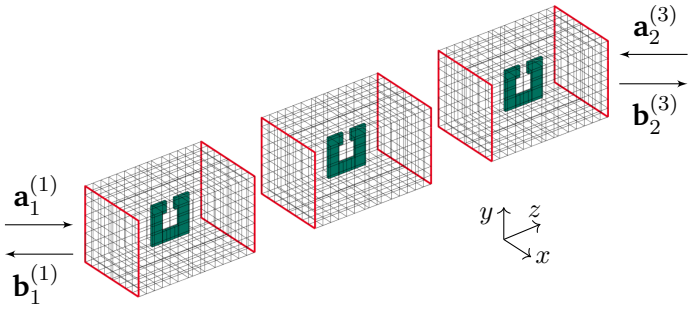
In summary, the considered grating is found to be rather robust with respect to small geometry deformations and, hence, has a reproducible performance. Systematic fabrication errors, i.e. periodic deformations of each grating period, lead to an acceptable mean deflection efficiency which is in agreement with the measurement result for a fabricated structure. It is found that for independent non-periodic variations of the different grating periods, the mean deflection efficiency is reduced by a few percent but the averaging interference effects of neighboring components stabilize the performance, which is of significant practical importance [181].

5.3.3 Split ring resonator array

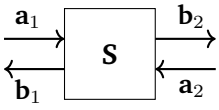
In this subsection, we now address shape uncertainties of periodic optical structures of finite size where we consider the benchmark problem of an SRR array [42, 87, 132]. Considering many elements of a periodic structure with independent shape deformations can lead to a large number of uncertain parameters, which was already seen in the previous section. Due to the curse-of-dimensionality, this is a major drawback for the application of spectral UQ methods, as discussed in Section 4.1.3. Hence, we propose a remedy by decoupling the uncertainty propagation for the individual unit cells by using an SMA [13, 95] such that a local SC



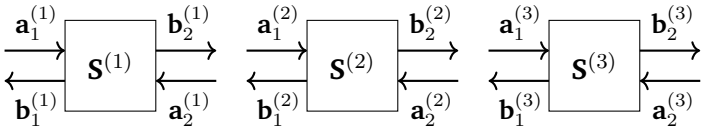
(a) Array of $N_{\text{cells}} = 3$ SRRs.



(b) Decomposed array of $N_{\text{cells}} = 3$ SRRs.



(c) SRR array is considered as a two port system with scattering matrix $\mathbf{S}$.



(d) Each SRR cell is considered as a two port system with scattering matrices $\mathbf{S}^{(i)}$.

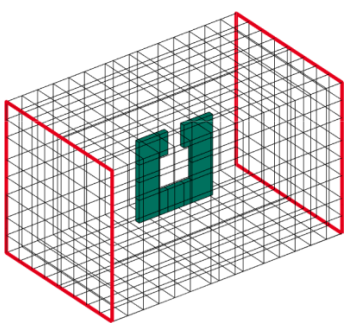
Figure 5.41: Illustration of SMA for an array of $N_{\text{cells}} = 3$ SRRs where the red boundaries indicate the ports: the structure is decomposed in individual cells which are coupled together using S-matrices.

approximation for a single unit cell can be employed. The so-obtained surrogate is then embedded as a low-fidelity model in a MFMC framework, as introduced in Section 4.1.2, to correct possible biasing errors. This procedure is found to be highly efficient for UQ, in particular speed-ups by several orders of magnitude w.r.t. the standard MC estimator (4.1) can be observed. The content and structure of this section follow our work [87].

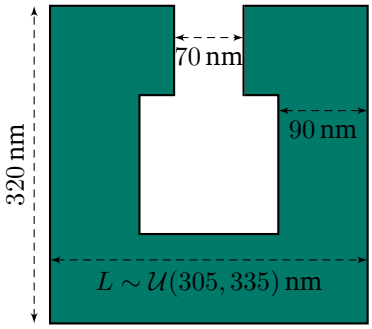
Our benchmark application is illustrated in Fig. 5.41a. It represents an array of N_{cells} coupled SRRs where each SRR is a nanoscale metallic structure corresponding to a resonance circuit. In particular, the ring and the small gap of each SRR element can be interpreted as an inductance and a capacitance, respectively. We note that this type of periodic structures has been investigated during the early research on optical metamaterials [42, 132]. The significant impact of manufacturing imperfections on the shape of such nanoscale SRRs is, for instance, illustrated in [42, Figure 54 and Figure 57] where it can be observed that the periodicity will not be perfect. To take these variations into account, we define parameters $\boldsymbol{\xi}_{\text{cell},j} \in \Gamma_{\text{cell}} \subset \mathbb{R}^{N_{\boldsymbol{\xi}}^{\text{cell}}}$, $j = 1, \dots, N_{\text{cells}}$ describing the shape or material variations in the j-th unit cell of the structure. The model then depends on the corresponding full input parameter vector which is obtained as

$$\boldsymbol{\xi} = \left[\boldsymbol{\xi}_{\text{cell},1}^\top, \dots, \boldsymbol{\xi}_{\text{cell},N_{\text{cells}}}^\top\right]^\top \in \Gamma \subset \mathbb{R}^{N_{\boldsymbol{\xi}}}, \quad \text{where } N_{\boldsymbol{\xi}} = N_{\text{cells}} \cdot N_{\boldsymbol{\xi}}^{\text{cell}}. \tag{5.35}$$

For brevity, we only give a few details on the numerical model and refer to [87] for details. The considered model is assumed to be loss-free and to be transversally terminated by perfect magnetic conducting (PMC) and PEC boundary conditions in x- and y-direction, respectively. The FIT time-domain method, see [124, 195, 203], is employed for the discretization of the wave equation which relies on a Cartesian mesh and computes broadband results based on a transient simulation by discrete Fourier transform (DFT) of the time-signals. In particular, we consider a plane wave excitation and evaluate the scattering parameters of the two-port structure, i.e. transmission and reflection coefficients. These coefficients are collected in the scattering matrix



(a) One unit cell (out of seven) of size $0.6\,\mu\text{m} \times 0.6\,\mu\text{m} \times 1\,\mu\text{m}$ where the ports are indicated by the red boundaries.



(b) Specification of the geometry of the SRR elements with a thickness of $20\,\text{nm}$. The longitudinal length L is considered as an uncertain parameter.

Figure 5.42: Illustration of the numerical model of the considered SRR array, based on [87, Figure 1].

$\mathbf{S}(\mathrm{i}\omega) \in \mathbb{C}^{2N_{\text{modes}} \times 2N_{\text{modes}}}$ such that

$$[\dots b_i(\mathrm{i}\omega) \dots]^\top = \mathbf{S}(\mathrm{i}\omega)\, [\dots a_i(\mathrm{i}\omega) \dots]^\top, \tag{5.36}$$

where $a_i \in \mathbb{C}^{2N_{\text{modes}}}$ and $b_i \in \mathbb{C}^{2N_{\text{modes}}}$ denote the amplitudes of suitably normalized incoming and outgoing waves, see [183] for details. Note that the reflection coefficient of the fundamental mode at the input port is here given by $S_{11}(\mathrm{i}\omega)$. In the following we suppress the frequency dependency for better readability.

An approach to enhance the computational efficiency in the simulation of finite-size periodic structures consists in decomposing the structure into single unit cells, as illustrated in Fig. 5.41b. One can then compute separate scattering matrix $\mathbf{S}^{(i)}$ for each cell as depicted in Fig. 5.41d and couple these single cell results together, see [13, 95] for details. In this work, we refer to this approach as SMA. Note that this procedure might introduce a systematic error, as the coupling between the cells is not necessarily fully captured by a low number of N_{modes} waveguide modes. As the a-priori selection of a suitable number of port modes as well as the computational cost to compute large scattering matrices can be challenging in practice, we consider the SMA approximation as low-fidelity data within the MFMC framework.

As explained in Section 4.1.2, MFMC is based on repeated model evaluations for different input parameter values. Since, even with SMA, the associated computational effort of employing the FIT for a large number of different parameter configurations can become prohibitive, we then further employ SC, see Section 4.1.3, to obtain a unit cell surrogate model. In particular, we construct a polynomial approximation of the map $\boldsymbol{\xi}_{\text{cell},j} \mapsto \mathbf{S}^{(j)}$ for the j-th unit cell of the periodic structure. Note that the same surrogate model can be used for all cells $j = 1, \dots, N_{\text{cells}}$. Employing the unit cell surrogate models for the SMA, we then obtain an approximation of the full structure. It shall be highlighted that the computational effort to evaluate the so-obtained approximation is negligible, as the scattering matrices have small dimensions compared to the FIT matrices, i.e. $N_h \gg 2N_{\text{modes}}$. Furthermore, it shall be emphasized that the proposed construction of a surrogate based on SMA and SC on the unit cell level is tremendously more efficient than the direct application of spectral UQ methods for the full structure. First, the computational cost of a FIT time-domain simulation for a single unit cell is significantly reduced compared to an evaluation of the full structure. Second, the number of parameters for a single unit cell is reduced by a factor of N_{cells} w.r.t. the full structure, which has a huge impact on the efficiency of an SC method due to the curse-of-dimensionality.

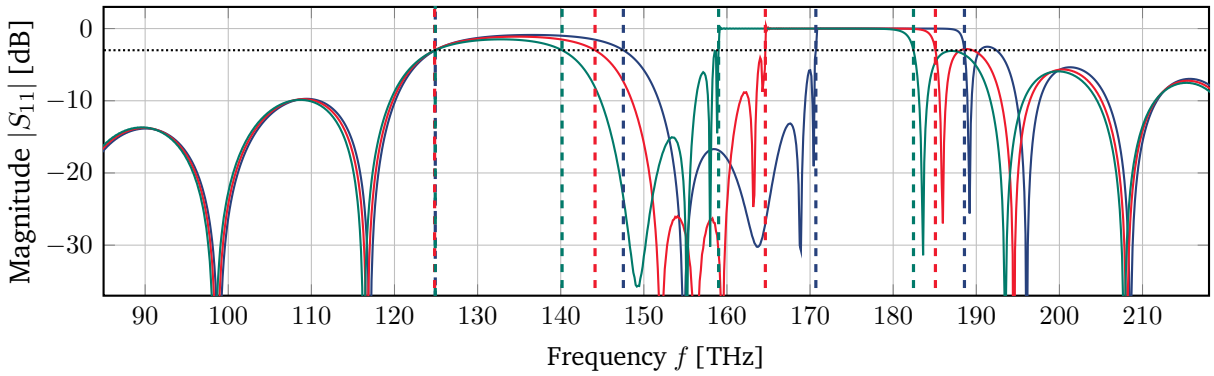


Figure 5.43: Reflection coefficient for different realizations of SRR array, based on [87, Figure 2]. The respective limit frequencies of the considered bandgaps are depicted by the dashed vertical lines. The dotted line indicates $-3\,\mathrm{dB}$.

Table 5.6: Considered multifidelity models of SRR array, based on [87, Table 1]. The columns specify the computational cost $C^{(i)}$ per model evaluation of $\mathcal{Q}^{(i)}$ and the respective estimated correlation coefficients $\rho_{1,i}$ with the high-fidelity model $\mathcal{Q}^{(i)}$ for both bandgap center frequencies $f_{\mathrm{c},1}, f_{\mathrm{c},2}$.

