universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„The reduced basis method for the Helmholtz problem“

submitted by

Mark Strempel BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 910
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Computational Science
Betreut von / Supervisor:	Univ.-Prof. Ilaria Perugia, PhD
Mitbetreut von / Co-Supervisor:	Dr. Lorenzo Mascotto

Abstract

In this thesis, we discuss the numerical performance of the reduced basis method (RBM) applied to several model problems. The reduced basis method (RBM) is a model order reduction (MOR) method for solving parametrized partial differential equations (PDEs). The RBM can be used in conjunction with the finite element method (FEM). The RBM aims to reduce the dimension of the finite element (FE) space by using finite element (FE) solutions as basis functions for constructing a reduced basis (RB) space. We review the RBM and its theoretical properties such as stability and error bounds. We also introduce the equi-logarithmic and greedy approaches for constructing reduced basis (RB) spaces. We apply the RBM to coercive and noncoercive parametrized model problems. The model problems are instances of the Poisson and Helmholtz equations. We report convergence and error rates of the RBM applied to the model problems. Furthermore, we compare RB spaces constructed using equi-logarithmic, equi-distant, random, and snapshots selected by the greedy algorithm. Moreover, we investigate how the greedy algorithm adapts to changes in parameters. Finally we consider three model problems where the RBM is applied in the context of optimal control and parameter estimation problems. Optimal control and parameter estimation problems are a natural fit for the RBM because of its improved efficiency. The examples include the estimation of a material property from measurements and the optimization of an airplane engine.

Zusammenfassung

Ziel der Arbeit ist es die reduced basis method (RBM) numerisch zu untersuchen. Die RBM ist eine model order reduction (MOR) Methode für das Lösen von parameterabhängigen Differentialgleichungen. Die RBM baut auf bestehenden Verfahren wie z.B. der finite element method (FEM) auf. Das Ziel der RBM ist es einen reduzierten reduced basis (RB) Raum auf Basis von bestehenden finite element (FE)-Lösungen, zu konstruieren. In dem ersten Teil der Arbeit werden die RBM und zwei Ansätze, der equi-logarithmic und der Greedy Ansatz, für das Konstruieren von RB-Räumen vorgestellt. Anschließend wird die RBM auf unterschiedliche Beispielprobleme angewendet. Die Beispielprobleme bestehen aus Beispielen für parameterisierte (coercive) Poisson und (noncoercive) Helmholtz Gleichungen. Die Numerischen Tests umfassen unter anderem Konvergenztests und Tests wie der Greedy Algorithmus auf Änderungen seiner Parameter reagiert. Außerdem werden RB-Räume die mithilfe unterschiedlicher Verfahren konstruiert wurden verglichen. Die Räume wurden durch die equi-logarithmic, equi-distant, random und Greedy Ansätze konstruiert. Abschließend wird die Anwendungen der RBM im Kontext von optimal control und parameter estimation Problemen vorgestellt. Die Anwendungsbeispiele umfassen die Bestimmung eines Materialparameters auf Basis von Messdaten und die Minimierung der Lärmemissionen einer Flugzeugturbine.

Contents

List of Tables

List of Figures

List of Algorithms

Chapter 1

Introduction

In many applications, the mathematical models that are used depend on parameters that characterize the concrete instance of a problem. In their numerical simulation, the fast and accurate evaluation of solutions at many values of these parameters is often required. Model order reduction (MOR) methods allow us to substitute multiple solves of the large dimensional problem with the solution to problems of small dimension. RBMs are an instance of MOR methods, which are constructed from precomputed solutions at few values of the parameters.

In this thesis, we focus on problems governed by partial differential equations (PDEs) discretized by finite element methods (FEMs), and the RBM as MOR.

In particular, we consider elliptic problems, whose variational formulation reads as follows: Given a Hilbert space V, a bilinear form $a : V \times V \to \mathbb{R}$, and a linear form $f : V \to \mathbb{R}$

$$\text{find } u \in V \text{ such that } a(u, v) = f(v) \quad \forall v \in V. \tag{1.1}$$

The FEM consists in restricting problem (1.1) to a finite dimensional space $V_h \subset V$:

$$\text{find } u \in V_h \text{ such that } a(u, v) = f(v) \quad \forall v \in V_h. \tag{1.2}$$

Assume that the problem is parametrized by a parameter $\mu \in \mathbb{P}$ from a parameter space $\mathbb{P}$. Hence, problem (1.1) becomes

$$\text{find } u(\mu) \in V(\mu) \text{ such that } a(\mu, u(\mu), v) = f(\mu, v) \quad \forall v \in V(\mu). \tag{1.3}$$

The solution to problem (1.3) can still be approximated by standard FEMs. Given a particular value $\mu_0 \in \mathbb{P}$ the FEM associated with (1.3) reads

$$\text{find } u(\mu_0) \in V_h(\mu_0) \text{ such that } a(\mu_0, u, v) = f(\mu_0, v) \quad \forall v \in V_h(\mu_0). \tag{1.4}$$

Assume we are interested in approximating solutions for many different values of the parameter. Then, the FEM has to be applied several times. The complexity of solving problem (1.4) depends on the dimension of the space V_h. Usually V_h is a high dimensional space. This makes solving (1.4) for many parameter values unfeasible.

RBMs offer a more efficient way to approximate solutions to problem (1.4). Instead of solving problem (1.4) in a high dimensional space V_h, RBMs reformulate the problem in a lower dimensional space V_N with $dim(V_N) = N << dim(V_h)$. This allows us to approximate solutions to (1.4) for many parameter values. Thus, the RBM reads

$$\text{find } u(\mu) \in V_N(\mu) \text{ such that } a(\mu, u(\mu), v) = f(\mu, v) \quad \forall v \in V_N(\mu). \tag{1.5}$$

The space V_N is constructed using solutions to (1.4) at selected parameter points, the so-called snapshots, μ_i for $i = 1 \dots N$. The snapshots can be computed by greedy algorithms or determined from a priori known distributions, e.g. the equi-logarithmic distribution.

The aim of the thesis is to investigate the numerical performance of the RBM. Noncoercive problems, in particular the Helmholtz equation, are the main focus.

In Chapter 2, we review the foundations of the FEM and of the RBM. We provide an overview of the theoretical properties of the RBM applied to coercive and weakly coercive, or inf-sup stable, problems. We also discuss the motivations behind the RBM and how the structure of (1.4) influences the reducibility of a problem. Moreover, we present an offline-online decomposition, originally introduced in [Pru+01], that allows for an efficient construction of RB spaces, in case the bilinear and linear forms in (1.3) depend affinely on the parameter μ. We discuss two approaches for constructing RB spaces: spaces obtained with equi-logarithmic distributions, introduced in [MPT02a; MPT02b], and the greedy algorithm, originally introduced in [Gre05]. In addition, we review an a-posteriori error estimator, from [Pru+01], that can be used in greedy algorithms.

In Chapter 3, we apply the RBM to coercive problems. We consider two examples. The first one is that of a thermal block that is heated from the inside, and whose shape depends on a parameter μ. The second one is a thermal block that is divided into different subregions with differing conductivity, and a heat source on the bottom boundary. We show that the RBM provides exponential convergence of the approximation error, and that the resulting RB spaces are very small ($N \simeq 10$).

In Chapter 4, the RBM is applied to a noncoercive problem, i.e. the Helmholtz equation. Again, we consider two examples. The first one is that of a plane wave that travels along a fixed direction. The problem is parametrized by the wavenumber k. The second problem is an example for the transmission or reflection of a wave that travels through two fluids with different refractive indices n_1 and n_2. The problem is parameterized by the wavenumber k, and the refractive indices n_1 and n_2.

The numerical experiments show that the approximation error decays exponentially with the dimension of the RB space V_N, albeit much slower than in the coercive case. Furthermore, the resulting RB spaces are larger than in the coercive case, i.e. the dimension N of the reduced space is of the order 10^2.

Besides showing the convergence of the RBM, we investigate certain aspects of the RBM, such as the greedy algorithm, in more detail. Furthermore, we investigate how the generated RB spaces change depending on the wavenumber interval $[k_{min}, k_{max}]$ for which they are constructed. We see that the dimension of RB spaces increases linearly with the interval length $k_{max} - k_{min}$.

In the subsequent experiments, we investigate how the greedy algorithm reacts to different initial snapshots and different choices for the training sets. In both cases, we find that the greedy algorithm is unaffected by them.

In a last experiment, we find that oscillations in solutions affect the resulting RB space. In the case of highly oscillating solutions, the resulting RB spaces are of higher dimension, compared to the case where the solutions do not oscillate.

Chapter 5 is devoted to practical examples on how the RBM can be used to solve parameter estimation or optimal control problems. An optimal control problem consists of a parameterized state problem and an output function. The output function maps a solution of the state problem to an output. The goal of solving an optimal control problem is finding the input parameter for the state problem that leads to a given output.

The input parameter might be material properties or parameters that define the geometry of the problem. The output is given by physical measurements from real world experiments. In order to estimate the material properties from the given output, we look for an input parameter that leads to an output that matches the given output as close as possible. This can be done by optimizing a cost function that measures the distance between the given output and an output derived from a given input parameter. Depending on the optimization procedure, the state problem must be solved several times.

We apply the RBM to three problems discretised by the FEM. The first one is the transmission/reflection problem taken from Chapter 4.

The second problem models the propagation of sound on the door of a car. A loud speaker

mounted on the door emits soundwaves that propagate through the domain. The bottom of the door consists of a damping material of unknown impedance. The aim is to estimate the impedance of the damping material based on measurements at six different points.

The third problem models the sound emitted from from an airplane engine. The engine is enclosed by a damping material of impedance i. The aim is to minimize the noise emitted from the engine by finding the optimal impedance i for the damping material.

In order to test the effectivity of the RBM, we select random inputs and calculate the output using the FEM. Afterwards we try to estimate the original input using the given output and the RBM. We find that the estimates accurately predict the original input. In case of the engine noise problem, no original input needs to be estimated. Instead, we first estimate the optimal impedance using the FEM. Afterwards we estimate the optimal impedance using the RBM and compare the results. We find that the RBM gives results close to the results computed using the FEM.

I would like to thank my supervisors Univ.-Prof. Ilaria Perugia, PhD and Dr. Lorenzo Mascotto for their numerous suggestions and continuing support of my thesis. I also would like to thank Paul Stocker MSc for his help with Ngsolve.

Chapter 2

The reduced basis method

RBMs provide a way to compute effectively FE solutions to parametrized problems. The core idea of the method consists in reducing the high dimensional (FE) discretisation of the function space to a lower dimensional RB-Space.

The goal of this chapter is to give a theoretical introduction of the RBM. In Section 2.1 we start by reviewing some relevant facts about the FEM. Next, we introduce RB spaces in Section 2.2. Sections 2.3 and 2.4 are devoted to the selection of parameter points. Section 2.5 describes the implementation details.

2.1 Review of the finite element method

In this section we review the FEM. For a thorough introduction see e.g. [Qua14]. The FEM is a method for approximating solutions to PDEs in weak formulation.

For example, consider the strong formulation of the Poisson equation, with homogeneous Dirichlet boundary conditions:

$$\begin{aligned} -\Delta u &= f \quad \text{in} \quad \Omega, \\ u &= 0 \quad \text{on} \quad \partial\Omega. \end{aligned}$$

Its weak formulation is obtained by multiplying the equation with a sufficiently smooth test function v and integrating by parts

$$\int_\Omega \nabla u \nabla v = \int_\Omega f v \quad \forall v \in V.$$

In symbols, the problem reads

$$\text{find } u \in V \text{ such that } \quad a(u,v) = f(v) \quad \forall v \in V, \tag{2.1}$$

where

$$V = H_0^1(\Omega), \qquad a(u,v) = \int_\Omega \nabla u \nabla v, \qquad f(v) = \int_\Omega f v.$$

Well posedness, i.e. existence, uniqueness and continuous dependence on the data of problem (2.1), is guaranteed by the Lax-Milgram theorem, see Theorem 1, or Nečas theorem, see Theorem 2.

2.1.1 Coercive problems and the finite element method

The bilinear form $a(u,v)$ is called coercive, with coercivity constant α, if

$$\exists \alpha > 0 \quad \text{such that} \quad a(v,v) \geq \alpha \|v\|_V^2 \quad \forall v \in V. \tag{2.2}$$

It is called continuous, with continuity constant γ, if

$$\exists \gamma > 0 \quad \text{such that} \quad a(u,v) \leq \gamma \left\|u\right\|_V \left\|v\right\|_V \quad \forall u, v : u, v \in V. \tag{2.3}$$

For bilinear forms satisfying both the coercivity and the continuity properties, well-posedness of problem (2.1) is ensured by the Lax-Milgram theorem.

Theorem 1. Lax-Milgram*: Let V be a Hilbert space with the norm $\left\|\cdot\right\|_V$. Let $a(\cdot,\cdot)$ be continuous and coercive with continuity constant γ and coercivity constant α. Let $l(\cdot) : V \to \mathbb{R}$ be a linear functional such that*

$$\exists C > 0 \quad \textit{such that} \quad |l(v)| \leq C \left\|v\right\|_V \quad \forall v \in V.$$

Then, there exists a unique $u \in V$ such that $a(u,v) = l(v)$ for all $v \in V$. Moreover, the solution u satisfies

$$\left\|u\right\|_V \leq \frac{1}{\alpha} \left\|l\right\|_{V'},$$

where

$$\left\|l\right\|_{V'} = \sup_{v \in V, v \neq 0} \frac{|l(v)|}{\left\|v\right\|_V}.$$

A proof can be found in [QV16, Section 5.1.1].
The FEM reads

$$\text{find } u_h \in V_h \text{ such that} \quad a(u_h, v_h) = f(v_h) \quad \forall v_h \in V_h, \tag{2.4}$$

where $V_h \subset V$ is of finite dimension.

Assuming the assumptions of Theorem 1 apply to the continuous problem (2.1), it follows:

1. $a(\cdot,\cdot)$ is continuous and coercive on $V_h \times V_h$, with continuity and coercivity constants γ_h and α_h. Moreover, $\gamma_h \leq \gamma$ and $\alpha_h \geq \alpha$.

2. The Galerkin orthogonality holds true

$$a(u - u_h, v_h) = 0 \quad \forall v_h \in V_h. \tag{2.5}$$

3. Céa's lemma guarantees

$$\left\|u - u_h\right\|_V \leq \frac{\gamma}{\alpha} \inf_{v_h \in V_h} \left\|u - v_h\right\|_V. \tag{2.6}$$

Thus Theorem 1 is valid for problem (2.4). To guarantee convergence of the FEM we require

$$\lim_{h \to 0} \inf_{v_h \in V_h} \left\|v - v_h\right\|_V = 0 \quad \forall v \in V. \tag{2.7}$$

2.1.2 Weakly coercive or inf-sup stable problems

For some problems the assumptions of Theorem 1 are too strong. Notwithstanding, for weakly coercive or inf-sup stable problems, the Nečas theorem provides similar results. Consider a problem more general than in the coercive case (2.1). Given two Hilbert spaces V and W, a bilinear form $a : V \times W \to \mathbb{R}$, and a linear form $f : W \to \mathbb{R}$ the problem reads

$$\text{find } u \in V \text{ such that} \quad a(u, w) = f(w) \quad \forall w \in W. \tag{2.8}$$

We assume that the bilinear form $a(\cdot,\cdot)$ is continuous with continuity constant γ. Further, we require that $a(\cdot,\cdot)$ is inf-sup stable (or weakly coercive):

$$\exists \beta > 0 \quad \text{such that} \quad \inf_{v\in V} \sup_{w\in W} \frac{a(v,w)}{\|v\|_V \|w\|_W} \geq \beta \tag{2.9}$$

and

$$\inf_{w\in W} \sup_{v\in V} \frac{a(v,w)}{\|v\|_V \|w\|_W} > 0. \tag{2.10}$$

Given a bilinear form $a(\cdot,\cdot)$ that is continuous and inf-sup stable the Nečas theorem ensures well-posedness of the problem, see [Neč67].

Theorem 2. Nečas*: Let V and W be two Hilbert spaces, let $a(\cdot,\cdot)$ be a bilinear form $a(\cdot,\cdot) : V \times W \to \mathbb{R}$ that is inf-sup stable and continuous with β and γ being the stability and continuity constants. Then, problem* (2.8) *has a unique solution, which satisfies*

$$\|u\|_V \leq \frac{1}{\beta} \|f\|_{W'}, \tag{2.11}$$

where

$$\|f\|_{W'} = \sup_{w\in W, w\neq 0} \frac{|f(w)|}{\|w\|_W}. \tag{2.12}$$

The FEM reads

$$\text{find } u_h \in V_h \text{ such that} \quad a(u_h, w_h) = f(w_h) \quad \forall w_h \in W_h, \tag{2.13}$$

where $V_h \subset V$ and $W_h \subset W$ are finite dimensional subspaces.

Unlike in the coercive case, well-posedness of problem (2.13) does not follow from the well-posedness of the continuous problem. Instead we assume that the discrete inf-sup condition holds for the discrete problem:

$$\beta_h = \inf_{v_h\in V_h} \sup_{w_h\in W_h} \frac{a(v_h,w_h)}{\|v_h\|_V \|w_h\|_V}. \tag{2.14}$$

The following theorem provides well-posedness of problem (2.8), see [Bab71, Theorem 2.1]:

Theorem 3. Babuška*: Let $V_h \subset V$ and $W_h \subset W$ be finite dimensional subspaces. Let $a(\cdot,\cdot) : V_h \times W_h \to \mathbb{R}$ denote a bilinear form and let f_h be a linear functional satisfying the assumptions of Theorem 1. Let a be continuous on $V_h \times W_h$ and satisfy the discrete inf-sup condition* (2.14).

Then, problem (2.13) *has a unique solution $u_h \in V_h$ $\forall h > 0$. The solution satisfies*

$$\|u_h\| \leq \frac{1}{\beta_h} \|f\|_{W'}. \tag{2.15}$$

Moreover, the error of the solution is bounded:

$$\|u - u_h\|_V \leq \frac{\gamma}{\beta_h} \inf_{v_h\in V_h} \|u - v_h\|_V. \tag{2.16}$$

As in the coercive case, the continuity of the bilinear form $a(\cdot,\cdot)$ over $V_h \times W_h$ follows from the continuity over $V \times W$. This is not the case for the discrete inf-sup stability. Instead, we must require that condition (2.14) is fulfilled in the discrete case.

In the following sections we will only cover the case, where the test and trial space coincide $V_N = W_N$. Other approaches are not considered. For more details see [QMN16].

2.1.3 The finite element linear system

Problem (2.4) can be solved by solving an equivalent linear system. Denote by $\{\phi_j\}_{j=1}^{N_h}$ a set of basis functions for the finite dimensional space V_h. Every function $v_h \in V_h$ is defined by:

$$v_h = \sum_{j=1}^{N_h} \mathbf{v}_j \phi_j \qquad \text{for some} \quad \mathbf{v} \in \mathbb{R}^{N_h}.$$

Let $\mathbf{u_h} \in \mathbb{R}^{N_h}$ denote the coefficient vector associated with the solution u_h to (2.4). Plugging this ansatz into (2.4) we derive:

$$\sum_{j=1}^{N_h} a(\phi_j, \phi_i)\mathbf{u_h}_j = f(\phi_i) \qquad \text{for} \quad i = 1, \ldots, N_h.$$

This is equivalent to the linear system:

$$\mathbf{A_h u_h} = \mathbf{f_h}, \tag{2.17}$$

where $\mathbf{A_h} \in \mathbb{R}^{N_h \times N_h}$, $(\mathbf{A_h})_{ij} = a(\phi_j, \phi_i)$ and $\mathbf{f_h} \in \mathbb{R}^{N_h}$, $(\mathbf{f_h})_i = f(\phi_i)$.

Thus the solution to problem (2.4) is obtained by solving the linear system (2.17). The complexity for solving a linear system depends on the dimension of the matrix $\mathbf{A_h}$ and on the solver.

Henceforth u denotes the exact solution to problem (2.1), $a(\cdot,\cdot)$ and $f(\cdot)$ denote the continuous forms, and V the continuous space. A subscript h indicates their FE versions and subscript N the RB counterparts. Thus u_h is the FE solution and u_N the RB solution. Ω denotes the domain. Unless otherwise noted the norm $\|\cdot\|_V$ is the norm of space V. Typically this is the H^1 norm: $\|v\|_V = \|v\|_{L^2(\Omega)} + \|\nabla v\|_{L^2(\Omega)}$.

2.2 The reduced basis ansatz

Let $\mathbb{P}$ denote a parameter space. Assume that problem (2.1) depends on a parameter $\mu \in \mathbb{P}$. More precisely, we want to solve:

$$\begin{aligned} \text{find } u_h(\mu) \in V_h(\mu) \text{ such that} \quad & a(\mu, u_h(\mu), v_h) = f(\mu, v_h) \\ & \forall v_h \in V_h(\mu), \forall \mu \in \mathbb{P}. \end{aligned} \tag{2.18}$$

Theorem 1 and Theorem 2 still apply, however all constants and bounds depend on μ (e.g. $\beta_h \to \beta_h(\mu)$ and $\gamma_h \to \gamma_h(\mu)$).

Consequently the linear system (2.17) becomes

$$\mathbf{A_h}(\mu)\mathbf{u_h}(\mu) = \mathbf{f_h}(\mu), \tag{2.19}$$

where

$$\begin{aligned} \mathbf{A_h}(\mu) &\in \mathbb{R}^{N_h \times N_h}, & (\mathbf{A_h}(\mu))_{ij} &= a(\mu, \phi_j, \phi_i), \\ \mathbf{f_h}(\mu) &\in \mathbb{R}^{N_h}, & (\mathbf{f_h}(\mu))_i &= f(\mu, \phi_i). \end{aligned}$$

Our aim is to solve problem (2.18) for a large number of parameters. This can become prohibitively expensive.

The key idea of the RBM ansatz consists in projecting the elements of V_h into a lower dimensional space $V_N \subset V_h$. For example: let $\{v_i\}_{i=1}^N \subset V_h$ denote a set of basis functions for V_h. Let $V_N \subset V_h$ be the space spanned by $\{v_i\}_{i=1}^N$. Denote by $\mathbf{v}^{(i)}$ the coefficient vector representing v_i. We define the matrix $\mathbf{A_N} \in \mathbb{R}^{N \times N}$,

$$(\mathbf{A}_N)_{i,j} = a(v_i, v_j) \quad \text{for} \quad i, j = 1 \ldots N. \tag{2.20}$$

Using the coefficient vectors $\mathbf{v_i}$ and (2.17) we obtain

$$(\mathbf{A_N})_{i,j} = a(v_i, v_j) = \sum_{k=1}^{N_h} \sum_{l=1}^{N_h} \mathbf{v}_k^{(i)} a(\phi_k, \phi_l) \mathbf{v}_l^{(j)}. \tag{2.21}$$

Let $\mathbb{V} \in \mathbb{R}^{N_h \times N}$ and $\mathbb{V} = [v^(1)| \dots |v^(N)]$ then (2.21) can be written as

$$\mathbf{A_N} = \mathbb{V}^T \mathbf{A_h} \mathbb{V}. \tag{2.22}$$

Analogously we derive

$$\mathbf{f_N} = \mathbb{V}^T \mathbf{f_h}. \tag{2.23}$$

Thus, the RBM reads:

$$\text{find } u_N(\mu) \in V_N \text{ such that} \quad a(\mu, u_N(\mu), v_N) = f(\mu, v_N) \quad \forall v_N \in V_N \tag{2.24}$$

and the linear system (2.19) becomes

$$\mathbf{A_N}(\mu)\mathbf{u}(\mu) = \mathbf{f_N}(\mu). \tag{2.25}$$

Note that $\mathbb{V}^T \mathbf{A_h} \mathbb{V} \in \mathbb{R}^{N \times N}$ and $\mathbb{V}^T \mathbf{f_h} \in \mathbb{R}^N$: Hence the dimension of this linear system is independent of N_h. If $N << N_h$, then the linear system (2.25) can be solved faster than (2.19).

Another way to look at RBM is as follows. Assume that the solution set $\mathbb{M}_h = \{u_h(\mu) \in V_h : \mu \in \mathbb{P}\}$ is such that it can be approximated by a linear combination of some of its elements, consequently these elements can be used as a basis for the space V_N.

This perspective gives us a way to form the projection matrix as follows: Let N denote the dimension of the reduced space V_N.

1. Select a set of N parameter points: $S_N = \{\mu^{(1)}, \dots, \mu^{(N)}\}$ and compute the corresponding FE solutions $\{u_h(\mu^{(1)}), \dots, u_h(\mu^{(N)})\}$.

2. Let $\{\zeta^{(1)}, \dots, \zeta^{(N)}\}$ be orthonormalized ($\langle \zeta^{(i)}, \zeta^{(j)} \rangle_V = \delta_{ij}$ and $\|\zeta_i\|_V = 1$) functions, such that

   $$V_N = \text{span}\{\zeta^{(1)}, \dots, \zeta^{(N)}\} = \text{span}\{u_h(\mu^{(1)}), \dots, u_h(\mu)^{(N)}\}.$$

   The $\{\zeta^{(i)}\}$ are called the reduced basis and V_N denotes the reduced basis space.

3. Define $\mathbb{V}$ by $\mathbb{V} = [\zeta^{(1)}| \dots |\zeta^{(N)}]$.

The orthogonalization, in step two, is necessary to improve the numerical condition of the problem, see [RHP07].

It remains open how to choose N and S_N. This will be discussed in Sections 2.3 and 2.4.

To calculate the solution at a parameter point we solve equation (2.25). The solution $u_N \in V_N$ can be used to calculate the coeffcients of the FE solution $\mathbf{u_h} = \mathbb{V}\mathbf{u_N}$.

