Reduced Basis Approximation and *A Posteriori* Error Estimation for Non-Coercive Elliptic Problems: Application to Acoustics

by

Sugata Sen

S.M. Civil and Environmental Engineering (2001)
Massachusetts Institute of Technology

B.Tech (Hons.) Civil and Environmental Engineering (1999)
Indian Institute of Technology, India

Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Civil and Environmental Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2007

Author .....
Department of Civil and Environmental Engineering
May 4, 2007

Certified by .....
Anthony T. Patera
Professor of Mechanical Engineering
Thesis Supervisor

Certified by .....
Franz-Josef Ulm
Professor of Civil and Environmental Engineering
Chairman, Thesis Committee

Accepted by .....
Daniele Veneziano
Professor of Civil and Environmental Engineering
Chairman, Departmental Committee for Graduate Studies

Reduced Basis Approximation and *A Posteriori* Error Estimation for Non-Coercive Elliptic Problems: Application to Acoustics

by

Sugata Sen

Submitted to the Department of Civil and Environmental Engineering
on May 4, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in the field of Computational Engineering

Abstract

Modern engineering problems often require accurate, reliable, and efficient evaluation of quantities of interest, evaluation of which demands the solution of a partial differential equation. We present in this thesis a general methodology for the predicition of outputs of interest of non-coercive elliptic partial differential equations. The essential ingredients are: (*i*) rapidly convergent reduced basis approximations — Galerkin projection onto a space W_N spanned by solutions of the governing partial differential equation at N selected points in parameter-time space; (*ii*) *a posteriori* error estimation — relaxations of the error-residual equation that provide rigorous and sharp bounds for the error in specific outputs of interest; and (*iii*) offline-online computational procedures — in the offline stage the reduced basis approximation is generated; in the online stage, given a new parameter value, we calculate the reduced basis output and associated error bound. The operation count for the online stage depends only on N (typically small) and the parametric complexity of the problem; the method is thus ideally suited for repeated, rapid, reliable evaluation of input-output relationships in the many-query or real-time contexts.

We consider the crucial ingredients for the treatment of acoustics problems — simultaneous treatment of non-coercive (and near–resonant), non-Hermitian elliptic operators, complex-valued fields, often unbounded domains, and quadratic outputs of interest. We introduce the successive constraint approach to approximate lower bounds to the inf-sup stability constant, a key ingredient of our rigorous *a posteriori* output error estimator. We develop a novel expanded formulation that enables treatment of quadratic outputs as linear *compliant* outputs. We also build on existing ideas in domain truncation to develop a radiation boundary condition to truncate unbounded domains.

We integrate the different theoretical contributions and apply our methods as proof of concept to some representative applications in acoustic filter design and characterization. In the *online* stage, we achieve $O(10)$ computational economies of cost while demonstrating both the rapid convergence of the reduced basis approximation, and the *sharpness* of our error estimators ($\approx O(20)$). The obtained computational economies are expected to be significantly greater for problems of larger size. We thus emphasize the feasibility of our methods in the many-query contexts of optimization, characterization, and control.

Thesis Supervisor: Anthony T. Patera
Title: Professor of Mechanical Engineering

Acknowledgments

I would first like to thank my thesis advisor, Professor Anthony T. Patera, for his support and guidance during my thesis work. Working with him and for him has been a great pleasure: I am truly grateful for his insights, humor, and the trust he showed me.

I am also very thankful to my thesis committee members, Professor Franz Josef Ulm and Professor Dennis McLaughlin, for their invaluable comments, suggestions, and encouragement throughout my studies.

I am very grateful to my current and former colleagues: George Pau, Ngoc Cuong Nguyen, Huynh Dinh Phuong, Simone Deparis, Gianluigi Rozza, Martin Grepl, Karen Veroy and Yuri Solodukhov. I am grateful for the many stimulating and enjoyable discussions I had with them, both work-related and otherwise. Most of all I have deeply enjoyed my interactions with Debra Blanchard (Debbie): her company lifted my spirits especially when I needed it the most. I also need to acknowledge the friends I have found in my time at MIT: in no particular order here is a *small* laundry list of people who have made a difference in my life — Messrs Debashish and Hemant Sahoo, Amit, Somani, Isha, Suddha, Anjali, Bala, Bharath, MV, and Sudarshan.

I also want to express my love and gratitude to my family: my parents, Rabindranath Sen and Madhuri Sen, my brother, Shamik, my sister, Suchira for their unquestioning support. However, this thesis would not have seen the light of day without the constant presence of my wife Joyopriya by my side: she has encouraged me always to pursue my dreams, given me belief that I can achieve them, and has been my tower of strength.

To them I dedicate this thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The optimization, control, and characterization of engineering systems requires the rapid (often real-time) evaluation of certain "metrics", or *outputs* ($s(\mu)$), such as deflections, stresses, maximum temperatures, average heat transfer rates, flowrates, or lifts and drags. These "quantities of interest" are typically functions of parameters reflecting variations in loading or boundary conditions, material properties, and geometry. Specifying the parameters, or *inputs* (μ), serves to identify a particular "configuration" of the system. Finally, the *field variable*, $u(\mu)$, that connects the input parameters, μ, to the outputs, $s(\mu)$, represents the displacement, temperature, pressure, velocity, or any such distribution function. The evaluation of the input-output relationship requires the solution of the parametrized partial differential equation (PDE) in $u(\mu)$ describing the underlying physics of the system. The outputs $s(\mu)$ are really functions of the field variable $u(\mu)$, $s(\mu) \equiv s(u(\mu))$, typically we suppress the explicit $u(\mu)$ dependence.

Our interest in the evaluation of the input-output relationship is in two particular contexts: the *real-time context*, and the *many-query context*. Both these contexts are crucial to engineering design and optimization, and to more widespread adoption and application of numerical methods for PDEs in practical engineering applications. The *real-time context* is relevant in many practical problems; for example, the application of parameter estimation techniques using real-time sensor data to predict the location and geometric characteristics of a submerged object (possibly an enemy submarine or a floating mine) in the sea in real-time. The *many-query* context is embodied by many

optimal design problems — consider evaluating the input-output dependence many many times to arrive at the optimal geometric attributes and configuration of a thermal fin array used in heat sinks to maximize heat loss. Clearly, many applications across engineering disciplines fall broadly into either of the two contexts highlighted here.

In most cases, the evaluation of the input-output relationship requires the use of numerical techniques to solve the underlying partial differential equation as analytical solutions cannot be obtained. However, classical numerical approaches (e.g. finite element (FEM), finite difference (FD), boundary element (BE)) are ill-suited to the real-time and many-query contexts embodied by many engineering problems. Virtually all classical numerical approaches consider large approximation subspaces for the underlying PDE: the computational cost for a particular input-output evaluation is typically too long especially in the context of real-time response or repeated queries. Engineering design and optimization typically requires thousands of evaluations of the PDE-constrained input-output relationship. Hence, even for problems with moderate complexity, the computational cost using conventional numerical techniques is unacceptably high.

The real-time and many-query contexts represent not only computational challenges, but also computational opportunities. We identify two key opportunities that are exploitable. First, in the parametric setting, the $\mu \rightarrow u(\mu) \rightarrow s(\mu)$ input-output map induces a low-dimensional parametric manifold: the set of fields $u(\mu)$ engendered as the input varies over the manifold. Thus generic rich approximation spaces used by classical numerical approaches are unnecessarily expensive in the parametric context. Second, in the real-time or many-query contexts, the premium is on *marginal* cost — thus considerable pre-processing (or *offline*) cost is acceptable as long as the response time is *fast* in the deployed (or *online*) stage.

The development of model order reduction techniques — principal orthogonal decomposition (POD) for linear time-dependent systems [99], balanced truncation in optimal control [69], related hybrid techniques [60, 112], and earlier work on reduced basis methods in structural analysis [5, 74] — that focus on generating an optimal low-order approximation space is then quite natural seen through the prism of the computational opportunities embodied by real-time and many-query contexts. However, these approaches are mostly adapted to the problem under consideration — thus the approximation spaces are rather local and also typically rather low-dimensional in parameter (often only one parameter); for example, the POD approaches mostly consider just time as the

parameter. While these techniques do enable rapid evaluation of the input-output map, quantification of the quality of the approximation is harder to obtain. Although *a priori* bounds to quantify the error due to model reduction do exist, these bounds are not sufficiently *sharp* in the *many-query* or *real-time* context. The low-order approximation of the solution is not enough, we need to quantify how good this approximation is relative to the solution obtained using classical numerical approaches.

The lack of *a posteriori* error estimation and adaptive sampling algorithms restricted earlier efforts in reduced-order modeling to local approximation spaces and small parameter ranges. In problems with more global, higher-dimensional parameter domains the reduced-order predictions "far" from sample points cannot be trusted for accuracy, primarily because the *a priori* bounds are often quite crude. *A posteriori* error estimation will quantify the quality of the reduced-order approximation. Furthermore, in higher parameter domains simple tensor-product/full-factorial sampling of the parameter space is not going to be practical. Sophisticated sampling strategies for the construction of the approximation sub-space will need to rely on the *a posteriori* error estimators as a mechanism to (i) cheaply explore much larger subsets of the parameter domain in search of the "richest" snapshots to include in the basis; and (ii) decide when the size of the basis is enough to satisfy some required tolerance criteria.

To further motivate the need for reduced-order modeling and associated *a posteriori* error estimation techniques, we now discuss several acoustics applications relevant to the *real-time* and *many-query* contexts.

1.1.1 Detection of Buried Objects

Acoustic waves are frequently used for the detection and identification of land mines [90, 115], unexploded ordnance, and other buried objects. The goal is to recover characteristics (like geometric measures, and location) of an interrogated object from experimental data (far-field pattern measured at distributed sensor locations) obtained by sending incident waves at the object. We consider the problem of detecting a object (e.g, land mine) buried in the land in *real-time.*

The detection of the buried object [32, 90] is an inverse problem [29, 30] because we do not know *a priori* the location and geometric properties of the buried object and if there *is* an object of the kind we are looking for — it is conceivable that our sensors only measured the far-field pattern

of acoustic waves scattered by a rock instead of the object.

Inverse problems are frequently encountered in other engineering disciplines as well: ranging from geophysics [103, 114], to ecology [13], image processing [23], heat transfer [4, 18, 19, 77], physiology [11], continuum mechanics [12], medicine (e.g., hyperthermia treatment) [27, 82], and nondestructive evaluation [48].

For this kind of problem, the computational domain consists typically of the finite domain of the object coupled with the unbounded exterior domain of the scattering field; the Helmholtz equation describes the scattered wave in the unbounded domain with the Laplace operator as the key operator. The unbounded domain is truncated by the prescription of radiation boundary conditions but the computational domain is still very large.

Many methods have been developed to solve inverse problems; see [104] for a fairly recent review. The most common approach [33] is to cast the inverse problem as an optimization problem: a cost functional that measures the difference between observed sensor data and numerically predicted data is defined. The buried object "parameters" (geometry characteristics, and location) are obtained by minimizing the cost functional subject to the governing equations being satisfied.

The inverse problem works by repeatedly solving a large number of "forward" problems: given a specific input parameter vector μ (geometry characteristics, location) in the set $\mathcal{D}$ (all admissible values for the different geometry parameters and locations), we evaluate the output of interest, $s(\mu)$ (the far-field pattern measured at sensor locations). Thus, the inverse problem can be characterized as follows: given the observed output, $s(\mu)$, what is the input parameter, μ, that resulted in this output?

Unfortunately, the solution of the "forward" problem using classical numerical techniques is too computationally intensive. Thus, reduced-order models of the governing equations are included in the optimization problem. However, instead of obtaining only *one* solution from the optimization problem, we now obtain a *set* of feasible inputs [42, 72], that could give us the observed $s(\mu)$. The size of this feasible *set* of inputs obtained from the optimization problem depends on the accuracy of the reduced-order model. *A posterori* error estimators by quantifying the accuracy of the reduced-order model play an important role in *regulating* the size of the feasible set of inputs that give us the observed outputs.

For example, let N denote the size of the reduced-order model, $s_N(\mu)$ the output obtained using

the reduced-order model, and $\Delta_N^s(\mu)$ the associated *a posterori* error estimator. Thus, $|s(\mu) - s_N(\mu)| \leq \Delta_N^s(\mu), \forall \mu \in \mathcal{D}$. The error in the reduced-order modeling acts as noise in the system: larger errors in modeling lead to a larger *feasible* set of *inputs*. Fortunately, $\Delta_N^s(\mu)$ quantifies the reduced-order modeling error; as N increases, $\Delta_N^s(\mu) \rightarrow 0$, and we obtain smaller and smaller *feasible* sets of *inputs*.

Thus, it is possible to introspect the properties of the buried object with a high degree of certainty in *real-time* provided we have access to a sophisticated reduced-order model with *a posteriori* error estimation. This acoustics application demonstrates both the *many-query* and *real-time* contexts we are interested in.

1.1.2 Speech Synthesis

The process of speech production in humans involves several simultaneous processes, one of which results in the production of the source sound for speech. This process, known as phonation, is performed by the vocal folds. The vocal tract acts as an acoustic converter that converts the source sound of the vocal folds into the desired speech sound emanating from the lips. Patients suffering from laryngeal cancer often have their vocal folds removed and thus lose the ability to produce the source sound for speech. The development of efficient artificial sound-producing articulators relies heavily on understanding and replicating the production of voiced sounds in laryngeal phonation.

The cyclic opening and closing of the glottis enables production of voiced sounds. Thus, significant research in the speech synthesis community has been focussed on developing both parametric [36] and physical [56] models of the glottis. It is also recognized that that the glottal excitation plays an important role in the quality of synthetic speech. There are many speech synthesis schemes based on extended models of the glottal excitation [26, 59].

The physical models of glottal excitation are very complex and their solution involves a large number of control parameters. Applications of the glottal models to enable articulatory speech synthesis will require the rapid, and real-time solution of the numerical idealization of the glottal system. Typical inputs will have to include parametrizations of the shape of the vocal tract, and the position of the speech articulators like the tongue, jaw, and lips.

Furthermore, many applications of the physical models often require the glottal model to fit glottal flow waveforms obtained by inverse filtering. While there are a number of studies dealing

with the parametric identification of analytical glottal models to match inverse filtered glottal flow waveforms [25, 38, 102]; parametric studies with numerically discretized physical glottal models are much rarer [35, 101].

Reduced order models can be gainfully introduced to study the effect of the different parameters in optimizing glottal efficiency (usually defined as the ratio of radiated acoustic power to aerodynamic power) by enabling rapid evaluations of the *quadratic* outputs of interest. *A posteriori* error estimation will be crucial in exploring the large parametric dependence optimally and constructing small reduced-order models; furthermore, they can serve as indicators quantifying the error in the underlying parametric models through comparison to actual experimental data. These reduced-order models with *a posteriori* error estimation can then be embedded in artificial speech-synthesizing articulators that need to work for a wide range of frequencies in *real-time*.

1.1.3 Noise Reduction

The rapid growth in ecological standards have imposed increased restrictions on the acceptable noise limit for industrial appliances. Thus, the use of filters to control the noise characteristics have become widespread. Along with the increasing use of filters to control the noise, there has been a push to design acoustic devices that work optimally in a wide range of settings for some specific contexts. For example, the need to control the intense sound fields generated within jet engines has seen extensive use of acoustic liners [24, 50] for noise suppression. Similarly, the use of expansion chambers as low-pass filters in custom-made acoustic mufflers is fairly common [34, 70, 85].

Acoustic liners are used in turbofan engine nacelles to suppress aircraft engine duct noise. They are located on the surface of the engine duct and provide an impedance boundary condition. Current single degree of freedom liners are comprised of a perforated face-sheet backed by honeycomb cells and a rigid plate: each liner cell behaves as a Helmholtz resonator [22, 45, 94, 97] comprised of a porous face-sheet, and honeycomb core with a rigid backing sheet. The liners present a very high impedance to acoustic wave propagation in the engine ducts at their natural frequency; the incident wave is largely blocked and there is significant noise attenuation. The impedance characteristics of these liners depends on the frequency in a non-linear fashion; in many cases they are represented as parameterized functions of the sound pressure level (SPL) [24] and Mach number. There are a number of existing nonlinear impedance models for micro-perforated liners [49, 68]; however,

many of the underlying parameters in such models need to be estimated by fitting curves to match experimental results.