Symbol	Model	Cost $C^{(i)}$	$\rho_{1,i}$ for $f_{\mathrm{c},1}$	$\rho_{1,i}$ for $f_{\mathrm{c},2}$
$\mathcal{Q}^{(1)}$	Full model (FIT, $2 \cdot 10^5$ time-steps)	197.50 s	1.000000	1.000000
$\mathcal{Q}^{(2)}$	Full model (FIT, $2 \cdot 10^4$ time-steps)	11.25 s	0.999236	0.998035
$\mathcal{Q}^{(3)}$	SMA (FIT, 1 port-mode)	9.64 s	0.999943	0.968376
$\mathcal{Q}^{(4)}$	SMA (FIT, 2 port-modes)	115.47 s	0.999998	0.999998
$\mathcal{Q}^{(5)}$	SMA + unit cell SC (1 port-mode)	0.006 s	0.999943	0.967540
$\mathcal{Q}^{(6)}$	SMA + unit cell SC (2 port-modes)	0.026 s	0.999998	0.999886

Next, we investigate this UQ methodology numerically for an SRR array with $N_{\mathrm{cells}} = 7$ cells which is considered as a benchmark problem. The Cartesian mesh and the SRR shape parameters, based on [132], are depicted in Fig. 5.42. Note that for simplicity, we only consider uncertainties in the longitudinal lengths $L^{(j)}$, $j = 1, \ldots, N_{\mathrm{cells}}$ of the SRR elements in cell j such that $L^{(j)} = 320\,\mathrm{nm} + \xi_j$, where

$$\Xi_j \sim \mathcal{U}(-15\,\mathrm{nm}, 15\,\mathrm{nm}),\ j = 1, \ldots, N_{\boldsymbol{\xi}},\ \text{with } N_{\boldsymbol{\xi}} = N_{\mathrm{cells}}, \tag{5.37}$$

are independent and identically distributed RVs.

The broadband behavior of the fundamental reflection coefficient is shown in Fig. 5.43 for different realizations of the structure. As expected, two bandgaps can be observed, that is frequency ranges where no transmission is possible, which both depend on the parameter variations. We define the corresponding limit frequencies of the bandgaps as the frequencies where the S_{11} parameter drops below $-3\,\mathrm{dB}$, which are depicted by dashed lines in Fig. 5.43 and can be computed in post-processing. As QoIs we then consider the respective center frequencies $f_{\mathrm{c},i}$, $i \in \{1, 2\}$ of the bandgaps, where the index i indicates the first or second bandgap, respectively. We note that for a few parameter configurations some additional resonances within the range of the second bandgap occur, which can be attributed to the slightly detuned resonances in the array of SRRs. However, we only consider the outer limit frequencies of this bandgap in the following MFMC study.

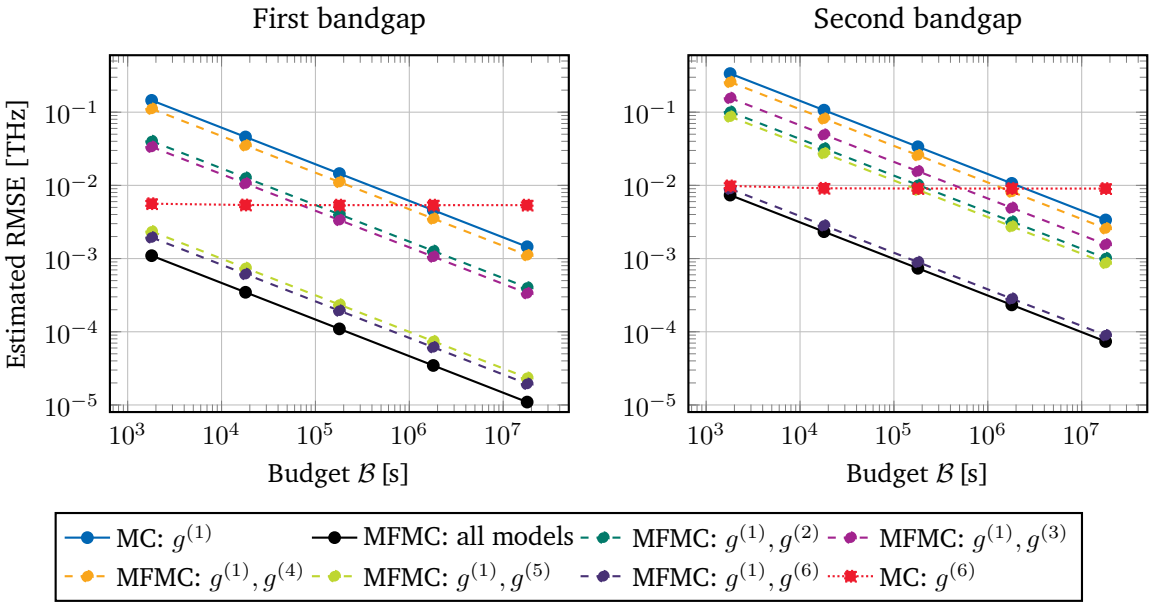


Figure 5.44: Convergence study of the estimated RMSEs for MC and MFMC based on the different models shown in Table 5.6. The illustration is based on [87, Figure 3].

We employ $N_{\text{mod}} = 6$ numerical models $\mathcal{Q}^{(j)}, j = 1, \dots, N_{\text{mod}}$ of different fidelity which are listed in Table 5.6 along with the associated cost per model evaluation $C^{(i)}$ measured in computation time on a standard workstation where an in-house Matlab implementation is employed for all models. As high-fidelity model $\mathcal{Q}^{(1)}$ we consider a FIT model of the full structure with a maximum number of $2 \cdot 10^5$ time-steps. However, for all FIT models the time-stepping is immediately terminated if the energy drops below $-120\,\text{dB}$. For the low-fidelity model $\mathcal{Q}^{(2)}$, which is again a FIT model of the full structure, the maximum number of time-steps is restricted to $2 \cdot 10^4$. Next, the SMA approach is employed, which yields the low-fidelity models $\mathcal{Q}^{(3)}$ and $\mathcal{Q}^{(4)}$ with different numbers of port modes N_{modes}. In particular, $\mathcal{Q}^{(3)}$ only employs the fundamental propagating transverse electric and magnetic (TEM) mode, while $\mathcal{Q}^{(4)}$ additionally considers the evanescent first TM mode. Finally, we construct the corresponding low-fidelity models $\mathcal{Q}^{(5)}, \mathcal{Q}^{(6)}$ by employing unit cell SC approximations which are here for simplicity obtained using univariate polynomial interpolation on $\#Z = 7$ Chebyshev nodes. We note that the computational cost for the surrogate modeling only needs to be invested once. Furthermore, in this case, even a single evaluation of the high-fidelity model $\mathcal{Q}^{(1)}$ requires a larger computation time than constructing the unit cell surrogate models. Hence, in the following we neglect this cost, for simplicity. It can also be observed that the cost $C^{(i)}$ scales roughly linear w.r.t. the number of cells N_{cells} of the simulation model while the offline-cost for the SC approximation of a single unit cell is obviously independent of N_{cells}. Hence, qualitatively similar results as shown in the following can also be expected for structures with $N_{\text{cells}} \neq 7$, which we confirmed numerically for $N_{\text{cells}} = 14$.

We investigate the performance of MFMC based on the so-obtained multifidelity models. To this end, we use a MFMC Matlab implementation based on [3], see [164]. First, we employ a random input sample $\{\boldsymbol{\Xi}_i\}_{i=1}^{\tilde{N}_{\text{MC}}}$ with $\tilde{N}_{\text{MC}} = 500$ and compute the respective output samples $\{\mathcal{Q}^{(j)}(\boldsymbol{\Xi}_i)\}_{i=1}^{\tilde{N}_{\text{MC}}}$, for each model $\mathcal{Q}^{(j)}$, $j = 1, \dots, N_{\text{mod}}$. These samples are used to estimate the respective correlation coefficients $\rho_{1,j}, j = 1, \dots, N_{\text{mod}}$ with the high-fidelity model which are given in Table 5.6. We note that all models have a strong correlation.

[3] `github.com/pehersto/mfmc`

In a next step, we compare the convergence of the estimated RMSEs for MC and MFMC w.r.t. the computational budget $\mathcal{B}$ which is presented in Fig. 5.44. These estimates for RMSEs are obtained based on the readily available output samples $\{\mathcal{Q}^{(j)}(\boldsymbol{\Xi}_i)\}_{i=1}^{\tilde{N}_{\mathrm{MC}}}$. In particular, for standard MC with the model $\mathcal{Q}^{(j)}$, we can employ (4.2) with $N_{\mathrm{MC}} = \mathcal{B}/C^{(j)}$ where the standard deviation of the model is replaced by its MC estimate. In a similar way, the RMSE of MFMC is accurately estimated using (4.17). It can be observed in Fig. 5.44 that the proposed MFMC approach reduces the computational cost to reach a fixed accuracy by orders of magnitude w.r.t. MC on the high-fidelity model. The MFMC approach excludes some low-fidelity models, since, for instance $\mathcal{Q}^{(2)}$ and $\mathcal{Q}^{(3)}$ have higher costs $C^{(2)}, C^{(3)}$ than the unit cell surrogate-based model $\mathcal{Q}^{(6)}$ but also a lower correlation with the high-fidelity model $\mathcal{Q}^{(1)}$. We note that the selection and ordering of suitable models can be done using a model selection strategy employing a small sample obtained in a pilot run, see also [164, 165]. To investigate the impact of the different low-fidelity models, we additionally present the convergence of MFMC when only $\mathcal{Q}^{(1)}$ and one low-fidelity model $\mathcal{Q}^{(j)}$, $j \in \{2, \dots, 6\}$ is included with dashed lines in Fig. 5.44. It can be observed that, as expected, this approach yields for all low-fidelity models better results than MC, but performs worse than the selection of models employed by the MFMC algorithm. We further note that, for both bandgaps, mainly the low-fidelity models based on the proposed unit cell SC approximations yield the huge improvements in computational efficiency. Comparing the results for the different bandgaps, it can be seen that considering $N_{\mathrm{modes}} = 2$ port-modes for the SMA is clearly necessary for the second bandgap while for the first bandgap also $N_{\mathrm{modes}} = 1$ port mode would already yield very good results. This is expected from a physical point of view, as the second bandgap is mainly caused by the mutual coupling between the cells while the first bandgap mainly depends on the fundamental resonance of the single SRR element.

Finally, we emphasize that the limited recurrences to the high-fidelity model of the MFMC method are indeed necessary to remove the biasing error of the low-fidelity models. In particular, the associated RMSE of standard MC based only on the unit cell SC-based model $\mathcal{Q}^{(6)}$ is depicted by the dotted red line in Fig. 5.44, where the sampling error and the biasing error are both estimated based on the MC sample. It can be seen that the sampling error already has a negligible magnitude in the considered range of computational budgets $\mathcal{B}$; however, due to the biasing error, no convergence can be observed.

In summary, the proposed UQ methodology combining SC on the unit cell level, SMA and MFMC, is found to be highly efficient for quasi-periodic optical structures with shape uncertainties. Compared to the direct application of spectral UQ methods, only a single unit cell SC approximation is required, which significantly reduces the number of uncertain parameters as well as the simulation effort in the offline-phase. The MFMC method then provides unbiased estimates where speed-ups by several orders of magnitude were observed for the considered benchmark problem.

5.3.4 Electric waveguide

In this subsection, we investigate the adjoint-based yield estimation procedure proposed in Section 4.4. To this end, we consider the model of a rectangular electric waveguide with a dielectric inset which is illustrated in Fig. 5.45. The particular model is chosen, as it was already established as a benchmark problem in [135, 137] and because it has a smooth input-to-output behavior as well as an available closed-form reference solution [135]. The content and structure of this section are based on our work [81].