2.2.1 Affine parameter dependence

A further ingredient of the RB approach is the assumption of affine parameter dependence of the bilinear and linear forms. Assume both the bilinear form $a(\mu, \cdot, \cdot)$ and $f(\mu, \cdot)$ are affine with respect to the parameter μ. In other words, assume:

$$a(\mu, u, v) = \sum_{q=1}^{Q_a} \theta_a^q(\mu) a_q(u, v), \tag{2.26}$$

$$f(\mu, v) = \sum_{q=1}^{Q_f} \theta_f^q(\mu) f_q(v). \tag{2.27}$$

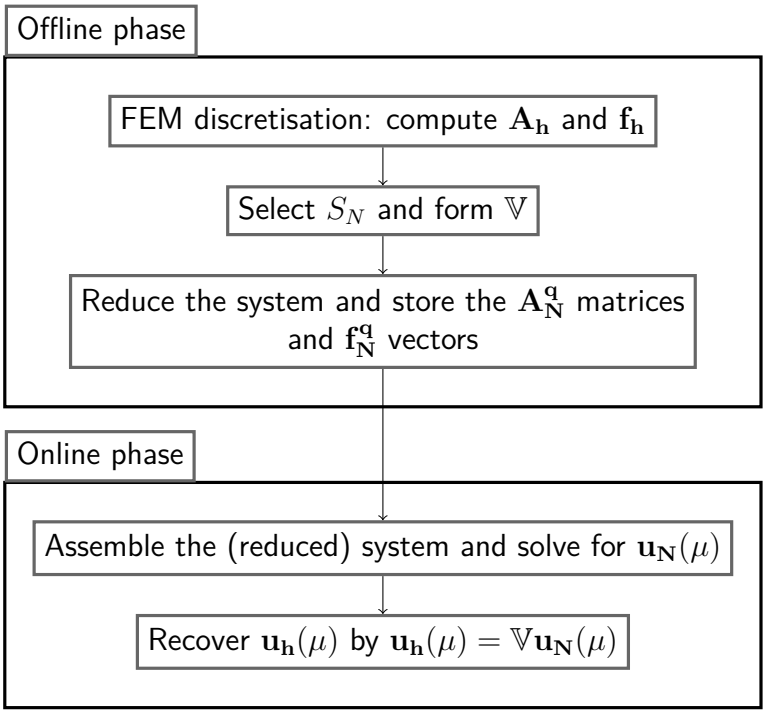


Figure 2.1: Overview of the offline-online phases. The (expensive) offline phase is only executed once, whereas the online phase is executed for every parameter point (μ) of interest.

The affine dependence carries through to the linear system:

$$\left(\sum_{q=1}^{Q_a} \theta_q^a(\mu)\mathbf{A_h^q}\right)\mathbf{u_N}(\mu) = \left(\sum_{q=1}^{Q_f} \theta_f^q(\mu)\mathbf{f_h^q}\right), \tag{2.28}$$

where

$$(\mathbf{A_h^q})_{ij} = a_q(\phi_i, \phi_j), \quad (\mathbf{f_h^q})_i = f_q(\phi_i), \quad 1 \leq i, j \leq N_h. \tag{2.29}$$

This allows us to computed all μ independent terms once and store them for further calculations. More precisely, we write:

$$\mathbf{A_N^q} = \mathbb{V}^T\mathbf{A_h^q}\mathbb{V}, \qquad \mathbf{f_N^q} = \mathbb{V}^\mathbb{T}\mathbf{f_h^q}. \tag{2.30}$$

To solve the system for a parameter point μ, we calculate

$$\mathbf{A_N}(\mu) = \sum_{q=1}^{Q_a} \theta_q^a(\mu)\mathbf{A_N^q}, \quad \mathbf{f_N}(\mu) = \sum_{q=1}^{Q_f} \theta_f^q(\mu)\mathbf{f_N^q}. \tag{2.31}$$

All operations in (2.31) involve only quantities, whose dimension depends on N. Typically $N << N_h$, hence this is faster than reducing the system for every parameter point.

The reduced basis ansatz and the affine parameter assumption allow us an efficient offline-online approach, see Figure 2.1.

In the following sections the stability of the RB approach is analysed.

2.2.2 Coercive problems

The stability analysis of the RBM for coercive problems is similar to the stability analysis of the FEM for coercive problems (see Section 2.1.1). The Lax-Milgram theorem provides well-posedness and the following bound for the solution u_N to problem (2.24):

Theorem 4. *Assume that the assumptions of Theorem 1 are fulfilled for any $\mu \in \mathbb{P}$. Then, the solution u_N to problem* (2.24) *is unique and satisfies:*

$$\|u_N(\mu)\|_V \leq \frac{1}{\alpha_N(\mu)} \|f(\mu, \cdot)\|_{V'}, \tag{2.32}$$

where

$$\alpha_N = \inf_{v \in V_N} \frac{a(\mu, v, v)}{\|v\|_V^2}. \tag{2.33}$$

For a symmetric bilinear form $a(\mu, \cdot, \cdot)$, it can be shown, see [Ver03]:

$$\|u_h(\mu) - u_N(\mu)\|_V \leq \sqrt{\frac{\gamma_h(\mu)}{\alpha_N(\mu)}} \inf_{w \in V_N} \|u_h(\mu) - w\|_V. \tag{2.34}$$

2.2.3 Weakly coercive problems

In case the bilinear form $a(\mu, \cdot, \cdot)$ is weakly coercive (or inf-sup stable) the Lax-Milgram theorem does not apply. Instead the well-posedness of the RB problem is ensured by the following theorem, which is a consequence of the Babuška theorem (Theorem 3):

Theorem 5. *Assume the bilinear form is continuous, with*

$$\gamma_h(\mu) = \sup_{v_h \in V_h} \sup_{w_h \in V_h} \frac{a(\mu, v_h, w_h)}{\|v_h\|_V \|w_h\|_V}. \tag{2.35}$$

Then, the following lower bound of the continuity constant is valid

$$\bar{\gamma}_F > \sup_{v \in V_N} \frac{f(\mu, w)}{\|w\|_W} \quad \forall \mu \in \mathbb{P}. \tag{2.36}$$

Moreover, the following lower bound of the inf-sup stability constant $\beta_N(\mu)$ is valid:

$$\beta_{0,N} \leq \beta_N(\mu) = \inf_{v_N \in V_N} \sup_{w_N \in V_N} \frac{a(\mu, v_N, w_n)}{\|v_N\|_V \|w_N\|_V} \qquad \forall \mu \in \mathbb{P}. \tag{2.37}$$

Consequently, the RB problem admits a unique solution $u_N(\mu) \in V_N$ satisfying

$$\|u_N(\mu)\|_V \leq \frac{1}{\beta_N(\mu)} \|f(\mu, \cdot)\|_{V'}. \tag{2.38}$$

Theorem 5 applies to the case where test and trial space are the same ($V_N = W_N$).

2.2.4 Reducibility and approximability

In order to successfully apply the RBM it is important that the problem is reducible. However it is not trivial to determine the reducibility of a problem. The reducibility of a problem depends on a number of factors. We will present two factors that play an important role in the reducibility of a problem. The first factor is the differentiability and regularity of the solution map. The second factor is approximability of the solution set by N dimensional subspaces. For a more detailed discussion see [QMN16, Chapter 5].

Define the solution maps

$$\phi : \mathbb{P} \to V, \qquad \mu \mapsto u(\mu), \tag{2.39}$$

$$\phi_h : \mathbb{P} \to V_h, \qquad \mu \mapsto u_h(\mu), \tag{2.40}$$

$$\phi_N : \mathbb{P} \to V_N, \qquad \mu \mapsto u_N(\mu), \tag{2.41}$$

and the associated solution sets

$$\mathcal{M} = \{\phi(\mu) : \mu \in \mathbb{P}\}, \tag{2.42}$$
$$\mathcal{M}_h = \{\phi_h(\mu) : \mu \in \mathbb{P}\}, \tag{2.43}$$
$$\mathcal{M}_N = \{\phi_N(\mu) : \mu \in \mathbb{P}\}. \tag{2.44}$$

Our goal is to identify a space $V_N \subset V$ with dimension N such that

$$\inf_{v \in V_N} \|u(\mu) - v\| < \epsilon \quad \forall \mu \in \mathbb{P}. \tag{2.45}$$

Moreover, we want to keep N as small as possible. As a first step the following two theorems are valid, see [QMN16, Chapter 5]:

Theorem 6. *Assume that the coercive bilinear form $a(\mu, \cdot, \cdot)$ and the linear form $f(\mu, \cdot)$ are Lipschitz continuous with respect to μ. Moreover, assume there exist two constants $L_a, L_f > 0$, such that*

$$|a(\mu, u, v) - a(\mu', u, v)| \leq L_a \|u\|_V \|v\|_V \|\mu - \mu'\| \quad \forall \mu, \mu' \in \mathbb{P}, \forall u, v \in V, \tag{2.46}$$
$$|f(\mu, v) - f(\mu', v)| \leq L_f \|v\|_V \|\mu - \mu'\| \quad \forall \mu, \mu' \in \mathbb{P}, \forall v \in V. \tag{2.47}$$

Then, there exists a positive constant $L_u > 0$ such that

$$\|u(\mu) - u(\mu')\|_V \leq L_u \|\mu - \mu'\| \quad \forall \mu, \mu' \in \mathbb{P}. \tag{2.48}$$

Further, ϕ is a continuous map, i.e., $\phi \in C^0(\mathbb{P}, V)$, where

$$C^0(\mathbb{P}, \mathbb{V}) = \{v : \mathbb{P} \to V : \mu \mapsto v(\mu) \text{ is continuous and } \max_{\mu \in \mathbb{P}} \|v(\mu)\| < +\infty\}. \tag{2.49}$$

The solution set $\mathcal{M}$ is compact in V.

Theorem 6 can be generalized to the case of inf-sup stable problems, see [QMN16, Chapter 5].

In case the bilinear form a and the linear form f fulfil the affine parametric dependence assumption (2.26) and (2.27) the Lipschitz continuity only depends on the Lipschitz continuity of θ_a^q and θ_f^q. If θ_a^q and θ_f^q are Lipschitz continuous

$$|\theta_a^q(\mu) - \theta_a^q(\mu')| \leq L_a^q \|\mu - \mu'\| \quad \forall \mu, \mu', q : \mu, \mu' \in \mathbb{P}, q = 1 \dots Q_a, \tag{2.50}$$
$$\left|\theta_f^q(\mu) - \theta_f^q(\mu')\right| \leq L_f^q \|\mu - \mu'\| \quad \forall \mu, \mu', q : \mu, \mu' \in \mathbb{P}, q = 1 \dots Q_f, \tag{2.51}$$

then one can show, see [QMN16, Section 5.3.1]

$$|a(\mu, u, v) - a(\mu', u, v)| \leq \sum_{q=1}^{Q_a} \gamma_f^q L_a^q \|\mu - \mu'\|,$$

and

$$|f(\mu, v) - f(\mu', v)| \leq \sum_{q=1}^{Q_f} \gamma^q L_f^q \|\mu - \mu'\|,$$

where

$$\gamma^q = \sup_{v \in V} \sup_{w \in W} \frac{a_q(v, w)}{\|v\|_V \|w\|_W}, \qquad \gamma_f^q = \sup_{v \in V} \frac{f_q(v)}{\|v\|_V}. \tag{2.52}$$

are the continuity constants.

Moreover, the following theorem is valid:

Theorem 7. *Assume that the bilinear form $a(\mu,\cdot,\cdot)$ and the linear form $f(\mu,\cdot)$ are C^k maps with respect to μ for some $k \geq 0$. Moreover, let $a(\mu,\cdot,\cdot)$ be continuous and inf-sup stable for all $\mu \in \mathbb{P}$ and let $f(\mu,\cdot,\mu)$ be continuous for all $\mu \in \mathbb{P}$. Then, the solution map ϕ is of class C^k as well.*

It can be shown that, see [QMN16, Section 5.3.2], given a point $\mu_0 \in \mathbb{P} = \mathbb{R}^P$

$$\left\| \frac{\partial u(\mu_0)}{\partial \mu_i} \right\|_V \leq \frac{1}{\beta(\mu_0)} \left(\sum_{i=1}^{P} \sum_{q=1}^{Q_f} \gamma_f^q \left| \frac{\partial \theta_f^q(\mu_0)}{\partial \mu_i} \right| + \sum_{i=1}^{P} \sum_{q=1}^{Q_a} \gamma^q \left| \frac{\partial \theta_a^q(\mu_0)}{\partial \mu_i} \right| \frac{1}{\beta(\mu_0)} \sum_{q=1}^{Q_f} \gamma_f^q(\mu_0) \right). \quad (2.53)$$

Theorems 6 and 7 are valid for ϕ_h as well, see [QMN16, Chapter 5].

Next we want to ensure that the dimension N of V_N can be small, while still maintaining good approximation properties. To this purpose, we introduce the Kolmogorov n-width, see [Mel00].

First, we introduce a measure for the approximation quality, of a subspace

$$d(K, S_n) = \sup_{k \in K} \inf_{s_n \in S_n} \|k - s_n\|_V . \quad (2.54)$$

K is a set of elements that we want to approximate by elements of the space s_n. We define the Kolmogorov n-width as

$$d_n(K, S) = \inf_{\substack{S_n \subset S \\ dim(S_n)=n}} d(K, S_n) = \inf_{\substack{S_n \subset K \\ dim(S_n)=n}} \sup_{k \in K} \inf_{s_n \in S_n} \|k - s_n\|_S . \quad (2.55)$$

The n-width measures how good elements of a subspace $K \subset S$ can be approximated by a discrete subspace S_n of dimension n of S. If we use the n-width together with the solution set $\mathcal{M}_h$ and V_h, then we get a measure of how good the solutions to problem (2.4) can be approximated by subspaces of a given dimension:

$$d_n(\mathcal{M}_h, V_h) = \inf_{\substack{V_n \subset V_h \\ dim(V_n)=n}} d(\mathcal{M}_h, V_n) = \inf_{\substack{V_n \subset V_h \\ dim(V_n)=n}} \sup_{u_h \in \mathcal{M}_h} \inf_{v_n \in V_n} \|u_h - v_n\|_V . \quad (2.56)$$

Moreover, we can define the n-width of a space by

$$d_n(V) = \inf_{\substack{V_n \subset V \\ dim(V_n)=n}} d_n(V, V_n) \quad (2.57)$$

In general it is not trivial to derive bounds for the n-width of general solution sets to problems of the form of (2.1). However given suitable assumptions, on the bilinear form $a(\mu,\cdot,\cdot)$, the n-width decays exponentially with n, see [QMN16, Chapter 5].

2.3 Selection of parameter points: equi-logarithmic spaces

Consider problems involving a bilinear form $a(\mu, w, v)$ of the form:

$$a(\mu, w, v) = a_0(w, v) + \mu a_1(w, v), \quad (2.58)$$

where $\mu \geq 0 \in \mathbb{R}$, $a_0 : V \times V \to \mathbb{R}$ and $a_1 : V \times V \to \mathbb{R}$ are continuous, symmetric and positive semi-definite, and a_0 coercive.

Then, the following theorem shows the exponential decay of the approximation error $\|u_h(\mu) - u_N(\mu)\|_V$, assuming a equi-logarithmic distribution of the parameter points.

Theorem 8. *Assume there exists a $\rho > 0 \in \mathbb{R}$ such that*

$$\frac{a_1(v, v)}{a_0(v, v)} \leq \rho \quad \forall v \neq 0 \in V. \quad (2.59)$$

Let the parameter points μ^k fulfil

$$\mu^k = e^{(-log\rho + \sum_{l=1}^{k} \delta_{lN})} - \frac{1}{\rho}, \tag{2.60}$$

where ρ is any lower bound for ρ_1, and

$$\sum_{l=1}^{N} \delta_{lN} = log(\rho\mu_{max} + 1), \tag{2.61}$$

$$c \geq \frac{\delta_{kn}}{\delta_N} \quad \textit{for } k = 1...N. \tag{2.62}$$

Let N satisfy

$$N \geq N_{crit} = c\, e\, log(\rho\mu_{max} + 1). \tag{2.63}$$

Then, the RB approximation error is bounded by

$$\|u_h(\mu) - u_N(\mu)\|_V \leq \sqrt{(1 + \mu_{max}\rho_1)}\, \|u_h(0)\|\, e^{-\frac{N}{N_{crit}}}. \tag{2.64}$$

The theorem is stated and proven in [MPT02b]. Besides, proving a uniform error bound for the RB approximation error Theorem 8 also gives an estimate for the dimension of the RB space. The dimension of the RB-space depends logarithmically on the exact bilinear form, through ρ, and on the length of the interval (μ_{max}).

Moreover, it is noted in [MPT02b] that the requirement of an exact logarithmic point distribution is rather weak and for example log random distributions perform almost as good. This can also be seen in the numerical tests in [Ver03]. Furthermore, the numerical tests in [Rov03] suggest that the logarithmic point distribution performs well also for problems not satisfying equation (2.58).

2.4 Selection of parameter points: greedy algorithm

The greedy algorithm is a different approach for the selection of parameter points. The greedy algorithm searches for an optimal set of parameters. Initially it was developed for time-dependent parabolic problems, see [GP05; Gre05]. Nonetheless, it can be applied to a more general setting.

In general, greedy algorithm construct a solution, by selecting the most advantageous option at every step. Applied to the construction of RB spaces, this reads: Given a current space V_N, of dimension N, select a new basis function $u_h(\mu^{N+1})$ by choosing the parameter μ^{N+1}, such that the approximation error is maximized. This is done until some stopping criterion is reached. For instance, the maximal error is lower than a given tolerance e_{max}.

Algorithm 1 contains the structure of a greedy algorithm for constructing RB spaces. Note that the $u_N(\mu)$ in equations (2.65) and (2.66) refers to the approximate RB solution in the RB space V_i.

The calculation of the approximation error in equations (2.65) and (2.66) can not be implemented directly. Discretising the parameter space $\mathbb{P}$ and calculating the approximation error at the grid points would be to expensive.

Error estimators offer a way out. If we have a way to predict the approximation error without using the FE solution $u_h(\mu)$, then we could use this error estimator as a surrogate for the true approximation error. The error estimator is given in the form of $r : \mathcal{P}(V_h) \times \mathbb{P} \to \mathbb{R}$. The function r estimates the RB error at a parameter point, given a current RB space. The design of an appropriate error estimator is the topic of Section 2.4.2.

Even with an error estimator calculating $\arg\max_{\mu\in\mathbb{P}} r(V_i, \mu)$ over the whole parameter space $\mathbb{P}$ is infeasible. Instead we choose a training set $\Xi \subset \mathbb{P}$ to calculate $\mu^{(i+1)}$ as $\arg\max_{\mu\in\Xi} r(V_i, \mu)$. The resulting algorithm is shown in Algorithm 2.

Algorithm 1 Greedy algorithm for constructing RB spaces.

Set a maximal error tolerance e_{max}

Set an initial $u_h(\mu^{(1)})$

Initialize $i \leftarrow 1$, $\epsilon \leftarrow \infty$, $\zeta_1 = \frac{u_h(\mu^{(1)})}{\left\|u_h(\mu^{(1)})\right\|_V}$ and $V_1 \leftarrow span(\zeta_1)$

while $\epsilon > e_{max}$ **do**

$$\epsilon \leftarrow \max_{\mu \in \mathbb{P}} \left\|u_h(\mu) - u_N(\mu)\right\|_V \tag{2.65}$$

$$\mu^{(i+1)} \leftarrow \arg\max_{\mu \in \mathbb{P}} \left\|u_h(\mu) - u_N(\mu)\right\|_V \tag{2.66}$$

$$u_h(\mu^{(i+1)}) \leftarrow (\mathbf{A_h}(\mu^{i+1}))^{-1}\mathbf{f_h}(\mu^{i+1}) \tag{2.67}$$

$$\zeta_{i+1} \leftarrow \text{ orthonormalize } u_h(\mu^{(i+1)}) \text{ w.r.t } \{\zeta_1, \ldots, \zeta_i\} \tag{2.68}$$

$$V_{i+1} \leftarrow span(\zeta_1, \ldots, \zeta_{i+1}) \tag{2.69}$$

$$i \leftarrow i + 1 \tag{2.70}$$

end while

$\mathbb{V} = [\zeta_1 | \ldots | \zeta_i]$

Algorithm 2 Greedy algorithm for constructing RB spaces, with an error estimator.

Set a maximal error tolerance e_{max}

Set an initial $u_h(\mu^{(1)})$

Initialize $i \leftarrow 1$, $\epsilon \leftarrow \infty$, $\zeta_1 = \frac{u_h(\mu^{(1)})}{\left\|u_h(\mu^{(1)})\right\|_V}$ and $V_1 \leftarrow span(\zeta_1)$

while $\epsilon > e_{max}$ **do**

$$\mu^{(i+1)} \leftarrow \arg\max_{\mu \in \Xi} r(V_i, \mu) \tag{2.71}$$

$$\epsilon \leftarrow r(V_i, \mu^{(i+1)}) \tag{2.72}$$

$$u_h(\mu^{(i+1)}) \leftarrow (\mathbf{A_h}(\mu^{i+1}))^{-1}\mathbf{f_h}(\mu^{i+1}) \tag{2.73}$$

$$\zeta_{i+1} \leftarrow \text{ orthonormalize } u_h(\mu^{(i+1)}) \text{ w.r.t } \{\zeta_1, \ldots, \zeta_i\} \tag{2.74}$$

$$V_{i+1} \leftarrow span(\zeta_1, \ldots, \zeta_{i+1}) \tag{2.75}$$

$$i \leftarrow i + 1 \tag{2.76}$$

end while

$\mathbb{V} = [\zeta_1 | \ldots | \zeta_i]$

The computational cost of the greedy algorithm is

$$O(NN_{\Xi}M_{est} + NM_{inv} + NM_{orth}), \tag{2.77}$$

where M_{est} denotes the complexity of the error estimator, N_{Ξ} the size of the training set Ξ, M_{inv} denotes the complexity of matrix inversion and M_{orth} of orthonormalization. Assuming that $O(M_{inv}) = O(NN_h^3)$, $O(M_{orth}) = O(N_h^2)$, $O(M_{est}) = O(1)$ and N const. with $N << N_h$ and $N_{\Xi} << N_h$. This leads to a global complexity of $O(N_h^3)$.

2.4.1 A priori convergence of the greedy algorithm

Let ζ_i denote the orthonormalized basis functions selected by the greedy algorithm. We define the projector $P_N : V \to V_N$ with respect to the V inner product

$$P_N s = \sum_{i=1}^{N} \langle s, \zeta_i \rangle_V \zeta_i. \tag{2.78}$$

Denote the projection error by $\sigma_N(s)$

$$\sigma_N(s) = \|s - P_N s\|_V . \tag{2.79}$$

For the greedy Algorithm 2 the following result is proved in [Bin+11]:

Theorem 9. *Assume that there exist two positive constants c_r and C_r independent of μ and N and that the error estimator r satisfies*

$$c_r r(V_N, \mu) \leq \|u(\mu) - u_N(\mu)\|_V \leq C_r r(V_N, \mu) \qquad \forall \mu \in \mathbb{P}. \tag{2.80}$$

Furthermore, given $M, a, \alpha > 0$ assume that the Kolmogorov n-width of the space V decays exponentially

$$d_n(V) \leq M e^{-aN^{\alpha}}. \tag{2.81}$$

Then, the following upper bound for the projection error holds

$$\max_{v \in V} \sigma_N(v) \leq CMe^{-cN^{\beta}}, \quad n \geq 0. \tag{2.82}$$

Here $0 < \theta \leq 1$:

$$\beta = \frac{\alpha}{\alpha + 1}, \qquad q = \lceil \frac{2C_r}{c_r \theta} \rceil^2, \qquad N_0 = \lceil (8q)^{\frac{1}{1-\beta}} \rceil,$$
$$c = min(|ln\theta|, (4q)^{\alpha} a), \qquad C = max(e^{cN_0^{\beta}}, q^{\frac{1}{2}}).$$

Although the theorem refers to the space V, it also applies to a sufficiently fine discretisation V_h of V, see [Bin+11]. Note that the performance of the greedy algorithm depends on the effectivity of the error estimator.

2.4.2 A posteriori error estimator

For Algorithm 2, it is crucial to estimate the RB approximation error $e_N = u_h(\mu) - \mathbb{V}u_N(\mu)$ in a fast and cheap way.

To calculate e_N directly we need to calculate the RB and the FE solutions. This is expansive. Therefore, we derive expressions that are independent of the FE solution u_h. The following derivations follow [QMN16, Chapter 3], [Sen+06] and [RHP07]: Define the residual $r(\mu, v)$ as:

$$\begin{aligned} r(\mu, v) &= f(\mu, v) - a(u_N(\mu), v, \mu), \\ &= a(u_h(\mu), v, \mu) - a(u_N(\mu), v, \mu), \\ &= a(e_N(\mu), v, \mu) \qquad \forall v \in V. \end{aligned}$$

By the continuity Property (2.3) of a and the definition of the dual norm, we get

$$\begin{aligned} a(e_h(\mu), v, \mu) &\leq \gamma_h(\mu) \left\| e_N(\mu) \right\|_V \left\| v \right\|_V, \\ |r(\mu, v)| &\leq \gamma_h(\mu) \left\| e_N(\mu) \right\|_V \left\| v \right\|_V, \\ \left\| r(\mu, \cdot) \right\|_{V'} &\leq \gamma_h(\mu) \left\| e_N(\mu) \right\|_V \qquad \forall v \in V_h. \end{aligned}$$

Using (2.14), we deduce:

$$\begin{aligned} \beta_h(\mu) \left\| e_N(\mu) \right\|_V &= \inf_{v \in V_h} \sup_{w \in V} \frac{a(\mu, v, w)}{\left\| v \right\|_V \left\| w \right\|_V} \left\| e_N(\mu) \right\|_V \\ &\leq \sup_{w \in V} \frac{a(\mu, e_N(\mu), w)}{\left\| e_N(\mu) \right\|_V \left\| w \right\|_V} \left\| e_N(\mu) \right\|_V \\ &\leq \left\| r(\mu, \cdot) \right\|_{V'}. \end{aligned}$$

Hence, we can derive the following bounds:

$$\frac{1}{\gamma_h(\mu)} \left\| r(\mu, \cdot) \right\|_{V_h'} \leq \left\| e_N(\mu) \right\|_V \leq \frac{1}{\beta_h(\mu)} \left\| r(\mu, \cdot) \right\|_{V_h'}. \tag{2.83}$$

This proves the reliability of the error estimator. Note that $\frac{1}{\gamma_h(\mu)}$ and $\frac{1}{\beta_h(\mu)}$ correspond to c_r and C_r respectively in Theorem 9. This implies that the closer $\frac{\beta_h(\mu)}{\gamma_h(\mu)}$ tends to 1, the better the greedy algorithms convergence will be.