The impedance distribution functions for optimal noise attentuation are not fixed. The optimal impedance condition is a function of the actual engine operating conditions, thus, the desired acoustic impedance of the acoustic liner is determined by geometry and the flight environment. Current efforts [24, 50] are dedicated to the development of electro-mechanical acoustic liners that can tune the acoustic impedance to optimize noise attenuation. These sophisticated liner constructions include Helmholtz resonators with compliant walls — by changing the acoustic compliance of the walls, the overall acoustic impedance characteristics can be tuned. However, the optimal design of these liners still require exhaustive understanding of the mechanisms of noise absorption to quantify the effect of the different acoustic elements.

A number of studies, both experimental and theoretical are dedicated to predicting the acoustic impedance (both resistive and reactive) of these custom-made acoustic liners for the range of operating frequencies. Lumped element models [62, 87] are extensively used at low frequencies [24] when an equivalent circuit representation is sufficient. However, at medium and high frequencies experimental validation of approximate theoretical results is necessary. Since the experimental setup and subsequent testing is quite expensive and time-consuming, preliminary computational modeling of acoustic filters to characterize liner behavior is particularly important.

In these problems, both parameter estimation (to determine impedance characteristics of acoustic liners) as well as geometry optimization (to obtain the most efficient acoustic liner configuration) are extremely important. Accurate reduced order modeling and *error estimation* procedures can be very useful in such contexts.

1.2 Representative Applications in Acoustics

The design of acoustic filters for industrial application requires tuning many parameters related to the geometry of the Helmholtz resonator or the expansion chamber, the acoustic compliance of the liner walls, and the range of operating frequency. In most cases, the repeated solution of the underlying PDE to optimize useful outputs like the sound pressure level (SPL) or the pressure intensity, is infeasible since the computational domain is fairly large. The situation is further complicated when the wave equation has to be solved repeatedly in the time-domain for problems

where the boundary conditions are frequency-dependent.

Here we consider two representative applications in acoustics. We consider the characterization of acoustic stop-band and low-pass filters for a particular range of operating frequencies. Similar to acoustic liners, we include general impedance boundary conditions in our formulations. In the first example (the band-stop filter), the Helmholtz resonator has a compliant backplate of fixed acoustic impedance. In the second example (the low-pass filter), the included expansion chamber has compliant walls; we consider variations in the impedance characteristics of the compliant walls — this is analogous to modeling active acoustic liners of varying impedance.

We consider general impedance boundary conditions in our formulations: the non-dimensional acoustic impedance, Z, can be written as $Z = Z_R + \mathrm{i}\, Z_I$; here Z_R is the (non-dimensional) acoustic resistance and Z_I the (non-dimensional) acoustic reactance. Typical ranges of non-dimensional acoustic impedance obtained from theoretical models and experimental studies [24, 50, 58, 105] of acoustic liners serve as guides in the selection of Z_R and Z_I. We now describe our representative filter problems in greater detail.

Band-Stop Filter

Our first example is the acoustic band-stop filter waveguide element shown in Figure 1-1. A band-stop filter is a device that rejects (attenuates) frequencies within a certain range and passes frequencies outside that range. The band-stop filter waveguide element under consideration consists of a rectangular acoustic waveguide coupled with a Helmholtz resonator [93, 94] with a fixed resistive impedance specified on its top wall.

Helmholtz resonators enable significant noise reduction in a narrow band around their natural or resonant frequency. The geometry of the Helmholtz resonator (the height of the cavity) and the small neck connecting the resonator to the main waveguide are important factors in the design of the band-stop filter — variations in the geometric properties of the cavity and the neck induce variations in the resonant frequency. The use of Helmholtz resonators as side-branches to tubes/waveguides to behave as band-stop filter elements is quite common. In many examples, multiple array resonators are used to broaden the narrow band characteristics of a single resonator to cover a low-frequency band [97].

In this example, we consider variations in the height of the resonator and examine its effect on

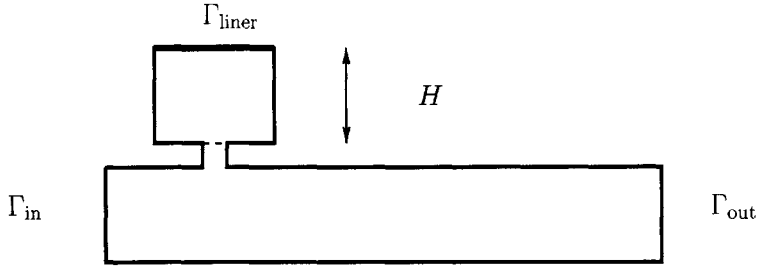


Figure 1-1: Acoustic Band-Stop Filter Geometry.

the the stop-band across a range of operating frequencies, in the low-frequency limit.

In particular, we are interested in computing the transmission coefficient (TC) for the band-stop filter which is given by

$$TC = -20\, log_{10}\left(\frac{\sqrt{\int_{\Gamma_{\text{out}}} p^* p}}{\sqrt{\int_{\Gamma_{\text{in}}} p^* p}}\right); \tag{1.1}$$

it is the logarithm of the ratio of the pressure intensities at the output and input faces respectively.

The aim is to characterize the transmission coefficient (TC) for a range of operating frequencies and for a number of different Helmholtz resonators with different cavity geometries. We present in Figure 1-2 a number of these transmission coefficient curves for different choices of the height of the Helmholtz resonator H across the frequency range under consideration. The Helmholtz resonator at its resonant frequency (depending on the geometry) reflects the acoustic wave upstream; for other frequencies it simply passes the signal.

The "inputs" or parameters for this problem are the (non-dimensional) frequency squared, $\mu_1 \equiv k^2$, and the height of the Helmholtz resonator cavity (relative to the reference cavity height $H_{\text{ref}} = 1$), $\mu_2 \equiv H$. The governing equation is the Helmholtz acoustic equation; we use an acoustic liner of fixed impedance at the top of the resonator cavity.

Low-Pass filter

The use of expansion chambers [79, 92] for broadband attenuation in exhaust mufflers of internal combustion engines is widespread. Automotive exhaust systems use either dissipative or reactive silencers for broadband frequency attentuation. In contrast to dissipative silencers which use sound absorbing material to take energy out of the acoustic motion in the wave, reactive mufflers work by reflecting the sound waves back to the source; both of these types use expansion chambers.

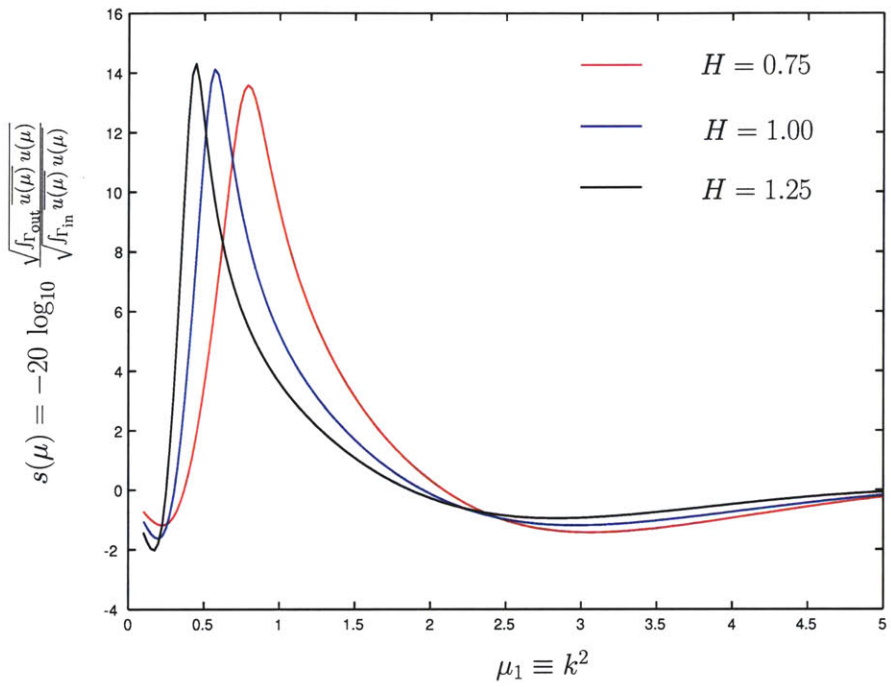


Figure 1-2: Transmission Coefficient curves for the Acoustic Band-Stop Filter Waveguide. The acoustic impedance of the liner is purely resistive, we fix the non-dimensional acoustic resistance $Z_R = 10.0$. We show the transmission coefficient for different H across the range of frequency. The peaks correspond to the resonant frequency of the Helmholtz resonator. $\mu_1 \equiv k^2$ is the square of the non-dimensional frequency. The TC curves are obtained from direct finite-element simulation.

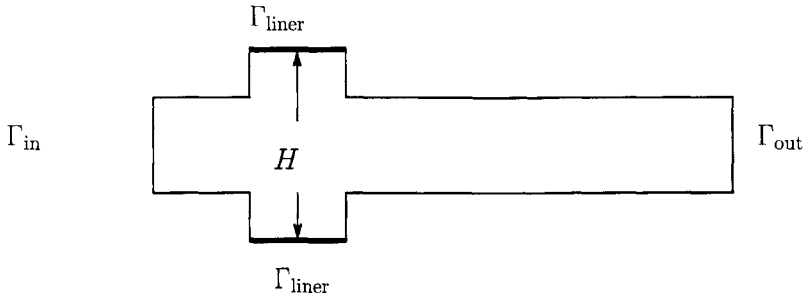


Figure 1-3: Acoustic Low-Pass Filter Geometry.

In our second example, we consider the acoustic low-pass filter waveguide element shown in Figure 1-3. The low-pass filter element consists of an expansion chamber inside the waveguide element; the sudden change in the cross-sectional area (the enlarged size of the expansion chamber relative to the waveguide cross-section) can generate reflection of sound even in cases where the medium is the same. The walls of the expansion chamber are compliant — we assume that the impedance characteristics can be changed within certain ranges.

We consider variations in the width of the expansion chamber (relative to the width of the waveguide) and observe its effect on our "output" of interest, the Transmission Coefficient (TC) (1.1) which represents the log of the ratio of the output pressure intensity to the input pressure intensity. We also allow small variations in the acoustic resistance and reactance of the compliant walls of the expansion chamber; this allows us to consider a range of impedance boundary conditions in our formulation similar to those obtained with adaptive acoustic liners.

We identify the compliant walls of the expansion chamber as "liners". The parameters for this problem are the (non-dimensional) frequency squared, $\mu_1 \equiv k^2$, the width of the low-pass expansion (relative to the nominal waveguide width), $\mu_2 \equiv H$, the non-dimensional acoustic resistance of the liner walls, $\mu_3 \equiv Z_R$, and the non-dimensional acoustic reactance of the liner walls, $\mu_4 \equiv Z_I$ respectively. The governing equation is again the Helmholtz equation.

In Figure 1-4 we present the variation of the transmission coefficient as a function of the non-dimensional wave-number squared for different choice of H. The filter element allows the signal to pass for low frequency, but at higher frequencies the signal is attenuated. Figure 1-4 shows the effect of the width of the expansion chamber on the transmission coefficient.

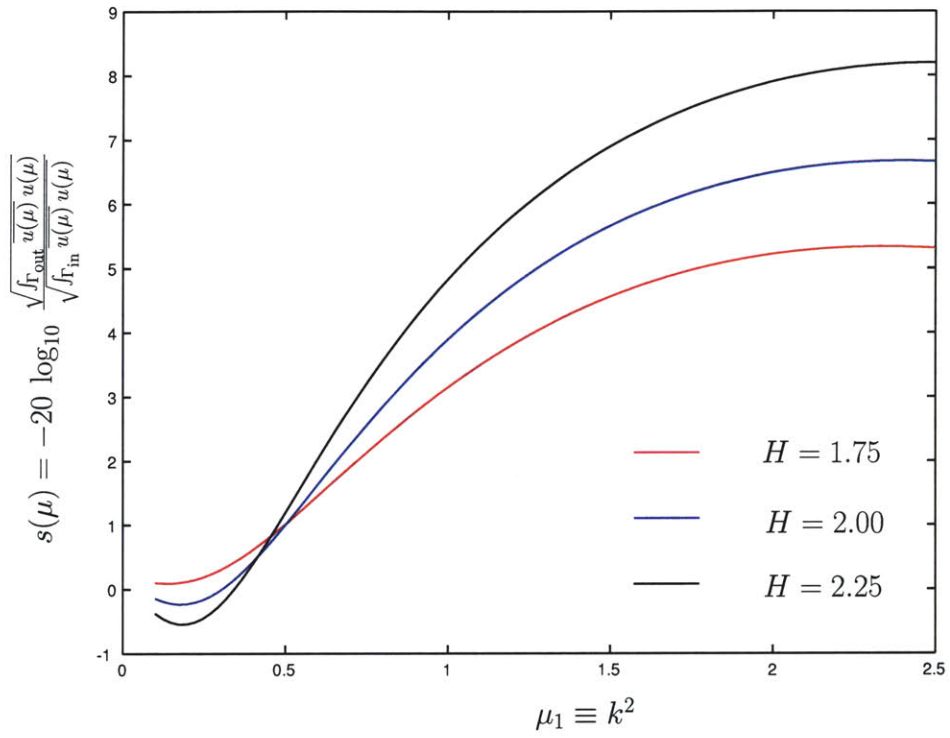


Figure 1-4: Low-pass filter transmission coefficient as a function of k^2 for different H; the non-dimensional acoustic resistance $Z_R = 4.8$ and reactance $Z_I = 1.8$. We show the transmission coefficient for different H across the range of frequency. At low frequencies, there is very little attenuation; attenuation increases with frequency. We see that changing H affects the degree of attenuation — increasing the width of the expansion chamber leads to greater attenuation. The TC curves are obtained from direct finite-element simulation.

1.2.1 Computational Challenge/Thesis Objectives

The acoustic filter applications described above have several common features. First, the governing equations are non-symmetric, non-coercive partial differential equations where the complex-valued field solution strongly depends on the parameters characterizing the problem. Second, the physical problems are of semi-infinite extent — truncation of the physical domain to make the numerical problem tractable requires the specification of an appropriate radiation boundary condition at the outflow. Finally, the efficient solution of the problem requires the fast (possibly real-time) and reliable evaluation of *quadratic* outputs of interest.

In actual practice, of course, we do not have access to the analytic solution of these problems and a discretization procedure such as the finite element method is used. The algebraic equations obtained using these procedures are, in general, very high-dimensional and their solution is expensive — applying these methods in the many-query context is thus prohibitive, and real-time solution infeasible.

We focus on the development of computational methods that permit *accurate*, *reliable*, and *rapid* evaluation of input-output relationships induced by non-coercive elliptic partial differential equations in *real-time* and *in the limit of many queries*. There are several aspects to non-coercive reduced basis approximation and error estimation: construction of appropriate spaces that ensure stability of the discrete approximation; treatment of resonance/singularity and in particular near-resonant systems; and development of inf-sup lower bounds for our *a posteriori* error estimators.

In addition, the acoustics abstraction introduces several other problems that need to be addressed: first, the introduction of radiation boundary conditions that enables the domain truncation; and second, the development of accurate error bounds for quadratic outputs (e.g., the transmission coefficient).

The generation of inf-sup lower bounds, a crucial ingredient in our output bounds, is complicated further by the complex nature of the problems. Furthermore, the choice of the parametrization and the nature of the problems (Helmholtz, non-coercive) require crucial improvements in the offline algorithms used to approximate the inf-sup lower bound.

To address the issue of quadratic outputs, we will need to develop a new expanded formulation and appropriate reduced basis approximation spaces that enables their treatment as linear outputs. The expanded formulation needs to be general enough to handle linear outputs or mixed linear and

quadratic outputs consistently.

For the highlighted problems, we seek to develop techniques for (*i*) accurate approximation of the relevant outputs of interest; (*ii*) inexpensive and rigorous error bounds yielding upper bounds for the error in the approximation; and (*iii*) a computational framework which allows rapid online calculation of the output approximation and associated error bounds.

1.3 Approach: The Reduced Basis Method

Our particular approach is based on the reduced basis method. The reduced basis method was first introduced in the late 1970s for nonlinear structural analysis [5, 74], and subsequently developed more broadly in the 1980s and 1990s [10, 15, 37, 46, 57, 80, 81, 86] for a much larger class of parametrized partial differential equations. The basic premise of the reduced basis method is that the field variable is not, in fact, some arbitrary member of the infinite-dimensional solution space associated with the partial differential equation; rather, it resides, or "evolves," on a much lower-dimensional manifold induced by the parametric dependence.