The corresponding numerical model (2.27) is, as in Section 5.3.1, again implemented using the FE library FEniCS [3]. As explained in Section 4.3.1.3, the corresponding adjoint problem (4.113) can be solved with negligible cost since we employ a sparse LU decomposition for the primal problem. The model has two

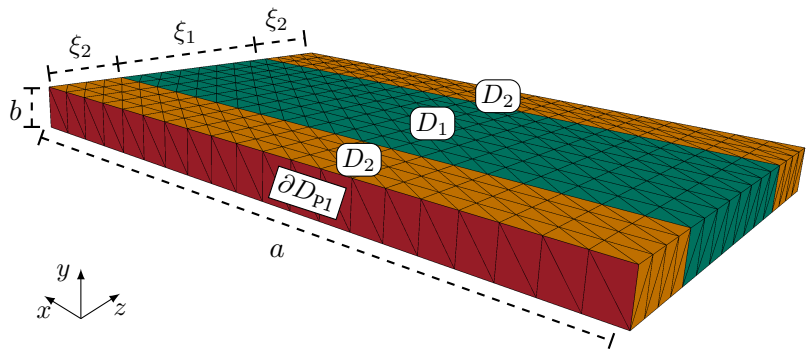


Figure 5.45: Numerical model of a rectangular electric waveguide with a dielectric inset D_1 of length y_1, based on [81, Figure 2]. As excitation an incident TE_{10} wave at the port ∂D_{P1} is considered.

geometrical parameters ξ_1 and ξ_2, which refer to the length of the dielectric inset D_1 and the vacuum offsets D_2 measured in millimeters, respectively. The material of the subdomain D_2 is assumed to have vacuum permeability and permittivity ε_0, μ_0, while different material properties are assumed for the inset D_1. In particular, we will study two cases with different material models and, hence, different numbers of parameters. In the first case, the permittivity and permeability of the dielectric inset are given as

$$\varepsilon_{\mathrm{r}}^{\{1\}}|_{D_1}(\boldsymbol{\xi}) = 1 + \xi_3 + (1-\xi_3)\,(1+\mathrm{i}\omega 4\tau)^{-1}\,, \tag{5.38}$$

$$\mu_{\mathrm{r}}^{\{1\}}|_{D_1}(\boldsymbol{\xi}) = 1 + \xi_4 + (2-\xi_4)\left(1+\mathrm{i}\omega\frac{10}{11}\tau\right)^{-1}\,, \tag{5.39}$$

where $\tau = \left(2\pi(20\,\mathrm{GHz})\right)^{-1}$ and $\boldsymbol{\xi} \in \mathbb{R}^4$. In the second case, we consider

$$\varepsilon_{\mathrm{r}}^{\{2\}}|_{D_1}(\boldsymbol{\xi}) = \xi_5 + (\xi_3-\xi_5)\,(1+\mathrm{i}\omega\xi_6\tau)^{-1} + (\xi_4-\xi_5)\,(1+\mathrm{i}\omega\xi_7\tau)^{-1}\,, \tag{5.40}$$

$$\mu_{\mathrm{r}}^{\{2\}}|_{D_1}(\boldsymbol{\xi}) = \xi_{10} + (\xi_8-\xi_{10})\,(1+\mathrm{i}\omega\xi_{11}\tau)^{-1} + (\xi_9-\xi_{10})\,(1+\mathrm{i}\omega\xi_{12}\tau)^{-1}\,, \tag{5.41}$$

where $\boldsymbol{\xi} \in \mathbb{R}^{12}$. In both cases, the parameters $\boldsymbol{\xi}$ are modeled as realizations of mutually independent RVs $\boldsymbol{\Xi}$ where the respective mean values are given as

$$\mathbb{E}\left[\boldsymbol{\Xi}^{\{1\}}\right] = [10.36, 4.76, 0.58, 0.64]^\top, \tag{5.42}$$

$$\mathbb{E}\left[\boldsymbol{\Xi}^{\{2\}}\right] = [8.6, 3.8, 2, 0.5, 0.7, 0.6, 1.4, 2.8, 1.7, 0.8, 0.3, 1.4]^\top. \tag{5.43}$$

To avoid unphysical parameter values, we employ truncated normal distributions for each RV. The respective standard deviations and truncation offset are given as $0.7\,\mathrm{mm}$ and $\pm 3\,\mathrm{mm}$, respectively, for the geometrical parameters Ξ_1 and Ξ_2. For each of the material parameters $\Xi_3, \ldots, \Xi_{N_{\boldsymbol{\xi}}}$ the standard deviation is given as $0.3\,\mathrm{mm}$ and the truncation offset as $\pm 0.3\,\mathrm{mm}$. As QoI $\mathcal{Q}$ we consider the fundamental reflection coefficient (2.28). The respective performance feature specification is here given as

$$|\mathcal{Q}(\boldsymbol{\xi},\omega)| \overset{!}{\leq} -24\,\mathrm{dB} \quad \forall \omega \in \Omega = [2\pi 6.5, 2\pi 7.5]\ \mathrm{GHz}. \tag{5.44}$$

Note that following [97], in contrast to (4.7), an additional range parameter in terms of the angular frequency ω is considered. This requires minor and straightforward generalizations of the yield estimation procedure described in Section 4.4, cf. [81]. In particular, the surrogate modeling and the hybrid decision process is

applied for each $\omega_i \in T_\Omega \subset \Omega$, where T_Ω denotes a set of 11 equidistant frequency points in Ω. A MC sample point $\boldsymbol{\xi}^{(i)}$ is then only classified as an element of the safe domain Γ_s if $\left|\mathcal{Q}(\boldsymbol{\xi}^{(i)}, \omega_i)\right| \leq -24\,\mathrm{dB}$ is fulfilled for all frequency points $\omega_i \in T_\Omega$. If the requirements are not fulfilled for a single arbitrary frequency value $\omega_i \in T_\Omega$, i.e. $\left|\mathcal{Q}(\boldsymbol{\xi}^{(i)}, \omega_i)\right| > -24\,\mathrm{dB}$, the MC sample point $\boldsymbol{\xi}^{(i)}$ can be immediately identified to be no element of the safe domain Γ_s and, hence, further computational effort for investigating the remaining frequency points can be avoided.

Next, we compare the adjoint-based hybrid yield estimation procedure with alternative approaches. To this end, we first calculate reference solutions for the yield by employing standard MC (4.11) on the closed-form reference solution of the electric waveguide. The MC sample size is chosen as $N_\mathrm{MC} = 2500$, i.e. the smallest number such that the RMSE error (4.13) is below 0.01 for all values of the yield. The reference solutions are then given as

$$\mathcal{Y}_\mathrm{Ref}^{\{1\}} = 95.44\,\%, \qquad \text{and} \qquad \mathcal{Y}_\mathrm{Ref}^{\{2\}} = 74.60\,\%, \tag{5.45}$$

for the case of $N_{\boldsymbol{\xi}} = 4$ and $N_{\boldsymbol{\xi}} = 12$ uncertain parameters, respectively. In Table 5.7 we compare different approaches for yield estimation in terms of computational effort and accuracy, in the case of $N_{\boldsymbol{\xi}} = 4$ parameters. Note that we employ for all methods the same MC sample which was already used for the reference solution. The accuracy of a yield estimate $\tilde{\mathcal{Y}}$ is quantified by the relative error with respect to the reference solution (5.45), i.e.

$$\mathcal{E}_\mathrm{rel} = \frac{\left|\mathcal{Y}_\mathrm{Ref} - \tilde{\mathcal{Y}}\right|}{\mathcal{Y}_\mathrm{Ref}}. \tag{5.46}$$

The computational costs are evaluated based on the number of evaluations of the different models, that is the respective number of sparse LU decompositions. As we consider FE models $\mathcal{Q}_h, \mathcal{Q}_{h/2}, \mathcal{Q}_{h/4}$ with different mesh sizes for the hybrid approach, the respective computational cost of each model varies. In this work, we consider three different refinement levels with mesh sizes $h, h/2, h/4$ where the respective number of evaluations are denoted by $N_h^\mathrm{FE}, N_{h/2}^\mathrm{FE}, N_{h/4}^\mathrm{FE}$, respectively. Here, we assume an optimal solver such that the computational effort scales linearly w.r.t. N_h, i.e. the number of DoFs. This implies an increase by a factor of 4 for 2D problems and a factor of 8 for 3D problems. For the considered numerical model, the y-component of the electric field is constant and, hence, a refinement in x- and z- direction is sufficient. Accordingly, we measure the effective computational cost as

$$C_\mathrm{eff} = N_h^\mathrm{FE} + 4N_{h/2}^\mathrm{FE} + 16N_{h/4}^\mathrm{FE}, \tag{5.47}$$

i.e. as equivalent cost in terms of multiples of the cost for a single evaluation of the FE model $\mathcal{Q}_h$. Note that the computational costs for evaluating the surrogate model $\mathcal{Q}_{h,Z}$ or the adjoint-error indicator are neglected here, for simplicity.

The first line of Table 5.7 refers to a standard MC approach (4.11) on the most refined FE model $\mathcal{Q}_{h/4}$, where the same accuracy as with the proposed hybrid approach can be expected. Note that this approach leads to the exact same result as $\mathcal{Y}_\mathrm{Ref}^{\{1\}}$ which was obtained with the closed-form solution of the rectangular waveguide. However, it can be seen that the computational cost is quite high. In contrast, in a next step we employ yield estimation based on MC on a surrogate model solely. To this end, the adjoint-based adaptive Leja algorithm 3 is employed with a computational budget of $\mathcal{B} = 30$ for each of the 11 frequency points. For simplicity we here employ the trivial conformal map $\mathbf{g} : \mathbf{s} \mapsto \mathbf{s}$ and unweighted Leja nodes obtained by using a uniform weight function on Γ in (4.33) in order to reach a high uniform accuracy. In this case, the computational cost is reduced drastically, but some sample points are wrongly classified, leading to a deviation of the estimated yield w.r.t. the reference value. This error can be avoided by employing the proposed adjoint-based hybrid scheme. In particular, we first examine the hybrid approach without surrogate modeling, that is we only

Table 5.7: Comparison of different approaches for yield estimation in the case of $N_{\boldsymbol{\xi}} = 4$ uncertain parameters, based on [81, Table 2]. The columns specify the number of model evaluations $N_h^{\mathrm{FE}}, N_{h/2}^{\mathrm{FE}}, N_{h/4}^{\mathrm{FE}}$ of the different refined FE models, the effective computational cost C_{eff} as well as the relative error $\mathcal{E}_{\mathrm{rel}}$ with respect to the reference solution $\mathcal{Y}_{\mathsf{Ref}}^{\{1\}}$.

Method	N_h^{FE} for SC	N_h^{FE} for MC	$N_{h/2}^{\mathrm{FE}}$ for MC	$N_{h/4}^{\mathrm{FE}}$ for MC	C_{eff}	$\mathcal{E}_{\mathrm{rel}}$ (%)
MC on $\mathcal{Q}_{h/4}$	0	0	0	26 300	421 760	0.00
MC on $\mathcal{Q}_{h,Z}$	330	0	0	0	330	0.13
Hybrid without SC	0	26 360	5	1	26 396	0.00
Hybrid with SC	330	165	5	1	531	0.00

Table 5.8: Comparison of different approaches for yield estimation in the case of $N_{\boldsymbol{\xi}} = 12$ uncertain parameters, based on [81, Table 1]. The columns specify the number of model evaluations $N_h^{\mathrm{FE}}, N_{h/2}^{\mathrm{FE}}, N_{h/4}^{\mathrm{FE}}$ of the different refined FE models, the effective computational cost C_{eff} as well as the relative error $\mathcal{E}_{\mathrm{rel}}$ with respect to the reference solution $\mathcal{Y}_{\mathsf{Ref}}^{\{2\}}$.

Method	N_h^{FE} for SC	N_h^{FE} for MC	$N_{h/2}^{\mathrm{FE}}$ for MC	$N_{h/4}^{\mathrm{FE}}$ for MC	C_{eff}	$\mathcal{E}_{\mathrm{rel}}$ (%)
MC on $\mathcal{Q}_{h/4}$	0	0	0	22 705	363 280	0.00
MC on $\mathcal{Q}_{h,Z}$	990	0	0	0	990	6.22
Hybrid without SC	0	22 705	25	6	22 901	0.00
Hybrid with SC	990	4812	25	6	5998	0.00

employ the adjoint-based mesh refinement strategy. In this case, all samples are correctly classified while the computational cost is reduced by more than one order of magnitude w.r.t. the MC approach on the FE model $\mathcal{Q}_{h/4}$. The efficiency is then further improved by incorporating the SC approximation. In this case, the hybrid approach still ensures a correct classification of all sample points while the computational cost is now reduced to only $0.13\,\%$ w.r.t. the cost of the MC approach on the FE model $\mathcal{Q}_{h/4}$. We emphasize that the hybrid approach evaluates mostly the cheapest models while only a few evaluations of the highly-resolved FE models are required.