Equation (2.83) also gives a way to implement an error estimator for the greedy algorithm. To compute the residual norm, note that:

$$\begin{aligned} \mathbf{r_h}(\mu, \mathbf{u_N}) &= \mathbf{f_h}(\mu) - \mathbf{A_h}(\mu) \mathbb{V} \mathbf{u_N} \\ &= \mathbf{A_h}(\mu) \mathbf{u_h} - \mathbf{A_h}(\mu) \mathbb{V} \mathbf{u_N} \\ &= \mathbf{A_h}(\mu)(\mathbf{u_h} - \mathbb{V} \mathbf{u_N}) \\ (\mathbf{A_h}(\mu))^{-1} \mathbf{r_h}(\mu, \mathbf{u_N}) &= \mathbf{u_h} - \mathbb{V} \mathbf{u_N} = \mathbf{e_N}(\mu), \end{aligned} \tag{2.84}$$

with $\mathbf{u_N}$ denoting the coefficients of the RB solution. By taking the norms on both sides:

$$\left\| \mathbf{e_N}(\mu) \right\|_2 \leq \left\| (\mathbf{A_h}(\mu))^{-1} \right\|_2 \left\| \mathbf{r_h}(\mu, \mathbf{u_N}) \right\|_2 = \frac{\left\| \mathbf{r_h}(\mu, \mathbf{u_N}) \right\|_2}{\sigma_{min}(\mathbf{A_h}(\mu))}, \tag{2.85}$$

where $\sigma_{min}(\mathbf{A_h}(\mu))$ denotes the smallest singular value of $\mathbf{A_h}(\mu)$.

To get a error bound in the norm of the space V, we define

$$(\mathbf{X_h})_{i,j} = (\phi_i, \phi_j)_V, \tag{2.86}$$

$$\left\| v \right\|_{\mathbf{X_h}}^2 = \left\| \mathbf{X_h}^{\frac{1}{2}} v \right\|_2^2 = v^T \mathbf{X_h} v, \tag{2.87}$$

where $\{\phi_i\}_{i=1}^{N_h}$ denotes a set of basis functions of V_h.

Left multiply (2.84) with $\mathbf{X_h}^{\frac{1}{2}}$ and use $\mathbf{I} = \mathbf{X_h}^{\frac{1}{2}} \mathbf{X_h}^{-\frac{1}{2}}$ to obtain

$$\mathbf{X_h}^{\frac{1}{2}} \mathbf{e_N}(\mu) = \mathbf{X_h}^{\frac{1}{2}} (\mathbf{A_h}(\mu))^{-1} \mathbf{X_h}^{\frac{1}{2}} \mathbf{X_h}^{-\frac{1}{2}} \mathbf{r_h}(\mu, \mathbf{u_N}). \tag{2.88}$$

By taking norms again, we deduce

$$\|\mathbf{e_N}(\mu)\|_{\mathbf{X_h}} \leq \frac{\|\mathbf{r_h}(\mu, \mathbf{u_N})\|_{\mathbf{X_h}^{-1}}}{\sigma_{min}(\mathbf{X_h}^{-\frac{1}{2}}\mathbf{A_h}(\mu)\mathbf{X_h}^{-\frac{1}{2}})}. \tag{2.89}$$

If a matrix S is symmetric, then $\sigma_{min}(S) = \lambda_{min}(S)$, with $\lambda_{min}(S)$ denoting the smallest eigenvalue of S. Moreover, $\sigma_{min}(\mathbf{X_h}^{-\frac{1}{2}}\mathbf{A_h}(\mu)\mathbf{X_h}^{-\frac{1}{2}})$ is the discrete inf-sup stability constant $\beta_h(\mu)$, see [QMN16].

To calculate the error bound for a given μ, the residual $r(\mu, v)$ and σ_{min} need to be evaluated. The residual can be evaluated in $O(N_h)$ operations. Let M denote the complexity of matrix-matrix multiplication. To evaluate σ_{min} at least $O(N_h M)$ are needed. This is can become too expensive for our purpose.

In order to avoid the calculation of singular/eigenvalue for every μ two approximation strategies are commonly used:

- (adaptive) interpolation,
- approximation by a sucessive constraint method (SCM).

More details about adaptive interpolation can be found in [QMN16; MN15; Neg+13]. For the SCM approach, see [Huy+07; RHP07; MN15; Man12].

In case of affine parametric dependency, see Section 2.2.1, the calculation of the residuals can be optimized further. Recall that

$$\begin{aligned}\|\mathbf{r_h}(\mu, \mathbf{u_N})\|^2_{\mathbf{X_h}^{-1}} =& \|\mathbf{f_h}(\mu) - \mathbf{A_h}(\mu)\mathbb{V}\mathbf{u_N}\|^2_{\mathbf{X_h}^{-1}} \\ =& \mathbf{f_h}(\mu)^T\mathbf{X_h}^{-1}\mathbf{f_h}(\mu) + (\mathbf{A_h}(\mu)\mathbb{V}\mathbf{u_N})^T\mathbf{X_h}^{-1}\mathbf{A_h}(\mu)\mathbb{V}\mathbf{u_N} \\ & - 2\mathbf{f_h}(\mu)\mathbf{X_h}^{-1}\mathbf{A_h}(\mu)\mathbb{V}\mathbf{u_N}.\end{aligned}$$

Given the affine parameter dependence, we obtain

$$\begin{aligned}\|\mathbf{r_h}(\mu, \mathbf{u_N})\|^2_{\mathbf{X_h}^{-1}} =& \sum_{q_1=1}^{Q_f}\sum_{q_2=1}^{Q_f}\theta^f_{q_1}(\mu)\theta^f_{q_2}(\mu)\underbrace{\mathbf{f_h^{q_1}}^T\mathbf{X_h}^{-1}\mathbf{f_h^{q_2}}}_{C_{q_1,q_2}} \\ &+ \sum_{q_1=1}^{Q_a}\sum_{q_2=1}^{Q_a}\theta^a_{q_1}(\mu)\theta^a_{q_2}(\mu)\mathbf{u_N}(\mu)^T\underbrace{\mathbb{V}^T\mathbf{A_h}^{q_1T}\mathbf{X_h}^{-1}\mathbf{A_h}^{q_2}\mathbb{V}}_{E_{q_1,q_2}}\mathbf{u_N}(\mu) \\ &- 2\sum_{q_1=1}^{Q_a}\sum_{q_2=1}^{Q_f}\theta^{q_1}_a(\mu)\theta^{q_2}_f(\mu)\mathbf{u_N}(\mu)^T\underbrace{\mathbb{V}^T\mathbf{A_h}^{q_1T}\mathbf{X_h}^{-1}\mathbf{f_h^{q_2}}}_{D_{q_1,q_2}}.\end{aligned} \tag{2.90}$$

The quantities $C_{q_1,q_2} \in \mathbb{R}$, $E_{q_1,q_2} \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ and $D_{q_1,q_2} \in \mathbb{R}^{\mathbb{N}}$ are μ independent and need only to be calculated once.

2.5 Implementation

For the experiments in Chapter 3 and 4 we implemented the RBM in Python using NGSolve ([Sch14]) and SciPy [JOP+01]. NGSolve is used to calculate FE solutions, SciPy to implement error estimators and to solve the RB problems. The implementation is based on the contents of the previous sections and the algorithms described in [QMN16].

The linear system (2.17) is solved using an iterative solver. The RB counterpart (2.25) is solved by using LU factorization, see [GL13]. The stability constant β_h is calculated using the

Implicitly restarted Arnoldi iteration, see [LS96] and [LSY98]. The appendix also includes a short introduction into the Arnoldi iteration, see Section A.3

To calculate the stability constant over the whole parameter space interpolation is used. Linear interpolation is used for 1D problems, see e.g. Chapter 4. Interpolation by radial basis function is used for higher dimensional problems, see e.g. Chapter 3 or Chapter 5.

Chapter 3

Reduced basis methods for the Poisson equation

In order to test the RBM we consider two instances of the Poisson problem. The Poisson problem reads

$$\text{find } u \in V \text{ such that } -\Delta u(\mu) = g(\mu) \quad \text{in} \quad \Omega(\mu). \tag{3.1}$$

Equation (3.1) can be endowed with either Dirichlet or Neumann boundary conditions. The weak formulation of (3.1) reads

$$\text{find } u \in V \text{ such that } a(\mu, u, v) = f(\mu, v) \quad \forall v \in V, \forall \mu \in \mathbb{P}, \tag{3.2}$$

where

$$a(\mu, u, v) = \int_{\Omega(\mu)} \nabla u \cdot \nabla v, \tag{3.3}$$

$$f(\mu, v) = \int_{\Omega(\mu)} g(\mu) v. \tag{3.4}$$

It can be shown that equation 3.2 fulfils the assumptions of Theorem 1, in particular the coercivity and continuity conditions; see equations (2.2) and (2.3).

3.1 Internal heating

The first example, taken from [Ver03, Model Example 3], is that of an rectangular domain Ω with an internal heat source. The shape of the domain is parametrized by $\mu \in \mathbb{R}$. For a given μ the problem reads

$$\text{find } u \in V \text{ such that } -\Delta u = \frac{1}{\mu} \quad \text{in} \quad \Omega(\mu). \tag{3.5}$$

Equation (3.5) is endowed with homogeneous Dirichlet boundary conditions

$$u = 0 \quad \text{on} \quad \partial\Omega(\mu).$$

The weak formulation of (3.5) reads

$$\text{find } u \in V \text{ such that } a(\mu, u, v) = f(\mu, v) \quad \forall v \in V(\Omega(\mu)). \tag{3.6}$$

with

$$a(\mu, u, v) = \int_{\Omega(\mu)} \nabla u \cdot \nabla v, \tag{3.7}$$

$$f(\mu, v) = \int_{\Omega(\mu)} \frac{v}{\mu}. \tag{3.8}$$

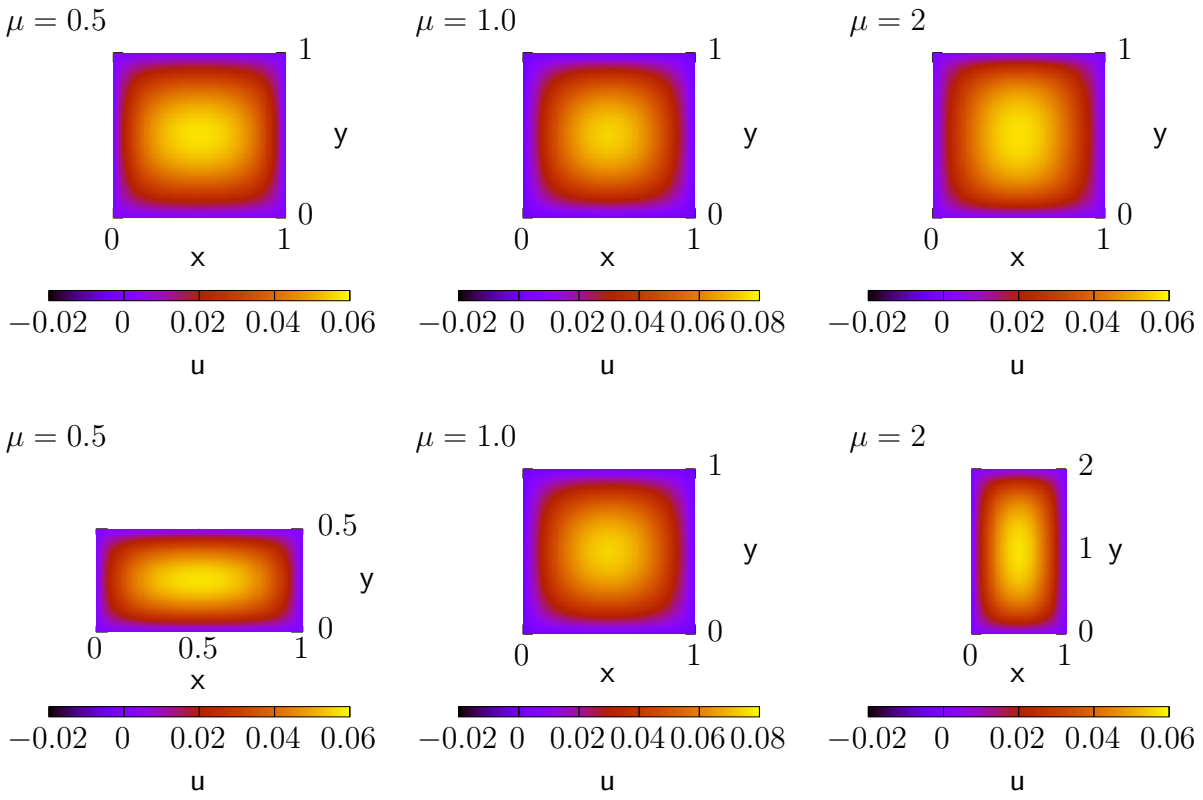


Figure 3.1: Solution to the internal heating problem for the parameter values $\mu = 0.5$, $\mu = 1$, and $\mu = 2$. The first three figures show the solution in the reference domain Ω. The last three in the original domain $\Omega(\mu)$.

The domain is parametrized by $\mu \in \mathbb{P}$ with the parameter space denoted by $\mathbb{P} = \mathbb{R}$. We reformulate the problem by mapping $\Omega(\mu)$ into a reference domain $\Omega = [0,1]^2$; for details see [Ver03, Chapter 3]. The weak formulation of the reformulated problem reads

$$\text{find } u \in V \text{ such that } a(\mu, u, v) = f(v) \quad \forall v \in V(\Omega). \tag{3.9}$$

with

$$a(\mu, u, v) = \mu \int_\Omega \frac{\partial u}{\partial x}\frac{\partial v}{\partial y} + \frac{1}{\mu}\int_\Omega \frac{\partial u}{\partial y}\frac{\partial v}{\partial y}, \tag{3.10}$$

$$f(v) = \int_\Omega v. \tag{3.11}$$

The bilinear form $a(\mu, \cdot, \cdot)$ depends on the parameter μ affinely with

$$a(\mu, u, v) = \theta_a^{(1)}(\mu) a_1(u, v) + \theta_a^{(2)} a_2(u, v), \tag{3.12}$$

where

$$\theta_a^{(1)}(\mu) = \mu, \qquad a_1(u, v) = \frac{\partial u}{\partial x}\frac{\partial v}{\partial y},$$

$$\theta_a^{(2)}(\mu) = \frac{1}{\mu}, \qquad a_2(u, v) = \frac{\partial u}{\partial y}\frac{\partial v}{\partial y}.$$

In this case the linear form f is independent of μ. In Figure 3.1 we depict the solution to problem (3.5).

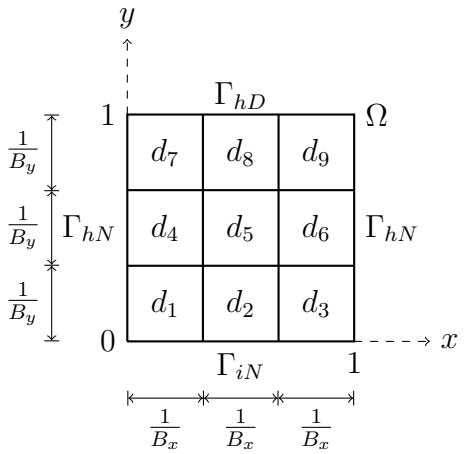


Figure 3.2: Thermal block domain Ω for $B_x = 3$ and $B_y = 3$.

3.2 Thermal blocks

The second example is that of a thermal block taken from [RHP07, Section 6.1.1]. The problem describes a square domain that is partitioned into $B_x B_y$ subdomains, see Figure 3.2. Each region d_i for $i = 2..B_x B_y$ is assigned a conductivity μ_i. The conductivity of d_1 is fixed at 1.

The problem is endowed with homogeneous Dirichlet boundary conditions on Γ_{hD}, homogeneous Neumann on Γ_{hN} and inhomogeneous Neumann

$$\frac{\partial u}{\partial n} = 1 \qquad \text{on } \Gamma_{iN}. \tag{3.13}$$

The weak formulation reads

$$\text{find } u \in V \text{ such that } a(\mu, u, v) = f(v) \quad \forall v \in V, \tag{3.14}$$

where

$$a(\mu, u, v) = \sum_{i=1}^{B_x B_y - 1} \mu_{i+1} \int_{d_{i+1}} \nabla u \cdot \nabla v + \int_{d_1} \nabla u \cdot \nabla v, \tag{3.15}$$

$$f(v) = \int_{\Gamma_{iN}} v. \tag{3.16}$$

Thus the problem is parametrized with $\mu \in \mathbb{R}^{B_x B_y - 1}$. Moreover, (3.15) depends affinely on μ, with

$$\theta_a^{(1)}(\mu) = 1, \qquad a_1(u, v) = \int_{d_1} \nabla u \cdot \nabla v, \tag{3.17}$$

$$\theta_a^{(i)}(\mu) = \mu_{i-1} \quad \text{for} \quad i = 2 \dots B_x B_x, \tag{3.18}$$

$$a(u, v)_i = \int_{d_{i-1}} \nabla u \cdot \nabla v \quad \text{for} \quad i = 2 \dots B_x B_x \quad \forall u, v \in V. \tag{3.19}$$

The linear form f is independent of μ. In Figure 3.3 we depict two solutions to the thermal block problem, for the parameter $\mu^{(1)}$ and $\mu^{(2)}$ with

$$\mu^{(1)} = (0.633, 0.003, 0.919, 12.789, 1.368, 6.832, 712.252, 10.077),$$
$$\mu^{(2)} = (0.049, 0.008, 35.889, 0.195, 4.427, 621.291, 252.26, 0.017).$$

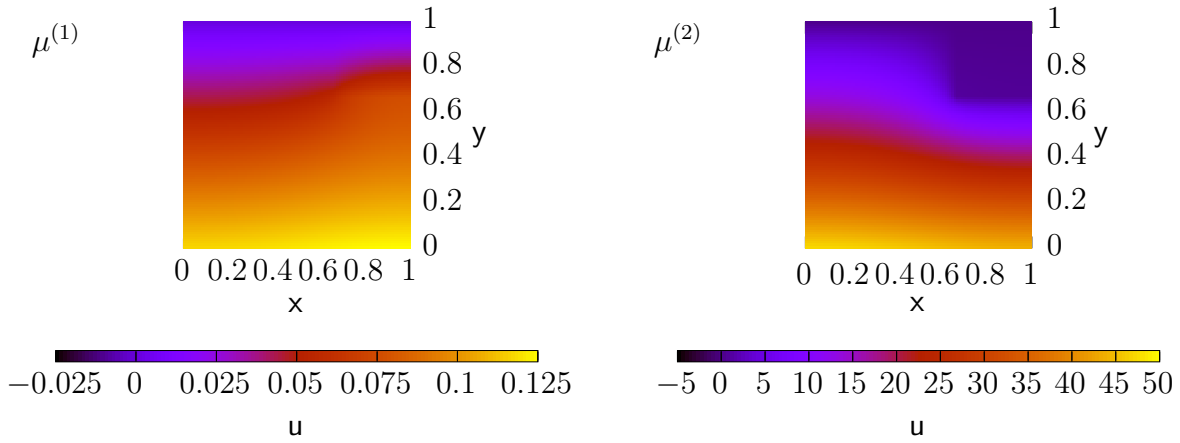


Figure 3.3: Two solutions to the thermal block problem for the parameter points $\mu^{(1)}$ and $\mu^{(2)}$. $\mu^{(1)} = (0.633, 0.003, 0.919, 12.789, 1.368, 6.832, 712.252, 10.077)$ and $\mu^{(2)} = (0.049, 0.008, 35.889, 0.195, 4.427, 621.291, 252.26, 0.017)$

Parameter	Internal heat	Thermal block
h	0.025	0.025
#d.o.f.	1943	1985
polynomial degree	1	1

Table 3.1: Parameter for the FE discretisation of (3.9) and (3.14).

Parameter	Internal heat	Thermal block
e_{max}	10^{-4}	10^{-4}
N	11	4
#samples for the interpolation of β_h	50	50
#samples in the training set Ξ	500	500

Table 3.2: Parameter for constructing the RB spaces for problem (3.9) and (3.14).

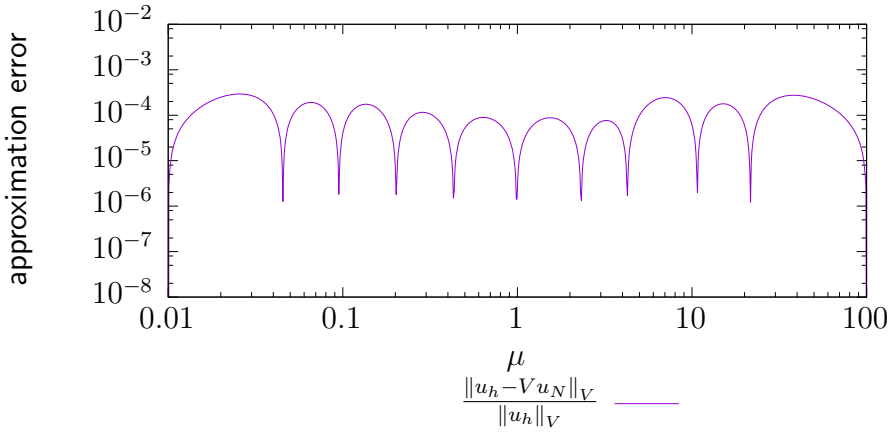


Figure 3.4: RB approximation error over the parameter space for the internal heat problem (3.9).

3.3 Numerical results

In order to test the performance of the RBM, we run several numerical experiments. The parameter of the underlying FE system are shown in Table 3.1. The RBM method is implemented as described in Chapter 2. The RB spaces are constructed using the greedy algorithm, see Section 2.4. The parameter that are used to construct the RB space are shown in Table 3.2.

Figure 3.4 shows the RB approximation error $\|u_h - Vu_N\|_V$ for the internal heat problem (3.9). The error is calculated over the whole parameter space $\mathbb{P} = [0.01, 100]$. The RB-space is constructed using the greedy algorithm, see Section 2.4.

The down spikes correspond to the selected snapshots. This highlights the local nature of RB spaces.

In order to compare strategies different RB spaces are constructed. The dimension of the space is kept fixed. The resulting spaces are used to approximate the solutions at different points in the parameter space. At each point, the approximation error $\frac{\|u_h - Vu_N\|_V}{\|u_h\|_V}$, is calculated and averaged. The results are shown in Figure 3.5 and 3.6.

For the internal heat problem the spaces are constructed using the greedy algorithm (greedy), snapshots at equidistant parameter points, snapshots at logarithmically distributed parameter points (equi-log.) and snapshots at random parameter points (random). For the thermal blocks problem only the greedy algorithm is used.

For some values of N the average approximation error of the equidistant space could not be calculated. This is due to the linear dependence of snapshots, which lead to singular A_N matrices.

The approximation errors for the greedy and equi-logarithmic spaces show similar convergence rates. For both spaces the error converges exponentially. This can be expected based on the theory presented in Sections 2.3 and 2.4. Both spaces reach a point where the approximation error does not decrease any further. For the equi-logarithmic spaces the lowest error is reached with dimension $N = 25$, for the greedy spaces the lowest error is reached at $N = 26$. The increase of the approximation error for higher values of N is due to the linear dependence of the selected snapshots.

The greedy algorithm converges even faster for the thermal blocks problem. This is surprising since the parameter space for the thermal blocks problem is of higher dimension compared to the internal heat problem. This shows that a higher dimension of the parameter space does not automatically lead to a higher dimensional RB space.

Table 3.3 shows the dimension of greedy RB spaces constructed for different parameter spaces. All spaces are constructed using the same parameter for the greedy algorithm. The results show that dimension of the RB space depends weakly on the length of the parameter space.

Based on the results we conclude that RBMs are an effective method to approximate solutions

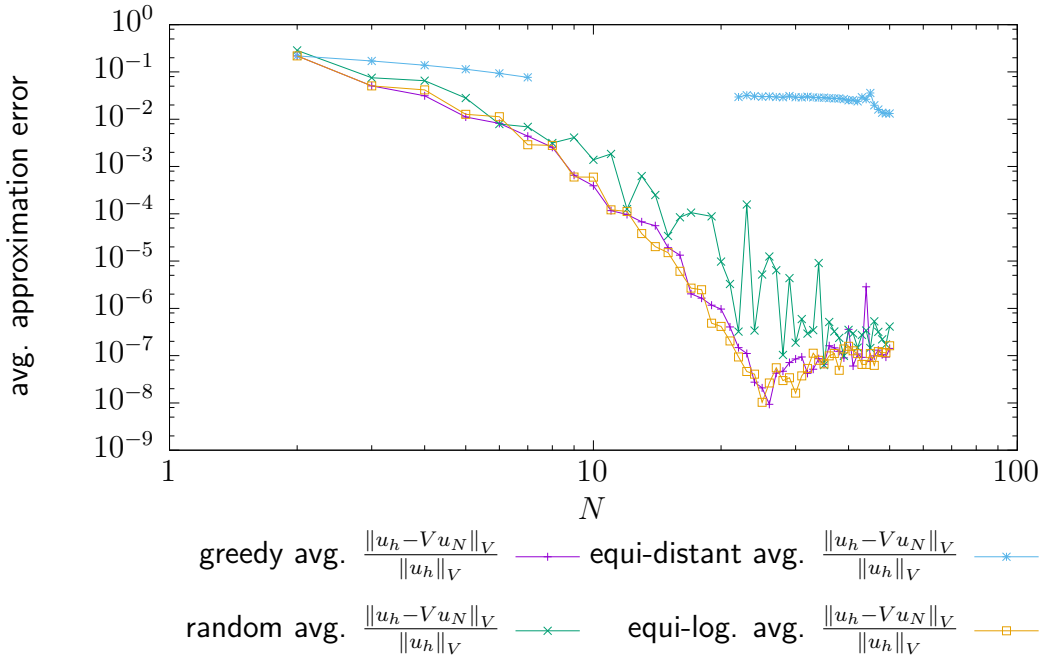


Figure 3.5: Average approximation error for different RB spaces for problem (3.9).