The earlier reduced basis approaches were in some sense local approximations in parameter space — the Lagrangian or Taylor approximation spaces for the low-dimensional manifold were typically defined relative to a particular parameter point; and the associated *a priori* convergence theory relied on asymptotic arguments in sufficiently small neighborhoods [37]. The local nature of the approximation meant that computational economies were never significant [81] when compared to conventional numerical procedures like the finite element method. Recent advances in the development of reduced basis methods [44, 64, 67, 73, 84, 109, 110, 111] differ from these earlier efforts in several important ways: first, *global* Lagrangian approximation spaces are developed; second, rigorous *a posteriori error estimators* are introduced; and third, *off-line/on-line* computational decompositions are exploited. These key ingredients allow us to decouple the generation and projection stages of reduced basis approximation for problems where the bilinear form a has an affine decomposition of the form $a(\cdot,\cdot;\mu) = \sum_{q=1}^{Q} \Theta_q(\mu)\, a^q(\cdot,\cdot)$, thereby effecting significant computational gains.

A posteriori error estimates for reduced basis approximations have been developed in a number of contexts. In particular, *a posteriori* error bounds have been successfully developed for (*i*) linear [44, 64, 67, 84, 111] and (*ii*) at most quadratically nonlinear [73, 109, 110] elliptic partial differential

equations that are affine in the parameter. These two assumptions enable the development of very efficient offline-online computational strategies relevant in the many-query and real-time contexts. The operation count for the online stage — in which, given a new parameter value, we calculate the reduced basis output and associated error bound — depends only on the dimension of the reduced basis space (typically small) and on the parametric complexity of the problem, but is *independent* of the dimension of the underlying "truth" finite element approximation space (typically very large). However, the requirement of a dual approximation and associated Petrov-Galerkin approximation spaces to deal with non-compliant outputs leave room for further improvement.

Non-affine parameter dependence can also be tackled in a rigorous manner. Locally non-affine parameter dependence in a small part of the domain has been treated in [100] while more general non-affine problems can be addressed using the empirical interpolation method [14]. These approaches enable the recovery of the efficient offline-online decomposition which is critical to effect computational savings. Finally, reduced basis approximations and error estimators have also been developed in parabolic partial differential equations [42, 88, 89] for and the theory extended to deal with non-affine and non-linear problems.

1.4 Scope

1.4.1 Thesis Contributions

In this thesis we focus on the development of reduced basis output bound methods and associated *a posteriori* error estimation for parametrized non-coercive elliptic partial differential equations with quadratic outputs. We improve and extend on earlier work [72, 89, 108] in this field in several directions.

First, we build on classical approaches to develop a radiation outflow boundary condition which is essential when considering acoustics waveguide applications on semi-infinite domains. The radiation outflow boundary conditon is very general and allows for the treatment of multiple propagating modes in a consistent manner. We demonstrate the efficacy of the developed Robin boundary condition in the acoustics applications described above. Furthermore, the condition allows us to truncate the domain considerably without sacrificing solution accuracy, thus reducing offline cost.

Second, we introduce a new expanded formulation for the treatment of quadratic outputs.

This formulation enables us to eliminate the nonlinearity of the output; the transformation into a "linear" output allows us to develop *sharp* error estimates. Furthermore, it eliminates the need to introduce a *dual* approximation by changing the quadratic output into a compliant output for a new problem. The new formulation leads to increased (mild) offline effort related to the generation of the reduced basis and the solution of eigen-problems for the inf-sup lower bound approximation; however the *online* complexity is sharply reduced by eliminating the need for the dual problem.

Third, we develop a new successive constraint formulation for treatment of inf-sup lower bounds for complex non-coercive elliptic problems. The formulation is very general and can be used to obtain lower bounds for coercive problems with very few modifications. Furthermore, the algorithm requires far less computational resources compared to existing approaches like the natural norm formulation [96]. We also outline a new greedy adaptive sampling algorithm used *offline* for the selection of a set of local inf-sup approximants; in the *online* stage, the formulation requires the solution of a standard Linear Program (LP) to obtain rigorous inf-sup lower bounds.

Finally, we bring the important pieces together — radiation boundary conditions, treatment of quadratic outputs and inf-sup lower bound approximation for complex-valued problems — and outline the reduced basis formulation and *a posteriori* error estimation procedure for complex non-coercive problems with quadratic outputs. We apply our methods to certain illustrative acoustic applications representative of solution contexts requiring repeated and rapid evaluations of the outputs of interest. We demonstrate in each case, the validity of the methods, the accuracy of the error bounds, and the applicability in *real-time* contexts.

1.4.2 Thesis Outline

Before proceeding with the development of the reduced basis method for non-coercive elliptic problems, in Chapter 2 we review the necessary mathematical background that will be required throughout the thesis. In Chapter 3, we present the abstract formulation for acoustics waveguide problems. We introduce the radiation boundary condition which will be repeatedly used in the applications and outline the expanded formulation which allows us to tackle quadratic outputs.

The reduced basis method is always applied in a parameter-independent reference domain. In Chapter 4, we introduce the affine mappings for our application problems and summarize the respective reference-domain abstract formulations associated with the "exact" problems. In Chap-

ter 5, we introduce the "truth" finite element approximation space and related linear and bilinear functionals. The accuracy of the reduced basis approximation will be measured relative to this "truth" approximation. We also present results to confirm the accuracy of the radiation boundary condition.

We summarize the reduced basis method formulation for non-coercive elliptic problems with quadratic outputs in Chapter 6. We describe the slightly modified reduced basis approximation spaces to handle the expanded formulation, and the construction of the space using a greedy algorithm for the selection of the bases.

The successive constraint formulation for the evaluation of inf-sup lower bounds for complex, non-coercive problems is discussed in Chapter 7. We apply the method to a model problem and the low-pass filter waveguide problem and present results for the inf-sup lower bound approximation.

We describe *a posteriori* error estimation procedures for the expanded weak formulation in Chapter 8. We particularly focus on the form of our output error bounds and on efficient offline-online computational procedures for key ingredients in the error bounds. We also remark upon the rigor and sharpness of our output bounds.

In Chapter 9, we apply the methods developed in this thesis to the acoustics waveguide applications. We briefly describe the formulation, present the "truth" solutions for the problems and then present results for the reduced basis approximation, the inf-sup lower bounds and the *a posteriori* error estimators for our quadratic outputs of interest.

Finally, in Chapter 10 we summarize our work and conclude with some suggestions for future work.

Chapter 2

Preliminaries

2.1 Introduction

Before proceeding with the development of reduced basis methods, in this chapter we provide some necessary mathematical background that will be used in the remaining chapters. In Section 2.2 we review real function spaces; we then extend the same ideas in Section 2.3 to deal with the complex function spaces used extensively throughout the thesis; in latter Sections we introduce the parametric linear and bilinear forms used in our reduced basis methods.

2.2 Real Function Spaces

In this section, we introduce some notation and review some basic definitions of real function spaces. The summary provided here is largely based on [3, 63, 71, 76, 89]. To begin, let $\Omega \subset \mathbb{R}^d$, $d = 1, \ldots, 3$ be an open bounded domain with Lipschitz-continuous boundary $\partial\Omega$; we denote the closed domain by $\overline{\Omega}$.

2.2.1 Linear Spaces

Definition 1. *Let* $\mathbb{R}$ *be a real* algebraic field. *A* linear vector space X *over the field* $\mathbb{R}$ *is a set of elements together with two operations,* $u, v \in X : u + v \in X$ *(addition) and* $\alpha \in \mathbb{R}, v \in X, v \in X :$ $\alpha v \in X$ *(scalar multiplication), if the following axioms hold*

(1) $u + v = v + u$ *(commutative) ;*

(2) $(u+v)+w = u+(v+w)$ *(associative) ;*

(3) $\exists \mathbf{0}$ *such that* $u+\mathbf{0} = u$ *for all* $u \in X$ *(null vector) ;*

(4) For each $u \in X$, $\exists -u \in X$ *such that* $u+(-u) = \mathbf{0}$ *(additive inverse vector) ;*

(5) $(\alpha\beta)u = \alpha(\beta u)$ *(associative) ;*

(6) $(\alpha+\beta)u = \alpha u + \beta u$ *(distributive) ;*

(7) $\alpha(u+v) = \alpha u + \alpha v$ *(distributive) ;*

(8) $1u = u$.

We restrict our attention to linear vector spaces of finite dimension, $\dim(X)$, finite. A basis set for a linear space X is a set of linearly independent elements $x_j \in X$, $1 \leq j \leq \dim(X)$, such that for all $u \in X$ there exists a *unique* set of real numbers, $\alpha_j \in \mathbb{R}$, $1 \leq j \leq \dim(X)$, such that

$$u = \sum_{j=1}^{\dim(X)} \alpha_j \, x_j; \tag{2.1}$$

note however that the choice of the basis set is not unique. We can describe X in terms of any basis set: $X = \text{span}\{x_j,\ 1 \leq j \leq \dim(X)\}$.

2.2.2 Inner Product Space

A (real) inner product space X is a *linear* vector space equipped with an inner product $(w,v)_X$, $\forall w, v \in X$, and induced norm $\|w\|_X$. We now define the inner product and associated induced norm.

Inner Product

Definition 2. *An* inner product *on* X *is a scalar valued function on* $X \times X \rightarrow \mathbb{R}$*, whose values are denoted by* $(w,v)_X$*, that satisfies the following axioms (i)* $(w,v)_X = (v,w)_X, \forall\, w,v \in X$ *(symmetric); (ii)* $(w,w)_X \geq 0, \forall\, w \in X$ *and* $(w,w)_X = 0$ *if and only if* $w = 0$ *(positive definite); (iii)* $(w+v,u)_X = (w,u)_X + (w,v)_X, \forall\, w,v \in X$ *and* $(\alpha w, v)_X = \alpha(w,v)_X, \forall\, w,v \in X, \forall\, \alpha \in \mathbb{R}$ *(bilinear).*

Norm

Definition 3. *A function* $\|\cdot\|_X : X \to \mathbb{R}$ *is called a* norm *if and only if it has the following properties*

(i) (a) $\|w\|_X \geq 0$, *and (b)* $\|w\|_X = 0$ *if and only if* $w = 0$;

(ii) $\|\alpha w\|_X = |\alpha| \|w\|_X, \forall\, \alpha \in \mathbb{R}$;

(iii) $\|w + v\|_X \leq \|w\|_X + \|v\|_X, \forall\, w, v \in X$ *(the triangle inequality)* .

If $\|w\|_X$ *satisfies (ia), (ii), and (iii) only, we call it a* seminorm *of the vector* w, *and denote it by* $|w|$. *A linear vector space* X *together with a norm defined on itself is a* normed space.

It directly follows from the conditions on the inner product, that we can induce a valid norm associated with the inner product

$$\|w\|_X = \sqrt{(w, w)_X}. \tag{2.2}$$

The Cauchy-Schwarz inequality,

$$|(w, v)_X| \leq \|w\|_X \|v\|_X, \tag{2.3}$$

follows from the inner product definition.

A Hilbert space is a complete inner product space [3]. We use Hilbert spaces extensively in the subsequent chapters.

2.2.3 Cartesian Product Spaces

The Cartesian product of two inner product spaces X_1 and X_2 is defined as $X = X_1 \times X_2 \equiv \{(w_1, w_2) \mid w_1 \in X_1, w_2 \in X_2\}$. Given $w = (w_1, w_2) \in X$, $v = (v_1, v_2) \in X$, we define

$$w + v \equiv (w_1 + v_1, w_2 + v_2) \ ; \tag{2.4}$$

it directly follows that X is a *linear* space. X is associated with the inner product

$$(w, v)_X = (w_1, v_1)_{X_1} + (w_2, v_2)_{X_2} \tag{2.5}$$

and induced norm

$$\|w\|_X = \sqrt{(w,w)_X}\ . \tag{2.6}$$

$(\,\cdot\,,\,\cdot\,)_X$ is a valid inner product and hence X an inner product space, although the choice of the inner product is not unique.

2.2.4 Spaces of Continuous Functions

Definition 4. *For a non-negative integer m, the space $C^m(\overline{\Omega})$ is defined as*

$$C^m(\overline{\Omega}) \equiv \{v \mid D^\alpha v \text{ is bounded and uniformly continuous on } \Omega,\ \forall\, \alpha \text{ s.t. } 0 \le |\alpha| \le m\}\,;$$

where, for given multi-index $\alpha \equiv (\alpha_1, \ldots, \alpha_d)$, $\alpha_i \ge 0$, $1 \le i \le d$,

$$D^\alpha \equiv \frac{\partial^{|\alpha|}}{\partial_{x_1}^{\alpha_1} \cdots \partial_{x_d}^{\alpha_d}}, \qquad |\alpha| = \sum_{i=1}^{d} \alpha_i.$$

Then $C^m(\overline{\Omega})$ is a Banach space (i.e., a complete normed linear space) with a norm

$$\|v\|_{C^m(\overline{\Omega})} = \max_{0 \le |\alpha| \le m} \sup_{x \in \Omega} |D^\alpha v(x)|.$$

It is clear that $C^m(\Omega)$ defined above is a Banach space, i.e. a complete normed linear space. Also note that $C_0^m(\Omega)$ is the space of continuous, m^{th} differentiable functions with compact support, i.e. vanishing near the boundary of Ω. We shall use the subscript 0 to indicate spaces with functions of compact support.

2.2.5 Lebesgue Spaces

Definition 5. *For $1 \le p \le \infty$, the Lebesgue space $L^p(\Omega)$ is defined as*

$$L^p(\Omega) \equiv \left\{v \mid \|v\|_{L^p(\Omega)} < \infty\right\}$$

with an associated norm

$$\begin{aligned} \|v\|_{L^p(\Omega)} &\equiv \left(\int_\Omega |v|^p \, dx\right)^{\frac{1}{p}}, \quad 1 \le p < \infty, \\ \|v\|_{L^\infty(\Omega)} &\equiv \operatorname*{ess\,sup}_{x\in\Omega} |v(x)|, \qquad p = \infty. \end{aligned}$$

These spaces are also Banach spaces. The ess sup (essential supremum) in the above definition means the smallest supremum over $\Omega \backslash B$ for all sets B of zero measure.

2.2.6 Hilbert Space $H^m(\Omega)$

Definition 6. *For a non-negative integer m, the Hilbert space $H^m(\Omega)$ is defined as*

$$H^m(\Omega) \equiv \left\{ v \mid D^\alpha v \in L^2(\Omega), \ \forall \, \alpha \ s.t. \ |\alpha| \le m \right\},$$

with associated inner product

$$(w, v)_{H^m(\Omega)} \equiv \sum_{|\alpha|\le m} \int_\Omega D^\alpha w \, D^\alpha v \, dx,$$

and induced norm

$$\|v\|_{H^m(\Omega)} \equiv \left(\sum_{|\alpha|\le m} \int_\Omega |D^\alpha v|^2 \, dx \right)^{\frac{1}{2}}.$$

Hilbert spaces, which are the natural generalization of Euclidean spaces in the functional setting, will be used extensively in the subsequent chapters. We note that $L^2(\Omega)$ ($\equiv H^0(\Omega)$) is the only Lebesgue space that is a Hilbert space. Finally, since the Hilbert norm is induced by an inner-product, the Cauchy-Schwarz inequality holds:

$$|(w, v)_{H^m(\Omega)}| \le \|w\|_{H^m(\Omega)} \|v\|_{H^m(\Omega)}.$$

These spaces are important not only in understanding well-posedness of weak statements, but also in expressing the convergence rate of the finite element method. In addition, we introduce $H^m(\Omega)$ *semi-norm* as

$$|v|_{H^m(\Omega)} = \left(\int_\Omega |D^m v|^2 \, dx \right)^{\frac{1}{2}}, \tag{2.7}$$

which include only the m^{th} derivative.

2.2.7 Dual Hilbert Spaces

Definition 7. *Given a functional f, Hilbert space X, and associated inner product and norm, $(\cdot,\cdot)_X$ and $\|\cdot\|_X$, respectively, we define the corresponding dual space X' as*

$$X' \equiv \{f \mid \|f\|_{X'} < \infty\},$$

where the dual norm $\|\cdot\|_{X'}$ is given by

$$\|f\|_{X'} \equiv \sup_{v \in X} \frac{f(v)}{\|v\|_X}. \tag{2.8}$$

The space X' is also a Hilbert space, and for $X = H^m(\Omega)$, we denote $X' = H^{-m}(\Omega)$; in general:

$$H^m(\Omega) \subset \cdots \subset H^1(\Omega) \subset H^0(\Omega) \subset H^{-1}(\Omega) \subset \cdots \subset H^{-m}(\Omega).$$

From the Riesz representation theorem we know that for every $f \in X'$ there exists a unique $u_f^X \in X$ such that

$$(u_f^X, v)_X = f(v), \quad \forall\, v \in X.$$

It follows from the Cauchy-Schwarz inequality applied to the X-inner product that

$$\|f\|_{X'} = \sup_{v \in X} \frac{(u_f^X, v)_X}{\|v\|_X} = \|u_f^X\|_X.$$

This result is widely used in subsequent chapters.