Next, we repeat the same study in the $N_{\boldsymbol{\xi}} = 12$ dimensional setting with reference solution $\mathcal{Y}_{\mathsf{Ref}}^{\{2\}}$ where we now employ a computational budget of $\mathcal{B} = 90$ model evaluations for the surrogate construction at each frequency point. The results are presented in Table 5.8. Reaching a high accuracy with an SC approximation becomes more challenging for a larger number of uncertain parameters due to the curse-of-dimensionality. Accordingly, the associated error of employing MC on the surrogate models significantly increases w.r.t. the previous study to $\mathcal{E}_{\mathrm{rel}} = 6.22\,\%$. Furthermore, it can also be seen that the speed-up in computation time, which can be attributed to the use of SC in the hybrid approach, slightly decreases. However, we emphasize that the computational cost is still reduced to only $1.7\,\%$ compared to the MC approach on the FE model $\mathcal{Q}_{h/4}$ while maintaining the exact same accuracy.

It can be concluded that the proposed adjoint-based hybrid yield estimation is highly efficient while ensuring a correct classification of the sample points, as it combines the efficiency of SC with the reliability of a high-fidelity MC analysis. Finally, it shall be mentioned that it can be efficiently incorporated in a yield optimization workflow, see [81] for details, which was excluded here for brevity.

5.4 Summary

In this chapter, we have investigated the application of the different methods suggested in this thesis to a number of benchmark problems that comprise both academic and real-world applications. In most cases, the proposed surrogate modeling techniques have been a key tool to reduce the computational effort significantly. In particular, they have enabled efficient approximations with up to approximately 20 sensitive parameters while we still relied on MC-based approaches for the studies with up to 640 independent uncertain parameters. In addition, several further challenges arising in UQ for RF and optical applications have been successfully addressed, e.g. the homotopy-based eigenvalue tracking method has been shown to enable spectral UQ for Maxwell's eigenproblem with uncertain input data. Furthermore, important UQ results have been obtained for the considered optical gratings as well as the TESLA cavities manufacturing process, where especially the estimated sensitivities have given valuable insights. In particular, it has been found that the considered optical grating coupler is highly sensitive to its shape parameters, as even small variations within $\pm 1.5\,\mathrm{nm}$ have a significantly larger impact than the considered measurement uncertainty in the material data.

6 Conclusion and prospects for future research

In this thesis, parametric problems derived from Maxwell's equations in frequency domain have been addressed where we have focused on radio frequency and optical applications. In particular, we have distinguished two types of parametric problems where we have suggested different surrogate modeling techniques for the frequency variable on the one side and the uncertain material and shape parameters on the other side. This distinction was not only motivated by the different problem settings but based on the usually significantly enlarged parameter range for the frequency variable compared to typical uncertain parameters. Furthermore, the uncertain shape and material parameters often feature more smoothness but pose another challenge due to the increased number of parameters and the curse-of-dimensionality.

For the first case, we have considered in Chapter 3 the approximation of stable frequency response functions and suggested a rational kernel-based interpolation method. It has been confirmed that, although all considered FRFs are holomorphic on the frequency axis, it is beneficial to treat the frequency variable differently than the uncertain parameters for which polynomial-based methods are employed. The suggested method combines Szegö kernel interpolation with a suitable pseudo-kernel and a few rational basis functions inspired from vector fitting as well as a dedicated tuning and model selection procedure. It has shown impressive results for a number of benchmark problems from different fields. In particular, the method has reached at least a comparable accuracy as the established state-of-the-art methods adaptive Antoulas–Anderson and vector fitting for the considered benchmarks while in some cases significant improvements could be observed.

For the second case, we have addressed in Chapter 4 the forward uncertainty propagation problem. A major part has been dedicated to the conformally mapped spectral uncertainty quantification methods. They have been shown to feature an enhanced convergence order with respect to standard generalized polynomial chaos or stochastic collocation methods, for functions which are analytic in an ϵ-neighborhood of the approximation interval. It has been shown that the conformally mapped generalized polynomial chaos method allows the direct computation of stochastic moments and variance-based sensitivity indices based on the expansion coefficients. This is not the case for the conformally mapped stochastic collocation method where further post-processing of the approximation would be required. However, it has been demonstrated that the collocation approach allows to realize very efficient dimension-adaptive schemes based on granular mapped Leja nodes. We emphasize that both proposed conformally mapped spectral methods can be realized non-intrusively, i.e. they only require access to the map from input to output parameters. Hence, they should be generally applied for models with high parametric sensitivities, where the underlying function is holomorphic in a small ϵ-neighborhood. Then, different extensions employing adjoint-based error indicators have been proposed. They are based on a number of (mild) assumptions, specified at the beginning of Section 4.3, e.g. regarding the general form of the parametric model, and require some access to the underlying numerical model, for instance in order to evaluate residuals. These error indicators have been employed to steer the adaptivity in the dimension-adaptive sparse stochastic collocation scheme based on mapped Leja nodes as well as for adjoint error correction such that a further improved convergence can be observed. It has been found that these enhanced surrogate modeling techniques combining dimension-adaptivity, adjoint error estimation,

and conformal maps allow to successfully address moderately high dimensional problems and delay the curse-of-dimensionality. In particular, the grating coupler model with 17 uncertain parameters could not be addressed with standard sparse approximation techniques. For example, total-degree generalized polynomial chaos based on sparse quadrature has not even reached a root-mean-square error below 10^{-1}, employing more than 7000 model evaluations. On the contrary, the enhanced surrogate modeling techniques have achieved a suitable error magnitude of 10^{-4} with smaller computational cost, enabling an efficient spectral uncertainty quantification. Next, a reliable yield, or equivalently failure probability, estimation has been addressed. To this end, a multifidelity scheme has been proposed, which is, again, based on stochastic collocation and adjoint error estimation, and takes all relevant error sources into account. In particular, it combines the sampling of the stochastic collocation approximation with evaluations of finite element models of different fidelity whenever required. This has lead to tremendous computational savings without sacrificing any accuracy with respect to a standard Monte Carlo method on the high-fidelity model for the considered benchmark problem of an electric waveguide. Finally, addressing uncertainty quantification for Maxwell's eigenproblem where eigenvalue crossings with respect to parameter variations can occur, an eigenvalue tracking method has been introduced. Numerical investigations have indicated that the method not only ensures a correct matching of eigenmodes but can also improve the computational efficiency.

Thorough uncertainty quantification studies have been presented for the 9-cell TeV-Energy Superconducting Linear Accelerator cavities employed in the European X-ray Free Electron Laser at Deutsches Elektronen-Synchrotron, which have been substantiated by the available measurement data. It has been found that the eccentric deformations only have a small impact on the fundamental mode spectra, while it could be confirmed that certain equatorial and iris radius parameters are highly relevant where the estimated sensitivity indices have given valuable insights. Furthermore, manufacturing imperfections for two different optical gratings [167, 181] have been addressed and various uncertainty quantification results have been presented. In this context, it has been found that for the considered optical grating coupler [89], shape uncertainties outweigh the impact of the measurement uncertainty in the material data of the noble metals. It has also been shown that the averaging interference effects of neighboring components in the considered gradient index metasurface stabilize the performance, which is of significant practical importance [181]. Finally, we have addressed a split ring resonator array, i.e. a periodic structure with an uncertain shape of finite size. For this type of problem, we have proposed a decoupled uncertainty propagation methodology, by combining stochastic collocation on the unit cell level, a scattering matrix approach and multifidelity Monte Carlo. This approach has been found to be highly efficient for the considered benchmark application. In summary, for the majority of the UQ studies, the suggested enhanced surrogate modeling techniques have been a key tool for keeping the computational effort manageable and enabling moderately high-dimensional approximations. For instance, in the considered applications up to about 20 sensitive uncertain parameters could be efficiently treated with the described methods.

To further investigate, employ or improve the proposed methods for surrogate modeling and uncertainty quantification, we suggest to address the following topics in further research works:

- There are several ideas how the rational kernel-based interpolation method for frequency response functions could be possibly extended to further improve the efficiency. For instance, adaptive sequential sampling strategies could be investigated. Further research could also address the efficient incorporation of derivative data which can in many cases be obtained with reduced computational effort, for instance by employing adjoint techniques. Additionally, alternative kernels tailored to frequency response functions could be investigated and compared to the here employed Szegö kernel. In particular, a detailed comparison and analysis of the kernels proposed in [125, Section 5] is still subject to further research.
- In order to address surrogate modeling for even larger numbers of uncertain parameters, it should

be investigated if the efficiency of the proposed adjoint-based mapped spectral stochastic collocation method could be even further enhanced. To this end, the sequential scheme could be extended to more efficiently utilize high-performance computing resources. In particular, for this method we assumed that all computing power is already exploited within the numerical model. However, if additional parallelization resources are available, the dimension-adaptive method should be adapted such that it simultaneously evaluates the model at different collocation points. Alternatively or complementary, a combination with the multilevel or multifidelity stochastic collocation approaches [148, 197] seems promising.

- As explained at the beginning of this section, we have separated the surrogate modeling task for the frequency dependence and the uncertain shape and material parameters. In particular, the rational kernel-based interpolation method yields good results for the approximation of frequency response functions, while the mapped spectral approximations have been shown to perform well for uncertainty quantification problems in fixed frequency settings as well as for eigenvalues. However, there are also various uncertainty quantification tasks where broadband quantities are of interest. Two particular examples have been presented in Section 5.3.3 and Section 5.3.4, where it has been possible to construct unit cell surrogate models at each frequency point separately. Future work should investigate if the different proposed surrogate modeling techniques can be combined efficiently to address surrogate modeling with respect to the frequency variable and the uncertain parameters simultaneously. Alternatively, a generalization of the rational kernel-based interpolation method to the multivariate case could be investigated, which could eventually be based on the recently suggested multivariate adaptive Antoulas–Anderson algorithm [169]. These methods should then be compared against the parameterized model order reduction schemes [20, 153, 170] which were also recently proposed for broadband uncertainty quantification in the frequency domain.
- Note that in this thesis, we have focused on the forward uncertainty propagation problem. However, the suggested surrogate modeling techniques could also be employed in an inverse uncertainty quantification context. In this regard, the estimation of all shape parameters for the operational TeV-Energy Superconducting Linear Accelerator cavities at cryogenic temperatures based on the available measurement data of the eigenmode spectra would be of high practical relevance. This topic has so far only been briefly addressed in Section 5.2.3 for the iris radius by applying significant simplifications. However, a more sophisticated study would require dedicated numerical tools [191] and eventually also higher order mode measurements, to avoid ill-conditioning.
- In this work, we have suggested several methods for the quantification of uncertainties. These methods could be further embedded in robust optimization routines in order to obtain more reliable designs. In this context, it shall be noted that the adjoint-based yield estimation method has already been successfully combined with an adaptive and efficient yield optimization procedure, see our work [81].
- The presented eigenmode tracking technique could eventually also be employed for automatic mode recognition [34] of higher order modes in cavity simulation. In particular, it is currently under investigation if tracking the eigenmodes during the shape morphing of a TeV-Energy Superconducting Linear Accelerator cavity to a cylindrical pillbox cavity, where the eigenmodes are known analytically, allows for a reliable classification.

List of acronyms

AAA	adaptive Antoulas–Anderson
AIC	Akaike information criterion
ANOVA	analysis of variances
BEM	boundary element method
BIC	Bayesian information criterion
BVP	boundary value problem
CAD	computer-aided design
CDF	cumulative distribution function
DESY	Deutsches Elektronen-Synchrotron
DFG	German Research Foundation
DFT	discrete Fourier transform
DoF	degree of freedom
EBM	electron-beam welding
EXFEL	European X-ray Free Electron Laser
FE	finite element
FEM	finite element method
FIT	finite integration technique
FORM	first-order reliability method
FRF	frequency response function
GaN	Gallium Nitride
GLS	generalized least squares
GP	Gaussian process
GPC	generalized polynomial chaos
HOM	higher order mode
IGA	isogeometric analysis
KDE	kernel density estimation
LOO	leave-one-out

MC Monte Carlo
MFMC multifidelity Monte Carlo
MIM metal-insulator-metal
MOR model order reduction

NURBS non-uniform rational B-splines

PDE partial differential equation
PDF probability density function
PEC perfect electric conducting
PMC perfect magnetic conducting
PML perfectly matched layer

QoI quantity of interest

RBF radial basis function
RCWA rigorous coupled wave analysis
RF radio frequency
RHS right-hand side
RKHS reproducing kernel Hilbert space
RKI rational kernel-based interpolation
RMSE root-mean-square error
RV random variable

SC stochastic collocation
SMA scattering matrix approach
SORM second-order reliability method
SQP sequential quadratic programming
SRR split ring resonator

TE transverse electric
TEM transverse electric and magnetic
TESLA TeV-Energy Superconducting Linear Accelerator
TM transverse magnetic

UQ uncertainty quantification

VF vector fitting

Bibliography

[1] A. Adelmann, "On nonintrusive uncertainty quantification and surrogate model construction in particle accelerator modeling", *SIAM/ASA Journal on Uncertainty Quantification*, vol. 7, no. 2, pp. 383–416, 2019.