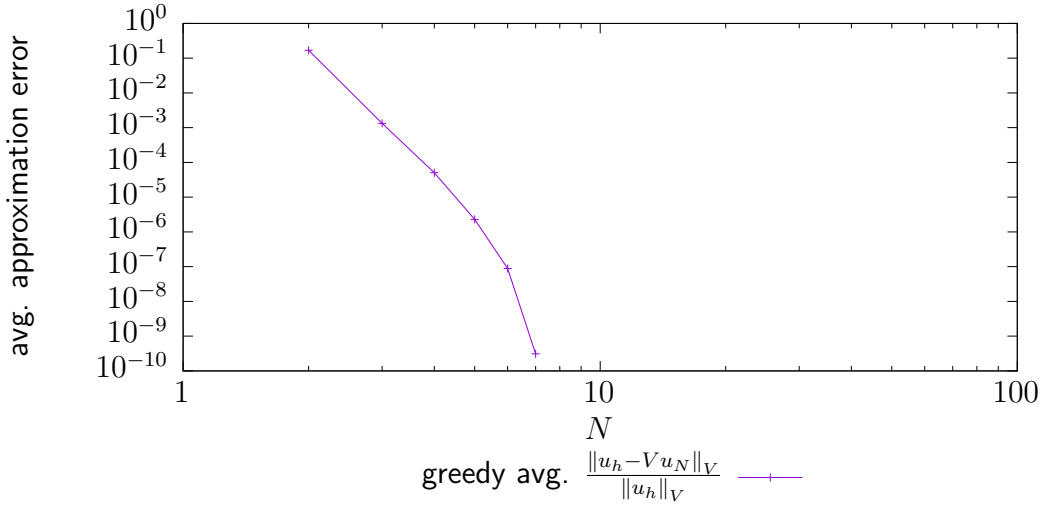


Figure 3.6: Average approximation error for the RBM using the greedy algorithm for problem (3.14).

parameter space	dimension
$[0.1 : 10]$	6
$[0.01 : 100]$	7
$[0.001 : 1000]$	7

Table 3.3: Parameter spaces and RB dimensions for problem (3.9)

to coercive problems. This is in line with the theory and is also confirmed by other works [Ver03, Chapter 3 and 4] or [Rov03, Chapter 3]. We acknowledge that the problems we have dealt with so far are very simple. The RBM applied to a more involved problem might show different behaviour and thus needs higher dimensional RB spaces. Furthermore, depending on the problem the RBM could be more sensitive to changes of the parameter space; see for example [QMN16, Section 3.8].

Chapter 4

Reduced basis methods for the Helmholtz equation

In order to test the performance of the RBM on noncoercive problems, two examples of the Helmholtz equation are considered. One models the propagation of a plane wave, the second the transmission/reflection of a wave in two fluids. The examples are taken from [BNP18] and [KMW15].

Let $k \in \mathbb{C}$ and denote the domain by $\Omega \subset \mathbb{R}^2$. The general form of the Helmholtz equation is

$$\text{find } u \in V \text{ such that } -\Delta u - k^2 u = g(k) \text{ in } \Omega. \tag{4.1}$$

Equation (4.1) can be endowed with either Dirichlet or Robin boundary conditions. V is either $H_0^1(\Omega)$ or $H^1(\Omega)$ depending on the choice of the boundary conditions. Also the functions in V are complex valued. Further, we introduce the following norm (see [Mel95]): given $w > 0$,

$$\|v\|_{V,w}^2 = \|\nabla v\|_{L^2(\Omega)}^2 + w^2 \|v\|_{L^2(\Omega)}^2 \quad \forall v \in V. \tag{4.2}$$

Define the set $\Lambda = \{\lambda_i\}_{i\in\mathbb{N}}$, where λ_i is the i-th Dirichlet-Laplace eigenvalue on Ω. Let u_i be the corresponding eigenfunction. We have

$$\int_\Omega \nabla u_i \overline{\nabla v} = \lambda_i \langle u_i, v\rangle_{L^2(\Omega)} \quad \forall v \in V. \tag{4.3}$$

Given a function $g \in L^2(\Omega)$, the weak formulation of (4.1) reads

$$\text{find } u \in V \text{ such that } \int_\Omega \nabla u \cdot \overline{\nabla v} - k^2 \int_\Omega u\bar{v} = \int_\Omega g(k,\cdot)\bar{v} \quad \forall v \in V. \tag{4.4}$$

This defines the following bilinear and linear forms:

$$a(k,u,v) = \int_\Omega \nabla u \cdot \overline{\nabla v} - k^2 \int_\Omega u\bar{v}, \tag{4.5}$$

$$f(k,v) = \int_\Omega g(k,\cdot)\bar{v}. \tag{4.6}$$

Different from (3.2), the bilinear form $a(k,\cdot,\cdot)$ is not coercive. Rather (4.5) fulfils the inf-sup and continuity condition, see (2.10) and (2.3). Theorem 2 still ensures well posedness of the weak formulation (4.4) for $k^2 \notin \Lambda$; see [QMN16].

Unlike in the coercive case, the inf-sup condition for the discrete space does not automatically follow from the continuous space (see Section 2.1.2). Instead, it must be required explicitly that the discrete FE space fulfils the inf-sup condition (see [Sch74]). This means that the space V_h cannot be chosen arbitrary. It must be fine enough to allow the FE solutions to converge.

The sesquilinear form in (4.5) fulfils the affine parameter dependence. Hence,

$$a(k,u,v) = \theta_a^{(1)} a_1(u,v) + \theta_a^{(2)} a_2(u,v), \tag{4.7}$$

with

$$Q_a = 2, \qquad Q_f = 1,$$
$$\theta_a^{(1)}(k) = 1, \qquad \theta_a^{(2)}(k) = -k^2.$$

As for the linear form f in (4.6) it depends on the function $g(k,\cdot)$ whether the problem depends affinely on the parameter or not.

Remark. *When applying the RBM to the Helmholtz equation we have to take care that the parameter space $\mathbb{P}$ does not contain any Dirichlet Laplace eigenvalue ($\mathbb{P} \cap \Lambda = \emptyset$). Further, we to need ensure that the underlying FE discretisation is fine enough to approximate the FE solutions for the max. wavenumber $k_{max} = max(k)$ accurately. See [EM12], for the interplay between k and h.*

4.1 Plane wave

The first model problem for the Helmholtz equation is that of a travelling plane wave, where the parameter is the wave number k. A wave travels through the domain $\Omega = (0,\pi) \times (0,\pi)$ along the direction $d = (d_1, d_2) \in \mathbb{R}^2$. We set the exact solution to $u_{ex} = b(x)w(x)$, where $w(x) = e^{-iv\langle d,x\rangle_2}$ and $b(x) = \frac{16}{\pi^4} x_1 x_2 (x_1 - \pi)(x_2 - \pi)$. The right-hand side is defined accordingly:

$$\begin{aligned} g(k,x) =& -\Delta u_{ex}(x) + k^2 u_{ex}(x) \\ =& \frac{16}{4\pi^4} e^{-iv\langle d,x\rangle_2} \left(2ivd_1 \left(2x_1x_2^2 - 2\pi x_1 x_2 - \pi x_2^2 + \pi^2 x_2 \right) \right. \\ & + 2ivd_2 \left(2x_1^2 x_2 - \pi x_1^2 - 2\pi x_1 x_2 + \pi^2 x_1 \right) \\ & \left. - \left(2x_2^2 - 2\pi x_1 x_2 + 2x_1^2 - 2\pi x_1 \right) \right). \end{aligned} \tag{4.8}$$

We endow the Helmholtz equation with homogeneous Dirichlet boundary conditions and set $d = (cos(\frac{\pi}{6}), sin(\frac{\pi}{6}))$. In Figure 4.1 we depict the exact solution. The parameter interval is $\mathbb{P} = [1, 12]$.

Since (4.8) does not exhibit affine parameter dependence, the offline-online procedure (see Figure 2.1) can not be applied to the right-hand side. Instead, the right-hand side has to be projected in the online stage.

4.2 Transmission/Reflection

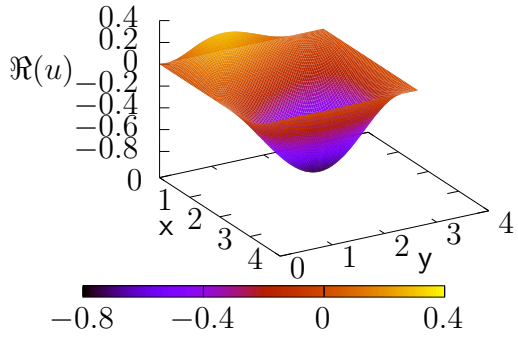
The second problem is that of transmission/reflection of a plane wave through a fluid-fluid interface. A plane wave $e^{ik\langle d,(x,y)\rangle_2}$, travelling along $d = (cos(\theta), sin(\theta))$, through an interface between two fluids with different refractive indices $n_1 < n_2$. We assume $\Omega = (-1,1)^2$ to be the domain. The fluids divide the domain horizontally in two different parts. One fluid occupies the upper half, the other the lower half. The domain is shown in Figure 4.2.

The problem is parametrized through k, n_1, n_2. The resulting parameter space $\mathbb{P}$ is given as: $\mathbb{P} = \mathbb{K} \times \mathbb{N}_1 \times \mathbb{N}_2$ where $\mathbb{K} \subset \mathbb{C}$ and $\mathbb{N}_1, \mathbb{N}_2 \subset \mathbb{N}$ and $k = p_1$, $n_1 = p_2$ and $n_2 = p_3$ for $p \in \mathbb{P}$. The problem we are interested in is:

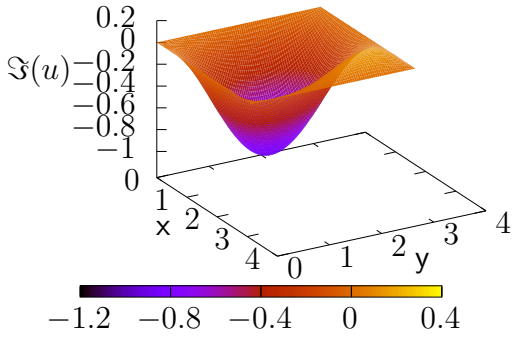
$$\text{find } u \in V \text{ such that } -\Delta u - p_1^2 \epsilon_r^2 u = 0 \qquad \text{with} \quad \epsilon_r(p,x,y) = \begin{cases} p_2 & \text{if} \quad y < 0 \\ p_3 & \text{if} \quad y > 0 \end{cases}. \tag{4.9}$$

For any angle $0 \le \theta \le \frac{\pi}{2}$ the exact solution to (4.9) is given by:

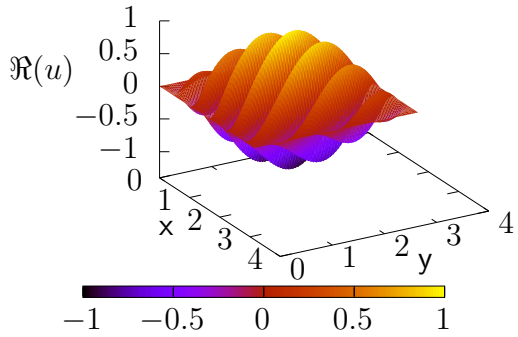
$$u_{ex}(p,x,y) = \begin{cases} T(p) e^{i\langle K(p),(x,y)\rangle_2} & \text{if} \quad y > 0 \\ e^{ip_1p_2\langle d,(x,y)\rangle_2} + R(p) e^{ip_1p_2\langle d,(x,-y)\rangle_2} & \text{if} \quad y < 0 \end{cases}, \tag{4.10}$$



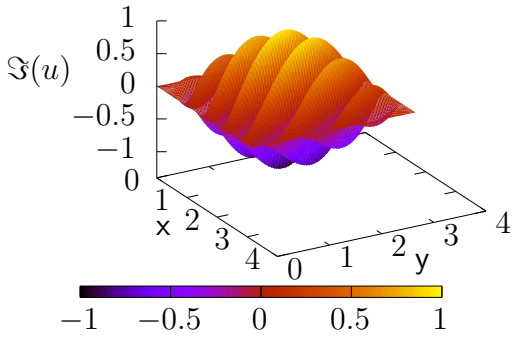
(a) Real part for $k = 1$



(b) Imaginary part for $k = 1$



(c) Real part for $k = 12$



(d) Imaginary part for $k = 12$

Figure 4.1: Exact solutions to (4.1) for the plane wave problem for different values of k.

with

$$K(p) = (p_1 p_2 d_1, \sqrt{(kp_3)^2 - (p_1 p_2 d_1)^2}),$$
$$R(p) = \frac{p_2 d_2 - K_2}{p_2 d_2 + K_2},$$
$$T(p) = 1 + R(p).$$

Depending on the angle θ, the wave either decays for $x_2 > 0$ or is refracted.

Equation (4.9) is endowed with Dirichlet boundary conditions that are taken from the exact solution (4.10) i.e. $u|_{\partial\Omega} = u_{ex}|_{\partial\Omega}$. The exact solution, for different values of k, is shown in Figure 4.3.

The weak formulation of (4.9) reads:

$$\text{find } \dot{u}(p) \in V \text{ such that } \int_\Omega \nabla \dot{u}(p,\cdot)\overline{\nabla v} - p_1^2 \int_\Omega \epsilon_r^2(p,\cdot)\dot{u}(p,\cdot)\bar{v} = p_1^2 \int_\Omega \epsilon_r^2(p,\cdot) w_g(p,\cdot)\overline{v}$$
$$\forall v \in V, \forall p \in \mathbb{P}, \quad (4.11)$$

where $g_\Omega(p) = u_{ex}(p)|_{\partial\Omega}$, $w_g(p) \in H^1(\Omega)$ is the unique harmonic extension of $g_\Omega(p)$, i.e. $\Delta w_g(p) = 0$ in Ω and $w_g(p)|_{\partial\Omega} = g_\Omega(p)|_{\partial\Omega}$, and $\dot{u}(p) = u(p) - w_g(p)$. Equation (4.11) defines the following bilinear and linear forms:

$$a(p, \dot{u}, v) = \int_\Omega \nabla \dot{u}(p,\cdot) \cdot \overline{\nabla v} - p_1^2 \int_\Omega \epsilon_r^2(p,\cdot)\dot{u}(p,\cdot)\bar{v}, \quad (4.12)$$
$$f(p, v) = p_1^2 \int_\Omega \epsilon_r^2(p,\cdot) w_g(p,\cdot)\overline{v}. \quad (4.13)$$

The linear form in 4.13 does not depend on the parameter p affinely .

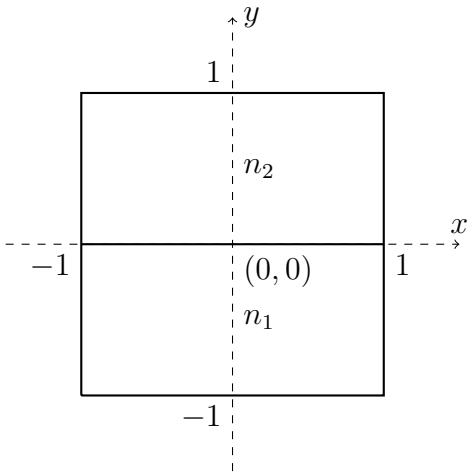
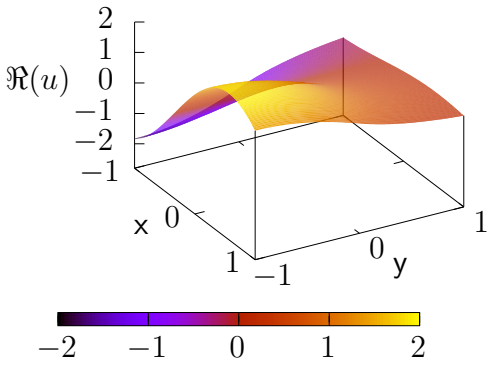
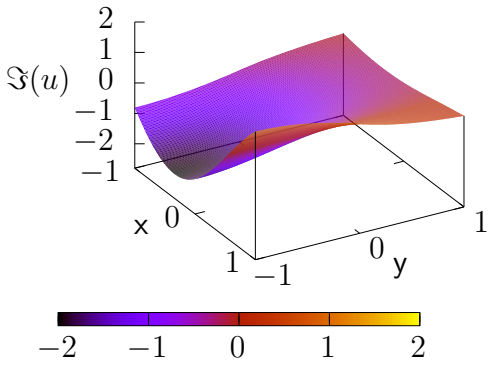


Figure 4.2: Domain of the transmission/reflection problem.

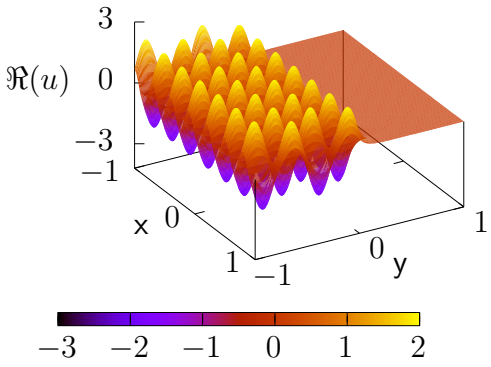
.



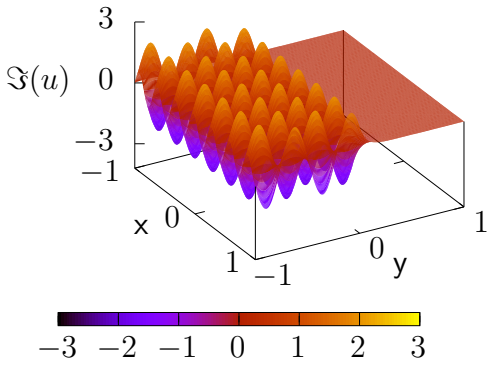
(a) Real part for $k = 1$



(b) Imaginary part for $k = 1$



(c) Real part for $k = 12$



(d) Imaginary part for $k = 12$

Figure 4.3: Exact solutions to problem (4.9).

Parameter	Plane wave	Transmission/Reflection
h	0.1	0.05
d.o.f.	10483	18376
polynomial degree	3	3

Table 4.1: Parameter for the FE discretisation of the plane wave and the transmission/reflection problem.

Parameter	Plane wave	Transmission/Reflection
e_{max}	10^{-4}	10^{-4}
N	76	119
#samples for the interpolation of β_h	50	50
#samples in the training set Ξ	500	500

Table 4.2: Parameter for constructing the RB spaces for the plane wave and the transmission/reflection problem.

4.3 Numerical results

The problems described in Sections 4.1 and 4.2 are discretised using the techniques described in Chapter 2. For the FE approximation a triangular mesh is used. The parameter of the FE system are shown in Table 4.1. The RB spaces are constructed using the greedy algorithm described in Section 2.4. The parameter that are used are shown in Table 4.2. In case the problem depends on more than one parameter, all parameter except k, are fixed and only k is varying. This means for the transmission/reflection problem n_1 and n_2 are fixed as $n_1 = 1$, $n_2 = 2$. Unless otherwise noted, k takes values from $[k_{min}, k_{max}] = [1, 12]$.

For the analysis the weighted norm defined in (4.2) is used. The weight w is chosen to be the center of the parameter interval $[k_{min}^2, k_{max}^2]$, $w^2 = \frac{k_{max}^2 - k_{min}^2}{2} + k_{min}^2$.

4.3.1 Approximation error

Figure 4.4 and 4.5 show the error of the FE and RB approximations. We can see that the RB solution (u_V) is close to the FE solution (u_h). The points with the lowest error coincide with the selected snapshots. This highlights the local nature of RB spaces. In Figure 4.4 we can see that the approximation error stays significantly below $e_{max} = 10^{-4}$. This hints that either the error estimator is not effective enough, see Section 2.4.1, or that the maximum error is dominated by high errors at only few parameter points. Interestingly this effect is less pronounced in the case of the transmission/reflection problem, see Figure 4.5.

4.3.2 Comparison of different strategies

In order to compare the greedy algorithm with other strategies for selecting parameter points, different RB spaces are constructed. Each RB space is used to calculate the approximation error on a validation set. Figure 4.6 and 4.7 show the approximation error and the selected snapshots. The error clearly shows a „local" approximation effect. For all strategies, the error is smallest in proximity of the selected snapshots. Moreover, for all strategies the approximation error increases with higher wave numbers.

Figure 4.8 and 4.9 show the convergence of different RB spaces. Each construction strategy is used to construct a RB space of a given dimension. In order to measure the approximation quality the average approximation error $\frac{\|u_h - V u_N\|_{V,w}}{\|u_h\|_{V,w}}$ is calculated on a validation set.

In general all RB spaces show exponential convergence, as long as the dimension is greater than a certain critical value. The equidistant snapshots perform almost as good as the snapshots

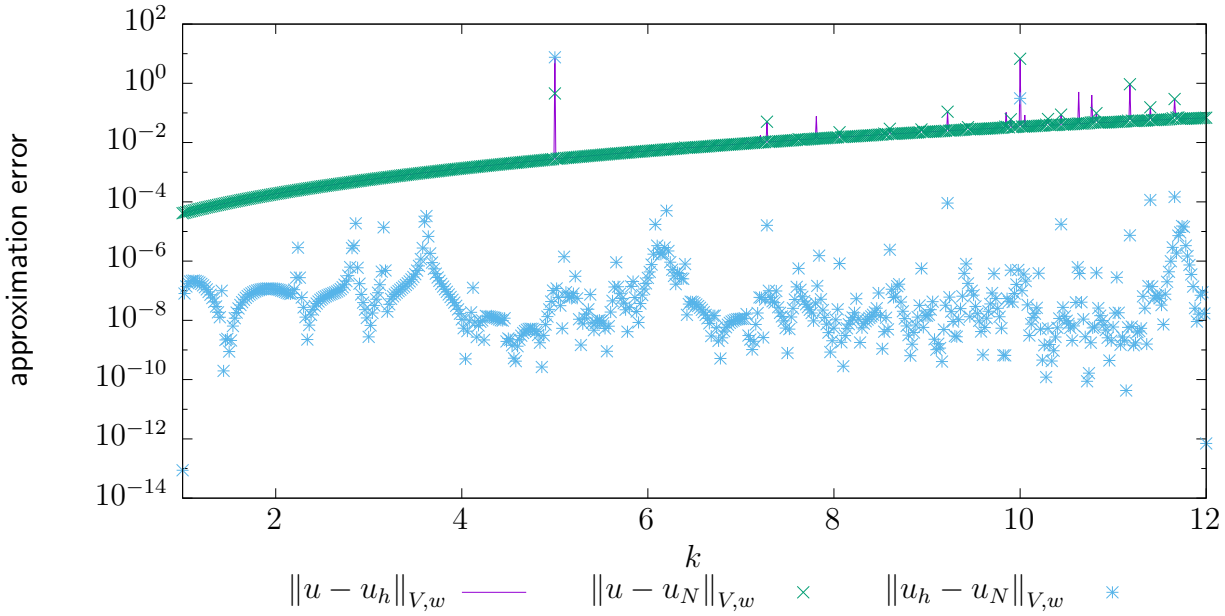


Figure 4.4: Error of solutions to the plane wave problem from Section 4.1.

chosen by the greedy algorithm. Furthermore, the convergence of the greedy algorithm behaves similar to a step function. The approximation error remains on the same order of magnitude, for few dimensions, and then drops in order.

Surprisingly the approximation error for the equidistant, equi-logarithmic, and greedy spaces is not monotonically decreasing. In the case of the greedy algorithm this is an indicator that selecting the snapshot with the highest approximation error does not lead to a lower approximation error for other solutions. In fact it might even increase the approximation error. The greedy algorithm, as described in Section 2.4, tries to extend a current RB space V_N by adding a new snapshot. The new snapshot is chosen to be the snapshot where the approximation error using the space V_N is maximized. As the Figure 4.8 and 4.9 show this does not always lead to a decrease in the average approximation error. This indicates that a snapshot where the approximation error is high is not automatically a good choice for the next snapshot. An ideal greedy algorithm would chosen the next snapshot in a way that the approximation error of V_{N+1}, see Algorithm 2, is minimized the most compared to all other choices for the next snapshot.

Compared with the results from Chapter 3 Figure 4.9 shows that noncoercive problems need higher dimensional RB spaces. Interestingly the approximation error for the equidistant RB space shows faster convergence than in the coercive case.

4.3.3 Robustness of the greedy algorithm

As detailed in Section 2.4 the greedy algorithm only provides a starting point for constructing RB spaces. The choice of the error estimator, the training set, or the parameter space might influence the resulting RB space. In order to test how each of these choices influences the resulting RB spaces we run several tests.

In the first test we construct different RB spaces using training sets Ξ from different parameter spaces $\mathbb{P} \subset [k_{min}, k_{max}]$. For each RB space we enlarge the parameter interval by increasing k_{max}. The intervals always start at $k_{min} = 2$. After the construction of the RB spaces we record their dimension and compute the RB approximation error over the largest parameter interval ($[2, 30]$). The result is shown in Figure 4.10 and Figure 4.11. The intervals boundaries are designated by the vertical black lines. For this experiment the RB spaces are constructed using a maximal error tolerance of $e_{max} = 10^{-3}$.

The results confirms the local nature of RB approximations. As soon as the parameter is

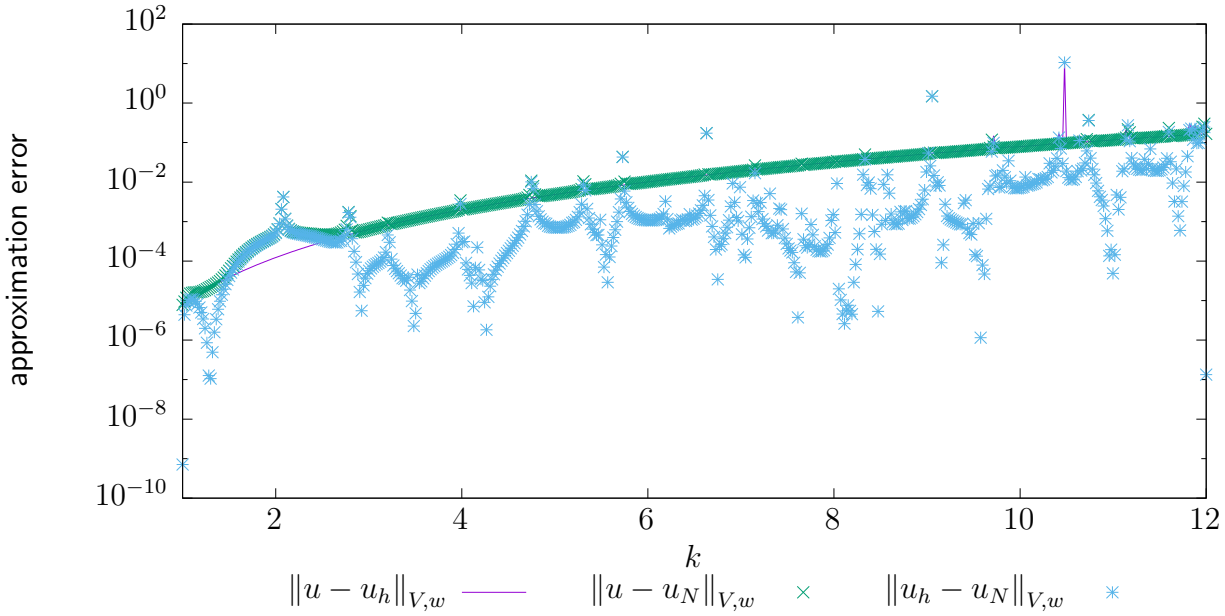


Figure 4.5: Error of solutions to problem (4.9).

outside the initial interval, the solutions start to diverge.