2.3 Complex Function Spaces

The acoustics problems of interest require the use of complex vector spaces. Here, we we discuss necessary extensions. The complex case is treated as a "conjugation" of the real case with $\mathbb{R}$ replaced by $\mathbb{C}$ wherever needed; for more elaboration on this material, please see [76].

2.3.1 Preliminaries

We first recall the preliminary notations. For $y \in \mathbb{C}$ (a complex number), $y = y_R + \mathrm{i}\, y_I$, where subscript "$R$" (respectively, "$I$") refers to the real (respectively, imaginary) part, and $\mathrm{i} = \sqrt{-1}$. We denote the complex conjugate of y as $\overline{y} = y_R - \mathrm{i} y_I$, and the modulus of y as $|y| = (y_R^2 + y_I^2)^{1/2} = (y\,\overline{y})^{1/2}$.

2.3.2 Inner Product Space

We can form a complex inner product space X_C from a corresponding real inner product space X_R [76] by considering pairs of vectors from X_R. We extend the underlying vector space as $X_C = X_R \times X_R$; for any two of members of X_C, $w_1 = (w_{R,1} \in X_R,\ w_{I,1} \in X_R)$, $w_2 = (w_{R,2} \in X_R,\ w_{I,2} \in X_R)$, we have

$$w_1 + w_2 \equiv (w_{R,1} + w_{R,2}, w_{I,1} + w_{I,2});$$

for any complex number $\alpha = \alpha_R + \mathrm{i}\,\alpha_I \in \mathbb{C}$,

$$\alpha\, w_1 \equiv (\alpha_R\, w_{R,1} - \alpha_I\, w_{I,1}, \alpha_I\, w_{R,1} + \alpha_R\, w_{I,1}). \tag{2.9}$$

Assuming that the real vector space X_R is a Hilbert space satisfying $H_0^1(\Omega) \subset X_R \subset (H^1(\Omega))^d$, we define our space X_C as

$$X_C \;=\; \{v = v_R + \mathrm{i}\, v_I \,|v_R \in X_R, v_I \in X_R\}. \tag{2.10}$$

The complex inner product space X_C is equipped with an inner product $(\cdot,\cdot)_{X_C}$ and an induced norm; X_C reduces to X_R for real elements.

Inner Product and Norm

We introduce our inner product as $(\cdot,\cdot)_{X_C}$: $X_C \times X_C \rightarrow \mathbb{C}$ as

$$(w_1, w_2)_{X_C} = ((w_{R,1}, w_{R,2})_{X_R} + (w_{I,1}, w_{I,2})_{X_R}, (w_{I,1}, w_{R,2})_{X_R} - (w_{R,1}, w_{I,2})_{X_R}); \tag{2.11}$$

note that now $(w_2, w_1) = \overline{(w_1, w_2)}$. As usual, our inner product induces a (well-defined) norm $\|w\|_{X_C}$ which is *real-valued*.

If $(w,v)_{X_R} \equiv (w,v)_{H^1(\Omega)} = \int_\Omega \nabla w \cdot \nabla v + wv$, we can extend the definition for the inner product for X_C as

$$(w,v)_{X_C} = \int_\Omega \nabla w \cdot \nabla \overline{v} + w\,\overline{v}; \tag{2.12}$$

and induced norm

$$\|w\|_{X_C} = \sqrt{(w,w)_{X_C}}. \tag{2.13}$$

2.3.3 Complex Hilbert Space $Z^m(\Omega)$

Definition 8. *For a non-negative integer m, the complex Hilbert space $Z^m(\Omega)$ is defined as*

$$Z^m(\Omega) = \{v = v_R + \mathrm{i}v_I \mid v_R \in H^m(\Omega), v_I \in H^m(\Omega)\}\,; \tag{2.14}$$

with associated inner product

$$(w,v)_{Z^m(\Omega)} = \sum_{|\alpha|\le m} \int_\Omega D^\alpha w \cdot D^\alpha \bar{v}\, dx\ , \tag{2.15}$$

and induced norm

$$\|v\|_{Z^m(\Omega)} = \left(\sum_{|\alpha|\le m} \int_\Omega |D^\alpha v|^2\, dx \right)^{\frac{1}{2}}\ . \tag{2.16}$$

Here and throughout this thesis subscript R and I denote the real and imaginary part, respectively, that is, $v_R = \mathrm{Re}(v)$ and $v_I = \mathrm{Im}(v)$; $\bar{v}$ and $|v|$ shall denote the complex conjugate and modulus of v, respectively.

2.4 Linear Forms and Dual Spaces

We say $\ell\colon X_C \to \mathbb{C}$ is antilinear (linear in the real case) if and only if it satisfies

$$\ell(\alpha u + \beta v) = \overline{\alpha}\,\ell(u) + \overline{\beta}\,\ell(v), \quad \forall\, u,v \in X_C, \forall\, \alpha,\beta \in \mathbb{C}\ .$$

The linear form ℓ is *bounded*, or continuous, over X_C if

$$|\ell(v)| \le C\|v\|_{X_C}, \qquad \forall\, v \in X_C\ , \tag{2.17}$$

for some finite real constant C.

The *dual space* of X_C, X'_C, is the space of all linear bounded functionals over X_C. If ℓ is a member of X'_C, we write $\ell(v) = (\ell, v)$ where $(\cdot, \cdot)$ denotes the duality pairing on $X'_C \times X_C$. X'_C is a normed space equipped with the norm

$$\|\ell\|_{X'_C} = \sup_{v \in X_C} \frac{|\ell(v)|}{\|v\|_{X_C}}, \qquad \forall\, \ell \in X'_C\ , \tag{2.18}$$

which we shall refer to as the "dual norm".

According to the Reisz Representation Theorem [76], there exists a unique element $u \in X_C$ such that

$$\ell(v) = (v, u)_{X_C},\ \forall v \in X_C. \tag{2.19}$$

Moreover,

$$\|\ell\|_{X'_C} = \|u\|_{X_C}; \tag{2.20}$$

thus the dual norm $\|\ell\|_{X'_C}$ is unique. This result is essentially unchanged from the real case.

2.5 Bilinear Forms

Let X_C and Y_C be two complex inner product spaces over the field $\mathbb{C}$, an operator $a : X_C \times Y_C \to \mathbb{C}$ that maps $(u, v), u \in X_C, v \in Y_C$ into $\mathbb{C}$ is called a sesquilinear (bilinear in the real case) form if and only if it satisfies

$$a(\alpha u_1 + \beta u_2, \gamma v_1 + \lambda v_2) = \alpha\overline{\gamma} a(u_1, v_1) + \alpha\overline{\lambda} a(u_1, v_2) + \beta\overline{\gamma} a(u_2, v_1) + \beta\overline{\lambda} a(u_2, v_2)$$

for all $u_1, u_2 \in X_C, v_1, v_2 \in Y_C, \alpha, \beta, \gamma, \lambda \in \mathbb{C}$.

A bilinear form a: $Y_C \times Y_C \to \mathbb{C}$ is *symmetric* or *Hermitian* if, for any $w, v \in Y_C$, $a(w, v) = \overline{a(v, w)}$. A bilinear form a: $Y_C \times Y_C \to \mathbb{C}$ is *skew-symmetric* or *skew-Hermitian* if, for any $w, v \in Y_C$, $a(w, v) = -\overline{a(v, w)}$. We define the symmetric and skew-symmetric parts of a general bilinear form

a: $Y_C \times Y_C \rightarrow \mathbb{C}$ as

$$a_{\mathrm{H}}(w,v) = \frac{1}{2}\,(\,a(w,v) + \overline{a(v,w)}\,),$$
$$a_{\mathrm{SH}}(w,v) = \frac{1}{2}\,(\,a(w,v) - \overline{a(v,w)}\,),$$

$\forall\, w, v, \in Y_C$, respectively.

a: $Y_C \times Y_C \rightarrow \mathbb{C}$ is *positive definite* if, for any $v \in Y_C, v \neq 0$, $a(v,v)$ $(= a_{\mathrm{H}}(v,v)) > 0$. a: $Y_C \times Y_C \rightarrow \mathbb{C}$ is *positive semidefinite* if, for any $v \in Y_C$, $a(v,v) \geq 0$. An inner product is simply a *symmetric positive-definite* (SPD) bilinear form.

A bilinear form a: $Y_C \times Y_C \rightarrow \mathbb{C}$ is said to be *coercive* over Y_C if

$$\alpha \equiv \inf_{w \in Y_C} \frac{a(w,w)}{\|w\|^2_{Y_C}} \tag{2.21}$$

is positive. We say that a bilinear form a: $Y_C \times Y_C \rightarrow \mathbb{C}$ is *continuous* over Y_C if

$$\gamma \equiv \sup_{w \in Y_C} \sup_{v \in Y_C} \frac{|a(w,v)|}{\|w\|_{Y_C}\,\|v\|_{Y_C}} \tag{2.22}$$

is finite. For a coercive, continuous bilinear form α is denoted the coercivity constant and γ is denoted the continuity constant.

For non-coercive bilinear forms a: $Y_{C,1} \times Y_{C,2} \rightarrow \mathbb{C}$, we consider the more general "inf-sup" constant [7] as

$$\beta \equiv \inf_{w \in Y_{C,1}} \sup_{v \in Y_{C,2}} \frac{|a(w,v)|}{\|w\|_{Y_{C,1}}\,\|v\|_{Y_{C,2}}} \tag{2.23}$$

and require β to be positive for stability.

2.6 Parametric Linear and Bilinear Forms

Let $\mathcal{D} \subset \mathbb{R}^P$ be a closed bounded parameter domain. We denote a typical parameter vector, or P-tuple, in $\mathcal{D}$ by $\mu \equiv (\mu_1, \ldots, \mu_P)$.

We shall say that ℓ: $Y_C \times \mathcal{D} \rightarrow \mathbb{C}$ is a *parametric linear form* if, for all $\mu \in \mathcal{D}$, $\ell(\,\cdot\,;\mu) : Y_C \rightarrow \mathbb{C}$ is a linear form. $\ell(\cdot;\mu)$ is bounded (or continuous) if, for all $\mu \in \mathcal{D}$, $\ell(\,\cdot\,;\mu) \in Y_C'$.

Similarly, we shall say that a: $Y_{C,1} \times Y_{C,2} \times \mathcal{D} \rightarrow \mathbb{C}$ is a *parametric bilinear form* if, for all

$\mu \in \mathcal{D}$, $a(\cdot, \cdot; \mu)$: $Y_{C,1} \times Y_{C,2} \rightarrow \mathbb{C}$ is a bilinear form. The parametric bilinear form a: $Y_C \times Y_C \rightarrow \mathbb{C}$ is *symmetric* if $a(w, v; \mu) = \overline{a(v, w; \mu)}$, $\forall\, w, v \in Y_C$, $\forall\, \mu \in \mathcal{D}$. We define the symmetric part as $a_{\rm H}(w, v; \mu) \equiv \frac{1}{2}(a(w, v; \mu) + \overline{a(v, w; \mu)})$, $\forall\, w, v \in Y_C$, $\forall\, \mu \in \mathcal{D}$. a is considered to be coercive over Y_C if

$$\alpha(\mu) \equiv \inf_{w \in Y_C} \frac{a(w, w; \mu)}{\|w\|^2_{Y_C}} \tag{2.24}$$

is positive for all $\mu \in \mathcal{D}$; we can then define $(0 <)\ \alpha_0 \equiv \min_{\mu \in \mathcal{D}} \alpha(\mu)$.

The parametric bilinear form a: $Y_C \times Y_C \times \mathcal{D} \rightarrow \mathbb{C}$ is continuous over Y_C if

$$\gamma(\mu) \equiv \sup_{w \in Y_C} \sup_{v \in Y_C} \frac{|a(w, v; \mu)|}{\|w\|_{Y_C} \|v\|_{Y_C}} \tag{2.25}$$

is finite for all $\mu \in \mathcal{D}$; we can then define $\gamma_0 = \max_{\mu \in \mathcal{D}} \gamma(\mu)$ $(< \infty)$.

Given a parametric bilinear form a: $Y_{C,1} \times Y_{C,2} \times \mathcal{D} \rightarrow \mathbb{C}$, we define the inf-sup constant [7] as

$$\beta(\mu) = \inf_{w \in Y_{C,1}} \sup_{v \in Y_{C,2}} \frac{|a(w, v; \mu)|}{\|w\|_{Y_{C,1}} \|v\|_{Y_{C,2}}} \ . \tag{2.26}$$

2.7 Affine Parameter Dependence

We shall say that the parametric bounded linear form ℓ: $Y_C \times \mathcal{D} \rightarrow \mathbb{C}$ is *affine* in the parameter if

$$\ell(v; \mu) = \sum_{q=1}^{Q_\ell} \Theta^q_\ell(\mu)\, \ell^q(v), \qquad \forall\, v \in Y_C \ , \tag{2.27}$$

for some finite Q_ℓ; here the Θ^q_ℓ: $\mathcal{D} \rightarrow \mathbb{C}$, $1 \le q \le Q_\ell$, are (typically very smooth) parameter-dependent functions, and the $\ell^q(v)$: $Y_C \rightarrow \mathbb{C}$, $1 \le q \le Q_\ell$, are parameter-independent bounded linear forms.

Similarly, we shall say that the parametric bilinear form a: $Y_{C,1} \times Y_{C,2} \times \mathcal{D} \rightarrow \mathbb{C}$ is *affine* in the parameter if

$$a(w, v; \mu) = \sum_{q=1}^{Q_a} \Theta^q_a(\mu)\, a^q(w, v), \qquad \forall\, w \in Y_{C,1},\ \forall\, v \in Y_{C,2} \ , \tag{2.28}$$

for some finite Q_a; here the Θ^q_a: $\mathcal{D} \rightarrow \mathbb{C}$, $1 \le q \le Q_a$, are (typically very smooth) parameter-dependent functions, and the $a^q(w, v)$: $Y_{C,1} \times Y_{C,2} \rightarrow \mathbb{C}$, $1 \le q \le Q_a$, are parameter-independent continuous bilinear forms.

Chapter 3

Abstract Formulation for Acoustic Waveguides

3.1 Introduction

In this chapter, we introduce the abstract formulation for acoustic waveguide problems. First we present a special treatment of the outflow boundary condition which allows us to handle propagating and evanescent modes correctly. In the acoustics community, the transfer matrix [31] method is used in conjuntion with mode matching techniques to impose impedance boundary conditions [75, 91] for waveguide acoustics problems. More recently, the perfectly matched layer (PML) [1] techniques for domain truncation have been developed for finite-element time-domain methods [2]. However, it was reported in [113] that the PML method is ineffective in absorbing evanescent waves. The simplest approach consists in using a truncated expansion of the trace of the solution on the boundary in terms of eigenfunctions of the waveguide [21] — the radiation boundary condition is derived from similar ideas. The radiation boundary condition is then included in the formulation of our "exact" problem statement. Next, we introduce a special formulation to handle quadratic outputs for non-coercive problems. The expanded formulation is the setting for our non-coercive elliptic acoustic waveguide problems with quadratic "outputs" of interest.

3.2 Derivation of the Acoustic Helmholtz Equation

Our starting point is the equations of state for lossless fluids described extensively in [22]. In vector notation, the three-dimensional continuity and momentum equations for lossless fluids are

$$\text{Continuity} \qquad \tilde{\rho}_t + \nabla \cdot (\tilde{\rho}\,\tilde{\mathbf{U}}) = 0 \tag{3.1}$$

$$\text{Momentum} \qquad \tilde{\rho}\,[\tilde{\mathbf{U}}_t + (\tilde{\mathbf{U}} \cdot \nabla)\tilde{\mathbf{U}}] + \nabla \hat{P} = 0. \tag{3.2}$$

where $\hat{P}$ denotes the total pressure, $\tilde{\rho}$ the density, and $\tilde{U}$ the particle velocity in the medium. In the case of acoustics, when losses remain negligible, the entropy remains constant and the pressure is only a function of the density. For any fluid (liquid or gas) we can express the isoentropic equation of state as a Taylor's series in the condensation $(\tilde{\rho} - \tilde{\rho}_0)/\tilde{\rho}_0$,

$$\hat{P} = \hat{P}_0 + A\frac{\tilde{\rho} - \tilde{\rho}_0}{\tilde{\rho}_0} + \frac{B}{2!}(\frac{\tilde{\rho} - \tilde{\rho}_0}{\tilde{\rho}_0})^2 + \frac{C}{3!}(\frac{\tilde{\rho} - \tilde{\rho}_0}{\rho_0})^3 + \ldots; \tag{3.3}$$

of the coefficients A, B, C, ..., A is the most important. To evaluate A we introduce the sound speed, $\tilde{c}$,

$$\tilde{c}^2 \equiv \frac{\partial \hat{P}}{\partial \tilde{\rho}} \text{ for an isoentropic process.} \tag{3.4}$$

In the limit of vanishing condensation, $\tilde{\rho} \to \tilde{\rho}_0$, $\tilde{c}^2$ becomes a constant, which we denote $\tilde{c}_0^2$. The coefficient A is then given by, $A = \tilde{\rho}_0\,\tilde{c}_0^2$, which is the adiabatic bulk modulus of the fluid.