[2] M. Ainsworth and J. T. Oden, "A posteriori error estimation in finite element analysis", *Computer Methods in Applied Mechanics and Engineering*, vol. 142, no. 1-2, pp. 1–88, 1997.

[3] M. S. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes, and G. N. Wells, "The FEniCS project version 1.5", *Archive of Numerical Software*, vol. 3, no. 100, pp. 9–23, 2015.

[4] M. Altarelli, R. Brinkmann, M. Chergui, W. Decking, B. Dobson, S. Düsterer, G. Grübel, W. Graeff, H. Graafsma, J. Hajdu, J. Marangos, J. Pflüger, H. Redlin, D. Riley, I. Robinson, J. Rossbach, A. Schwarz, K. Tiedtke, T. Tschentscher, I. Vartaniants, H. Wabnitz, H. Weise, R. Wichmann, K. Witte, A. Wolf, M. Wulff, and M. Yurkov, *The Technical Design Report (TDR) of the European XFEL*, DESY 2006-097, 2007.

[5] A. C. Antoulas, S. Lefteriu, A. C. Ionita, P. Benner, and A. Cohen, "A tutorial introduction to the Loewner framework for model reduction", *Model Reduction and Approximation: Theory and Algorithms*, vol. 15, p. 335, 2017.

[6] D. N. Arnold, R. S. Falk, and R. Winther, "Finite element exterior calculus from Hodge theory to numerical stability", *Bulletin of the American Mathematical Society*, vol. 47, no. 2, pp. 281–354, 2010.

[7] S.-K. Au and J. L. Beck, "Estimation of small failure probabilities in high dimensions by subset simulation", *Probabilistic Engineering Mechanics*, no. 16, pp. 263–277, 2001.

[8] B. Aune, R. Bandelmann, D. Bloess, B. Bonin, A. Bosotti, M. Champion, C. Crawford, G. Deppe, B. Dwersteg, D. A. Edwards, H. T. Edwards, M. Ferrario, M. Fouaidy, P.-D. Gall, A. Gamp, A. Gössel, J. Graber, D. Hubert, M. Hüning, M. Juillard, T. Junquera, H. Kaiser, G. Kreps, M. Kuchnir, R. Lange, M. Leenen, M. Liepe, L. Lilje, A. Matheisen, W.-D. Möller, A. Mosnier, H. Padamsee, C. Pagani, M. Pekeler, H.-B. Peters, O. Peters, D. Proch, K. Rehlich, D. Reschke, H. Safa, T. Schilcher, P. Schmüser, J. Sekutowicz, S. Simrock, W. Singer, M. Tigner, D. Trines, K. Twarowski, G. Weichert, J. Weisend, J. Wojtkiewicz, S. Wolff, and K. Zapfe, "Superconducting TESLA cavities", *Physical Review Accelerators and Beams*, vol. 3, no. 9, p. 092 001, 2000.

[9] A. C. M. Austin and C. D. Sarris, "Efficient Analysis of Geometrical Uncertainty in the FDTD Method Using Polynomial Chaos With Application to Microwave Circuits", *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 12, pp. 4293–4301, 2013.

[10] I. Babuška, F. Nobile, and R. Tempone, "A stochastic collocation method for elliptic partial differential equations with random input data", *SIAM Review*, vol. 2, pp. 317–355, 2010.

[11] I. Babuška, R. Tempone, and G. Zouraris, "Galerkin finite element approximations of stochastic elliptic partial differential equations", *SIAM Journal on Numerical Analysis*, vol. 42, no. 2, pp. 800–825, 2004.

[12] J. Bäck, F. Nobile, L. Tamellini, and R. Tempone, "Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: a numerical comparison", *Spectral and High Order Methods for Partial Differential Equations: Selected papers from the ICOSAHOM '09 conference, June 22-26, Trondheim, Norway*, pp. 43–62, 2011.

[13] B. Bandlow, R. Schuhmann, G. Lubkowski, and T. Weiland, "Analysis of single-cell modeling of periodic metamaterial structures", *IEEE Transactions on Magnetics*, vol. 44, no. 6, pp. 1662–1665, 2008.

[14] V. Barthelmann, E. Novak, and K. Ritter, "High dimensional polynomial interpolation on sparse grids", *Advances in Computational Mathematics*, vol. 12, no. 4, pp. 273–288, 2000.

[15] S. Basak, S. Petit, J. Bect, and E. Vazquez, "Numerical issues in maximum likelihood parameter estimation for Gaussian process regression", *arXiv preprint arXiv:2101.09747*, 2021.

[16] R. Becker and R. Rannacher, "An optimal control approach to a posteriori error estimation in finite element methods", *Acta Numerica*, vol. 10, pp. 1–102, 2001.

[17] J. Bect, L. Li, and E. Vazquez, "Bayesian subset simulation", *SIAM/ASA Journal on Uncertainty Quantification*, vol. 5, no. 1, pp. 762–786, 2017.

[18] R. E. Bellman, *Dynamic Programming*. Princeton University Press, 1957.

[19] S. Belomestnykh and V. Shemelin, "High-β cavity design–a tutorial", in *The Proceedings of 12th International Workshop on RF Superconductivity (SRF2005)*, 2005.

[20] P. Benner, S. Gugercin, and K. Willcox, "A survey of projection-based model reduction methods for parametric dynamical systems", *SIAM Review*, vol. 57, no. 4, pp. 483–531, 2015.

[21] J. Berrut and L. Trefethen, "Barycentric Lagrange interpolation", *SIAM Review*, vol. 46, no. 3, pp. 501–517, 2004.

[22] A. K. Bhattacharyya, *Phased array antennas: Floquet analysis, synthesis, BFNs and active array systems*. Hoboken: John Wiley & Sons, 2006, vol. 179.

[23] G. Blatman and B. Sudret, "Adaptive sparse polynomial chaos expansion based on least angle regression", *Journal of Computational Physics*, vol. 230, no. 6, pp. 2345–2367, 2011.

[24] D. Boffi, "Finite element approximation of eigenvalue problems", *Acta Numerica*, vol. 19, pp. 1–120, 2010.

[25] R. Boloix-Tortosa, J. J. Murillo-Fuentes, F. J. Payán-Somet, and F. Pérez-Cruz, "Complex Gaussian processes for regression", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 11, pp. 5499–5511, 2018.

[26] R. Boloix-Tortosa, J. J. Murillo-Fuentes, I. Santos, and F. Pérez-Cruz, "Widely linear complex-valued kernel methods for regression", *IEEE Transactions on Signal Processing*, vol. 65, no. 19, pp. 5240–5248, 2017.

[27] F. Bonizzoni, F. Nobile, I. Perugia, and D. Pradovera, "Least-squares Padé approximation of parametric and stochastic Helmholtz maps", *Advances in Computational Mathematics*, vol. 46, no. 3, pp. 1–28, 2020.

[28] J. Bonyo, "Reproducing kernels for Hardy and Bergman spaces of the upper half plane", *Communications in Advanced Mathematical Sciences*, vol. 3, no. 1, pp. 13–23, 2020.

[29] C. de Boor, *A Practical Guide to Splines*, rev., ser. Applied Mathematical Sciences. New York: Springer, 2001, vol. 27.

[30] E. Borgonovo, "A new uncertainty importance measure", *Reliability Engineering & System Safety*, vol. 92, no. 6, pp. 771–784, 2007.

[31] E. Borgonovo and E. Plischke, “Sensitivity analysis: a review of recent advances”, *European Journal of Operational Research*, vol. 248, no. 3, pp. 869–887, 2016.

[32] A. Bossavit, “Whitney forms: A class of finite elements for three-dimensional computations in electromagnetism”, *IEE Proceedings A - Physical Science, Measurement and Instrumentation, Management and Education - Reviews*, vol. 135, no. 8, pp. 493–500, 1988.

[33] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*, 2nd ed. Dover Publications, 2001.

[34] K. Brackebusch, T. Galek, and U. van Rienen, “Automated mode recognition algorithm for accelerating cavities”, in *Proceedings of IPAC 2014, Dresden, Germany*, 2014, pp. 409–411.

[35] V. Braibant and C. Fleury, “Shape optimal design using B-splines”, *Computer Methods in Applied Mechanics and Engineering*, vol. 44, no. 3, pp. 247–267, 1984.

[36] K. Breitung, “Asymptotic approximations for multinormal integrals”, *Journal of Engineering Mechanics*, vol. 110, no. 3, pp. 357–366, 1984.

[37] A. Buffa and R. Hiptmair, “Galerkin boundary element methods for electromagnetic scattering”, in *Topics in computational wave propagation*, M. Ainsworth, P. Davies, D. Duncan, B. Rynne, and P. Martin, Eds., Springer, 2003, pp. 83–124.

[38] A. Buffa, G. Sangalli, and R. Vázquez, “Isogeometric analysis in electromagnetics: B-splines approximation”, *Computer Methods in Applied Mechanics and Engineering*, vol. 199, pp. 1143–1152, 2010.

[39] M. D. Buhmann, “Radial basis functions”, *Acta Numerica*, vol. 9, pp. 1–38, 2000.

[40] H.-J. Bungartz and M. Griebel, “Sparse grids”, *Acta Numerica*, vol. 13, pp. 147–269, 2004.

[41] K. P. Burnham and D. R. Anderson, “Multimodel inference: understanding AIC and BIC in model selection”, *Sociological Methods & Research*, vol. 33, no. 2, pp. 261–304, 2004.

[42] K. Busch, G. Von Freymann, S. Linden, S. Mingaleev, L. Tkeshelashvili, and M. Wegener, “Periodic nanostructures for photonics”, *Physics reports*, vol. 444, no. 3-6, pp. 101–202, 2007.

[43] T. Butler, P. G. Constantine, and T. M. Wildey, “A posteriori error analysis of parameterized linear systems using spectral methods”, *SIAM Journal on Matrix Analysis and Applications*, vol. 33, no. 1, pp. 195–209, 2012.

[44] T. Butler, C. Dawson, and T. M. Wildey, “Propagation of uncertainties using improved surrogate models”, *SIAM/ASA Journal on Uncertainty Quantification*, vol. 1, no. 1, pp. 164–191, 2013.

[45] T. Butler and T. Wildey, “Utilizing adjoint-based error estimates for surrogate models to accurately predict probabilities of events”, *International Journal for Uncertainty Quantification*, vol. 8, no. 2, pp. 143–159, 2018.

[46] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, “A limited memory algorithm for bound constrained optimization”, *SIAM Journal on Scientific Computing*, vol. 16, no. 5, pp. 1190–1208, 1995.

[47] R. E. Caflisch, “Monte Carlo and quasi-Monte Carlo methods”, *Acta Numerica*, vol. 7, pp. 1–49, 1998.

[48] W. Cai and V. M. Shalaev, *Optical Metamaterials: Fundamentals and Applications*. New York: Springer, 2010.

[49] R. H. Cameron and W. T. Martin, “The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals”, *Annals of Mathematics*, vol. 48, no. 2, pp. 385–392, 1947.

[50] C. Chauvière, J. S. Hesthaven, and L. Lurati, “Computational modeling of uncertainty in time-domain electromagnetics”, *SIAM Journal on Scientific Computing*, vol. 28, no. 2, pp. 751–775, 2006.

[51] H.-T. Chen, A. J. Taylor, and N. Yu, "A review of metasurfaces: physics and applications", *Reports on Progress in Physics*, vol. 79, no. 7, p. 076 401, 2016.

[52] A. Chkifa, A. Cohen, and C. Schwab, "Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs", *Journal de Mathématiques Pures et Appliquées*, vol. 103, no. 2, pp. 400–428, 2015.