Figure 4.12 and 4.13 show the selected snapshots. The snapshots are always spread across the whole initial interval. Furthermore, most of the snapshots seem to be nested, in the sense that in the overlapping regions the same snapshots are chosen as for short intervals. With increasing wavenumber the snapshots are chosen in increasingly shorter intervals.

In Figure 4.14 and 4.15, the dimension of the RB spaces is shown for each interval length. The dimension increases almost linearly with the interval length. This stands in contrast to Theorem 8 and the results in Chapter 3. This reaffirms that noncoercive problems are more difficult to treat.

At the beginning of the greedy algorithm, see Algorithm 2, an initial snapshot must be chosen. In Figure 4.16 and 4.17, the selected snapshots for different choices of the initial snapshot are shown. The initial snapshots are marked with a 1. The choice of latter snapshots is slightly influenced by the initial snapshot.

In the next test, we fix a parameter interval $[1, 12]$ and construct multiple RB spaces using different training sets. The training sets are chosen by discretising the parameter interval using different step sizes h. The constructed RB spaces are used to compute the average approximation error over a validation set. In Figure 4.18 and 4.19 we depict the average error for the different step sizes. There is no correlation between the grid size and the average error.

In order to test how the underlying FE discretisation influences the resulting RB space, we construct two RB spaces, $V_{N,1}$ and $V_{N,2}$, using different FE spaces $V_{h,1}$ and $V_{h,2}$. The first FE space $V_{h,1}$ is constructed with a maximal grid size of $h_1 = 0.1$. The second FE space $V_{h,2}$ is constructed using a maximal grid size $h_2 = 0.05$. The resulting FE spaces are of dimension $dim(V_{h,1}) = 10483$ and $dim(V_{h,2}) = 41770$. The considered parameter interval is $k \in [2, 20]$. Figure 4.20 shows the norm of the exact u and the FE solution for both FE spaces. We see that the first FE space shows oscillations in the solutions for higher wave numbers.

Both FE spaces are used to construct RB spaces for the given parameter interval. The resulting RB spaces $V_{N,1}$ and $V_{N,2}$ are of dimension $dim(V_{N,1}) = 231$ and $dim(V_{N,2}) = 145$ for the same tolerance $e_{max} = 10^{-4}$. This shows that the dimension of RB spaces is not necessarily correlated with the dimension of the underlying FE space. Furthermore, for the Helmholtz problem the underlying FE space should accurately represent the solutions at all wavenumbers of interest.

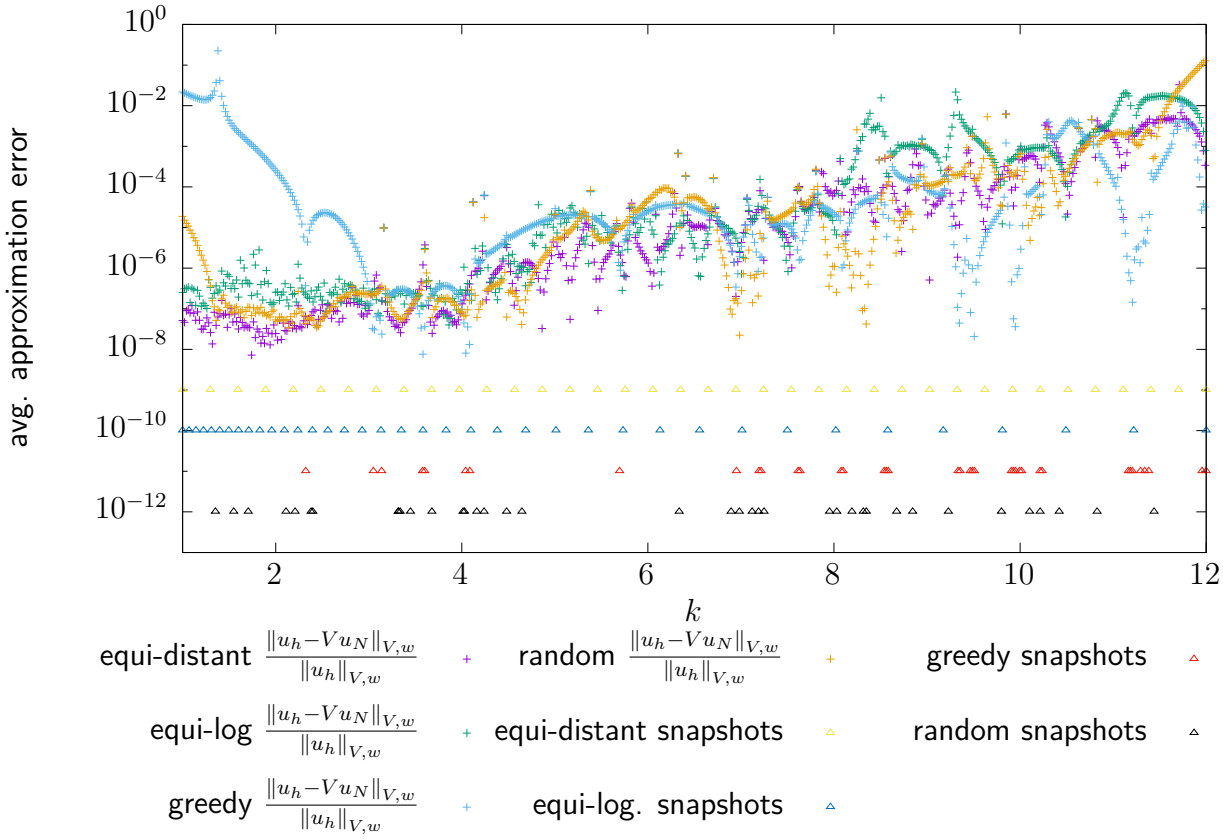


Figure 4.6: Comparison of the different construction strategies for RB spaces for the plane wave problem from Section 4.1. The approximation errors are marked by crosses. The snapshots are marked using triangles.

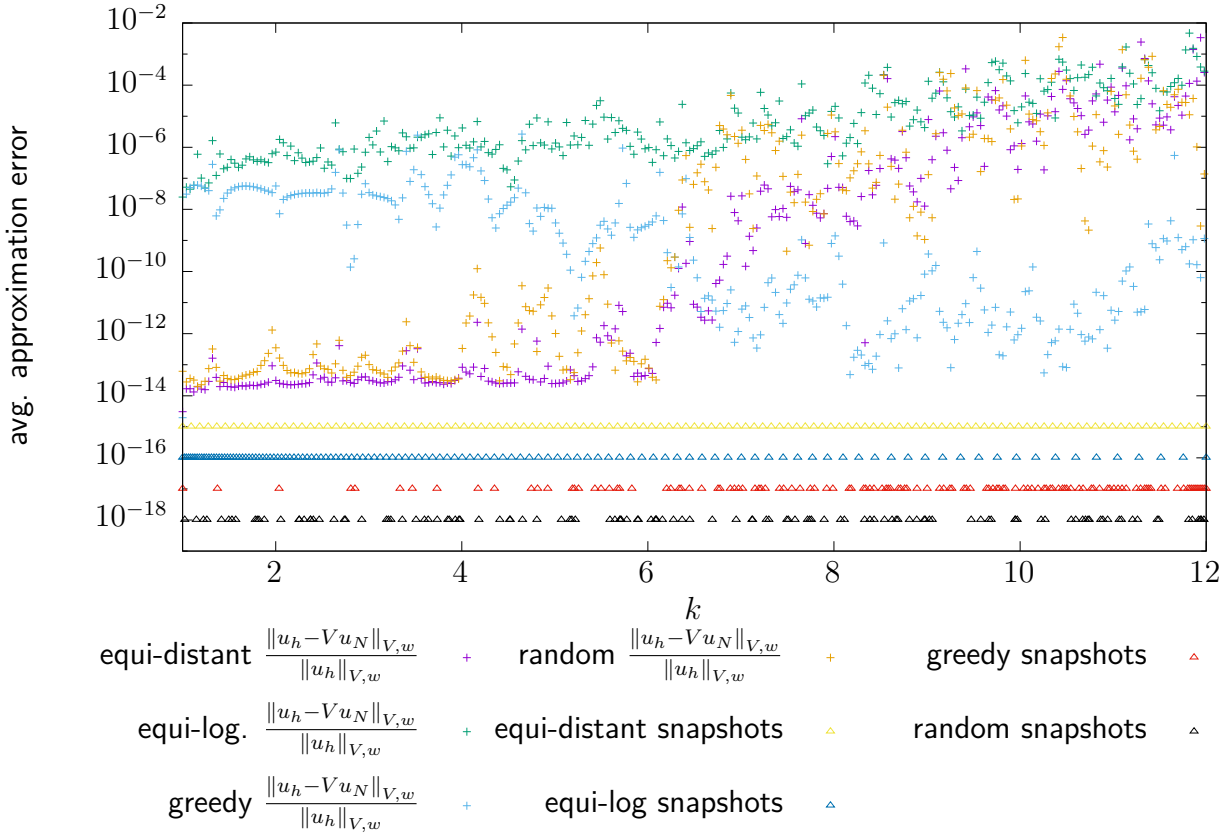


Figure 4.7: Comparison of the different construction strategies for RB spaces for the transmission/reflection problem (4.9). The approximation errors are marked by crosses. The snapshots are marked using triangles.

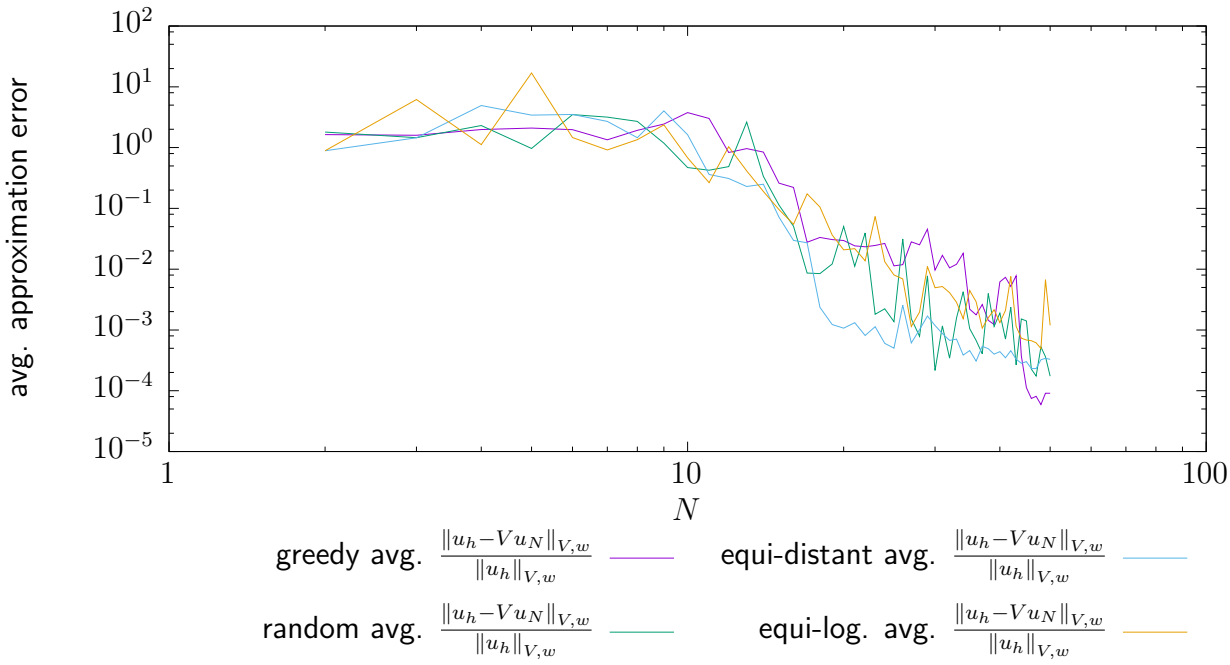


Figure 4.8: Convergence of RB spaces using different construction strategies for the plane wave problem from Section 4.1.

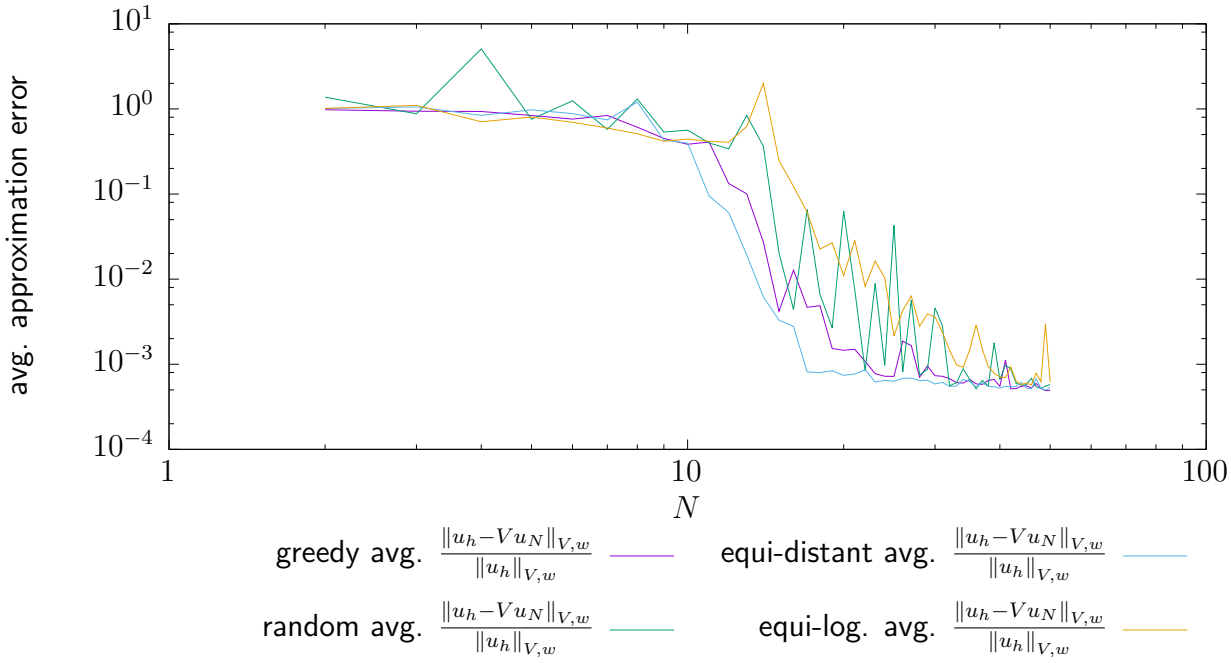


Figure 4.9: Convergence of RB spaces using different construction strategies using for the transmission/reflection problem (4.9).

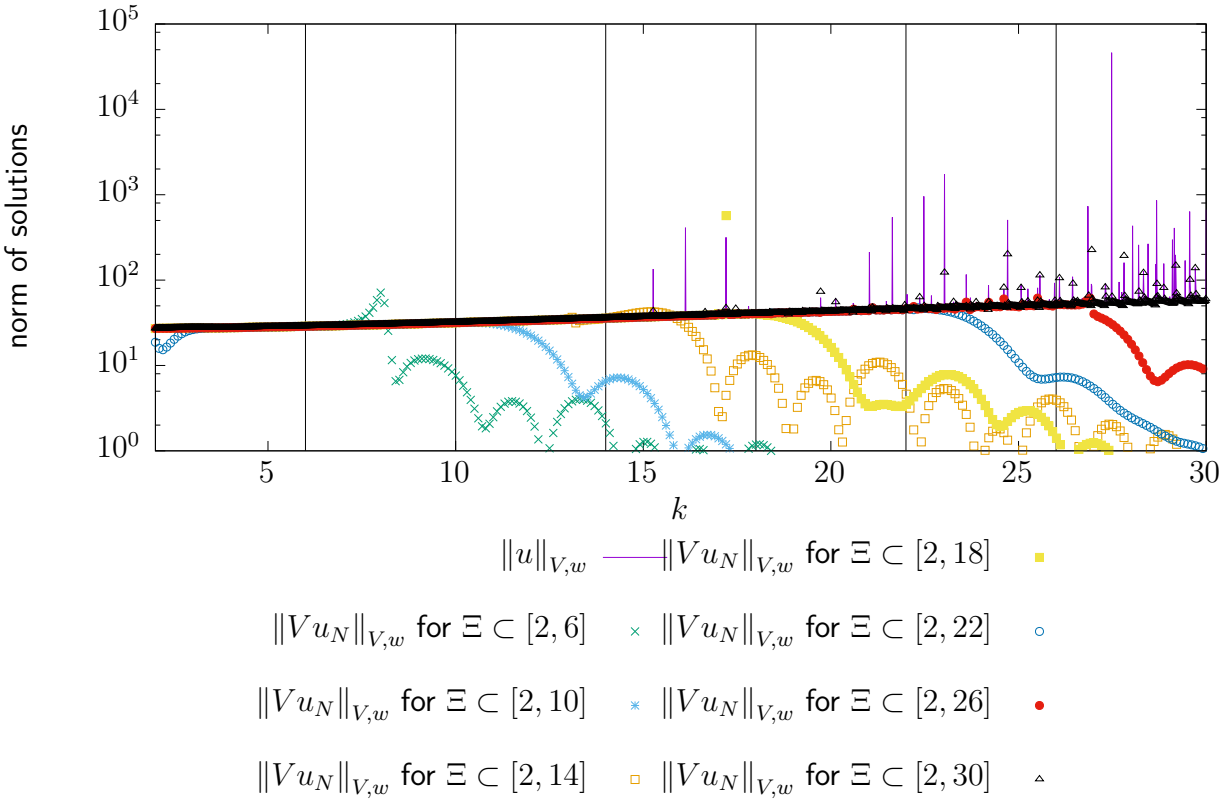


Figure 4.10: Comparison of different RB spaces for the plane wave problem from Section 4.1. The RB are constructed using training sets Ξ taken from different intervals.

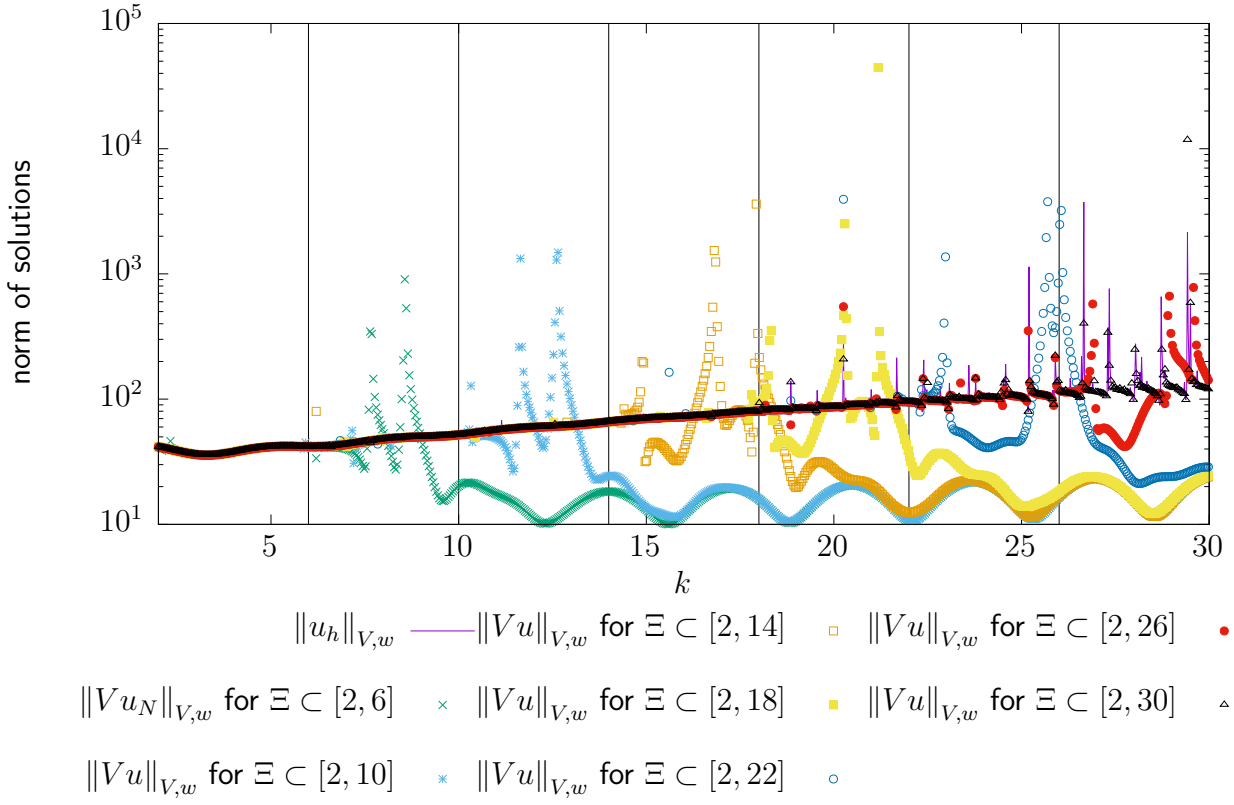


Figure 4.11: Comparison of different RB spaces for the transmission/reflection problem (4.9). The RB are constructed using training sets Ξ taken from different intervals.

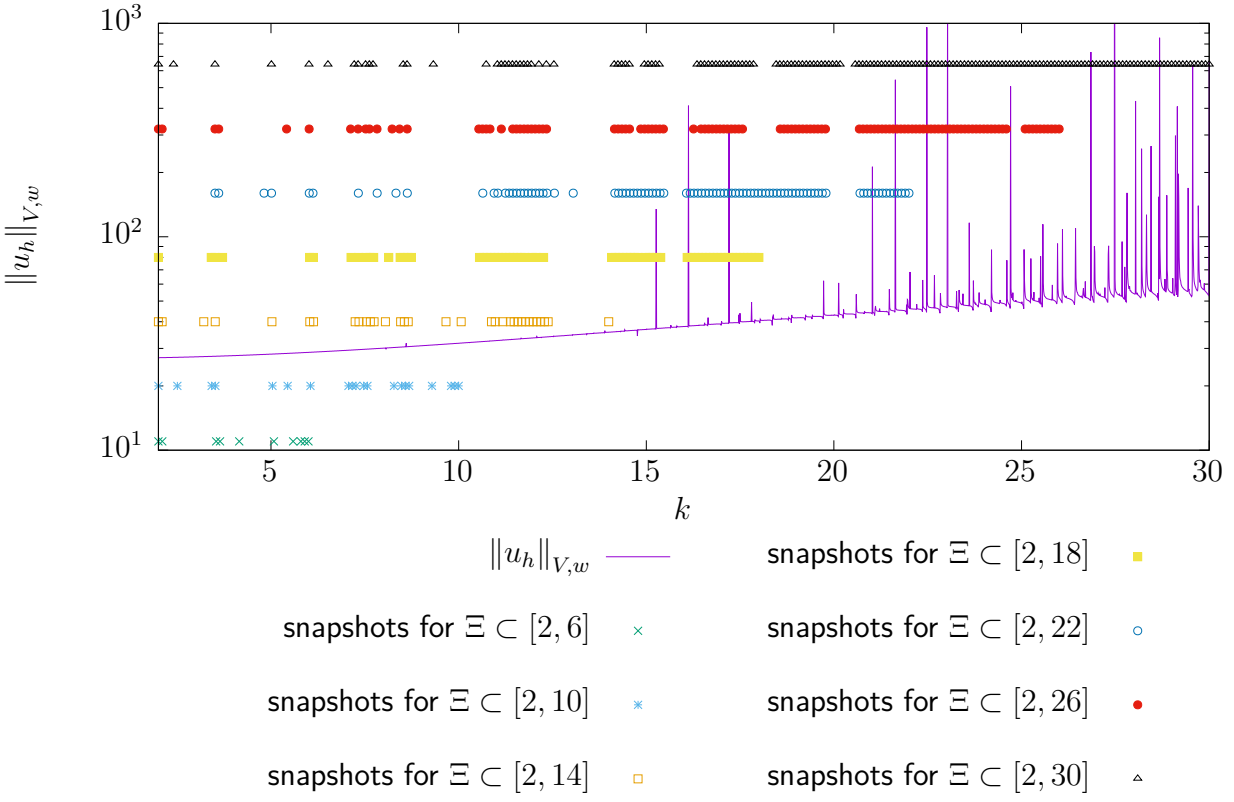


Figure 4.12: Selected snapshots to construct RB spaces for the plane wave problem from Section 4.1. The spaces are constructed using training sets Ξ taken from different intervals.