We can simplify (3.1) and (3.2) by eliminating the static components. We write the total pressure $\hat{P}$ as the sum of the ambient pressure $\hat{P}_0$ and the acoustic or excess pressure, $\tilde{P}$, given by

$$\tilde{P} \equiv \hat{P} - \hat{P}_0. \tag{3.5}$$

We can similarly identify the excess density $\delta\rho$ as

$$\delta\tilde{\rho} \equiv \tilde{\rho} - \tilde{\rho}_0 \tag{3.6}$$

In the absence of sound the fluid is assumed to be quiet, we set the pressure $P = P_0$, the density $\rho = \rho_0$ (assumed homogenenous) and the particle velocity $U = 0$. The isoentropic equation of state

is given by

$$\tilde{P} = c_0^2 \, \delta\rho \tag{3.7}$$

We then linearize (3.1) and (3.2) using the small-signal approximation

$$\begin{aligned} |\delta\tilde{\rho}| &\ll \tilde{\rho}_0 \\ |\tilde{P}| &\ll \tilde{\rho}_0 \, \tilde{c}_0^2 \\ |\tilde{\mathbf{U}}| &\ll \tilde{c}_0. \end{aligned}$$

The linearized equations of state are

$$\text{Continuity} \qquad \delta\tilde{\rho}_t + \tilde{\rho}_0 \, \nabla \cdot \tilde{\mathbf{U}} = 0 \tag{3.8}$$

$$\text{Momentum} \qquad \tilde{\rho}_0 \, \tilde{\mathbf{U}}_t + \nabla\tilde{P} = 0. \tag{3.9}$$

We eliminate $\tilde{\mathbf{U}}$ from (3.8) and (3.9) to obtain the acoustic wave equation

$$\tilde{c}_0^2 \, \nabla^2\tilde{P} - \tilde{P}_{tt} = 0. \tag{3.10}$$

We seperate the time and space dependence, assuming a time-harmonic pressure, $\tilde{P} = e^{\mathrm{i}\,\tilde{\omega}\,t} \, P$, to obtain the acoustic Helmholtz equation in dimensional form

$$\nabla^2 \, P + \frac{\tilde{\omega}^2}{\tilde{c}_0^2} \, P = 0. \tag{3.11}$$

We also define the acoustic impedance $\tilde{Z}$ as

$$\tilde{Z} \equiv \frac{\tilde{P}}{\tilde{U}} \equiv \frac{P}{U}; \tag{3.12}$$

where $\tilde{U} = e^{\mathrm{i}\,\tilde{\omega}\,t} \, U$.

Boundary Conditions

We also discuss some typical boundary conditions encountered for acoustic problems. A general impedance boundary condition can be derived from the momentum equation of state (3.9) for a

time-harmonic pulse $\tilde{U} = e^{\mathrm{i}\,\tilde{\omega}\,t}\,U$, as

$$\mathrm{i}\,\tilde{\rho}_0\,\tilde{\omega}\,U + \nabla P = 0. \tag{3.13}$$

We use (3.13) to specify two kinds of boundary conditions in the context of this thesis. The first kind is obtained by specifying a known velocity $U = U_{\text{known}}$ on a particular boundary, we thus obtain

$$\nabla P = -\mathrm{i}\,\tilde{\rho}_0\,\tilde{\omega}\,U_{\text{known}}; \tag{3.14}$$

the second kind is a boundary where we specify a known impedance, $\tilde{Z} = \tilde{Z}_{\text{known}}$ (remembering that $U = \tilde{Z}_{\text{known}}\,P$)

$$\nabla P = -\frac{\mathrm{i}\,\tilde{\rho}_0\,\tilde{\omega}}{\tilde{Z}_{\text{known}}}\,P. \tag{3.15}$$

3.2.1 Non-dimensional Form

We now derive the non-dimensional form of (3.11). We non-dimensionalize (3.11) using some characteristic length scale, denoted by L_c, and some fixed (known) velocity U_{known}, for the specific acoustical system under consideration. We identify the non-dimensional pressure p as

$$p = \frac{P}{\tilde{\rho}_0\,\tilde{c}_0\,U_{\text{known}}}; \tag{3.16}$$

the acoustic Helmholtz equation can then be stated as

$$\nabla^2 p + k^2 p = 0; \tag{3.17}$$

where we denote k as the non-dimensional wavenumber

$$\frac{\tilde{\omega}\,L_c}{\tilde{c}_0}. \tag{3.18}$$

In the problems under consideration, we vary the frequency, $\tilde{f}$, of the time-harmonic pulse, ($\tilde{\omega} = 2\,\pi\,\tilde{f}$) and therefore the non-dimensional wavenumber k of the system. It is important from the numerical standpoint to ensure that our mesh is sufficiently dense. If $k \in [k_{\min}, k_{\max}]$, then we have to ensure that we have adequate mesh resolution to ensure accuracy for the shortest wavelength

$\lambda_{\min} = \frac{2\pi}{k_{\max}}$; typically, 10 points per wavelength is sufficient from accuracy considerations.

We also recognize that the characteristic length plays a very important role through k — if relative to the dimensional domain, the characteric length L_c is very small then the size of the mesh will be correspondingly large. Furthermore, if we consider high frequencies, to provide for adequate resolution the mesh density will have to be increased further to respect the approximate rule of ten points per wavelength. Thus, both the characteristic length and the frequency, restrict the kind of problems we can consider: if we consider large length scales relative to L_c, our frequency range will have to be limited; on the other hand, if we want to increase the range of frequencies, it is better to consider small domains.

These effects of geometric scale L_c and range of frequencies under consideration add to the complexity of dealing with unbounded domains for acoustic problems. Many problems in acoustics scattering from rigid bodies consist of solving the Helmholtz equation in unbounded domains; the truncation of the domain is a necessary and important step in treating such problems. Consideration of large wavenumbers k invalidate the use of traditional FEM, BEM approaches because of the large size of the truncated domains. In the problems described in this thesis, we shall require the specification of a radiation boundary condition to effectively truncate the computational domain. We discuss the development of a radiation boundary condition in Section 3.3.

Boundary Conditions

We now present the non-dimensional analogues for the boundary conditions (3.14) and (3.15). We substitute $P = \tilde{\rho}_0\, \tilde{c}_0\, U_{\text{known}} p$ in (3.14) to obtain

$$\begin{aligned} \tilde{\rho}_0\, \tilde{c}_0\, U_{\text{known}}\, \frac{1}{L_c}\, \nabla p &= -\mathrm{i}\, \tilde{\rho}_0\, \tilde{\omega}\, U_{\text{known}} \\ \Rightarrow \nabla p &= -\mathrm{i}\, k; \end{aligned} \tag{3.19}$$

where L_c appears because of the non-dimensionalization of the ∇ operator.

Similarly, from (3.15) we obtain

$$\begin{aligned} \tilde{\rho}_0\, \tilde{c}_0\, U_{\text{known}}\, \frac{1}{L_c}\, \nabla p &= -\frac{\mathrm{i}\, \tilde{\rho}_0\, \tilde{\omega}}{\tilde{Z}_{\text{known}}}\, \tilde{\rho}_0\, \tilde{c}_0\, U_{\text{known}}\, p \\ \Rightarrow \nabla p &= -\frac{\mathrm{i}\, k}{Z_{\text{known}}}; \end{aligned} \tag{3.20}$$

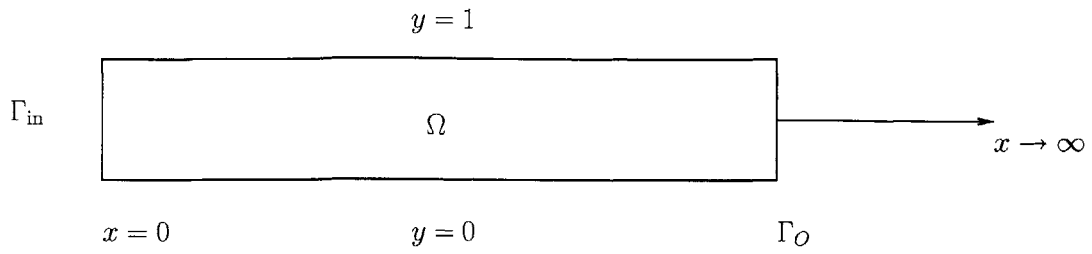


Figure 3-1: Semi-infinite Acoustic Waveguide

where we have replaced the dimensional acoustic impedance $\tilde{Z}_{\text{known}}$, ($\tilde{Z}_{\text{known}} = \tilde{\rho}_0\, \tilde{c}_0\, Z_{\text{known}}$) by its non-dimensional counterpart, Z_{known}.

In the examples discussed in this thesis, we shall always solve for the non-dimensional pressure, thus all the quantities used as inputs — for example, the wavenumber k, and acoustic impedance Z — are necessarily non-dimensional. The derivation of the non-dimensional Helmholtz equation from the actual equations of state and associated boundary conditions thus allows us to relate the non-dimensional quantities of interest to their dimensional counterparts.

3.3 Radiation Boundary Condition for the Straight Channel Waveguide

We consider the motion of free waves in the semi-infinite homogeneous waveguide shown in Figure 3-1. A constant velocity is applied at $x = 0$, the walls at $y = 0$ and $y = 1$ are rigid. The steady-state non-dimensional pressure distribution, p, satisfies

$$\nabla^2 p + k^2 p = 0 \tag{3.21}$$

on the domain Ω with illustrative boundary conditions

$$\frac{\partial p}{\partial x} = -\mathrm{i}\, k, \text{ at } x = 0, \tag{3.22}$$

$$\frac{\partial p}{\partial y} = 0, \text{ at } y = 0,\ 1. \tag{3.23}$$

Here k is the non-dimensional wave-number (3.18) obtained in the derivation of the non-dimensional acoustics Helmholtz equation in Section 3.2.1. The first boundary condition (3.22) is the result of

specifying a fixed velocity at the boundary $x = 0$ and can be derived from the momentum equation (see (3.19)); the second boundary condition is a zero velocity boundary condition specified at $y = 0$ and $y = 1$. The development of the radiation boundary condition for truncation of the domain is not dependent on the choice of the boundary conditions (3.22) and (3.23); rather they are typical of boundary conditions used for the problems described in this thesis and also make the analytical development easier to follow.

The choice of the non-dimensional wavenumber k, as discussed before in Section 3.2.1, is physically relevant. For a fixed characteristic geometric length scale L_c and sound speed $\tilde{c}_0$, $k \equiv \frac{\tilde{\omega} L_c}{\tilde{c}_0}$, depends on the frequency $\tilde{\omega}$. The frequency, $\tilde{\omega}$, and therefore, the wavenumber k, specifies the number of modes of the system that are propagating. We now discuss the effect of k on the number of active modes.

We are interested in generating appropriate boundary conditions for a truncated domain of *finite* extent in the x direction which accurately replicates the actual solution on the semi-infinite domain. To that end, we now present a formulation for $\frac{\partial p}{\partial x}$ on Γ_O.

3.3.1 Dispersion Relation

We obtain the solution of the wave equation by the method of the seperation of variables. We write p as a modal sum of the product of the modes in the x (i.e., $e^{-\mathrm{i}\,\alpha_n\,x}$) and y (i.e., $\Xi_n(y)$) directions, as

$$p(x,y) = \sum_{n=1}^{\infty} c_n\, e^{-\mathrm{i}\,\alpha_n\,x}\, \Xi_n(y). \tag{3.24}$$

The homogeneous Neumann boundary conditions on $y = 0$ and $y = 1$ help us identify $\Xi_n(y)$ as

$$\Xi_n(y) = \cos(\lambda_n\, y), \tag{3.25}$$

where λ_n is given by

$$\lambda_n = n\,\pi, n = 0, 1, \ldots. \tag{3.26}$$

Inserting the modal expansion (3.24) into (3.60), gives us the dispersion relation for each mode as

$$\alpha_n^2 + \lambda_n^2 = k^2. \tag{3.27}$$

When $k^2 > \lambda_n^2$, the "n"-th mode is propagating ($e^{-\mathrm{i}\,|\alpha_n|\,x}$) with

$$\alpha_n = \sqrt{k^2 - \lambda_n^2}; \tag{3.28}$$

if $k^2 <= \lambda_n^2$, then the "n"-th mode is evanescent ($e^{-|\alpha_n|\,x}$) where

$$\alpha_n = -\mathrm{i}\,\sqrt{\lambda_n^2 - k^2}. \tag{3.29}$$

For a given frequency, $\tilde{\omega}$, and associated non-dimensional wavenumber k, we can identify two kinds of modes. The x-directional modes ($e^{-\mathrm{i}\,\alpha_n\,x}$) with α_n given by (3.28) are propagating modes; while modes with α_n given by (3.29) are "evanescent" or decaying. The idea is to choose the length of the truncated domain judiciously such that the evanescent modes will have decayed. The evanesscent modes having all decayed, we are left with only the propagating modes which will be present even as $x \to \infty$.

3.3.2 Outflow Boundary Condition

Considerable efforts have been addressed to the issue of domain truncation of unbounded domains for wave propagation problems.

In the acoustics community, development of impedance boundary conditions for domain truncation have been the subject of significant research [107]. The development of the outflow boundary condition reduces to the specification of a impedance boundary condition of the form

$$p|_{\Gamma_O} = Z|_{\Gamma_O}\,\frac{\partial}{\partial x}\,p|_{\Gamma_O}. \tag{3.30}$$

Therefore, the specification of the condition only requires knowing the impedance at the outlet $Z|_{\Gamma_O}$.The central idea is the development of a generalised *transmission matrix*, H, relating the modal expansion coefficients for the acoustic pressure and velocity at the input and output ends of a waveguide region of fixed impedance; different segments give rise to similar transmission matrices which can be multiplied together to form the transfer matrix of the whole structure. Thus, knowing the transfer matrix relating the modal expansion coefficients of the pressure and the velocity at the inlet and outlet would allow us to solve for the pressure/velocity at the inlet and outlet without

necessarily solving the Helmholtz equation in the domain. However, numerical implementations of the ideas often run into problems, because the elements of H corresponding to evanescent modes grows exponentially with the length of the segment and the problem becomes ill-posed. More recently, [75] has outlined the development of *impedance matrices* which do not suffer from the defects of the transmission-matrix approaches.

A different and quite standard solution procedure is based on characterizing the propagating wave in terms of the modes of the waveguide. Each mode can be expressed in terms of an eigenfunction of a related boundary-valued problem on the fixed waveguide cross-section at outflow. For the Helmholtz equation, the eigenvalue problem is related to the transverse Laplacian in the interior of the domain. Truncation of the semi-infinite domain leads to a boundary condition at outflow involving only the eigenmodes of the previous eigenproblem. The simplest approach involves using a truncation expansion of the trace of the solution at outflow in terms of eigenfunctions, the unknowns are the coefficients associated with each mode, as in [40, 78]. [21] introduces a new class of boundary conditions to remove the necessity of calculating the eigenfunctions on the boundary.

Our own method uses effectively the first step in [21] by recognizing that evanescent modes do not contribute to the solution far from their origination. Given the wavenumber k, we use the same modal superposition approach to identify the eigen-modes that are going to propagate. The difference lies in that we explicitly calculate the eigenvectors associated with the propagating modes; thanks to the offline-online decomposition embedded in the reduced-basis framework we push the computationally intensive work of calculating the eigenvectors of interest and associated matrices to the *offline* stage. We now describe the manner in which we truncate the domain and derive the outflow boundary condition.