[53] ——, "High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs", *Foundations of Computational Mathematics*, vol. 14, no. 4, pp. 601–633, 2014.

[54] C. K. Choi and H. H. Yoo, "Uncertainty analysis of nonlinear systems employing the first-order reliability method", *Journal of Mechanical Science and Technology*, vol. 26, no. 1, pp. 39–44, 2012.

[55] S. Clénet, "Uncertainty quantification in computational electromagnetics: The stochastic approach", *International Compumag Society Newsletter*, vol. 13, pp. 3–13, 2013.

[56] C. W. Clenshaw and A. R. Curtis, "A method for numerical integration on an automatic computer", *Numerische Mathematik*, vol. 2, no. 1, pp. 197–205, 1960.

[57] P. G. Constantine, *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*. Philadelphia: Society for Industrial and Applied Mathematics, 2015.

[58] J. Corno, C. de Falco, H. De Gersem, and S. Schöps, "Isogeometric analysis simulation of TESLA cavities under uncertainty", in *Proceedings of the International Conference on Electromagnetics in Advanced Applications (ICEAA) 2015*, R. D. Graglia, Ed., IEEE, 2015, pp. 1508–1511.

[59] ——, "Isogeometric simulation of Lorentz detuning in superconducting accelerator cavities", *Computer Physics Communications*, vol. 201, pp. 1–7, 2016.

[60] J. Corno, N. Georg, S. Gorgi Zadeh, J. Heller, V. Gubarev, T. Roggen, U. Römer, C. Schmidt, S. Schöps, J. Schultz, A. Sulimov, and U. van Rienen, "Uncertainty modeling and analysis of the European X-ray free electron laser cavities manufacturing process", *Nuclear Instruments & Methods in Physics Research Section A – Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 971, p. 164 135, 2020.

[61] CST AG, *Optical Applications with CST Microwave Studio*, `https://www.cst.com/content/events/downloads/euc2012/talk_5-3-1_cst_euc_2012.pdf` (Accessed: 2018-03-12), 2012.

[62] M. D'Elia and M. Gunzburger, "Coarse-grid sampling interpolatory methods for approximating Gaussian random fields", *SIAM/ASA Journal on Uncertainty Quantification*, vol. 1, no. 1, pp. 270–296, 2013.

[63] S. Da Veiga, "Global sensitivity analysis with dependence measures", *Journal of Statistical Computation and Simulation*, vol. 85, no. 7, pp. 1283–1305, 2015.

[64] R. L. Dailey, "Eigenvector derivatives with repeated eigenvalues", *AIAA Journal*, vol. 27, no. 4, pp. 486–491, 1989.

[65] Dassault Systèmes, *CST STUDIO SUITE 2019*, Darmstadt, 2019.

[66] B. J. Debusschere, H. N. Najm, P. P. Pébay, O. M. Knio, R. G. Ghanem, and O. P. Le Maître, "Numerical challenges in the use of polynomial chaos representations for stochastic processes", *SIAM Journal on Scientific Computing*, vol. 26, no. 2, pp. 698–719, 2004.

[67] D. Deschrijver, M. Mrozowski, T. Dhaene, and D. De Zutter, "Macromodeling of multiport systems using a fast implementation of the vector fitting method", *IEEE Microwave and Wireless Components Letters*, vol. 18, no. 6, pp. 383–385, 2008.

[68] P. Deuflhard, *Newton methods for nonlinear problems: affine invariance and adaptive algorithms*. Berlin: Springer, 2004.

[69] J. Dölz, "A higher order perturbation approach for electromagnetic scattering problems on random domains", *SIAM/ASA Journal on Uncertainty Quantification*, vol. 8, no. 2, pp. 748–774, 2019.

[70] T. A. Driscoll, N. Hale, and L. N. Trefethen, *Chebfun guide*, 2014.

[71] D. Edwards, "TESLA test facility linac – design report", *DESY Print March*, pp. 95–01, 1995.

[72] V. A. Epanechnikov, "Non-parametric estimation of a multivariate probability density", *Theory of Probability & Its Applications*, vol. 14, no. 1, pp. 153–158, 1969.

[73] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, "Introduction to adaptive methods for differential equations", *Acta Numerica*, vol. 4, p. 105, 1995.

[74] O. G. Ernst, A. Mugler, H.-J. Starkloff, and E. Ullmann, "On the convergence of generalized polynomial chaos expansions", *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 46, no. 2, pp. 317–339, 2012.

[75] D. Estep, "A short course on duality, adjoint operators, Green's functions, and a posteriori error analysis", *Lecture Notes*, 2004.

[76] C. de Falco, A. Reali, and R. Vázquez, "GeoPDEs: A research tool for isogeometric analysis of PDEs", *Advances in Engineering Software*, vol. 42, pp. 1020–1034, 2011.

[77] O. Farle, V. Hill, P. Nickel, and R. Dyczij-Edlinger, "Multivariate finite element model order reduction for permittivity or permeability estimation", *IEEE Transactions on Magnetics*, vol. 42, no. 4, pp. 623–626, 2006.

[78] J. Feinberg and H. P. Langtangen, "Chaospy: an open source tool for designing methods of uncertainty quantification", *Journal of Computational Science*, vol. 11, pp. 46–57, 2015.

[79] D. R. Fokkema, G. L. Sleijpen, and H. A. Van der Vorst, "Jacobi–Davidson style QR and QZ algorithms for the reduction of matrix pencils", *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 94–125, 1998.

[80] J. N. Fuhg, A. Fau, and U. Nackenhorst, "State-of-the-art and comparative review of adaptive sampling methods for kriging", *Archives of Computational Methods in Engineering*, pp. 1–59, 2020.

[81] M. Fuhrländer, N. Georg, U. Römer, and S. Schöps, "Yield optimization based on adaptive Newton-Monte Carlo and polynomial surrogates", *International Journal for Uncertainty Quantification*, vol. 10, no. 4, pp. 351–373, 2020.

[82] M. Fuhrländer and S. Schöps, "Yield optimization using hybrid Gaussian process regression and a genetic multi-objective approach", *Advances in Radio Science*, 2021. arXiv: 2010.04028, forthcoming.

[83] R. Gaignaire, S. Clénet, B. Sudret, and O. Moreau, "3-D spectral stochastic finite element method in electromagnetism", *IEEE Transactions on Magnetics*, vol. 43, no. 4, pp. 1209–1212, 2007.

[84] P. Gall, V. Gubarev, S. Yasar, A. Sulimov, and J. Iversen, "A database for the European XFEL", in *Proceedings of The 16th International Conference on RF Superconductivity, SRF 2013, Paris, France*, 2013, pp. 205–7.

[85] L. Gallimard, "Adaptive reduced basis strategy for rare-event simulations", *International Journal for Numerical Methods in Engineering*, vol. 120, no. 3, pp. 283–302, 2019.

[86] N. Georg, J. Bect, U. Römer, and S. Schöps, "Rational kernel-based interpolation for complex-valued frequency response functions", *In preparation.*, 2021.

[87] N. Georg, C. Lehmann, U. Römer, and R. Schuhmann, “Multifidelity uncertainty quantification for optical structures”, in *Scientific Computing in Electrical Engineering SCEE 2020*, M. van Beurden, N. Budko, and W. Schilders, Eds., ser. Mathematics in Industry, Berlin: Springer, 2021.

[88] N. Georg, W. Ackermann, J. Corno, and S. Schöps, “Uncertainty quantification for Maxwell’s eigenproblem using isogeometric analysis and mode tracking”, *Computer Methods in Applied Mechanics and Engineering*, vol. 350, pp. 228–244, 2019.

[89] N. Georg, D. Loukrezis, U. Römer, and S. Schöps, “Enhanced adaptive surrogate models with applications in uncertainty quantification for nanoplasmonics”, *International Journal for Uncertainty Quantification*, vol. 10, no. 2, pp. 165–193, 2020.

[90] N. Georg and U. Römer, “Conformally mapped polynomial chaos expansions for Maxwell’s source problem with random input data”, *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 33, no. 6, e2776, 2020.

[91] T. Gerstner and M. Griebel, “Dimension-adaptive tensor-product quadrature”, *Computing*, vol. 71, no. 1, pp. 65–87, 2003.

[92] C. Geuzaine and J.-F. Remacle, *Gmsh reference manual*, 2009.

[93] R. G. Ghanem and P. D. Spanos, *Stochastic finite elements: a spectral approach*, rev., ser. Civil, Mechanical and Other Engineering Series. Dover Publications, 2003.

[94] M. B. Giles, “Multilevel Monte Carlo methods”, *Acta Numerica*, vol. 24, pp. 259–328, 2015.

[95] H.-W. Glock, K. Rothemund, and U. van Rienen, “CSC – A procedure for coupled S-parameter calculations”, *IEEE Transactions on Magnetics*, vol. 38, no. 2, pp. 1173–1176, 2002.

[96] S. Gorgi Zadeh, T. Flisgen, and U. van Rienen, “Eigenmode computation of cavities with perturbed geometry using matrix perturbation methods applied on generalized eigenvalue problems”, *Journal of Computational Physics*, vol. 364, pp. 347–364, 2018.

[97] H. E. Graeb, *Analog Design Centering and Sizing*. Dordrecht: Springer, 2007.

[98] B. Gustavsen, “Improving the pole relocating properties of vector fitting”, *IEEE Transactions on Power Delivery*, vol. 21, no. 3, pp. 1587–1592, 2006.

[99] B. Gustavsen and A. Semlyen, “Rational approximation of frequency domain responses by vector fitting”, *IEEE Transactions on Power Delivery*, vol. 14, no. 3, pp. 1052–1061, 1999.

[100] W. Hackbusch, “Numerical Tensor Calculus”, *Acta Numerica*, vol. 23, pp. 651–742, 2014.

[101] N. Hale, “On the use of conformal maps to speed up numerical computations”, Ph.D. dissertation, Oxford University, 2009.

[102] N. Hale and L. N. Trefethen, “New quadrature formulas from conformal maps”, vol. 46, no. 2, pp. 930–948, 2008.

[103] J. M. Hammersley and D. C. Handscomb, *Monte Carlo methods*. London: Methuen, 1964.

[104] W. Hendrickx and T. Dhaene, “A discussion of "Rational approximation of frequency domain responses by vector fitting"”, *IEEE Transactions on Power Systems*, vol. 21, no. 1, pp. 441–443, 2006.

[105] R. Hiptmair, “Finite elements in computational electromagnetism”, *Acta Numerica*, vol. 11, pp. 237–339, 2002.

[106] T. Homma and A. Saltelli, “Importance measures in global sensitivity analysis of nonlinear models”, *Reliability Engineering & System Safety*, vol. 52, no. 1, pp. 1–17, 1996.

[107] M. Hornikx, M. Kaltenbacher, and S. Marburg, "A platform for benchmark cases in computational acoustics", *Acta Acustica United With Acustica*, vol. 101, no. 4, pp. 811–820, 2015.

[108] T. J. R. Hughes, J. A. Cottrell, and Y. Bazilevs, "Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement", *Computer Methods in Applied Mechanics and Engineering*, vol. 194, pp. 4135–4195, 2005.

[109] J. P. Hugonin and P. Lalanne, *Reticolo software for grating analysis*, www.lp2n.institutoptique.fr, Online; accessed 29 January 2014.

[110] J. D. Jackson, *Classical Electrodynamics*, 3rd. New York, NY, USA: Wiley & Sons, 1998.

[111] J. D. Jakeman and T. Wildey, "Enhancing adaptive sparse grid approximations and improving refinement strategies using adjoint-based a posteriori error estimates", *Journal of Computational Physics*, vol. 280, pp. 54–71, 2015.

[112] R. Jankoski, U. Römer, and S. Schöps, "Stochastic modeling of magnetic hysteretic properties by using multivariate random fields", *International Journal for Uncertainty Quantification*, vol. 9, no. 1, pp. 85–102, 2019.

[113] P. Jantsch and C. Webster, "Sparse grid quadrature rules based on conformal mappings", in *Sparse Grids and Applications-Miami 2016*, Springer, 2018, pp. 117–134.

[114] Y. Jiang, S. Pillai, and M. A. Green, "Re-evaluation of literature values of silver optical constants", *Optics Express*, vol. 23, no. 3, pp. 2133–2144, 2015.