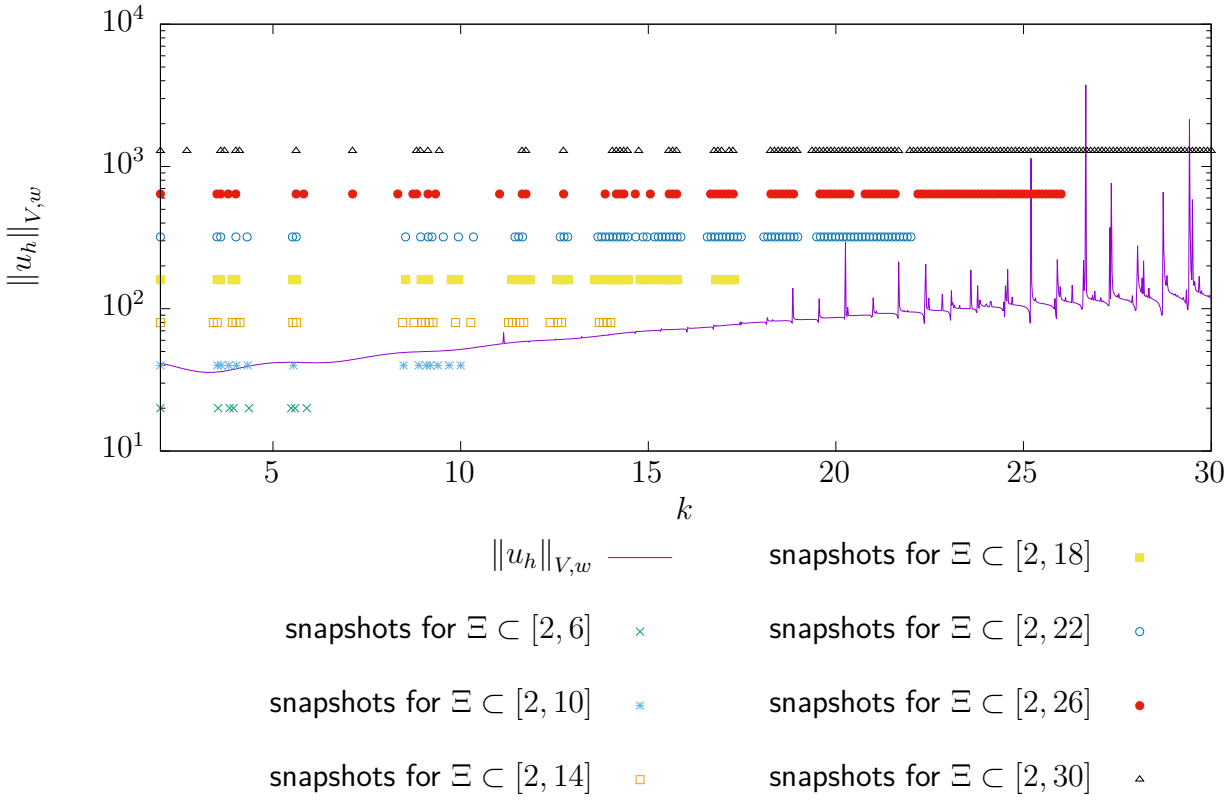


Figure 4.13: Selected snapshots to construct RB spaces for the transmission/reflection problem (4.9). The spaces are constructed using training sets Ξ taken from different intervals.

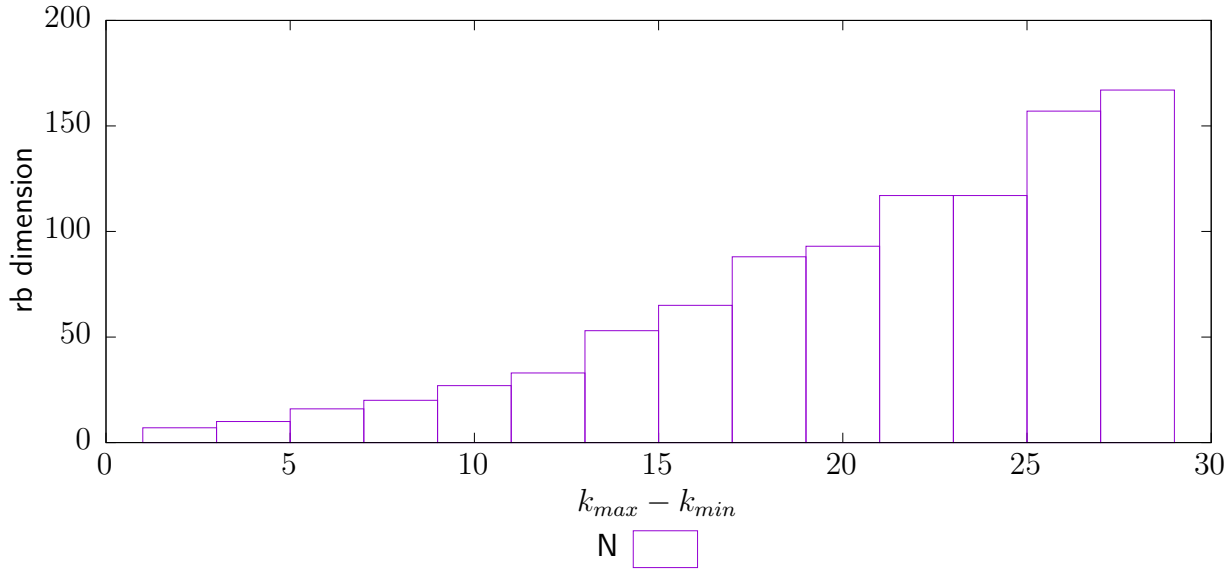


Figure 4.14: Dimensions of RB spaces depending on the interval length $k_{max} - k_{min}$ for the plane wave problem from Section 4.1.

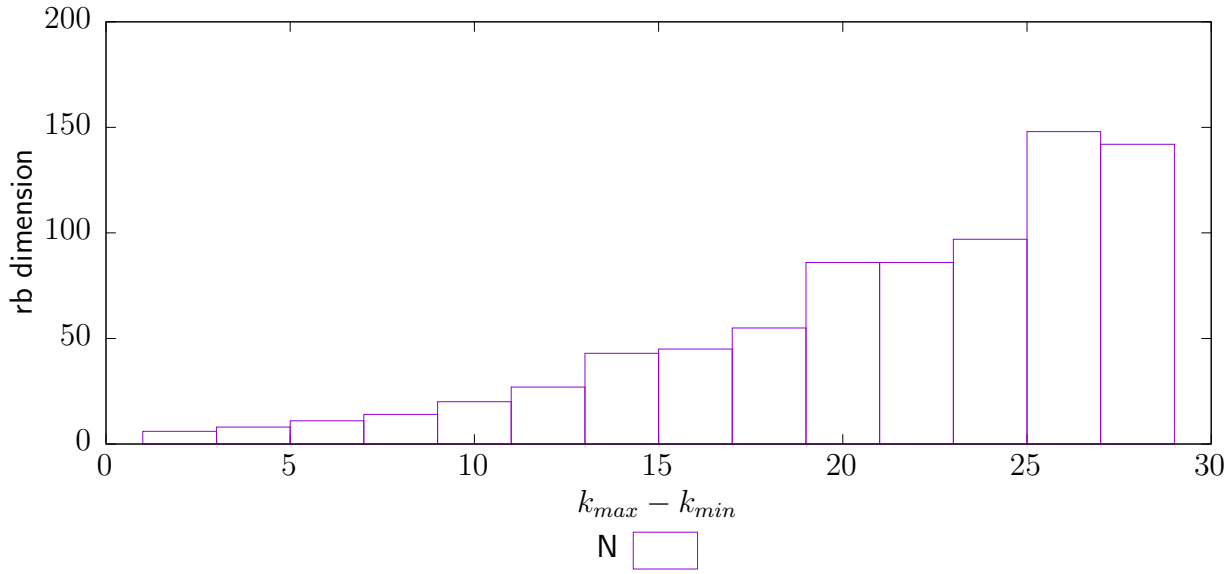


Figure 4.15: Dimensions of RB spaces depending on the interval length $k_{max} - k_{min}$ for the transmission/reflection problem (4.9).

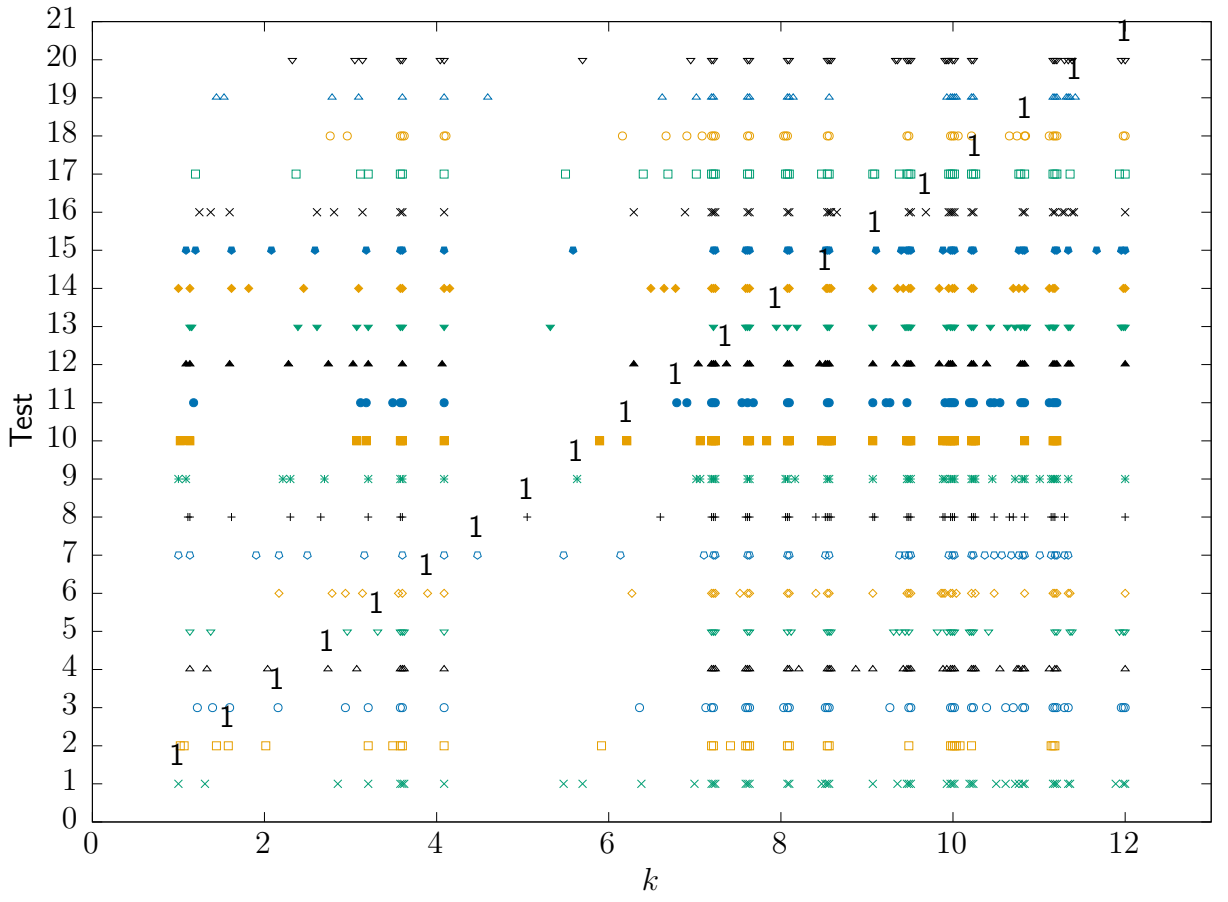


Figure 4.16: Snapshots selected by the greedy algorithm using different starting points for the plane wave problem from Section 4.1. The initial snapshot is marked by 1.

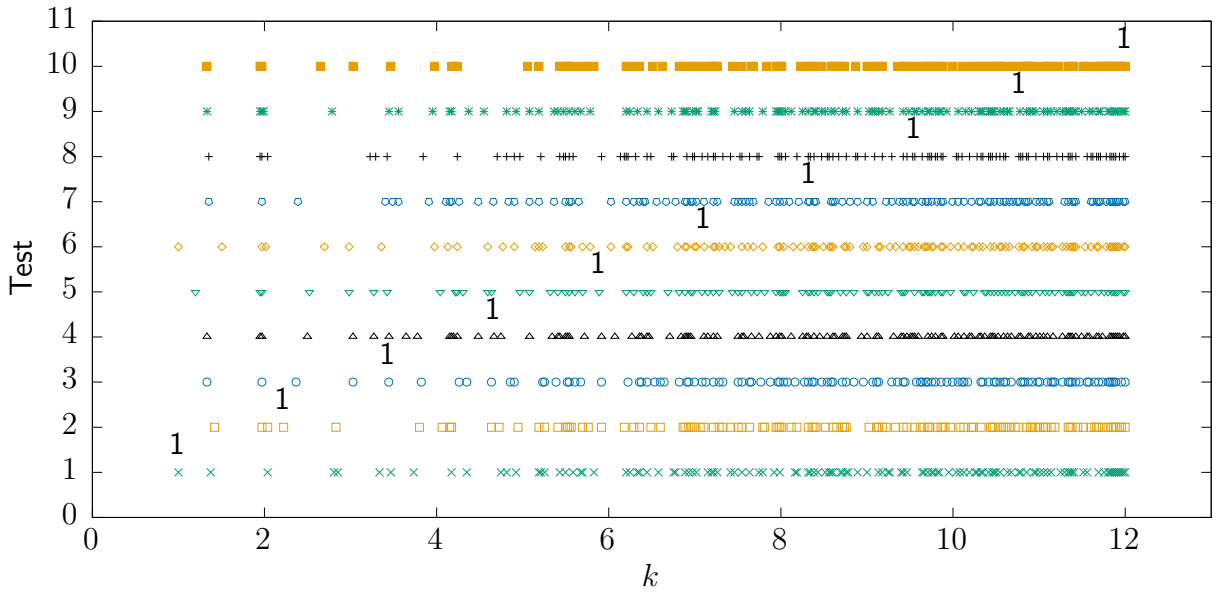


Figure 4.17: Snapshots computed by the greedy algorithm using different starting points for the transmission/reflection problem (4.9). The initial snapshot is marked by 1.

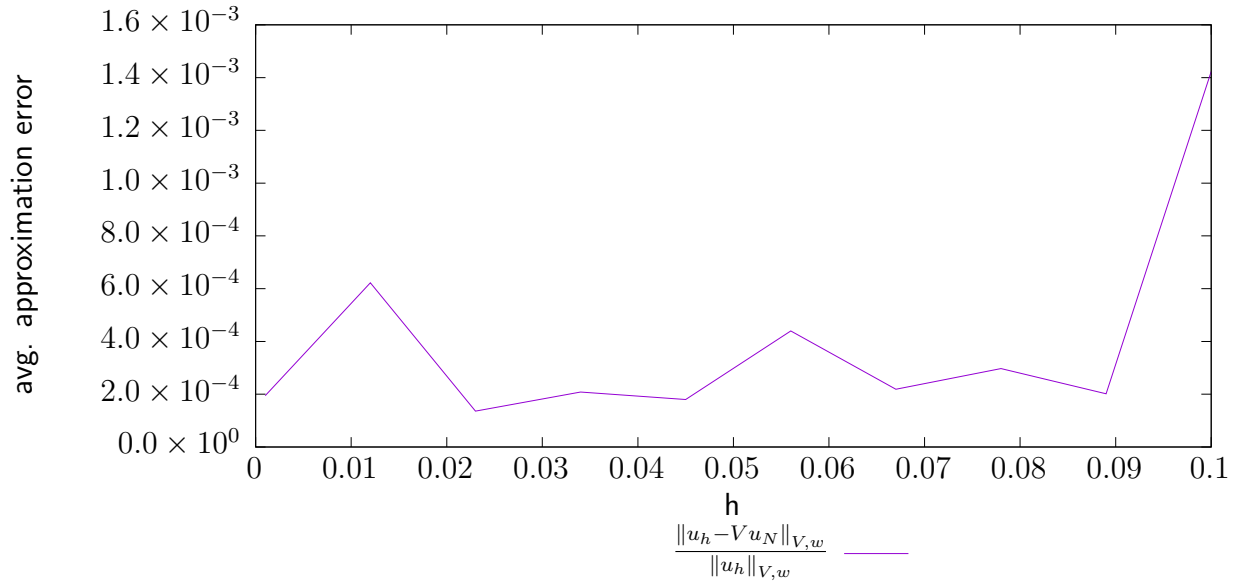


Figure 4.18: Average error for different RB spaces for the plane wave problem from Section 4.1. The RB spaces are constructed using training sets that are discretisations of the interval $[1, 12]$ with differing step sizes h.

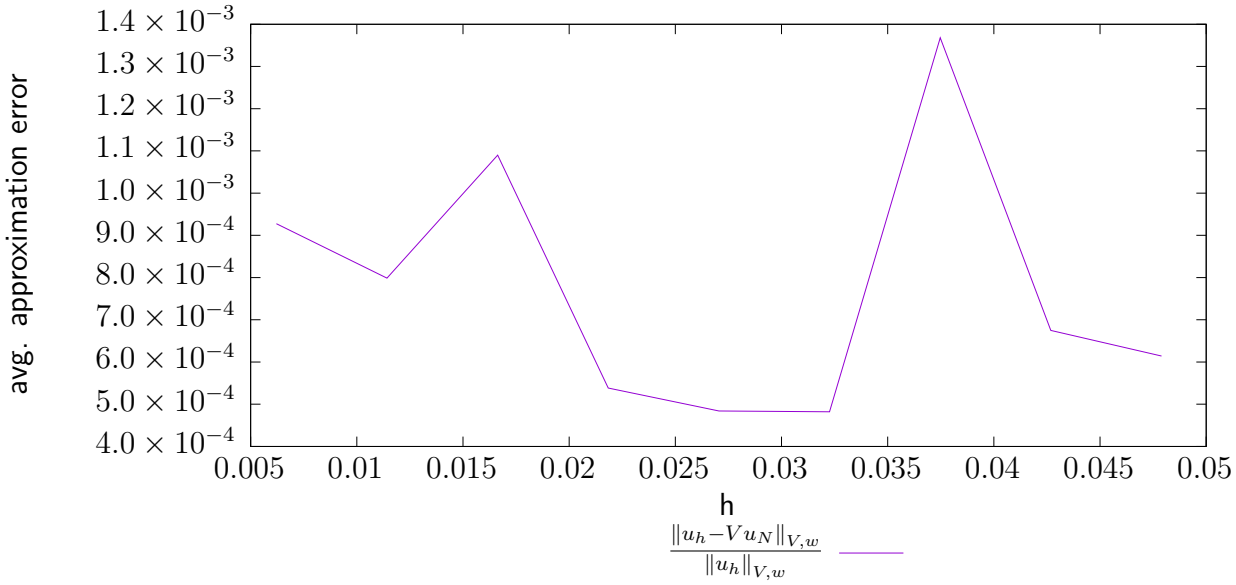
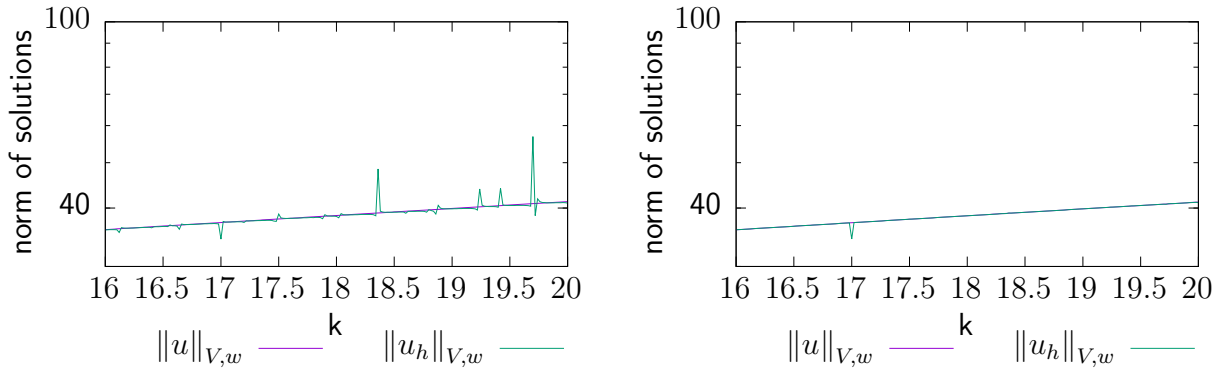


Figure 4.19: Average error for different RB spaces for the transmission/reflection problem (4.9). The RB spaces are constructed using training sets that are discretisations of the interval $[1, 12]$ with step sizes h.



(a) Norm of solutions for a FE space with grid size $h = 0.1$.

(b) Norm of solutions for a FE space with grid size $h = 0.05$.

Figure 4.20: Norm of the exact solution and the FE solutions to the plane wave problem from Section 4.1.

The numerical results presented in this chapter confirm that the RBM can be applied to noncoercive problems. Compared to the results for coercive problems, see Chapter 3, higher dimensional spaces are needed and the approximation error is generally higher. This is in line with the theory and the results in [QMN16, Chapter 2] and [Rov03, Chapter 5]. Moreover, the results show that in the case of the Helmholtz equation the underlying FE space should be chosen in such a way that all solutions can be represented accurately. It should be noted that there exist extensions of the RBM that might improve the mentioned problems e.g. the least-squares reduced basis method. For more details we refer to [Rov03, Chapter 3] and [QMN16, Chapter 5].

Chapter 5

Reduced basis methods for inverse and control problems

Two scenarios in which RBMs are useful are inverse and optimal control problems. An optimal control problem consists of a control $\mu \in \mathbb{P}$, a state $u \in V$, and a output $z \in Z$. The state is given as solution to the state problem that is parametrized by the input. The output is a function of the state $z = o(u)$.

Take the Helmholtz equation as an example. The control μ is the wavenumber k taken from the control space $\mathbb{P} = \mathbb{R}$ or $\mathbb{P} = \mathbb{C}$. The state problem is given by equation (4.1). The state space V is either $V = H^1$ or $V = H^1_0$. The state u_k is a solution to (4.1). The output could be any function depending on u_k, i.e. $o : V \to Z$. Let

$$s(k) = u_k \quad \text{such that } u_k \text{ solves (4.1) for wavenumber } k = \mu \tag{5.1}$$

be the solution map of the Helmholtz problem. The solution map connects the control to the state. Given a value of the control variable the output can be obtained by composition of s and o. Hence the input-output mapping c is given by $c = o \circ s$. Given some data z the control problem reads

$$\text{find } \mu_z \text{ such that} \quad c(\mu_z) = z. \tag{5.2}$$

Thus we want to find $c^{-1} = s^{-1} \circ o^{-1}$. Depending on s and o c^{-1} does not exist or is not well-defined. To circumvent this problem, optimal control problems are treated as an optimization problem. Instead of inverting c, we look for an μ_z that minimizes a cost function $J : \mathbb{P} \to \mathbb{R}$

$$\mu^* = \arg\min_{\mu \in \mathbb{P}} J(\mu). \tag{5.3}$$

Common choices for J, see [FZ03] or [Qua14, Chapter 17], are

$$J(\mu) = \frac{1}{2} \|c(\mu) - z\|^2 \quad \forall \mu \in \mathbb{P} \tag{5.4}$$

or for a $\alpha \leq 0$

$$J(\mu) = \frac{1}{2} \|c(\mu) - z\|^2 + \frac{\alpha}{2} \|\mu\|^2_{\mathbb{P}} \quad \forall \mu \in \mathbb{P}. \tag{5.5}$$

One approach to solve (5.3) numerically is the Model - Discretisation - Control approach, see [FZ03]. In this approach, the state problem is discretised and the discretised problem is solved when evaluating the cost function. Equation (5.3) can then be solved by a numerical optimization method, e.g. Quasi-Newton, see [NW06] or CG, see [HS52]. For a short introduction into the Quasi-Newton and CG method see Appendix A.1 and A.2.

If the state problem is a PDE, RBMs can be used to solve the state problem more efficiently. Instead of solving a problem in a high dimensional space the RBM solves the problem in a lower dimensional space. This leads to the following procedure:

1. discretise the state problem using the FEM;
2. construct a RB space based on the FE space;
3. find the optimal control μ_z for the given data.

Optimal control problem arise in a wide range of applications. A frequent example is parameter estimation. Given an output find an input that produces a similar output. This can be used to determine material properties. The control represents the unknown material properties. The state problem describes the inner dynamics of the system. The known output corresponds to physical measurements. By solving the optimal control problem, the material properties can be determined from the measurements.

5.1 Transmission/Reflection

In order to demonstrate the outlined procedure we considered three model problems. The first model problem is the transmission/reflection problem, see Section 4.2, with the domain $\Omega = (-0.5, 0.5)^2$. The state problem is given by equation (4.9). The control parameter corresponds to the parameter of the problem. Hence the control is given by $\mu = (k, n_1, n_2)$. The solution to (4.9) is taken as the output.

The parameter n_1 and n_2 could be seen as material properties that need to be determined. The parameter k fixes the frequency. Taking the coefficient vector as the output is an idealistic assumption. In any real context it would be impossible to measure the complete solution of the state problem.

5.2 Admittance identification

The second example is taken from [VH09]. It models the (simplified) interior of a car. Inside the car, a loudspeaker emits sound waves that interact with a damping material at the bottom (Γ_D) of the car door. The loudspeaker is located at the point x_q. The interaction is modelled by the Helmholtz equation in the domain Ω. The damping material makes up a part of the boundary designated as Γ_D. The domain is depicted in Figure 5.1.

The damping material is characterized by its impedance $z \in \mathbb{C}$. In order to have a linear model we try to estimate the admittance $a \in \mathbb{C}$. The admittance is given by $a = \frac{1}{z}$. The outward normal vector is denoted by n. The speed of sound is denoted by c and the density of the ambient air by $\rho_0 \in \mathbb{R}$. Given the (material) constants

$$c = 343.799\frac{m}{s}, \qquad k = \frac{2\pi f}{c}, \qquad \rho_0 = 1.19985\frac{kg}{m^3}, \qquad \omega = 2\pi f = ck, \tag{5.6}$$

the state problem is defined by

$$\text{find } u \in V \text{ such that } -\Delta u - k^2 u = g \quad \text{in } \Omega, \tag{5.7}$$

$$\frac{i}{\rho_0 \omega}\frac{\partial u}{\partial n} = 0 \quad \text{on } \Gamma_H, \tag{5.8}$$

$$\frac{i}{\rho_0 \omega}\frac{\partial u}{\partial n} = au \quad \text{on } \Gamma_D. \tag{5.9}$$

The source term g is given by

$$g(f, x) = \frac{1}{5} e^{\frac{i\pi(f-200)}{50}} e^{-50\|x - x_q\|_2^2}. \tag{5.10}$$

Hence, the problem is parametrized by $\mu = (k, a)$.

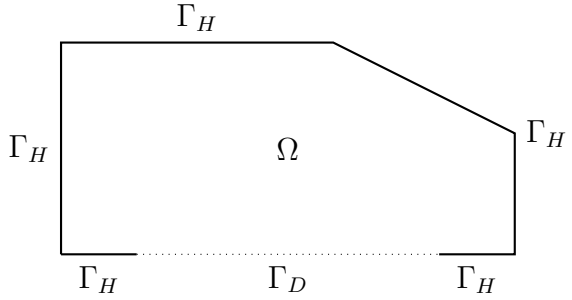


Figure 5.1: The domain and boundaries for problem (5.7).