The two relations (3.28), (3.29) inform our choice of the length of the truncated domain. We choose the length of the truncated domain L_{trunc} such that we are guaranteed that the evanescent modes have all decayed. Let us assume that there are N_{prop} propagating modes. We can then specify the boundary condition at Γ_O ($x = L_{\text{trunc}}$) as

$$\frac{\partial p}{\partial x}|_{\Gamma_O} = \sum_{n=1}^{\infty} -c_n \, \mathrm{i} \, \alpha_n e^{-\mathrm{i}\,\alpha_n L_{\text{trunc}}} \, \Xi_n(y) \tag{3.31}$$

$$= \sum_{n=1}^{N_{\text{prop}}} -c_n \, \mathrm{i} \, \alpha_n e^{-\mathrm{i}\,\alpha_n L_{\text{trunc}}} \, \Xi_n(y). \tag{3.32}$$

We now exploit the orthogonality of the y-directional eigen modes at the outlet Γ_O,

$$\int_{\Gamma_O} \Xi_m \overline{\Xi}_n = \delta_{mn}, \tag{3.33}$$

to re-write (3.24) as

$$\int_{\Gamma_O} p\, \overline{\Xi}_n = c_n\, e^{-\mathrm{i}\,\alpha_n L_{\text{trunc}}}; \tag{3.34}$$

$\int_{\Gamma_O} p\,\overline{\Xi}_n$ is really the projection of the acoustic pressure p onto the n-th mode and is representative of the amount of acoustical energy in the n-th mode.

We can now write our boundary condition at the outlet Γ_O as

$$\frac{\partial p}{\partial x}|_{\Gamma_O} = \sum_{n=1}^{N_{\text{prop}}} -\mathrm{i}\,\alpha_n \Big(\int_{\Gamma_O} p\,\overline{\Xi}_n\Big)\,\Xi_n. \tag{3.35}$$

We now discuss the solution of the eigen problem to calculate the terms in (3.35).

3.3.3 Eigen Problem on Boundary

While we know the form of the outflow boundary conditon (3.35) we still need to determine $\Xi_n(y)$. In this section, we outline the weak form of the eigen problem on the boundary.

Parameter Independent Case

We consider (3.60) on the semi-infinite domain with Neumann boundary conditions $\frac{\partial p}{\partial n} = 0$ on all the boundaries (here n is the outward normal on the respective boundaries). Note that although we do allow k^2 to vary, for our purposes it does not change the form of the eigen-problem or the eigen vectors that we obtain. We will return to this point again.

We multiply (3.60) by an arbitrary function $v \in X^e$, a complex vector space (see Chapter 2 for a discussion on complex vector spaces), and integrate over the domain Ω to obtain

$$\int_{\Omega} \Big(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + k^2 p\Big)\,\overline{v} = 0; \tag{3.36}$$

we now apply Green's Theorem and drop the boundary terms (homogeneous Neumann) and get

$$\int_{\Omega} \frac{\partial p}{\partial x}\frac{\partial \overline{v}}{\partial x} + \frac{\partial p}{\partial y}\frac{\partial \overline{v}}{\partial y} - k^2 p\,\overline{v} = 0,\ \forall v \in X^e. \tag{3.37}$$

Replacing the integral over Ω in (3.37) with the double-integral, we obtain

$$\int_{x=0}^{\infty}\int_{y=0}^{1} \frac{\partial p}{\partial x}\frac{\partial \overline{v}}{\partial x} + \frac{\partial p}{\partial y}\frac{\partial \overline{v}}{\partial y} - k^2 p\,\overline{v} = 0,\ \forall v \in X^e. \tag{3.38}$$

We now consider modal expansions for p and v

$$p = \sum_{\alpha} e^{-\mathrm{i}\,\alpha x}\,\hat{p}^{\alpha}(y), \tag{3.39}$$

$$v = \sum_{\beta} e^{-\mathrm{i}\,\beta x}\,\hat{v}^{\beta}(y), \tag{3.40}$$

note that the x-directional eigen modes satisfy the following orthogonality condition

$$\int_{x=0}^{\infty} e^{-\mathrm{i}\alpha x}\overline{e^{-\mathrm{i}\beta x}} = \delta_{\alpha\,\beta}. \tag{3.41}$$

Inserting the expansions (3.39), (3.40) into (3.38) we obtain

$$\begin{aligned}\int_{x=0}^{\infty}\int_{y=0}^{1} \Bigg(&\sum_{\alpha} -\mathrm{i}\alpha\, e^{-\mathrm{i}\alpha x}\,\hat{p}^{\alpha}(y) \sum_{\beta} \mathrm{i}\beta\,\overline{e^{-\mathrm{i}\beta x}}\,\overline{\hat{v}^{\beta}(y)} \\ &+ \int_{x=0}^{\infty}\int_{y=0}^{1} \sum_{\alpha} e^{-\mathrm{i}\alpha x}\frac{\partial \hat{p}^{\alpha}(y)}{\partial y} \sum_{\beta} \overline{e^{-\mathrm{i}\beta x}}\,\frac{\partial \overline{\hat{v}^{\beta}(y)}}{\partial y} \\ -k^2 \int_{x=0}^{\infty}\int_{y=0}^{1} &\sum_{\alpha} e^{-\mathrm{i}\alpha x}\,\hat{p}^{\alpha}(y) \sum_{\beta} \overline{e^{-\mathrm{i}\beta x}}\,\overline{\hat{v}^{\beta}(y)} \Bigg) = 0;\end{aligned} \tag{3.42}$$

we use the orthogonality of the x-directional eigen modes (3.41) and write out just the y-integral in (3.42) as

$$\int_{y=0}^{1} \frac{\partial \hat{p}^{\alpha}}{\partial y}\frac{\partial \overline{\hat{v}^{\alpha}}}{\partial y} = -(\alpha^2 - k^2)\int_{y=0}^{1} \hat{p}^{\alpha}\overline{\hat{v}^{\alpha}},\ \forall\, \hat{v}^{\alpha} \in X^e|_{\Gamma_O};$$

here $X^e|_{\Gamma_O}$ is the restriction of the complex vector space X^e to the outflow boundary Γ_O.

We can now define our eigen-problem on the boundary: find $(\Xi_i(\mu), \gamma_i(\mu)) \in X^e|_{\Gamma_O} \times \mathbb{R}, i =$

$1, \ldots,$ such that

$$\int_{y=0}^{1} \frac{\partial \Xi_i}{\partial y} \frac{\partial \overline{v}}{\partial y} = \lambda_i \int_{y=0}^{1} \Xi_i \, \overline{v}, \; \forall v \in X^e|_{\Gamma_O} \tag{3.43}$$

and

$$\int_{y=0}^{1} \Xi_i \, \overline{\Xi}_j = \delta_{ij}; \tag{3.44}$$

we then calculate

$$\gamma_i = \lambda_i - k^2; \tag{3.45}$$

and order the solutions as $\gamma_1 \leq \gamma_2 \leq \gamma_{N_{\text{prop}}} \leq 0 \leq \gamma_{N_{\text{prop}}+1}, \ldots$. The first N_{prop} modes correspond to propagating modes with behavior $e^{-\mathrm{i}\sqrt{-\gamma_i}\, x}$, while modes $N_{\text{prop}}+1, \ldots, \infty$ correspond to evanescent modes with behavior $e^{-\sqrt{\gamma_i}\, x}$.

Note that the formulation of the eigen-problem does not change when we change k^2, this is expected since the obtained eigen values and eigen vectors are independent of the choice of k^2. Choosing a particular k^2 simply determines the number of modes that propagate, and not the form of these modes. We expect the eigen-modes to change shape only when the eigen-problem defined on the boundary changes which would happen if our input vector μ includes changes in the physical parameters that enter into the form of the eigen-problem or from geometric variations on the boundary. For example, we would expect different eigen-vectors and associated eigen-values for a two-medium waveguide with different densities for the two media that are allowed to vary — a water and sand waveguide would behave differently from a water and rock waveguide.

Parameter-Dependent Case

We now discuss a more general problem where the form of the eigen-problem does depend on our input vector μ. We consider a μ-dependent form of the Helmholtz equation (3.60) on the semi-infinite domain shown in Figure 3-1 as

$$\sigma_x(y;\mu)\frac{\partial^2 p}{\partial x^2} + \sigma_y(y;\mu)\frac{\partial^2 p}{\partial y^2} + \phi(y;\mu)p = 0; \tag{3.46}$$

here $\sigma_x(y;\mu), \sigma_y(y;\mu) \in L^\infty(]0,1[) > 0$, $\phi(y;\mu) \in L^\infty(]0,1[$, and $\mu \in \mathcal{D} \in \mathbb{R}^P$. We assume that we have homogeneous Neumann conditions

$$\frac{\partial p}{\partial n} = 0 \tag{3.47}$$

on all the boundaries (here n is the outward normal on the respective boundaries).

To begin, we multiply (3.46) by an arbitrary function $v \in X^e$, a complex vector space (see Chapter 2 for a discussion on complex vector spaces), and integrate over the domain Ω to obtain

$$\int_\Omega \left(\sigma_x(y;\mu)\frac{\partial^2 p}{\partial x^2} + \sigma_y(y;\mu)\frac{\partial^2 p}{\partial y^2} + \phi(y;\mu)p\right)\overline{v} = 0; \tag{3.48}$$

applying the Green's Theorem and dropping the boundary terms we obtain

$$\int_\Omega \sigma_x(y;\mu)\frac{\partial p}{\partial x}\frac{\partial \overline{v}}{\partial x} + \sigma_y(y;\mu)\frac{\partial p}{\partial y}\frac{\partial \overline{v}}{\partial y} - \phi(y;\mu)p\,\overline{v} = 0,\ \forall v \in X^e. \tag{3.49}$$

We can replace the integral over Ω in (3.49) with the double-integral

$$\int_{x=0}^{\infty}\int_{y=0}^{1} \sigma_x(y;\mu)\frac{\partial p}{\partial x}\frac{\partial \overline{v}}{\partial x} + \sigma_y(y;\mu)\frac{\partial p}{\partial y}\frac{\partial \overline{v}}{\partial y} - \phi(y;\mu)p\,\overline{v} = 0,\ \forall v \in X^e; \tag{3.50}$$

and consider modal expansions for p (3.39), and for v (3.40); the x-directional eigen modes satisfy the orthogonality condition given by (3.41).

Inserting the expansions (3.39), (3.40) into (3.50) we obtain

$$\begin{aligned}
&\int_{x=0}^{\infty}\int_{y=0}^{1} \sigma_x(y;\mu)\sum_\alpha -\mathrm{i}\alpha\, e^{-\mathrm{i}\alpha x}\,\hat{p}^\alpha \sum_\beta \mathrm{i}\beta\,\overline{e^{-\mathrm{i}\beta x}}\,\overline{\hat{v}^\beta} \\
&+\int_{x=0}^{\infty}\int_{y=0}^{1} \sigma_y(y;\mu)\sum_\alpha e^{-\mathrm{i}\alpha x}\frac{\partial \hat{p}^\alpha}{\partial y}\sum_\beta \overline{e^{-\mathrm{i}\beta x}}\,\frac{\partial \overline{\hat{v}^\beta(y)}}{\partial y} \\
&-\int_{x=0}^{\infty}\int_{y=0}^{1} \phi(y;\mu)\sum_\alpha e^{-\mathrm{i}\alpha x}\,\hat{p}^\alpha \sum_\beta \overline{e^{-\mathrm{i}\beta x}}\,\overline{\hat{v}^\beta} = 0.
\end{aligned} \tag{3.51}$$

We now exploit the orthogonality of the x-directional eigen modes (3.41) and write out just the y-integral in (3.51) as

$$\int_{y=0}^{1} \sigma_y(y;\mu)\frac{\partial \hat{p}^\alpha}{\partial y}\frac{\partial \overline{\hat{v}^\alpha}}{\partial y} - \phi(y;\mu)\hat{p}^\alpha\overline{\hat{v}^\alpha} = -\alpha^2 \int_{y=0}^{1} \sigma(x;\mu)\hat{p}^\alpha\overline{\hat{v}^\alpha},\ \forall\, \hat{v}^\alpha \in X^e|_{\Gamma_O},$$

as before, $X^e|_{\Gamma_O}$ is the restriction of the complex vector space X^e to the outflow boundary Γ_O.

We can now define our eigen-problem on the boundary: find $(\Xi_i(\mu), \gamma_i(\mu)) \in X^e|_{\Gamma_O} \times \mathbb{R}, i =$

$1, \ldots,$ such that

$$\int_{y=0}^{1} \sigma_y(y;\mu)\frac{\partial \Xi_i}{\partial y}\frac{\partial \overline{v}}{\partial y} - \phi(y;\mu)\Xi_i\,\overline{v} = \gamma_i \int_{y=0}^{1} \sigma(x;\mu)\Xi_i\,\overline{v},\ \forall v \in X^e|_{\Gamma_O} \tag{3.52}$$

and

$$\int_{y=0}^{1} \sigma(x;\mu)\Xi_i\,\overline{\Xi}_j = \delta_{ij}; \tag{3.53}$$

order solutions as $\gamma_1 \leq \gamma_2 \leq \gamma_{N_{\text{prop}}} \leq 0 \leq \gamma_{N_{\text{prop}}+1}, \ldots$. The first N_{prop} modes correspond to propagating modes with behavior

$$e^{-\mathrm{i}\sqrt{-\gamma_i}\,x}, \tag{3.54}$$

while modes $N_{\text{prop}}+1, \ldots, \infty$ correspond to evanescent modes with behavior

$$e^{-\sqrt{\gamma_i}\,x}. \tag{3.55}$$

For the problems under consideration in this thesis, the eigen problems are not μ-dependent, this is important in the sense that we only need to calculate our eigen values and eigen vectors only once. We will not address the more general μ-dependence requiring the solution of the μ-dependent eigen problem, while it is certainly feasible it requires the use of a collateral eigen reduced-basis space which is beyond the scope of discussion here. Please see [64] for a treatment of the μ-dependent eigen problem.

3.3.4 Weak Formulation

We can now identify the terms in the weak formulation coming from the specification of the radiation outflow boundary condition. For the sake of completeness, we outline the different steps again. The radiation boundary condition is specified on the boundary as

$$\frac{\partial p}{\partial x}|_{\Gamma_O} = \sum_{j=1}^{N_{\text{prop}}} -\mathrm{i}\,\alpha_j\Big(\int_{\Gamma_O} p\,\overline{\Xi}_j\Big)\,\Xi_j, \tag{3.56}$$

where Ξ_j, α_j can be solved from the eigen-problem (3.43), (3.44).

The bilinear forms arising from the boundary condition in the weak formulation can then be

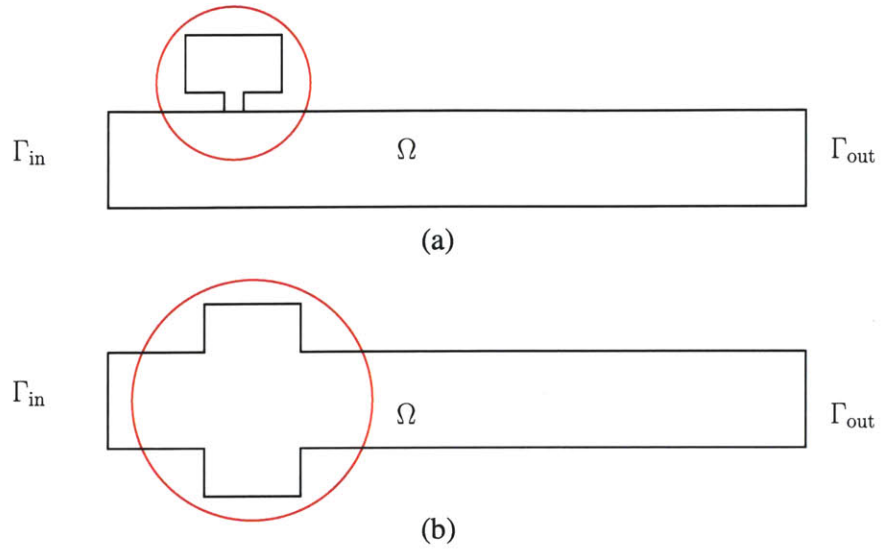


Figure 3-2: Waveguide with (a) Helmholtz resonator, and (b) Expansion chamber

identified as

$$\int_{\Gamma_O} \overline{v} \frac{\partial p}{\partial n} = \int_{\Gamma_O} \overline{v} \frac{\partial p}{\partial x} \tag{3.57}$$

$$= \int_{\Gamma_O} \overline{v} \Big(\sum_{j=1}^{N_{\text{prop}}} -\mathrm{i}\,\alpha_j \; \Xi_j \, (\int_{\Gamma_O} p\, \overline{\Xi}_j) \Big) \tag{3.58}$$

$$= \sum_{j=1}^{N_{\text{prop}}} -\mathrm{i}\,\alpha_j \, (\int_{\Gamma_O} \Xi_j \, \overline{v}) \, (\int_{\Gamma_O} p\, \overline{\Xi}_j) \tag{3.59}$$

The bilinear forms on the boundary are outer-products. Depending on the choice of problem, given k^2 and N_{prop}, Ξ_j, α_j can be solved from the eigen problem and the terms assembled. For problems where σ_x, σ_y are functions of μ, the eigen-problem becomes a μ-dependent eigen-problem.For our acoustics applications, the Ξ_j, α_j will be μ-independent and thus *a priori* computable.