[115] J.-M. Jin, *The Finite Element Method in Electromagnetics*. Hoboken: John Wiley & Sons, 2015.

[116] R. Jin, W. Chen, and A. Sudjianto, "On sequential sampling for global metamodeling in engineering design", in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 36223, 2002, pp. 539–548.

[117] P. B. Johnson and R.-W. Christy, "Optical constants of the noble metals", *Physical Review B*, vol. 6, no. 12, p. 4370, 1972.

[118] P. Jorkowski and R. Schuhmann, "Mode tracking for parametrized eigenvalue problems in computational electromagnetics", in *2018 International Applied Computational Electromagnetics Society (ACES) Symposium*, B. M. Notaros, Ed., Denver, Colorado: IEEE, 2018, p. 17 803 096.

[119] M. Kanagawa, P. Hennig, D. Sejdinovic, and B. K. Sriperumbudur, "Gaussian processes and kernel methods: A review on connections and equivalences", *arXiv preprint arXiv:1807.02582*, 2018.

[120] B. Kim, Y. Lee, and D.-H. Choi, "Construction of the radial basis function based on a sequential sampling approach using cross-validation", *Journal of mechanical science and technology*, vol. 23, no. 12, pp. 3357–3365, 2009.

[121] P. Koosis, *Introduction to Hp Spaces*, 2nd ed., ser. Cambridge Tracts in Mathematics. Cambridge University Press, 1999.

[122] D. Kosloff and H. Tal-Ezer, "A modified Chebyshev pseudospectral method with an $\mathcal{O}(N-1)$ time step restriction", *Journal of Computational Physics*, vol. 104, no. 2, pp. 457–469, 1993.

[123] G. Kreps, J. Sekutowicz, and D. Proch, "Tuning of the TESLA superconducting cavities and the measurement of higher order mode damping", in *Proceedings of the XVth Conference on Charged Particle Accelerators, Dubna, Russia*, 1996.

[124] K. S. Kunz and R. J. Luebbers, *The Finite Difference Time Domain Method for Electromagnetics*. Boca Raton: CRC Press, 1993.

[125] J. Lataire and T. Chen, "Transfer function and transient estimation by Gaussian process regression in the frequency domain", *Automatica*, vol. 72, pp. 217–229, 2016.

[126] O. Le Maitre and O. Knio, *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*. Springer Netherlands, 2010.

[127] R. Lebrun and A. Dutfoy, "Do Rosenblatt and Nataf isoprobabilistic transformations really differ?", *Probabilistic Engineering Mechanics*, vol. 24, no. 4, pp. 577–584, 2009.

[128] R. B. Lehoucq, D. C. Sorensen, and C. Yang, *Arpack User's Guide: Solution of Large-Scale Eigenvalue Problems With Implicityly Restorted Arnoldi Methods (Software, Environments, Tools)*. Philadelphia: Society for Industrial and Applied Mathematics (SIAM), 1998.

[129] R. B. Lehoucq and D. C. Sorensen, "Deflation techniques for an implicitly restarted Arnoldi iteration", *SIAM Journal on Matrix Analysis and Applications*, vol. 17, no. 4, pp. 789–821, 1996.

[130] C. Lemieux, *Monte Carlo and quasi-Monte Carlo sampling*, ser. Springer Series in Statistics. Dordrecht: Springer, 2009.

[131] J. Li and D. Xiu, "Evaluation of failure probability via surrogate models", *Journal of Computational Physics*, vol. 229, no. 23, pp. 8966–8980, 2010.

[132] S. Linden, C. Enkrich, M. Wegener, J. Zhou, T. Koschny, and C. M. Soukoulis, "Magnetic response of metamaterials at 100 terahertz", *Science*, vol. 306, no. 5700, pp. 1351–1353, 2004.

[133] A. Litvinenko, A. C. Yucel, H. Bagci, J. Oppelstrup, E. Michielssen, and R. Tempone, "Computation of electromagnetic fields scattered from objects with uncertain shapes using multilevel Monte Carlo method", *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, vol. 4, pp. 37–50, 2019.

[134] Z. Liu, D. Lesselier, B. Sudret, and J. Wiart, "Surrogate modeling based on resampled polynomial chaos expansions", *Reliability Engineering & System Safety*, vol. 202, p. 107 008, 2020.

[135] D. Loukrezis, "Adaptive approximations for high-dimensional uncertainty quantification in stochastic parametric electromagnetic field simulations", Ph.D. dissertation, Technische Universität Darmstadt, 2019.

[136] D. Loukrezis, U. Römer, T. Casper, S. Schöps, and H. De Gersem, "Application of high dimensional uncertainty quantification methods to electrothermal field problems", in *10th International Symposium on Electric and Magnetic Fields (EMF 2016)*, Lyon, France, 2016.

[137] D. Loukrezis, U. Römer, and H. De Gersem, "Assessing the performance of Leja and Clenshaw-Curtis collocation for computational electromagnetics with random input data", *International Journal for Uncertainty Quantification*, vol. 9, no. 1, pp. 33–57, 2019.

[138] S. H. Lui, H. B. Keller, and T. W. C. Kwok, "Homotopy method for the large, sparse, real nonsymmetric eigenvalue problem", *SIAM Journal on Matrix Analysis and Applications*, vol. 18, no. 2, pp. 312–333, 1997.

[139] S. A. Maier, *Plasmonics: fundamentals and applications*. New York: Springer Science & Business Media, 2007.

[140] H. Matthies and A. Keese, "Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations", *Computer Methods in Applied Mechanics and Engineering*, vol. 194, no. 12-16, pp. 1295–1331, 2005.

[141] J. C. Maxwell, "A dynamical theory of the electromagnetic field", *Royal Society Transactions*, vol. CLV, pp. 459–512, 1864.

[142] N. Meinshausen and P. Bühlmann, "Stability selection", *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 72, no. 4, pp. 417–473, 2010.

[143] G. Migliorati, F. Nobile, E. von Schwerin, and R. Tempone, "Analysis of discrete L2 projection on polynomial spaces with random evaluations", *Foundations of Computational Mathematics*, vol. 14, no. 3, pp. 419–456, 2014.

[144] P. Monk, *Finite Element Methods for Maxwell's Equations*. Oxford: Oxford University Press, 2003.

[145] S. A. Murphy and A. W. Van der Vaart, "On profile likelihood", *Journal of the American Statistical Association*, vol. 95, no. 450, pp. 449–465, 2000.

[146] D. G. Myakishev and V. P. Yakovlev, "An interactive code SUPERLANS for evaluation of RF-cavities and acceleration structures", in *Conference Record of the 1991 IEEE Particle Accelerator Conference*, 1991, 3002–3004 vol.5.

[147] Y. Nakatsukasa, O. Sète, and L. N. Trefethen, "The AAA algorithm for rational approximation", *SIAM Journal on Scientific Computing*, vol. 40, no. 3, A1494–A1522, 2018.

[148] A. Narayan, C. Gittelson, and D. Xiu, "A stochastic collocation algorithm with multifidelity models", *SIAM Journal on Scientific Computing*, vol. 36, no. 2, A495–A521, 2014.

[149] A. Narayan and J. D. Jakeman, "Adaptive Leja sparse grid constructions for stochastic collocation and high-dimensional approximation", *SIAM Journal on Scientific Computing*, vol. 36, no. 6, 2014.

[150] J. C. Nédélec, "Mixed finite elements in R^3", *Numerische Mathematik*, vol. 35, no. 3, pp. 315–341, 1980.

[151] J. Nitzler, J. Biehler, N. Fehn, P.-S. Koutsourelakis, and W. A. Wall, "A generalized probabilistic learning approach for multi-fidelity uncertainty propagation in complex physical simulations", *arXiv preprint arXiv:2001.02892*, 2020.

[152] F. Nobile, R. Tempone, and C. G. Webster, "An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data", *SIAM Journal on Numerical Analysis*, vol. 46, no. 5, pp. 2411–2442, 2008.

[153] F. Nobile and D. Pradovera, "Non-intrusive double-greedy parametric model reduction by interpolation of frequency-domain rational surrogates", *arXiv preprint arXiv:2008.10864*, 2020.

[154] F. Nobile and R. Tempone, "Analysis and implementation issues for the numerical approximation of parabolic equations with random coefficients", *International Journal for Numerical Methods in Engineering*, vol. 80, no. 6-7, pp. 979–1006, 2009.

[155] F. Nobile, R. Tempone, and C. G. Webster, "A sparse grid stochastic collocation method for partial differential equations with random input data", *SIAM Journal on Numerical Analysis*, vol. 46, no. 5, pp. 2309–2345, 2008.

[156] J. Nocedal and S. J. Wright, *Numerical optimization*, 2nd ed., ser. Springer Series in Operations Research and Financial Engineering. New York: Springer, 2006.

[157] E. Novak and K. Ritter, "High dimensional integration of smooth functions over cubes", *Numerische Mathematik*, vol. 75, no. 1, pp. 79–97, 1996.

[158] I. Ojalvo, "Gradients for large structural models with repeated frequencies", SAE Technical Paper, Tech. Rep., 1986.

[159] H. Padamsee, "Design topics for superconducting RF cavities and ancillaries", *CERN Yellow Report CERN-2014-005*, pp. 141–169, 2015.

[160] H. Padamsee, J. Knobloch, and T. Hays, *RF Superconductivity for Accelerators*. John Wiley & sons., 1998, p. 544.

[161] I. Papaioannou, M. Daub, M. Drieschner, F. Duddeck, M. Ehre, L. Eichner, M. Eigel, M. Götz, W. Graf, L. Grasedyck, R. Gruhlke, D. Hömberg, M. Kaliske, D. Moser, Y. Petryna, and D. Straub, "Assessment and design of an engineering structure with polymorphic uncertainty quantification", *GAMM*, vol. 42, no. 2, e201900009, 2019.

[162] I. Papaioannou, M. Ehre, and D. Straub, "PLS-based adaptation for efficient PCE representation in high dimensions", *Journal of Computational Physics*, vol. 387, pp. 186–204, 2019.

[163] V. I. Paulsen and M. Raghupathi, *An introduction to the theory of reproducing kernel Hilbert spaces*. Cambridge University Press, 2016, vol. 152.

[164] B. Peherstorfer, K. Willcox, and M. Gunzburger, "Optimal model management for multifidelity Monte Carlo estimation", *SIAM Journal on Scientific Computing*, vol. 38, no. 5, A3163–A3194, 2016.

[165] ——, "Survey of multifidelity methods in uncertainty propagation, inference, and optimization", *SIAM Review*, vol. 60, no. 3, pp. 550–591, 2018.

[166] L. Piegl and W. Tiller, *The NURBS Book*, 2nd ed. Springer, 1997.

[167] M. J. Preiner, K. T. Shimizu, J. S. White, and N. A. Melosh, "Efficient optical coupling into metal-insulator-metal plasmon modes with subwavelength diffraction gratings", *Applied Physics Letters*, vol. 92, no. 11, p. 113 109, 2008.

[168] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge: The MIT Press, 2006.

[169] A. C. Rodriguez and S. Gugercin, "The p-AAA algorithm for data driven modeling of parametric dynamical systems", *arXiv preprint arXiv:2003.06536*, 2020.

[170] U. Römer, M. Bollhöfer, H. Sreekumar, C. Blech, and S. Langer, "An adaptive sparse grid rational Arnoldi method for uncertainty quantification of dynamical systems in the frequency domain ", *International Journal for Numerical Methods in Engineering*, 2021.

[171] U. Römer, T. Casper, R. Pulch, and E. J. W. ter Maten, "Algorithm for computation of failure probabilities and ageing effects", nanoCOPS FP7-ICT-2013-11/619166, Report D2.6, 2015.

[172] C. J. Roy and W. L. Oberkampf, "A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing", *Computer Methods in Applied Mechanics and Engineering*, vol. 200, no. 25, pp. 2131–2144, 2011.

[173] A. Saltelli, "Making best use of model evaluations to compute sensitivity indices", *Computer Physics Communications*, vol. 145, no. 2, pp. 280–297, 2002.

[174] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.