A weak formulation of (5.7) together with (5.8) and (5.9) reads, see [Vol10],

$$\text{find } u \in V \text{ such that } a(\mu, u, v) = f(\mu, v) \qquad \forall v \in V, \forall \mu \in \mathbb{P}, \tag{5.11}$$

where

$$a(\mu, u, v) = \int_\Omega \nabla u \overline{\nabla v} - \mu_1^2 \int_\Omega u\overline{v} - i\rho_0\omega\mu_2 \int_{\Gamma_D} u\overline{v}, \tag{5.12}$$

$$f(\mu, v) = \int_\Omega g(c\mu_1, \cdot)\overline{v}. \tag{5.13}$$

Both (5.12) and (5.13) depend on the parameter μ affinely. Thus the offline-online approach for building RB spaces can be used, see Section 2.2.1.

The output $c(\mu)$ is the evaluation of the solution to (5.7) at six different measurement points. Hence, given the measurement points $p_i \in \mathbb{R}^2$ for $i = 1 \dots 6$ the output $c(\mu)$ is given as

$$c(\mu) = (u(p_1), \dots, u(p_6)), \tag{5.14}$$

where u solves (5.7) for $k = \mu_1$ and $a = \mu_2$.

5.3 Reduction of the engine noise

The third example is taken from [CHY07]. The problem describes the propagation of noise from an airplane engine. We assume the problem to be axisymmetric along the x-axis. The problem is described by

$$\Delta u + k^2 u = 0 \quad \text{in } \Omega, \tag{5.15}$$

$$u = g \quad \text{on } \Gamma_1, \tag{5.16}$$

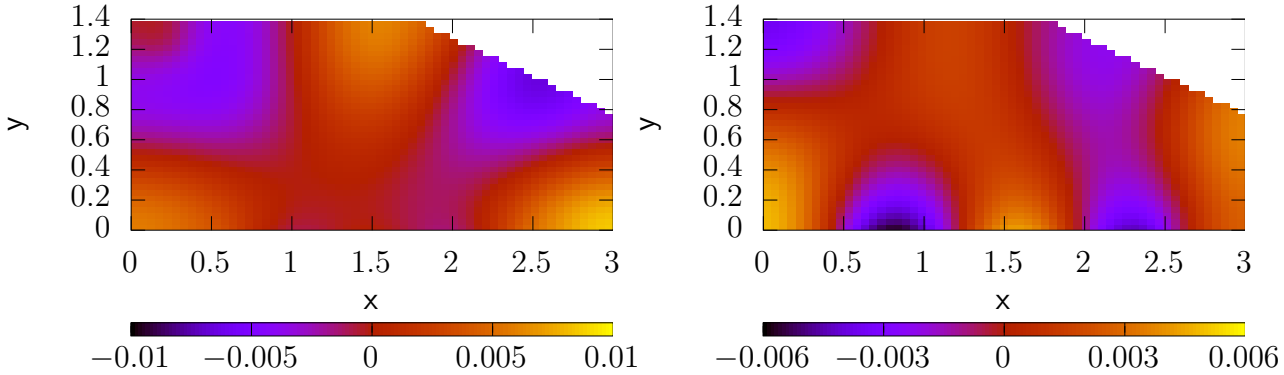
$$\left(\frac{\partial u}{\partial n} + \chi u\right) = 0 \quad \text{on } \Gamma_2, \tag{5.17}$$

$$\frac{\partial u}{\partial n} = 0, \quad \text{on } \Gamma_3 \cup \Gamma_5, \tag{5.18}$$

$$\left(\frac{\partial u}{\partial n} + iku\right) = 0 \quad \text{on } \Gamma_4. \tag{5.19}$$

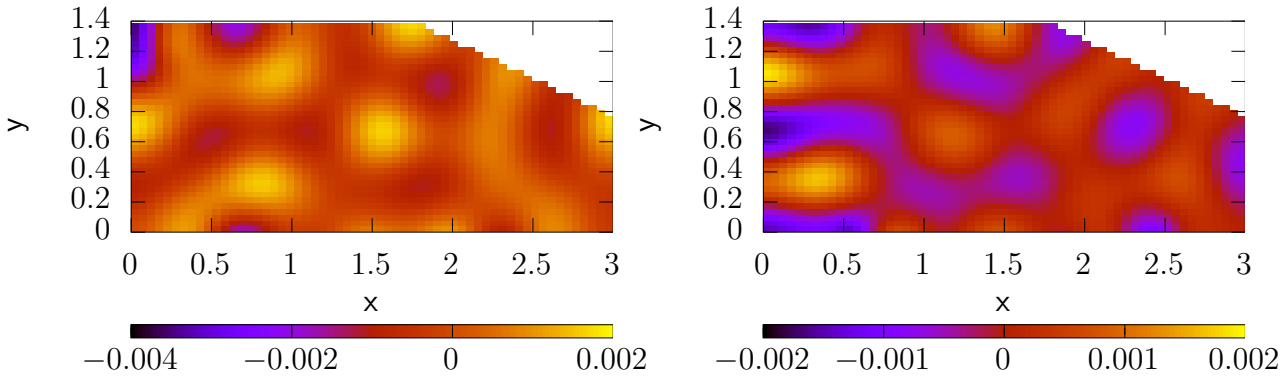
$$\tag{5.20}$$

The engine noise is modelled by the Dirichlet boundary conditions (5.16). The boundary parts Γ_2 Γ_3 represent the engine enclosing. The material of the engine enclosing is represented by $\chi \in \mathbb{C}$ in (5.17). The boundary Γ_4 is assumed to be far enough from the noise source such that Sommerfeld radiation conditions can be imposed. The boundary Γ_5 is along the symmetry axis. The domain of the problem is depicted in Figure 5.3. The well-posedness of (5.15) is proven in [CHY07]. The engine noise g is given by $g(x, y) = 20.0 + e^{3ysin(10\pi y)}$.



(a) Real part for $f = 200Hz$

(b) Imaginary part for $f = 200Hz$



(c) Real part for $f = 500Hz$

(d) Imaginary part for $f = 500Hz$

Figure 5.2: Exact solutions to problem (5.7).

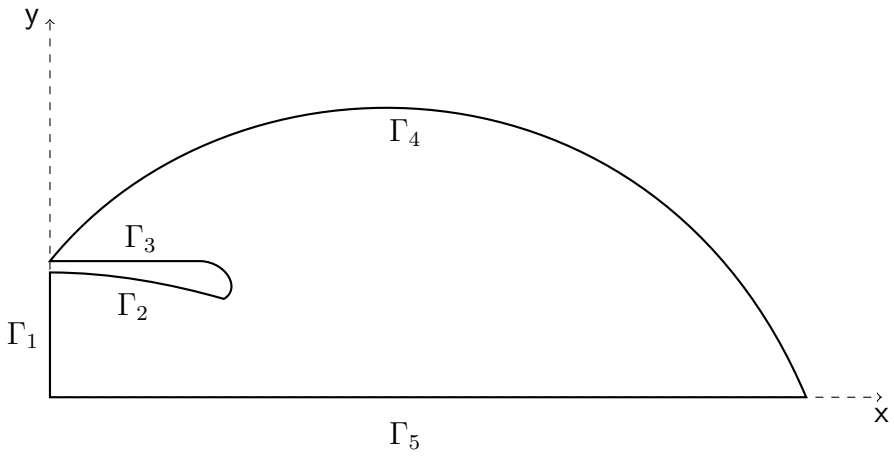


Figure 5.3: The domain and boundaries for problem (5.15).

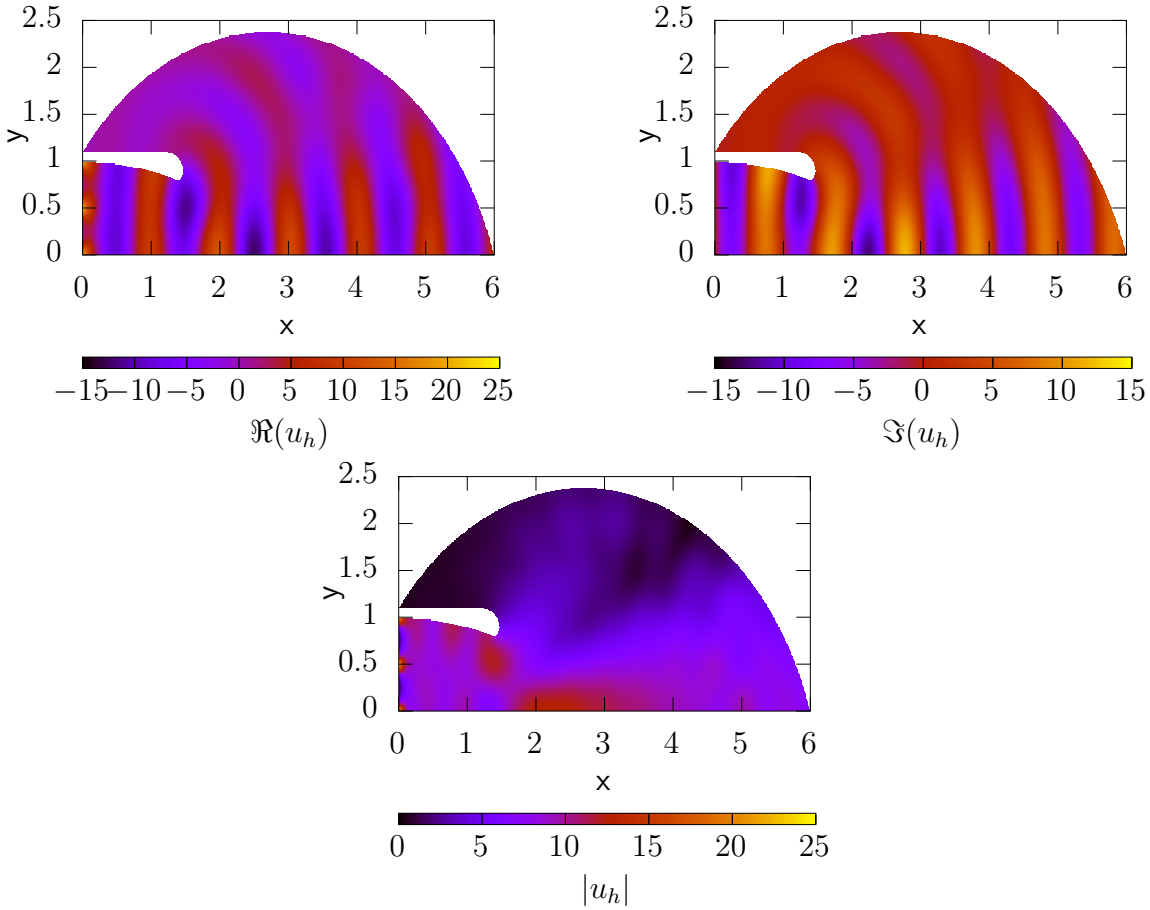


Figure 5.4: FE solution u_h to (5.21) for $\chi = 0$ and $k = 2\pi$.

The problem is parametrized by $\mu = (k, \chi) \in \mathbb{C}^2$. The weak formulation of (5.7) reads

$$\text{find } u \in V \text{ such that } a(\mu, u, v) = -a(\mu, w_g, v) \quad \forall v \in V, \forall \mu \in \mathbb{P}, \tag{5.21}$$

where

$$a(\mu, u, v) = \int_\Omega \nabla u \overline{\nabla v} - \mu_1^2 \int_\Omega u\overline{v} + \chi \int_{\Gamma_2} u\overline{v} + ik \int_{\Gamma_4} u\overline{v}, \tag{5.22}$$

$$w_g|_{\Gamma_2} = g. \tag{5.23}$$

Since the bilinear form $a(\mu, \cdot, \cdot)$ in (5.22) depends affinely on the parameter μ we use the offline-online approach for building RB spaces. In Figure 5.4 we depict the solution to (5.21) for $\chi = 0$.

The aim of the current problem is not to find the best control given an observation but to find a value for χ such that the amount of noise emitted from the engine is minimal. Hence, the cost function $J(\mu)$ is identical to the output. Given two constants $\alpha \leq 0$ and $\beta \leq 0$, the output $c(\mu)$ for the optimal control problem is given by

$$c(\mu) = J(\mu) = \alpha \int_\Omega u^2 + \beta \int_\Omega |\nabla u|^2, \tag{5.24}$$

with u such that u solves the state problem (5.15) for $k = \mu_1$ and $\chi = \mu_2$. In Figure 5.5 we depict the cost function for $k = 2\pi$ and $\chi = a + bi$ for $a, b \in [-6 : 6]$.

5.4 Numerical results

In order to test whether RBMs can be used instead of the FE discretisation to solve the state problem during the optimisation of the cost function, we run several tests. At the beginning of

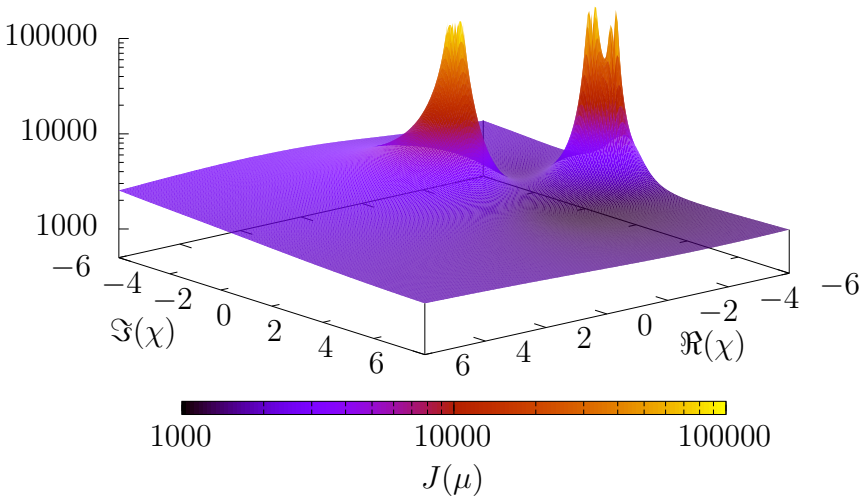


Figure 5.5: Value of the cost function $J(\mu)$ defined in (5.24) for $\mu = (2\pi, \chi)$.

Parameter	Transmission/reflection	Admittance identification	Engine noise
h	0.05	0.05	0.1
d.o.f.	4729	7142	11356
polynomial degree	3	2	3

Table 5.1: Parameter for the FE discretisation of the transmission/reflection, admittance identification and engine noise problem.

each test, we construct a FE discretisation of the state problem. The FE discretisation is used to construct a RB space. We construct the RB space using the greedy algorithm described in Section 2.4. Once the RB space is constructed, we select the control inputs $\mu^\star \in \mathbb{P}$ randomly. For each selected control input $\mu^\star$, we calculate the output $c(\mu^\star)$ using the FE discretisation. Using the previously calculated output $c(\mu^\star)$, we minimize the cost function 5.4 and calculate an estimated control input μ. Afterwards, we calculate the error $\frac{\|\mu^\star - \mu\|_{\mathbb{P}}}{\|\mu\|_{\mathbb{P}}}$ between the selected control input $\mu^\star$ and the estimated μ.

The parameter of the FE discretisation and the dimension of the computed RB space are shown in Table 5.1. Note that the RB space for the transmission/reflection and the admittance identification problem are constructed for varying wavenumbers, whereas the RB space for the engine noise problem is constructed for a fixed wavenumber $k = 2\pi$.

Usually the cost function (5.4) is not convex and very flat, see Figure 5.5 for an example. In order to improve the convergence, when using gradient based optimization methods, the cost function needs to be normalized. One possible normalization consists in adding a term to the cost function that measures the distance to an initial guess $\mu' \in \mathbb{P}$ for the optimal solution. The resulting cost function is

$$J(\mu) = \frac{1}{2} \|c(\mu) - z\|^2 + \|\mu' - \mu\|_{\mathbb{P}}^2 \quad \forall \mu \in \mathbb{P}. \tag{5.25}$$

In the numerical tests of the transmission/reflection and admittance identification problem the initial guess is derived from the selected control input.

In Figure 5.6 we depict the error in the estimated values for n_1 and n_2 from the transmission/reflection problem. It can be seen that the n_1 and n_2 can be estimated accurately. Further, we can see that the error increases with higher wavenumbers.

In contrast the error for the admittance identification problem shown in Figure 5.7 does not increase with higher frequencies (wavenumber). Instead is only spikes at certain frequencies. However, the tests are run at lower wavenumbers, i.e. $k_{min} = \frac{2\pi 200Hz}{c} \approx 3.66$ and $k_{max} = \frac{2\pi 500Hz}{c} \approx 9.14$.

Parameter	Transmission Reflection	Admittance identification	Engine noise
e_{max}	10^{-4}	10^{-4}	10^{-4}
N	95	105	13
#samples for the interpolation of β_h	100	76	50
#samples in the training set Ξ	1100	2709	500

Table 5.2: Parameter for constructing the RB space for the transmission/reflection, admittance identification and engine noise problem.

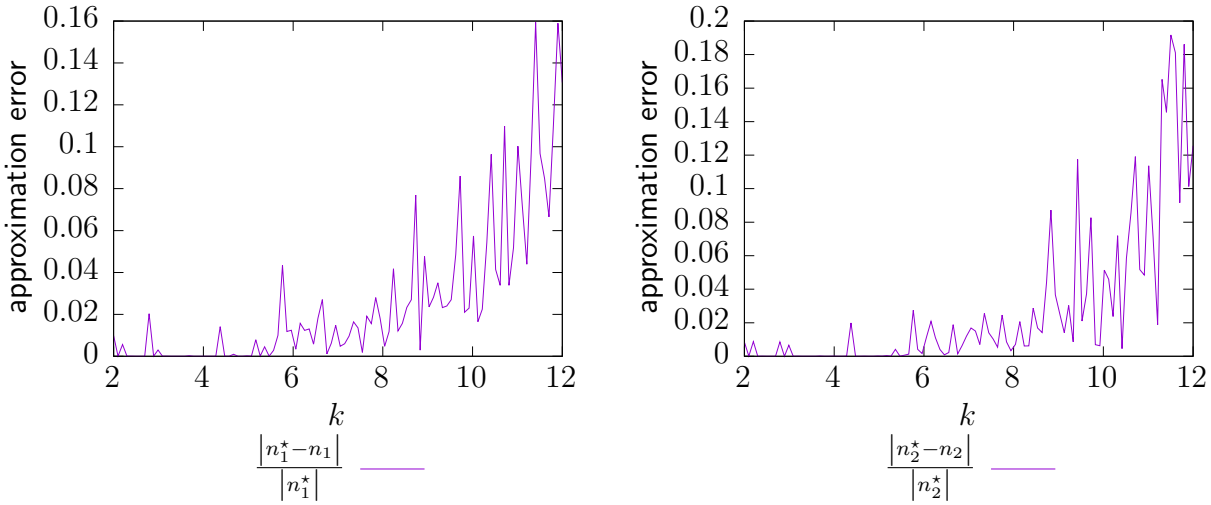


Figure 5.6: Error of the predicted values n_1 and n_2 for the inverse transmission/reflection problem from Section 5.1 at different wavenumbers. The exact solution is denoted by $n_1^\star$ and $n_2^\star$.

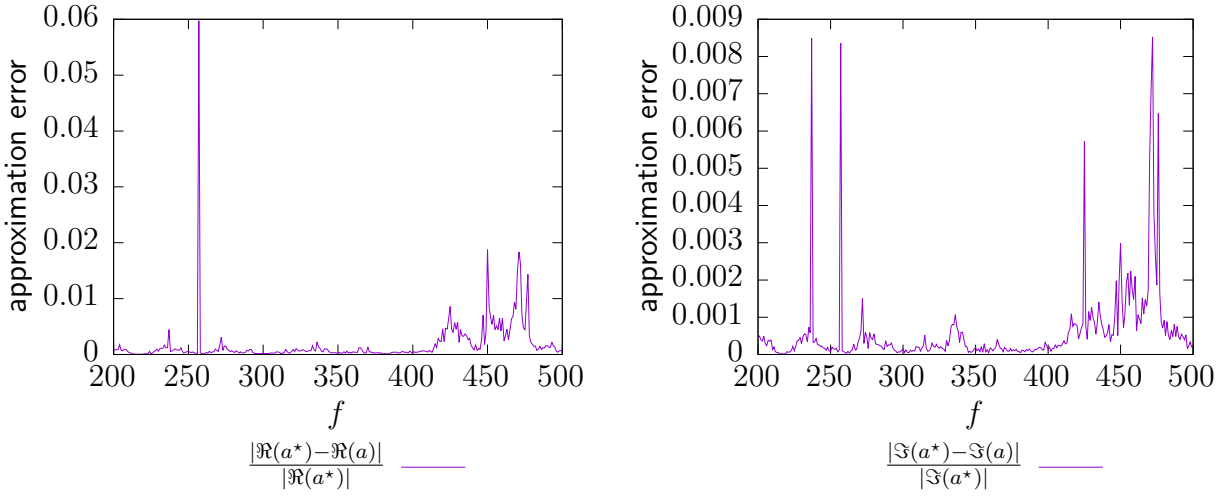


Figure 5.7: Error of the predicted admittance a for the admittance identification problem (5.7) at different frequencies. The exact solution is denoted by $a^\star$.

	$\alpha = 1, \beta = 0$	$\alpha = 1, \beta = 1$
χ_h	$-3.51236919 + 1.3803579i$	$-3.3137602 + 1.561295i$
χ_N	$-3.51237423 + 1.3803595i$	$-3.3137628 + 1.561293i$
$\frac{\lvert\chi_h - \chi_N\rvert}{\lvert\chi_h\rvert}$	1.309×10^{-6}	8.25×10^{-7}
$J(\chi_h)$	1159.5394896	53355.9240432
$J(\chi_N)$	1159.5394850	53355.9238456

Table 5.3: Results of the minimization of the cost function (5.21) for the engine noise problem.

Both experiments show that the RBM can be used in parameter estimation problems or much more general in optimization procedures that involve FE discretisations. Moreover, the improved efficiency of RBM allow for the use of global optimisation methods eliminating the need for normalization of the cost functional.

In case of the engine noise problem, we minimize the cost function (5.24) in two rounds. In the first one, we use the FEM to solve the state problem (5.21) and calculate the minimal χ_h. In the second one, we use the RBM to solve (5.21) and calculate the minimal χ_N. The cost function in (5.24) is parametrized by α and β. The test are run for $\alpha = 1$ and β either 0 or 1. The initial guess for both optimizations is $\chi_0 = 0$. The results are shown in Table 5.3. The error introduced by the RBM is negligible.

Chapter 6

Summary and conclusions

The aim of this thesis was to test the reduced basis method (RBM) with particular focus on noncoercive problems. In Chapter 2 we reviewed the relevant theory about RBMs. We introduced the offline-online procedure for the efficient implementation of the RBM. Furthermore, we introduced equi-logarithmic reduced spaces and the greedy algorithm, and discussed the implementation of an error estimator for the greedy algorithm. Based on the theoretical results, we expect that RBMs can be applied to coercive as well as noncoercive problems. We also expect that RBMs offer better convergence when applied to coercive problems. The numerical results in Chapter 3 and Chapter 4 support these assumptions.

In Chapter 3, we tested two parametric Poisson problems. The first problem is that of a thermal block that is heated from the inside. The geometry of the block is parametrized by the parameter μ. The second problem is that of a thermal block that is divided into subregions of different conductivity μ_i. The block is heated from the bottom side.

The presented tests confirm that the RBM is effective when applied to coercive problems. The approximation error is small and the RB spaces show rapid exponential convergence. Moreover, the dimension of the RB space V_N is largely unaffected by the size of the parameter space $\mathbb{P}$.

In Chapter 4, we tested the RBM on two noncoercive problems governed by the Helmholtz equation. The first one models a travelling plane wave where the parameter is the wavenumber. The second one models the transmission or reflection of a wave that travels through an interface between two fluids with different refractive indices n_1 and n_2.

The results suggest that RBMs can be used with noncoercive problems. The observed convergence is still exponential, but slower than in the coercive case, resulting in the need of using higher dimensional reduced spaces to reach the same accuracy. The greedy algorithm itself performs not as well as in the coercive case. The snapshots chosen by the greedy algorithm sometimes increase the maximum approximation error instead of decreasing it. This suggest that the greedy algorithm could be improved.

We compared different RB spaces that are constructed for problems parametrized by the wavenumber varying in different intervals $[k_{min}, k_{max}]$. We keep k_{min} fixed and increase k_{max}. In contrast to the results from Chapter 3, we see that the dimension of the reduced basis space V_N increases linearly with the interval length $k_{max} - k_{min}$. In addition, similar snapshots are chosen by the greedy algorithm for the overlapping regions.

Moreover, the investigation of the greedy algorithm showed that the algorithm is unaffected by different starting points and the selected training set.

However, the RBM is affected by oscillations of the underlying FE discretisation. Oscillations exhibited by the FE solutions also affect the resulting reduced basis (RB) spaces. While the resulting RB space is able to approximate the oscillations the resulting RB space is of higher dimension. Thus diminishing the efficiency of the RBM.

The results we obtained in Chapter 5 show that the RBM is an effective method that can be employed in the context of optimal control or parameter estimation. Optimal control and parameter estimation problems consist of a parametrized state problem and an output function.

The output function maps the solution of the state problem to an output. Given an output, we want to estimate the input to the state problem that lead to the given output. The procedures used to solve such problems present a good opportunity to use RBMs. In order to determine the original input, we need to solve the state problem many times.

In order to test how the RBM performs in parameter estimation and optimal control problems, we considered three problems. The first one is the transmission/reflection problem from Chapter 4. The goal is to estimate the refractive indices n_1 and n_2. The second one is an idealized model of car door and a loudspeaker. The bottom of the car door consists of a damping material of unknown impedance/admittance. The loudspeaker emits sound that interacts with the damping material. The sound pressure is measured at six different points. The aim is to estimate the impedance/admittance of the damping material from the measurements. The third problem models the sound emitted from an airplane engine. The engine is enclosed by a damping material of unknown impedance/admittance. The goal is to determine an impedance/admittance such that the noise emitted from the engine is minimized.