3.4 "Exact" Problem Statement

We consider a suitably regular (smooth) domain $\Omega \subset \mathbb{R}^d, d = 1, 2,$ or 3 with Lipschitz-continuous boundary Γ. Note that Ω is a *reference domain* and hence does *not* depend on the parameter. In Chapter 4, we discuss the formulation of the problem in the reference domain.

3.4.1 Strong Form

Consider a semi-infinite acoustic waveguide coupled with other acoustic elements — for example, a pressure release tube, Helmholtz resonator, or an expansion chamber, as shown in Figure 3-2. The governing equations are

$$\nabla^2 p + k^2 p = 0 \tag{3.60}$$

on the domain Ω.

The boundary Γ $(= \Gamma_{\text{in}} \cup \Gamma_{\text{out}} \cup \Gamma_N)$ has an imposed velocity condition specified at the inlet (please see Section 3.2.1 for a derivation of this condition),

$$\frac{\partial p}{\partial n}|_{\Gamma_{\text{in}}} = \mathrm{i}\, k, \tag{3.61}$$

and homogeneous Neumann (zero velocity) boundary conditions on Γ_N

$$\frac{\partial p}{\partial n}|_{\Gamma_N} = 0; \tag{3.62}$$

here n is the outward normal to the boundaries Γ_{in} and Γ_{out}. Hence, n points in the direction of x on Γ_{out} and in the $-x$ direction on Γ_{in}.

Truncation of the Domain

We truncate the semi-infinite domain choosing a truncated domain length L, and solve the Helmholtz problem (3.60) as before on the truncated domain with boundary conditions specified on Γ_{in} (3.61) and (3.62). We also introduce the radiation boundary condition at the outlet Γ_{out} as

$$\frac{\partial p}{\partial n}|_{\Gamma_{\text{out}}} = -\sum_{j=1}^{N_{\text{prop}}} \mathrm{i}\, \alpha_j \left(\int_{\Gamma_{\text{out}}} p\, \overline{\Xi}_j\right) \Xi_j; \tag{3.63}$$

N_{prop} and the length of the truncated domain, L, are functions of the range of k^2 for a particular problem and need to be chosen correctly.

It is important to understand why the radiation boundary condition is still applicable even when we have additional acoustic elements (i.e, the Helmholtz resonator of Figure 3-2(a) or the expansion chamber of Figure 3-2(b)) coupled to the semi-infinite acoustic waveguide. The addition

of the acoustic elements definitely changes the solution relative to the simple semi-infinite waveguide — it is quite feasible that these acoustical elements excites some higher order modes previously not present in the solution. However given k, as long as the truncated domain is long enough we expect *all* evanescent modes to decay. This effect is only magnified for higher-order modes: these modes will decay even faster because the γ_i associated with the x-directional mode $e^{-|\gamma_i|x}$ will be larger. If we had N_{prop} propagating modes, then any other evanescent mode which might be excited will decay. Thus, the choice of the truncated domain length, L, and the accuracy of the solution depends on the choice of N_{prop} and the range of frequency k — as long N_{prop} is correct and L is long enough the radiation boundary condition in (3.63) will be accurate.

Note that the eigenvectors and eigenvalues are obtained from the eigen problem on the boundary Γ_{out} of the waveguide — the eigen problem is not affected by the addition of afore-mentioned acoustic elements to the domain. The eigen problem being unaffected by our choice of domain; the radiation boundary condition simply tells us given k, what N_{prop} should be — from the solution of the eigen problem and for modes where $\gamma_i \leq 0$ ($\alpha_i = \sqrt{-\gamma_i}$) from (3.45); the choice of L is ours to make to ensure accuracy.

An important consequence of the modal expansion is that we could conceivably shorten the truncated domain even further and include those evanescent modes which we do *not* expect to decay for the chosen truncated length in (3.63) and still expect our solution to be accurate. As long as the boundary condition prescribed on the boundary includes all the modes which are expected to be present in the modal expansion of the pressure; our solution will be accurate. The radiation boundary condition makes no distinction between propagating and evanescent modes; we make the distinction by *only* including the N_{prop} propagating modes in the expansion; thus L needs to be carefully chosen. These issues are explored in more detail in Chapter 5, where we discuss the effect of the truncated domain length, L, and the number of propagating modes, N_{prop}, on the accuracy of the solution.

Outputs

Given our input parameters μ (which may be related to the geometry of the waveguide or the coupled acoustic elements, damping parameters, and range of wavenumber), we are interested in

"outputs" like the average pressure at the inlet

$$s_1(\mu) = \frac{1}{|\Gamma_{\text{in}}|} \int_{\Gamma_{\text{in}}} p(\mu) \tag{3.64}$$

or the average pressure intensity at the outlet

$$s_2(\mu) = \frac{1}{|\Gamma_{\text{out}}|} \int_{\Gamma_{\text{out}}} \overline{p(\mu)}\, p(\mu) = \frac{1}{\Gamma_{\text{out}}} \int_{\Gamma_{\text{out}}} |p(\mu)|^2. \tag{3.65}$$

3.4.2 Weak Form

We now consider the abstract form of the "exact" (superscript e) problem posed in the weak form on the *reference* domain Ω: for any $\mu \in \mathcal{D} \subset \mathbb{R}^P$, calculate outputs

$$s_1^e(\mu) = \ell(u^e(\mu)), \tag{3.66}$$

and/or

$$s_2^e(\mu) = Q(u^e(\mu), u^e(\mu)), \tag{3.67}$$

where $u^e(\mu)$ satisfies the weak form of the μ-parametrized PDE

$$a(u^e(\mu), v; \mu) = f(v), \quad \forall\, v \in X^e. \tag{3.68}$$

Here μ and $\mathcal{D}$ are the input and (closed) input domain, respectively. Depending on the choice of problem, our "output" of interest can be linear ($s_1^e(\mu)$) or quadratic ($s_2^e(\mu)$). Here $u^e(x;\mu)$ is our field variable and X^e is a complex vector space.

The abstract statement needs to be posed in the *reference* domain, since the reduced basis recipe requires that Ω be *parameter-independent.* We identify the weak form $a(\cdot,\cdot;\mu)$ (3.68) associated with the acoustic Helmholtz equation (3.60) and associated boundary conditions for $\mu \equiv k^2$ as

$$a(w, v; \mu) = \int_\Omega \nabla w \nabla \overline{v} - \mu \int_\Omega w\, \overline{v} + \sum_{j=1}^{N_{\text{prop}}} \mathrm{i}\, \alpha_j(\mu) \left(\int_{\Gamma_{\text{out}}} \Xi_j(\mu)\, \overline{v} \right) \left(\int_{\Gamma_{\text{out}}} w\, \overline{\Xi_j(\mu)} \right); \tag{3.69}$$

and $f(\cdot;\mu)$ as

$$f(v;\mu) = \mathrm{i}\sqrt{\mu} \int_{\Gamma_{\text{in}}} \overline{v}; \tag{3.70}$$

here we make the μ-dependence of α_j and Ξ_j explicit.

Similarly, the quadratic functional $Q(\cdot,\cdot;\mu)$ associated with the average pressure intensity on Γ_{out} is given by

$$Q(w,v;\mu) = \frac{1}{|\Gamma_{\text{out}}|}\int_{\Gamma_{\text{out}}} w\,\overline{v}. \tag{3.71}$$

3.5 Expanded Formulation for Quadratic Outputs

We consider the problem of calculating quadratic outputs like the Transmission Coefficient (TC) or the average pressure intensity for our acoustics waveguide problems. We build on earlier work [52] on an expanded formulation for quadratic outputs — relevant in the real case for symmetric operators $a(w,v;\mu) = a(w,v;\mu), \forall w,v \in X$ — that reduced quadratic outputs to simple linear outputs associated with a different weak statement. Prior to [52], the treatment of quadratic outputs was based on a primal-dual formulation which required the development of a primal and dual reduced basis aproximation spaces. We first develop the formulation, and then briefly remark on the benefits in the reduced basis context. Our extension allows for the treatment of non-coercive and non-symmetric complex operators in a similar manner.

3.5.1 Abstraction

We start with our abstract μ-dependent problem statement specified in (3.68). Our abstract problem (**A**) can be stated as follows: given $\mu \in \mathcal{D} \in \mathbb{R}^P$, we are interested in solving for general quadratic outputs of the form

$$\begin{aligned} s^e(\mu) &= \ell(u^e(\mu;\mu) + Q(u^e(\mu), u^e(\mu);\mu), \\ &= s_1^e(u^e(\mu);\mu) + s_2^e(u^e(\mu);\mu), \end{aligned} \tag{3.72}$$

where $Q(\cdot,\cdot;\mu)$ and $\ell(\cdot;\mu)$ are bilinear and linear forms, and $u^e(\mu) \in X^e$ is the solution of

$$a(u^e(\mu), v;\mu) = f(v;\mu), \quad \forall v \in X^e. \tag{3.73}$$

Here $a(\cdot,\cdot;\mu)$ and $f(\cdot;\mu)$ are μ- parametrized forms, and X^e is a complex inner-product space. Moreover, we further require that all $a(\cdot,\cdot;\mu)$, $Q(\cdot,\cdot;\mu)$ and $f(\cdot;\mu)$, $\ell(\cdot;\mu)$ are "affine" in the pa-

rameter: it may be decomposed as an affine decomposition of the form

$$a(w,v;\mu) = \sum_{q=1}^{Q_a} \Theta_a^q(\mu) a^q(w,v), \tag{3.74}$$

$$f(v;\mu) = \sum_{q=1}^{Q_f} \Theta_f^q(\mu) f^q(v), \tag{3.75}$$

$$Q(w,v;\mu) = \sum_{q=1}^{Q_p} \Theta_Q^q(\mu) Q^q(w,v), \tag{3.76}$$

$$\ell(v;\mu) = \sum_{q=1}^{Q_\ell} \Theta_\ell^q(\mu) \ell^q(v), \tag{3.77}$$

where $\Theta_a^q(\mu), \Theta_f^q(\mu), \Theta_Q^q(\mu), \Theta_\ell^q(\mu) : \mathcal{D} \to \mathbb{C}$ and $a^q(w,v), Q^q(w,v) : X^e \times X^e \to \mathbb{C}$, $f^q(v), \ell^q(v) : X^e \to \mathbb{C}$ are parameter-dependent and parameter-independent continuous bilinear/linear forms, respectively. We shall further assume that the parameter $\Theta_a^q(\mu)$, $\Theta_Q^q(\mu)$ and $\Theta_f^q(\mu)$, $\Theta_\ell^q(\mu)$ are continuous: $\Theta_a^q(\mu) \in \mathcal{C}^1(\mathcal{D})$, $1 \le q \le Q_a$, $\Theta_Q^q(\mu) \in \mathcal{C}^1(\mathcal{D})$, $1 \le q \le Q_p$ and $\Theta_f^q(\mu) \in \mathcal{C}^1(\mathcal{D})$, $1 \le q \le Q_f$, $\Theta_\ell^q(\mu) \in \mathcal{C}^1(\mathcal{D})$, $1 \le q \le Q_\ell$.

3.5.2 Expanded Formulation

We next introduce another problem **(B)**: given $\mu \in \mathcal{D} \in \mathbb{R}^P$, we are interested to find the "compliance" output

$$\mathcal{S}^e(\mu) = \overline{\mathcal{F}(\mathcal{U}^e(\mu);\mu)}, \tag{3.78}$$

where $\mathcal{U}^e(\mu) \in \mathcal{X}^e$ satisfies

$$\mathcal{A}(\mathcal{U}^e(\mu), \mathcal{V};\mu) = \mathcal{F}(\mathcal{V};\mu), \quad \forall \mathcal{V} \in \mathcal{X}^e. \tag{3.79}$$

Here $\mathcal{A}(\cdot,\cdot;\mu)$ and $\mathcal{F}(\cdot;\mu)$ are also μ-parametrized forms, and $\mathcal{X}^e$ is the function space based on X^e that will be defined later. We also require $\mathcal{A}(\cdot,\cdot;\mu)$ and $\mathcal{F}(\cdot;\mu)$ are "affine" in the parameter

$$\mathcal{A}(\mathcal{W},\mathcal{V};\mu) = \sum_{q=1}^{Q_\mathcal{A}} \Theta_\mathcal{A}^q(\mu) \mathcal{A}^q(\mathcal{W},\mathcal{V}), \tag{3.80}$$

$$\mathcal{F}(\mathcal{V};\mu) = \sum_{q=1}^{Q_\mathcal{F}} \Theta_\mathcal{F}^q(\mu) \mathcal{F}^q(\mathcal{V}), \tag{3.81}$$

where $\Theta^q_{\mathcal{A}}, \Theta^q_{\mathcal{F}} : \mathcal{D} \to \mathbb{C}$ and $\mathcal{A}^q(\mathcal{W}, \mathcal{V}) : \mathcal{X}^e \times \mathcal{X}^e \to \mathbb{C}$, $\mathcal{F}^q(\mathcal{V}) : \mathcal{X}^e \to \mathbb{C}$ are parameter-dependent and parameter-independent continuous bilinear/linear forms, respectively. We shall further assume that the parameter $\Theta^q_{\mathcal{A}}(\mu)$ and $\Theta^q_{\mathcal{F}}(\mu)$ are continuous: $\Theta^q_{\mathcal{A}}(\mu) \in \mathcal{C}^1(\mathcal{D})$, $1 \le q \le Q_{\mathcal{A}}$ and $\Theta^q_{\mathcal{F}}(\mu) \in \mathcal{C}^1(\mathcal{D})$, $1 \le q \le Q_{\mathcal{F}}$.

We now may make the claim that

Proposition 1. *The problems* ***(A)*** *and* ***(B)*** *are equivalent by a proper choice of* $\mathcal{A}(\cdot, \cdot; \mu)$ *and* $\mathcal{F}(\cdot; \mu)$.

Proof. We start by considering the following system of equations

$$a(u^e(\mu), v; \mu) = f(v; \mu), \forall v \in X^e \tag{3.82}$$

$$\overline{a(v, z^e(\mu); \mu)} = Q(u^e(\mu), v; \mu) + \frac{1}{2}\ell(v; \mu), \forall v \in X^e. \tag{3.83}$$

Now we let

$$u^e(\mu) = U^{+e} + U^{-e}, \tag{3.84}$$

$$z^e(\mu) = U^{+e} - U^{-e}, \tag{3.85}$$

then (3.82)and (3.83) become

$$a(U^{+e}, v; \mu) + a(U^{-e}, v; \mu) = f(v; \mu), \tag{3.86}$$

$$\overline{a(v, U^{+e}; \mu)} - \overline{a(v, U^{-e}; \mu)} = \frac{1}{2}\ell(v; \mu) + Q(U^{+e}, v; \mu) + Q(U^{-e}, v; \mu). \tag{3.87}$$

Adding (3.86) and (3.87), we get

$$\begin{aligned} &a(U^{+e}, V^+; \mu) + \overline{a(V^+, U^{+e}; \mu)} + a(U^{-e}, V^+; \mu) \\ &-\overline{a(V^+, U^{-e}; \mu)} - Q(U^{+e}, V^+; \mu) - Q(U^{-e}, V^+; \mu) \\ &= f(V^+; \mu) + \frac{1}{2}l(V^+; \mu), \forall V^+ \in X^e. \end{aligned} \tag{3.88}$$

Subtracting (3.86) from (3.87), we get

$$\begin{aligned}
&\overline{a(V^-, U^{+e}; \mu)} - a(U^{+e}, V^-; \mu) - \overline{a(V^-, U^{-e}; \mu)} \\
&- a(U^{-e}, V^-; \mu) - Q(U^{+e}, V^-; \mu) - Q(U^{-e}, V^-; \mu) \\
&\qquad = -f(V^-; \mu) + \frac{1}{2} l(V^-; \mu), \; \forall V^- \in X^e.
\end{aligned} \tag{3.89}$$

We now bring together (3.88) and (3.89) to form a Hermitian (equivalent to symmetric in the real case) non-coercive expanded system. We define the "big" operator $\mathcal{A}$ as

$$\begin{aligned}
\mathcal{A}(\mathcal{W}, \mathcal{V}; \mu) \quad &= a(W^+, V^+; \mu) + \overline{a(V^+, W^+; \mu)} \\
&+ a(W^-, V^+; \mu) - \overline{a(V^+, W^-; \mu)} \\
&+ \overline{a(V^-, W^+; \mu)} - a(W^+, V^-; \mu) \\
&- \overline{a(V^-, W^-; \mu)} - a(W^-, V^-; \mu) \\
&- Q(W^+, V^+; \mu) - Q(W^-, V^+; \mu) \\
&- Q(W^+, V^-; \mu) - Q(W^-, V^-; \mu),
\end{aligned} \tag{3.90}$$

and $\mathcal{F}$ as

$$\mathcal{F}(\mathcal{V}; \mu) = f(V^+; \mu) - f(V^-; \mu) + \frac{1}{2} l(V^+; \mu) + \frac{1}{2} l(V^-; \mu), \tag{3.91}$$

where $\mathcal{W} = [W^+, W^-]^T$ and $\mathcal{V} = [V^+, V^-]^T$. Note that $\mathcal{A}$ is symmetric: $\mathcal{A}(\mathcal{W}, \mathcal{V}; \mu) = \overline{\mathcal{A}(\mathcal{V}, \mathcal{W}; \mu)}$.