[175] A. Saltelli, S. Tarantola, and F. Campolongo, "Sensitivity anaysis as an ingredient of modeling", *Statistical Science*, vol. 15, no. 4, pp. 377–395, 2000.

[176] W. Sauerbrei and M. Schumacher, "A bootstrap resampling procedure for model building: application to the Cox regression model", *Statistics in Medicine*, vol. 11, no. 16, pp. 2093–2109, 1992.

[177] S. A. Sauter and C. Schwab, *Boundary Element Methods*. Berlin-Heidelberg: Springer, 2011.

[178] L. Scarabosio, "Shape uncertainty quantification for scattering transmission problems", Ph.D. dissertation, ETH Zurich, 2016.

[179] B. Schieche, “Unsteady adaptive stochastic collocation methods on sparse grids”, Ph.D. dissertation, Technische Universität Darmstadt, Darmstadt, 2012.

[180] C. Schmidt, T. Flisgen, J. Heller, and U. van Rienen, “Comparison of techniques for uncertainty quantification of superconducting radio frequency cavities”, in *Proceedings of the International Conference on Electromagnetics in Advanced Applications (ICEAA) 2014*, R. D. Graglia, Ed., IEEE, 2014, pp. 117–120.

[181] N. Schmitt, N. Georg, G. Brière, D. Loukrezis, S. Héron, S. Lanteri, C. Klitis, M. Sorel, U. Römer, H. De Gersem, S. Vézian, and P. Genevet, “Optimization and uncertainty quantification of gradient index metasurfaces”, *Optical Materials Express*, vol. 9, no. 2, pp. 892–910, 2019.

[182] N. Schmitt, C. Scheid, S. Lanteri, A. Moreau, and J. Viquerat, “A DGTD method for the numerical modeling of the interaction of light with nanometer scale metallic structures taking into account non-local dispersion effects”, *Journal of Computational Physics*, vol. 316, pp. 396–415, 2016.

[183] R. Schuhmann, M. Clemens, P. Thoma, and T. Weiland, “Frequency and time domain computations of S-parameters using the finite integration technique”, in *Proceedings of the 12th Annual Review of Progress in Applied Computational Electromagnetics*, 1996, pp. 1295–1302.

[184] C. Schwab and R. A. Todor, “Karhunen–Loève approximation of random fields by generalized fast multipole methods”, *Journal of Computational Physics*, vol. 217, no. 1, pp. 100–122, 2006.

[185] G. Silva-Oelker, R. Aylwin, C. Jerez-Hanckes, and P. Fay, “Quantifying the impact of random surface perturbations on reflective gratings”, *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 2, pp. 838–847, 2017.

[186] W. Singer, A. Brinkmann, R. Brinkmann, J. Iversen, A. Matheisen, W. D. Moeller, A. Navitski, D. Reschke, J. Schaffran, A. Sulimov, N. Walker, and H. Weise, “Production of superconducting 1.3-GHz cavities for the European X-ray Free Electron Laser”, *Physical Review Accelerators and Beams*, vol. 19, p. 092 001, 9 2016.

[187] G. L. G. Sleijpen and H. A. van der Vorst, “A Jacobi-Davidson iteration method for linear eigenvalue problems”, *SIAM Journal on Matrix Analysis and Applications*, vol. 17, no. 2, pp. 401–425, 1996.

[188] S. A. Smolyak, “Quadrature and interpolation formulas for tensor products of certain classes of functions”, *Doklady Akademii Nauk SSSR*, vol. 4, pp. 240–243, 1963.

[189] I. M. Sobol’, “Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates”, *Mathematics and Computers in Simulation*, vol. 55, no. 1, pp. 271–280, 2001, The Second IMACS Seminar on Monte Carlo Methods.

[190] C. Soize and R. Ghanem, “Physical systems with random uncertainties: chaos representations with arbitrary probability measure”, *SIAM Journal on Scientific Computing*, vol. 26, no. 2, pp. 395–410, 2004.

[191] A. M. Stuart, “Inverse problems: a Bayesian perspective”, *Acta Numerica*, vol. 19, pp. 451–559, 2010.

[192] A. Sulimov, “RF analysis of equator welding stability for the European XFEL cavities”, in *Proceedings of The 17th International Conference on RF Superconductivity, SRF 2015, Whistler, BC, Canada*, 2015.

[193] A. Sulimov, G. Kreps, and J. Sekutowicz, “Estimation of small geometry deviation for TESLA-shape cavities due to inner surface polishing”, in *Proceedings of The 16th International Conference on RF Superconductivity, SRF 2013, Paris, France*, 2013, p. 537.

[194] A. Sulimov, A. Visentin, M. Giaretta, and A. Gresele, “RF analysis of electropolishing for EXFEL cavities production at Ettore Zanon Spa”, in *Proceedings of The 28th Linear Accelerator Conference, LINAC16, East Lansing, MI, USA*, 2017.

[195] A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain-Method*. Norwood, MA, USA: Artech House, 1995.

[196] A. L. Teckentrup, R. Scheichl, M. B. Giles, and E. Ullmann, "Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients", *Numerische Mathematik*, vol. 125, no. 3, pp. 569–600, 2013.

[197] A. L. Teckentrup, P. Jantsch, C. G. Webster, and M. Gunzburger, "A multilevel stochastic collocation method for partial differential equations with random input data", *SIAM/ASA Journal on Uncertainty Quantification*, vol. 3, no. 1, pp. 1046–1074, 2015.

[198] F. Tisseur and K. Meerbergen, "The quadratic eigenvalue problem", *SIAM Review*, vol. 43, no. 2, pp. 235–286, 2001.

[199] L. N. Trefethen, *Approximation Theory and Approximation Practice*. SIAM, 2013, vol. 128.

[200] P. Tsilifis and R. G. Ghanem, "Reduced Wiener chaos representation of random fields via basis adaptation and projection", *Journal of Computational Physics*, vol. 341, pp. 102–120, 2017.

[201] T. P. Wangler, *RF Linear Accelerators*, ser. Physics Textbook. Wiley, 2008.

[202] T. Weiland, "A discretization method for the solution of Maxwell's equations for six-component fields", *International Journal of Electronics and Communications*, vol. 31, pp. 116–120, 1977.

[203] ——, "Time domain electromagnetic field computation with finite difference methods", *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 9, no. 4, pp. 295–319, 1996.

[204] T.-W. Weng, Z. Zhang, Z. Su, Y. Marzouk, A. Melloni, and L. Daniel, "Uncertainty quantification of silicon photonic devices with correlated and non-Gaussian random parameters", *Optical Materials Express*, vol. 23, no. 4, pp. 4242–4254, 2015.

[205] N. Wiener, "The homogeneous chaos", *American Journal of Mathematics*, vol. 60, no. 4, pp. 897–936, 1938.

[206] N. Wiener and R. Paley, *Fourier transforms in the complex domain*. American Mathematical Society, 1934.

[207] T. Wittig, R. Schuhmann, and T. Weiland, "Model order reduction for large systems in computational electromagnetics", *Linear Algebra and its Applications*, vol. 415, no. 2-3, pp. 499–530, 2006.

[208] *XFEL Cavity Database*, `https://cavity-database.desy.de/`, accessed: 2019-05-16.

[209] L. Xiao, C. Adolphsen, V. Akcelik, A. Kabel, K. Ko, L. Lee, Z. Li, and C. Ng, "Modeling imperfection effects on dipole modes in TESLA cavity", in *IEEE Particle Accelerator Conference (PAC) 2007*, 2007, pp. 2454–2456.

[210] D. Xiu, "Fast numerical methods for stochastic computations: A review", *Communications in Computational Physics*, vol. 5, no. 2-4, pp. 242–272, 2009.

[211] ——, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*. Princeton, NJ, USA: Princeton University Press, 2010.

[212] D. Xiu and J. S. Hesthaven, "High-order collocation methods for differential equations with random inputs", *SIAM Journal on Scientific Computing*, vol. 27, no. 3, pp. 1118–1139, 2005.

[213] D. Xiu and G. Karniadakis, "The Wiener–Askey polynomial chaos for stochastic differential equations", *SIAM Journal on Scientific Computing*, vol. 24, no. 2, pp. 619–644, 2002.

[214] D. Yang and V. Ajjarapu, “Critical eigenvalues tracing for power system analysis via continuation of invariant subspaces and projected Arnoldi method”, *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 324–332, 2007.

[215] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, “Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization”, *ACM Transactions on Mathematical Software*, vol. 23, no. 4, pp. 550–560, 1997.

[216] Y. Zhu and A. C. Cangellaris, *Multigrid finite element methods for electromagnetic field modeling*. Hoboken: John Wiley & Sons, 2006.

[217] H. Ziegelwanger and P. Reiter, “The PAC-MAN model: benchmark case for linear acoustics in computational physics”, *Journal of Computational Physics*, vol. 346, pp. 152–171, 2017.

[218] O. C. Zienkiewicz and J. Z. Zhu, “The superconvergent patch recovery and a posteriori error estimates. Part 1: The recovery technique”, *International Journal for Numerical Methods in Engineering*, vol. 33, no. 7, pp. 1331–1364, 1992.

Acknowledgments

First of all, I would like to thank my supervisors Prof. Ulrich Römer and Prof. Sebastian Schöps for the opportunity to work on this thesis, their ongoing extraordinary support, and excellent supervision. Furthermore, I would like to acknowledge the friendly and encouraging working atmosphere as well as the numerous and fruitful scientific discussions.

Second, I am thankful to the various colleagues and external researchers I collaborated with during my doctoral studies. In this context, I would like to express my gratitude to:

- Prof. Rolf Schuhmann for the cooperation in the German Research Foundation (DFG) project *Simulation of RF and Optical Components with Uncertainty Quantification* and for the advice and suggestions.
- All colleagues from the scientific DFG network *Uncertainty quantification techniques and stochastic models for superconducting radio frequency cavities* for the comprehensive discussions and cooperation. In particular, I would like to thank Prof. Ursula van Rienen for the organization and Dr. Alexey Sulimov for the detailed explanations on the TESLA cavity manufacturing process as well as providing access to the DESY cavity database.
- Dr. Nikolai Schmitt and Dr. Patrice Genevet for organizing the international research cooperation regarding the optimization and uncertainty quantification of the gradient index metasurfaces.
- Prof. Julien Bect for the advice and support on complex Gaussian process regression and the collaboration on the rational kernel-based interpolation method, in particular regarding the functional analysis for complex/real reproducing kernel Hilbert spaces.
- Dr. Dimitrios Loukrezis for various fruitful discussions about uncertainty quantification and the successful cooperation for our joint publications.
- Dr. Jacopo Corno and Dr. Wolfgang Ackermann for the fruitful joint work on uncertainty quantification for accelerator cavities.
- Alexander Krimm and Malena Kellermann for their contributions to the first implementation of the eigenvalue tracking method.
- Mona Fuhrländer for the collaboration on yield estimation and optimization.
- Armin Galetzka for various helpful discussions as well as continuing the work on the multi-element Leja collocation idea.
- Anna Ziegler for pursuing the shape morphing based eigenmode classification idea.
- Prem Ratan Mohan Ram and Julius Schultz for the inspiring discussions on various topics in the UQ group meetings.

I would further like to thank:

- Dr. Zeger Bontinck and Mona Fuhrländer for the successful cooperation in the organization and supervision of the *Erfolgreich CE Studieren I+II* program.
- Carina Schuster, Annette Struckmann, Heike Koch and Achim Wagner for the help with organizational issues at the Centre for CE, IDS and TEMF institute, respectively.
- Melina Merkel and Dr. Dimitrios Loukrezis for being great office mates and creating a very(!) comfortable atmosphere, for instance even organizing an office couch.
- All former and current colleagues of the CEM group, and TEMF in general, for making my time at the university also a great personal experience and a lot of fun, for example during the lunches, the coffee/cake meetings, the (short) breaks at the kicker table, the GSC events, the hiking in Kleinwalsertal, as well as the beers or games at various occasions.

Finally, I am deeply grateful to my girlfriend, my family, and my friends for their continuous support and encouragement.

This work is supported by the DFG grant RO4937/1-1, the Excellence Initiative of the German Federal and State Governments and the Graduate School of Computational Engineering at TU Darmstadt.