For the transmission/reflection and the admittance identification problem we select inputs randomly and use the finite element method (FEM) to solve the corresponding state problem. In the next step, we estimate the original input by solving the optimal control problem using the RBM. The results suggest that the RBM can be used as a substitute for the FEM, significantly improving the efficiency when solving optimal control and parameter estimation problems.

In case of the engine noise problem the goal is to determine an impedance value for the engine enclosing that minimizes the noise emitted by the engine. In the first step, we use the FEM to find the optimal value for the impedance, in the second step we use the RBM to find the optimal value for the impedance. Afterwards we compare the solutions and calculate the error. We see that the RBM estimate is close to the FEM estimate. Confirming that RBM can be used in the context of parameter estimation and optimal control.

The increased efficiency of RBMs might even allow the use of global optimization procedures that eliminate the need for initial guesses of solutions.

However, these results are only an indication for the performance of the RBM applied to different problems. The performance of the RBM highly depends on the particular problem. This is already evident in the theory. Theorem 9 relates the convergence of the greedy algorithm to the decay of so-called the n-width of the continuous spaces or discrete spaces. However, it is not trivial to show the n-width decay, see e.g. [QMN16, Section 5.4].

Furthermore, in Chapter 4 we only considered one dimensional parameter spaces. This is an idealistic assumption although the results from Chapter 3 and 5 suggest that RBMs can be used with higher dimensional parameter spaces.

There exist several extensions of the RBM that are not considered here. This includes Petrov Galerkin reduced basis methods or RBMs based on proper orthogonal decomposition. For more details we refer to [QMN16].

Appendix A

Review of numerical methods

In the following sections we review some of the numerical methods that are used in the Chapters 2, 3, 4, and 5. The descriptions are based on [NW06] and [GL13].

A.1 Quasi-Newton methods

Quasi-Newton methods are methods for minimizing functions $f : \mathbb{R}^n \to \mathbb{R}$. They are based on Newton methods. Newton methods are a variant of line search methods. Line search methods are iterative methods that generate successive points $x_k \in \mathbb{R}^n$ that converge to a (local) minimum. Given a step length $\alpha_k \in \mathbb{R}^n$ and a direction $p_k \in \mathbb{R}^n$, the x_k are generated by

$$x_{k+1} = x_k + \alpha_k p_k. \tag{A.1}$$

Assume the function f is twice continuously differentiable, i.e. $f \in C^2$, and denote the Hesse matrix at the point $x \in \mathbb{R}^n$ by $H_f(x) \in \mathbb{R}^{n\times n}$, $H(x)_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$. The Newton method corresponds to taking $p_k = -H_f(x_k)^{-1}\nabla f(x_k)$ and $\alpha_k = 1$. This leads to the following iteration

$$a_{k+1} = x_k - H_f(x_k)^{-1}\nabla f(x_k). \tag{A.2}$$

In practice the step length α_k is optimized by inexact line search methods based on the Wolfe conditions, see [NW06, Chapter 2].

One problem of standard Newton methods is that $H_f(x_k)$ needs to be calculated and inverted at each step of the method. This can pose a problem if the second derivatives are not available or the Hesse matrix is not invertible. Quasi-Newton methods circumvent this by approximating the Hessian $H_f(x_k)$ by a symmetric positive definite matrix $B_k \in \mathbb{R}^{n\times n}$. Using B_k the Quasi-Newton methods minimizes a quadratic model $m_k : \mathbb{R}^n \to \mathbb{R}$

$$m_k(p) = f(x_k) + \nabla f(x_k)^T p + \frac{1}{2} p^T B_k p. \tag{A.3}$$

of the function f. Instead of computing the matrix B_k at each step, B_k, is updated at each step to form B_{k+1}. In order keep the model function m_k accurate, we impose the following conditions for all B_k, $k = 1 \dots N$

$$\nabla m_{k+1}(-\alpha_k p_k) = \nabla f(x_k), \tag{A.4}$$
$$\nabla m_{k+1}(0) = \nabla f(x_{k+1}). \tag{A.5}$$

Since $\nabla m_{k+1}(p) = \nabla f(x_{k+1}) + B_{k+1}p$ condition (A.5) is fulfilled automatically. Condition (A.4) becomes

$$\nabla f(x_{k+1}) - \alpha B_{k+1} p_k = \nabla f(x_k). \tag{A.6}$$

Hence

$$\nabla f(x_{k+1}) - \nabla f(x_k) = \alpha B_{k+1} p_k. \tag{A.7}$$

Setting $s_k = x_{k+1} - x_k = \alpha_k p_k$ and $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$ we can rewrite (A.7) as

$$B_k s_k = y_k. \tag{A.8}$$

The equation is known as as the *secant* equation. In order for (A.8) to hold, s_k and y_k need to fulfil the curvature condition

$$s_k^T y_k > 0. \tag{A.9}$$

(A.9) is automatically fulfilled for convex functions. For nonconvex functions, (A.9) is fulfilled if the step length α_k fulfils the strong Wolfe conditions, see [NW06, Chapter 6].

The conditions (A.4) and (A.5) are not sufficient to define B_k uniquely. Thus, we require that B_{k+1} is as close as possible to B_k. This yields the following problem

$$B_{k+1} = \min_{B \in \mathbb{R}^{n \times n}} \|B - B_k\| \quad \text{such that } B = B^T \text{ and } B s_k = y_k. \tag{A.10}$$

Depending on the norms used in (A.10), different solutions arise. Given a weight matrix $W \in \mathbb{R}^{n \times n}$

$$W = \bar{G}^1, \tag{A.11}$$

where $\bar{G}$ denotes the average Hessian

$$\bar{G} = \int_0^1 H_f(x_k + \tau \alpha_k p_k) d\tau. \tag{A.12}$$

We define the norm

$$\|A\|_W = \left\| W^{\frac{1}{2}} A W^{\frac{1}{2}} \right\|, \tag{A.13}$$

where $\|\cdot\|_F$ denotes the Frobenius norm $\|A\|_F^2 = \sum_{i=1}^n \sum_{j=1}^n A_{ij}^2$. Given the norm (A.13), the unique solution to (A.10) is given by

$$B_{k+1} = \left(I - \frac{y_k s_k^T}{y_k^T s_k} \right) B_k \left(I - \frac{s_k y_k^T}{y_k^T s_k} \right) + \frac{y_k y_k^T}{y_k^T s_k}. \tag{A.14}$$

Equation (A.14) is also called the DFP update formula.

A further improvement on equation (A.14) is instead of updating B_k we update the inverse $H_k = B_k^{-1}$ directly. This leads to following problem

$$H_{k+1} = \min_{H \in \mathbb{R}^{n \times n}} \|H - H_k\| \quad \text{such that } H = H^T \text{ and } H y_k = s_k. \tag{A.15}$$

Reusing the norm defined in (A.13), we derive

$$H_{k+1} = \left(I - \frac{s_k y_k^T}{y_k^T s_k} \right) H_k \left(I - \frac{y_k s_k^T}{y_k^T s_k} \right) + \frac{s_k s_k^T}{y_k^T s_k}. \tag{A.16}$$

Equation (A.16) is also called the BFGS formula. The resulting algorithm is shown in Algorithm 3.

Algorithm 3 The quasi Newton algorithm.

Given f, ∇f, x_0, H_0 and a tolerance e_{tol}
Set $k \leftarrow 0$
while $\|\nabla f(x_k)\| > e_{tol}$ **do**

$$p_{k+1} \leftarrow H_k p_k \tag{A.17}$$
$$s_k \leftarrow x_{k+1} - x_k \tag{A.18}$$
$$y_k \leftarrow \nabla f(x_{k+1}) - \nabla f(x_k) \tag{A.19}$$
$$H_{k+1} \leftarrow \left(I - \frac{s_k y_k^T}{y_k^T s_k}\right) H_k \left(I - \frac{y_k s_k^T}{y_k^T s_k}\right) + \frac{s_k s_k^T}{y_k^T s_k}, \tag{A.20}$$
$$k \leftarrow k + 1 \tag{A.21}$$

end while

A.2 Conjugate gradient method

The CG method was developed originally to solve a linear system

$$Ax = b, \tag{A.22}$$

where $A \in \mathbb{R}^{n \times n}$ is symmetric and positive definite, and $x, b \in \mathbb{R}^n$. CG methods work by generating a conjugate set of vectors $p_k \in \mathbb{R}^n$ $\{p_0, p_1, \dots p_{n-1}\}$. A set $s \subset \mathbb{R}^n$ of vectors is said to be conjugate with respect to a symmetric positive definite matrix $A \in \mathbb{R}^{n \times n}$ if

$$p_i^T A p_j = 0 \qquad 1 \leq i, j \leq n, i \neq j. \tag{A.23}$$

Denote the residual of the linear system (A.22) by r_k:

$$r_k(x) = Ax_k - b. \tag{A.24}$$

Given a set $\{p_0, p_1, \dots p_{n-1}\}$ of conjugate vectors, we generate iterates $x_k \in \mathbb{R}^n$ by

$$x_{k+1} = x_k + \alpha_k p_k. \tag{A.25}$$

The step length α_k is given by

$$\alpha_k = -\frac{r(x_k)^T p_k}{p_k^T A p_k}. \tag{A.26}$$

The x_k generated by (A.25) converge to the solution to the linear system (A.22) in at most n steps; see [NW06, Theorem 5.1].

The important ingredient of the CG method is the ability to generate conjugate vectors p_k based on the previous vector p_{k-1}. The next conjugate direction is chosen by

$$p_k = -r(x_k) + \beta_k p_{k-1}, \tag{A.27}$$

where

$$\beta_k = \frac{r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}. \tag{A.28}$$

Noting that (A.26) and (A.28) can be rewritten as

$$\alpha_k = \frac{r(x_k)^T r(x_k)}{p_k^T A p_k}, \tag{A.29}$$
$$\beta_{k+1} = \frac{r(x_{k+1})^T r(x_{k+1})}{r(x_k)^T r(x_k)}, \tag{A.30}$$

Algorithm 4 The CG algorithm for solving linear systems.

Given x_0, A, and b
Set $p_0 \leftarrow -r(x_0)$, $k \leftarrow 0$
while $r_k \neq 0$ **do**

$$\alpha_k \leftarrow \frac{r(x_k)^T r(x_k)}{p_k^T A p_k}, \tag{A.31}$$

$$x_{k+1} \leftarrow x_k + \alpha_k p_k \tag{A.32}$$

$$\beta_{k+1} \leftarrow \frac{r(x_{k+1})^T r(x_{k+1})}{r(x_k)^T r(x_k)} \tag{A.33}$$

$$p_{k+1} \leftarrow -r(x_{k+1}) + \beta_{k+1} p_k \tag{A.34}$$

$$k \leftarrow k + 1 \tag{A.35}$$

end while

this leads to the Algorithm shown in 4.

The CG can also be viewed as an optimization algorithm. Solving the linear system (A.22) is equivalent to minimizing

$$\phi(x) = \frac{1}{2} x^T A x - b^T x. \tag{A.36}$$

Moreover, the residual r is infact the gradient of (A.36)

$$\nabla \phi(x) = r(x) = Ax - b. \tag{A.37}$$

Further, by noting that α_k is chosen to be the minimizer of ϕ along p_k we derive Algorithm 5. The minimization problem in (A.38) can be carried out by an inexpensive line search; see [NW06, Chapter 2 and Chapter 5].

Algorithm 5 The CG algorithm for solving minimization problems.

Given x_0, ϕ, and $\nabla \phi$
$p_0 \leftarrow -\nabla \phi(x_0)$, $k \leftarrow 0$
while $\nabla \phi(x_0) \neq 0$ **do**

$$\alpha_k \leftarrow \min_{\alpha \in \mathbb{R}} \phi(x_k + \alpha p_k) \tag{A.38}$$

$$x_{k+1} \leftarrow x_k + \alpha_k p_k \tag{A.39}$$

$$\beta_{k+1} \leftarrow \frac{\nabla \phi(x_{k+1})^T \nabla \phi(x_{k+1})}{\nabla \phi(x_k)^T \nabla \phi(x_k)} \tag{A.40}$$

$$p_{k+1} \leftarrow -\nabla \phi(x_{k+1}) + \beta_{k+1} p_k \tag{A.41}$$

$$k \leftarrow k + 1 \tag{A.42}$$

end while

A.3 Restarted Arnoldi iteration

The Arnoldi iteration is a method for calculating the extremal eigenvalues of a matrix $A \in \mathbb{R}^{n \times n}$. At every step the Arnoldi iteration computes a tridiagonal matrix $T_k \in \mathbb{R}^{k \times k}$ whose extremal eigenvalues approximate the extremal eigenvalues of the original matrix A. For brevity we consider

only the case where A is symmetric. In this case, the Arnoldi iteration reduces to the Lanczos iteration.

In order to introduce the Lanczos iteration, we introduce the Rayleigh quotient $r : \mathbb{R}^n \to \mathbb{R}$

$$r(x) = \frac{x^T Ax}{x^T x} \quad \forall x \neq 0. \tag{A.43}$$

The maximum and minimum of r correspond to maximum and minimum eigenvalues of A $\lambda_n(A)$ and $\lambda_1(A)$; see [GL13, Theorem 8.1.2]. Further, let $\{q_i\} \subset \mathbb{R}^n$ be a sequence of orthogonal vectors and define $Q_j = [q_1|\dots|q_j]$. Moreover, let

$$M_j = \lambda_1(Q_j^T AQ) = \max_{y \neq 0} \frac{y^T Q_j^T AQ_j}{y^T y} = \max_{\|y\|_2 = 1} r(Q_j y) \leq \lambda_1(A), \tag{A.44}$$

$$m_j = \lambda_n(Q_j^T AQ) = \min_{y \neq 0} \frac{y^T Q_j^T AQ_j}{y^T y} = \min_{\|y\|_2 = 1} r(Q_j y) \geq \lambda_n(A). \tag{A.45}$$

Our goal is to find q_i such that M_j and m_j become increasingly closer to $\lambda_1(A)$ and $\lambda_n(A)$. Let $u_j \in span(q_1, \dots, q_j)$ be such that $M_j = r(u_j)$. We want to ensure that q_{j+1} is chosen in such a way that $M_{j+1} > M_j$. This can be done by ensuring that

$$\nabla r(u_j) \in span(q_1, \dots, q_{j+1}). \tag{A.46}$$

The gradient $\nabla r(x)$ is given by

$$\nabla r(x) = \frac{2(Ax - r(x)x)}{x^T x}. \tag{A.47}$$

We note that $\nabla r(x) \in span(x, Ax)$. Hence (A.46) can always be fulfilled, if we choose

$$span(q_1, \dots, q_j) = span(q_1, Aq_1, \dots, A^{j-1} q_j) = \mathcal{K}(A, q_1, j). \tag{A.48}$$

The space $\mathcal{K}(A, q_1, j)$ is also called the Krylov subspace. We define the Krylov matrix

$$K(A, q_1, n) = [q_1, Aq_1, \dots A^{n-1} q_1]. \tag{A.49}$$

We define the tridiagonal matrix $T = Q^T AQ$ and note that

$$K(A, q_1, n) = Q[e_1, Te_1, \dots T^{n-1} e_1] \tag{A.50}$$

is the QR factorization of $K(A, q_1, n)$. This ensures that A can be tridiagonalized with an orthogonal matrix Q whose first column is q_1. By setting $T = Q^T AQ$ and

$$T = \begin{pmatrix} \alpha_1 & \beta_1 & & \cdots & 0 \\ \beta_1 & \alpha_2 & \ddots & & \vdots \\ & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & \beta_{n-1} \\ 0 & \cdots & & \beta_{n-1} & \alpha_n \end{pmatrix} \tag{A.51}$$

and writing $AQ = TQ$ we derive

$$Aq_k = \beta_{k-1} q_{k-1} + \alpha_k q_k + \beta_k q_{k+1}. \tag{A.52}$$

By rearranging and setting $\alpha_k = q_k^T Aq_k$ we derive the iteration shown in Algorithm 6. For details see [GL13, Chapter 9.]. The iteration produces a matrix T whose extremal eigenvalues converge to the extremal eigenvalues of A. The key benefit of Algorithm 6 is that not all n α_k and β_k need to be calculated. Instead the iteration will be terminate much earlier resulting in a $T \in \mathbb{R}^{n_t \times n_t}$ with $n_t << n$.

The restarted Arnoldi iteration improves on the Lanczos iteration by generalizing the Lanczos iteration to non symmetric A and by improving the starting vector q_1 through multiple iterations. For more details about how the Lanczos iteration can be generalized to nonsymmetric A and about the restarted Arnoldi iteration we refer to [GL13; LS96; LSY98].

Algorithm 6 The Lanczos iteration for computing the tridiagonal matrix T.

Given q_1
Set $r_0 \leftarrow q_1$, $q_0 \leftarrow 0$ and $k \leftarrow 0$
while $\beta_j \neq 0$ **do**

$$q_{k+1} \leftarrow \frac{r_k}{\beta_k} \tag{A.53}$$
$$\alpha_{k+1} \leftarrow q_{k+1}^T A q_{k+1} \tag{A.54}$$
$$r_{k+1} \leftarrow (A - \alpha_k I) q_k - \beta_k q_k \tag{A.55}$$
$$\beta_{k+1} \leftarrow \|r_k\|_2 \tag{A.56}$$
$$k \leftarrow k + 1 \tag{A.57}$$

end while

Acronyms

CG conjugate gradient. 4, 45, 57, 58

FE finite element. 1, 2, 4, 6, 8, 11, 12, 17, 20, 21, 26, 27, 30, 31, 34, 36, 44, 46, 49, 50, 52, 53, 61

FEM finite element method. 1, 5–10, 13, 46, 52, 54, 61

MOR model order reduction. 5, 61

PDE partial differential equation. 5, 8, 45, 61

PG-RB Petrov Galerkin reduced basis. 54

POD proper orthogonal decomposition. 54

RB reduced basis. 2–4, 6–8, 11–14, 17–21, 26–29, 34–44, 46, 47, 49–51, 53, 61

RBM reduced basis method. 5–8, 11–14, 21, 23, 24, 27, 29–31, 44, 45, 49, 52–54, 61

SCM sucessive constraint method. 21

Bibliography

[Bab71] I. Babuška. 'Error-bounds for finite element method'. In: *Numer. Math.* 16.4 (1971), pp. 322–333.

[Bin+11] P. Binev, A. Cohen, W. Dahmen, R. Devore, G. Petrova and P. Wojtaszczyk. 'Convergence rates for greedy algorithms in reduced basis methods'. In: *SIAM J. Math. Anal.* 43.3 (2011), pp. 1457–1472. DOI: 10.1137/100795772.

[BNP18] F. Bonizzoni, F. Nobile and I. Perugia. 'Convergence analysis of Padé approximations for Helmholtz frequency response problems'. In: *ESAIM Math. Model. Numer. Anal.* 52.4 (2018), pp. 1261–1284. DOI: 10.1051/m2an/2017050.

[CHY07] Y. Cao, M. Y. Hussaini and H. Yang. 'Estimation of optimal acoustic linear impedance factor for reduction of radiated engine noise'. In: *Int. J. Numer. Anal. Mod.* 4.1 (2007), pp. 116–126.

[EM12] S. Esterhazy and J. M. Melenk. 'On stability of discretizations of the Helmholtz equation'. In: *Numerical Analysis of Multiscale Problems*. Berlin, Heidelberg: Springer, 2012, pp. 285–324. ISBN: 9783642220616. DOI: 10.1007/978-3-642-22061-6_9.

[FZ03] E. Fernández-Cara and E. Zuazua. 'Control theory: history, mathematical achievements and perspectives'. In: *Bol. Soc. Esp. Mat. Apl.* 26 (2003), pp. 79–140.

[GL13] G. H. Golub and F. Van Loan. *Matrix computations*. 4th. Johns Hopkins series in the mathematical sciences. Baltimore: Johns Hopkins Univ. Press, 2013. ISBN: 9781421407944.

[GP05] M. A. Grepl and A. T. Patera. 'A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations'. In: *ESAIM: Math. Model. Numer. Anal.* 39.1 (2005), pp. 157–181. DOI: 10.1051/m2an:2005006.

[Gre05] M. A. Grepl. 'Reduced-basis approximation and a posteriori error estimation for parabolic partial differential equations'. PhD thesis. Massachusetts Institute of Technology, 2005.

[HS52] M.R. Hestenes and E. Stiefel. 'Methods of conjugate gradients for solving linear systems'. In: *J. Res. Natl. Inst. Stand. Technol.* 49.6 (1952), p. 409.

[Huy+07] D.B.P. Huynh, G. Rozza, S. Sen and A.T. Patera. 'A successive constraint linear optimization method for lower bounds of parametric coercivity and inf–sup stability constants'. In: *C. R. Math. Acad. Sci. Paris* 345.8 (2007), pp. 473–478. DOI: 10.1016/j.crma.2007.09.019.

[JOP+01] E. Jones, T. Oliphant, P. Peterson et al. *SciPy: Open source scientific tools for Python*. 2001. URL: http://www.scipy.org/.

[KMW15] S. Kapita, P. Monk and T. Warburton. 'Residual-based adaptivity and PWDG methods for the Helmholtz equation'. In: *SIAM J. Sci. Comput.* 37.3 (2015), A1525–A1553. DOI: 10.1137/140967696.

[LS96] R. B. Lehoucq and D. C. Sorensen. 'Deflation techniques for an implicitly restarted arnoldi iteration'. In: *SIAM J. Matrix Anal. and Appl.* 17.4 (1996), p. 789. DOI: 10.1137/S0895479895281484.

[LSY98] R. B. Lehoucq, D. C. Sorensen and C. Yang. *ARPACK users' guide: solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*. Vol. 6. SIAM, 1998. ISBN: 9780898714074.

[Man12] A. Manzoni. 'Reduced models for optimal control, shape optimization and inverse problems in haemodynamics'. PhD thesis. École polytechnique fédérale de Lausanne, 2012. DOI: 10.5075/epfl-thesis-5402.

[Mel00] J. M. Melenk. 'On n-widths for elliptic problems'. In: *J. Math. Anal. Appl.* 247.1 (2000), pp. 272–289. DOI: 10.1006/jmaa.2000.6862.

[Mel95] J. M. Melenk. 'On generalized finite element methods'. PhD thesis. University of Maryland, 1995.

[MN15] A. Manzoni and F. Negri. 'Heuristic strategies for the approximation of stability factors in quadratically nonlinear parametrized PDEs'. In: *Adv. Comput. Math.* 41.5 (2015), pp. 1255–1288. DOI: 10.1007/s10444-015-9413-4.

[MPT02a] Y. Maday, A. T. Patera and G. Turinici. 'Global a priori convergence theory for reduced-basis approximations of single-parameter symmetric coercive elliptic partial differential equations'. In: *C. R. Math. Acad. Sci. Paris* 335.3 (2002), pp. 289–294. DOI: 10.1016/S1631-073X(02)02466-4.

[MPT02b] Y. Maday, A. Patera and G. Turinici. 'A priori convergence theory for reduced-basis approximations of single-parameter elliptic partial differential equations'. In: *J. Sci. Comput.* 17.1 (2002), pp. 437–446. DOI: 10.1023/A:1015145924517.

[Neč67] J. Nečas. *Les méthodes directes en théorie des équations elliptiques*. Prague Paris: Academia Masson, 1967.

[Neg+13] F. Negri, G. Rozza, A. Manzoni and A. Quarteroni. 'Reduced basis method for parametrized elliptic optimal control problems'. In: *SIAM J. Sci. Comput.* 35.5 (2013), A2316–A2340. DOI: 10.1137/120894737.

[NW06] J. Nocedal and S. Wright. *Numerical optimization*. Second Edition. Springer series in operations research and financial engineering. New York, NY: Springer, 2006. ISBN: 9780387303031.

[Pru+01] C. Prud'homme, D. V. Rovas, K. P. L. Veroy, L. Machiels, Y. Maday, A. T. Patera and G. Turinici. 'Reliable real-time solution of parametrized partial differential equations: reduced-basis output bound methods '. In: *J. Fluids Eng.* 124.1 (Nov. 2001), pp. 70–80. DOI: 10.1115/1.1448332.

[QMN16] A. Quarteroni, A. Manzoni and F. Negri. *Reduced basis methods for partial differential equations : an introduction*. Cham: Springer, 2016. ISBN: 9783319154305.

[Qua14] A. Quarteroni. *Numerical Models for Differential problems*. Vol. 2. Mailand: Springer-Verlag, 2014. ISBN: 9788847058835.

[QV16] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*. Berlin, Heidelberg: Springer, 2016. ISBN: 9783540852674.

[RHP07] G. Rozza, D. B. P. Huynh and A. T. Patera. 'Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations'. In: *Arch. Comput. Methods Eng.* 15.3 (2007), p. 1. DOI: 10.1007/BF03024948.

[Rov03] D. V. Rovas. 'Reduced-basis output bound methods for parametrized partial differential equations'. PhD thesis. Massachusetts Institute of Technology, 2003.

[Sch14] J. Schöberl. *C++11 Implementation of Finite Elements in NGSolve*. Institute for analysis and scientific computing, Vienna University of Technology, 2014. URL: http://www.ngsolve.org/.

[Sch74] A. H. Schatz. 'An observation concerning Ritz-Galerkin methods with indefinite bilinear forms'. In: *Math. Comput.* 28.128 (1974), pp. 959–962. DOI: 10.2307/2005357.

[Sen+06] S. Sen, K. P. L. Veroy, D.B.P. Huynh, S. Deparis, N.C. Nguyen and A.T. Patera. '"Natural norm" a posteriori error estimators for reduced basis approximations'. In: *J. Comput. Phys.* 217.1 (2006), pp. 37–62. DOI: 10.1016/j.jcp.2006.02.012.

[Ver03] K. P. L. Veroy. 'Reduced-basis methods applied to problems in elasticity: analysis and applications'. PhD thesis. Massachusetts Institute of Technology, 2003.

[VH09] S. Volkwein and A. Hepberger. 'Impedance identification by POD model reduction techniques'. In: *at - Automatisi1erungstechnik* 56 (2009), pp. 437–446. DOI: 10.1524/auto.2008.0724.

[Vol10] S. Volkwein. 'Admittance identification from point-wise sound pressure measurements using reduced-order modelling'. In: *J. Optim. Theory Appl.* 147.1 (2010), pp. 169–193. DOI: 10.1007/s10957-010-9704-3.