Now consider the system

$$\mathcal{A}(\mathcal{U}^e(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}; \mu), \quad \forall \mathcal{V} \in \mathcal{X}^e, \tag{3.92}$$

and its compliance output

$$\mathcal{S}^e(\mu) = \overline{\mathcal{F}(\mathcal{U}^e(\mu); \mu)}, \tag{3.93}$$

where $\mathcal{X}^e \subset (X^e)^2$ and $\mathcal{U}^e(\mu) = [U^{+e}, U^{-e}]^T$.

We can show that

$$\begin{aligned}
\mathcal{S}^e(\mu) &= \overline{\mathcal{F}(\mathcal{U}(\mu);\mu)} \\
&= \overline{f(U^{+e};\mu)} + \frac{1}{2}\ell(U^{+e};\mu) - \overline{f(U^{-e};\mu)} + \frac{1}{2}\ell(U^{-e};\mu) \\
&= \overline{f(z^e(\mu);\mu)} + \frac{1}{2}\ell(u^e(\mu);\mu) \\
&= \overline{a(u^e(\mu), z^e(\mu);\mu)} + \frac{1}{2}\ell(u^e(\mu);\mu) \\
&= Q(u^e(\mu), u^e(\mu);\mu) + \ell(u^e(\mu);\mu) \\
&= s^e(\mu), \qquad (3.94)
\end{aligned}$$

which concludes the proof.

□

We can also see that the general expanded formulation can handle pure linear outputs $\ell(v)$: this is akin to including a dual/adjoint formulation. Similarly, pure quadratic outputs can be handled as well — in this case however the quadratic terms enter into the $\mathcal{A}$ bilinear form.

Benefits of Expanded Formulation

We now discuss some of the benefits from the expanded formulation. The expanded formulation enables the treatment of quadratic outputs as linear outputs; however, instead of solving a problem with $\dim(X^e)$ number of unknowns, we reformulate the problem such that we have to solve a system with twice the number of unknowns, $\dim(\mathcal{X}^e) = 2\dim(X^e)$.

This seemingly disadvantageous step actually is very useful in the reduced-basis context: the change allows us to calculate the quadratic output, $Q(u^e(\mu), u^e(\mu);\mu)$, as the linear *compliant* output, $\mathcal{S}(\mathcal{U};\mu)$, associated with (3.92). In the reduced basis context, the linear *compliant* reduced basis output $\mathcal{S}_N(\mathcal{U}_N(\mu;\mu))$, converges rapidly to $\mathcal{S}(\mathcal{U};\mu)$.

Furthermore, the *compliance* of the output eliminates the need for a dual formulation which would otherwise be required to guarantee the same order of convergence for a linear *non-compliant* output. The primal-dual formulation is *necessary* for the treatment of general non-compliant outputs, but the quadratic non-linearity of the output requires additional special handling. We have some empirical evidence of this claim in the treatment of quadratic outputs like the stress intensity

factor: we obtain sharper effectivities (a measure of the reduced basis approximation accuracy) using the expanded formulation [52] relative to the more complicated primal-dual approach [53]. The expanded formulation is a clear winner over the significantly more complicated primal-dual formulation.

Most importantly, the reduced basis *a posteriori* error estimators for the linear *compliant* case are extremely *sharp*: eliminating the non-linear nature of the output is thus worth the effort of solving a system with twice the number of unknowns. The error estimators for quadratic outputs, based on the primal-dual formulation are also much less *sharp* [52].

Finally, it is important to note that the higher number of unknowns has an effect *only* on the *offline* stage when the reduced basis space is generated. The *online* effort is untainted by this increase; instead the rapid convergence of the output helps reduce the size of the basis, N, to tackle the linear *compliant* output. We shall discuss all these issues in more detail in Chapter 6.

Chapter 4

Model Problems in Acoustics

In Chapter 3, we introduced the abstract formulation for the acoustics waveguide problems. The abstract formulation is a shorthand notation of the formalism of the weak statement. In this chapter, we start with a brief description of the μ-dependent weak formulation. Since the reduced basis method itself is framed on a parameter-independent reference domain, we next present the reference domain formulation and required affine mappings. We then describe our model acoustic applications and derive the weak formulations for the same.

4.1 Weak Statement

While the derivation of the governing equations for the acoustics problems is not too difficult, the exact solution often cannot be found using analytical approaches. In such cases, numerical procedures are used to obtain approximate solutions — the finite element method is our chosen method for the numerical approximation. The starting point for the finite element method is a weighted-integral statement of the differential equation, called the weak formulation. Stating the problem in the weak form allows us to consider more general solution spaces; moreover, the natural boundary conditions and continuity conditions are included in the developed equations.

There has been considerable effort expended in obtaining finite element solutions for the Helmholtz equation. It has been shown [47] that the quality of discrete numerical solutions of the Helmholtz equation depends greatly on the wavenumber k. The finite element mesh size h needs to be adjusted to the wavenumber k. An approximate rule of thumb is to choose h such that $k\,h$ = constant; it

has been shown that in computations with low wavenumber this rule gives reasonably accurate solutions. However, with increasing wavenumber, the quality of the numerical results deteriorates significantly. The finite element treatment of the Helmholtz equation with high wavenumbers requires the use of special (h and $h-p$) finite elements [54, 55]. Our own efforts are in the low wavenumber regime; we thus avoid the need to use these special elements by choosing a very fine mesh for the computational domain.

We presented the weak form (3.68) in Chapter 3 without proof. Here we derive the weak form of the non-dimensional Helmholtz equation; please see Section 3.2.1 for a derivation of the non-dimensional form of the equations. For convenience, we restate the problem and the boundary conditions described in Chapter 3. The choice of the boundary conditions are dicated by the two problems we tackle in detail in this thesis.

We consider a suitably regular (smooth) domain $\hat{\Omega} \subset \mathbb{R}^d$, $d = 1, 2$, or 3 with Lipschitz-continuous boundary $\hat{\Gamma} = \hat{\Gamma}_{\text{in}} \cup \hat{\Gamma}_N \cup \hat{\Gamma}_{\text{liner}} \cup \hat{\Gamma}_{\text{out}}$. Here $\hat{\Gamma}_{\text{in}}$ represents the waveguide inlet boundary, $\hat{\Gamma}_N$ represents the sound-hard walls of the boundary, $\hat{\Gamma}_{\text{liner}}$ is the acoustic liner boundary, and $\hat{\Gamma}_{\text{out}}$ is the waveguide outlet where we prescribe the radiation boundary condition.

We solve for the non-dimensional pressure p from the governing Helmholtz equation

$$\nabla^2 p + k^2 p = 0, \text{ in } \hat{\Omega}, \tag{4.1}$$

where k is the non-dimensional wavenumber.

We have prescribed boundary conditions on the boundary $\hat{\Gamma}$: $\hat{\Gamma}_N$ represents the part of the domain with a homogeneous Neumann (zero velocity) condition

$$\frac{\partial p}{\partial n}|_{\hat{\Gamma}_N} = 0; \tag{4.2}$$

$\hat{\Gamma}_{\text{in}}$ represents the inlet boundary where we prescribe a known velocity condition

$$\frac{\partial p}{\partial \hat{n}}|_{\hat{\Gamma}_{\text{in}}} = \mathrm{i}\, k; \tag{4.3}$$

$\hat{\Gamma}_{\text{liner}}$ is the damped liner boundary where we specify a known non-dimensional impedance, Z,

$$\frac{\partial p}{\partial \hat{n}}|_{\hat{\Gamma}_{\text{liner}}} = -\frac{\mathrm{i}\, k}{Z}\, p; \tag{4.4}$$

and $\hat{\Gamma}_{\text{out}}$ represents the outflow boundary where we impose the radiation boundary condition

$$\frac{\partial p}{\partial n}|_{\hat{\Gamma}_{\text{out}}} = \sum_{j=1}^{N_{\text{prop}}} -\mathrm{i}\,\alpha_j \,(\int_{\hat{\Gamma}_{\text{out}}} p\,\overline{\Xi}_j)\,\Xi_j; \tag{4.5}$$

here $\hat{n}$ represents the outward normal to $\hat{\Gamma}$.

To derive the weak form of the governing equations, we introduce the complex vector space

$$\hat{X} = \{v = v_R + \mathrm{i}\,v_I \mid v_R \in \hat{X}_R,\, v_I \in \hat{X}_R\}, \tag{4.6}$$

$$\hat{X}_R = \{v \in (H^1(\hat{\Omega}))^d \mid v = 0 \text{ on } \hat{\Gamma}_D\}; \tag{4.7}$$

and associated inner product and norm

$$(w,v)_{\hat{X}} = \int_{\hat{\Omega}} \nabla w\,\nabla \overline{v} + w\,\overline{v}, \tag{4.8}$$

$$\|w\|_{\hat{X}} = (w,w)_{\hat{X}}. \tag{4.9}$$

Next, multiplying by a test function $v \in \hat{X}$ and integrating by parts we obtain

$$\int_{\hat{\Omega}} \nabla p\,\nabla \overline{v} - k^2 \int_{\hat{\Omega}} p\,\overline{v} - \int_{\hat{\Gamma}} \overline{v}\,\frac{\partial p}{\partial \hat{n}} = 0. \tag{4.10}$$

We now insert the boundary conditions (4.2)– (4.5) in the integral on the boundary, $\int_{\hat{\Gamma}} \overline{v}\,\frac{\partial p}{\partial \hat{n}}$. It thus follows that the pressure $p \in \hat{X}$ satisfies

$$a(p,v) = f(v),\ \forall v \in \hat{X}, \tag{4.11}$$

where for all $w, v \in \hat{X}$,

$$a(w,v) = \int_{\hat{\Omega}} \nabla w\,\nabla \overline{v} - k^2 \int_{\hat{\Omega}} w\,\overline{v} + \frac{\mathrm{i}\,k}{Z} \int_{\hat{\Gamma}_{\text{liner}}} w\,\overline{v} + \sum_{j=1}^{N_{\text{prop}}} \mathrm{i}\,\alpha_j \,(\int_{\hat{\Gamma}_{\text{out}}} w\,\overline{\Xi}_j)\,(\int_{\hat{\Gamma}_{\text{out}}} \Xi_j\,\overline{v}) \tag{4.12}$$

and

$$f(v) = \mathrm{i}\,k \int_{\hat{\Gamma}_N} \overline{v}. \tag{4.13}$$

4.2 Reference Domain Formulation

In our model problems, some of our parameters often affect the geometry of the domain or the boundary. The reference domain formulation is obtained by re-writing (4.11) on a parameter-independent "reference" domain, Ω. The RB recipe requires that Ω be *parameter-independent* — the μ-dependence of the domain geometry $\hat{\Omega}(\mu)$ violates that requirement. Thus, to allow for geometric variations we need to look at Ω as the image of the parameter-dependent domain $\hat{\Omega}(\mu)$.

The geometric transformations yield parameter-dependent coefficients in the reference domain linear and bilinear forms. These transformations are obtained using piecewise continuous affine mappings. This allows us to model all possible configurations, corresponding to every $\mu \in \mathcal{D}$, on a single reference domain. Furthermore, the affine transformations manifest themselves in the reference domain formulation through μ-dependent coefficients that multiply μ-independent bilinear forms; the bilinear form on the reference domain is thus in the requisite affine separable form.

4.2.1 Affine Mappings

We first partition the domain $\hat{\Omega}(\mu)$ into R (finite) non-overlapping suitably regular (Lipschitz) *parameter-dependent* sub-domains $\hat{\Omega}^r(\mu)$ such that, for any $\mu \in \mathcal{D}$, $\hat{\Omega}^r(\mu) \cap \hat{\Omega}^{r'}(\mu) = \emptyset, 1 \leq r < r' \leq R$; furthermore, we require

$$\overline{\hat{\Omega}}(\mu) = \cup_{r=1}^{R} \overline{\hat{\Omega}^r}(\mu); \tag{4.14}$$

recall that $\overline{\hat{\Omega}}$ denotes the closure of $\hat{\Omega}$.

We also introduce the *parameter-independent* domain Ω consisting of R non-overlapping suitably regular (Lipschitz) *parameter-independent* sub-domains Ω^r such that, $\Omega^r \cap \Omega^{r'} = \emptyset$, $1 \leq r < r' \leq R$; again, we require

$$\overline{\Omega} = \cup_{r=1}^{R} \overline{\Omega^r}. \tag{4.15}$$

We next introduce parametric *bijective* and *affine* mappings from the parameter-dependent sub-domains to the parameter-independent sub-domains: for any $\mu \in \mathcal{D}$, and for $1 \leq r \leq R$, $\mathcal{G}^r(\cdot;\mu)$ is an invertible affine transformation from $\overline{\hat{\Omega}^r}(\mu)$ to $\overline{\Omega^r}$. These mappings must satisfy a global continuity condition, for all $\mu \in \mathcal{D}$,

$$\mathcal{G}^r(\hat{x};\mu) = \mathcal{G}^{r'}(\hat{x};\mu),\ \forall \hat{x} \in \overline{\hat{\Omega}^r} \cap \overline{\hat{\Omega}^{r'}},\ 1 \leq r < r \leq R. \tag{4.16}$$

The affine transformation $\mathcal{G}^r(\hat{x};\mu)$ maps $\hat{x} \in \hat{\Omega}^r(\mu)$ to $x \in \Omega^r$, and is given by

$$x_i = \mathcal{G}_i^r(\hat{x};\mu) = g_i^r(\mu) + R_{ij}^r(\mu)\,\hat{x}_j,\ 1 \le i \le d, \tag{4.17}$$

for given $g^r : \mathcal{D} \to \mathbb{R}^d$ and $R^r : \mathcal{D} \to \mathbb{R}^{d\times d}$. Here

$$\frac{\partial}{\partial \hat{x}_i} = \frac{\partial x_j}{\partial \hat{x}_i}\frac{\partial}{\partial x_j} = R_{ji}^r(\mu)\frac{\partial}{\partial x_j}. \tag{4.18}$$

We then define the associated Jacobians $\det R^r(\mu)^{-1},\ 1 \le r \le R$, where det denotes determinant; the Jacobian is constant in space over each subdomain.

We also interpret our local mappings in terms of a global mapping. For any $\mu \in \mathcal{D}$, the local mappings $\mathcal{G}^r$ induce a global bijective piecewise-affine transformation $\mathcal{G}(\cdot;\mu) : \hat{\Omega}(\mu) \to \Omega$; the uniqueness of the mapping is preserved by the condition (4.16). For any $\hat{x} \in \hat{\Omega}(\mu)$, the image $x \in \Omega$ is given by

$$x = \mathcal{G}(\mu;\hat{x}) = \mathbf{R}(\mu)\,\hat{x} + \mathbf{g}(\mu). \tag{4.19}$$

Here $\mathbf{R}(\mu) \in \mathbb{R}^{d\times d}$ is a piecewise-constant matrix with components R_{ij}, $i,j = 1,\ldots,d$; $\mathbf{g}(\mu) \in \mathbb{R}^d$ is a piecewise-constant vector with components g_i, $i = 1,\ldots,d$. It thus follows that $\mathcal{G}(\mu;\hat{x}) : \hat{\Omega}(\mu) \to \Omega$ is a piecewise-affine geometric mapping.

4.2.2 Formulation on the Reference Domain

We now define the function space X as $X(\Omega) = \hat{X}(\mathcal{G}^{-1}(\mu;\Omega)) = \hat{X}(\hat{\Omega}(\mu))$ such that

$$X = \{v \in \left(H^1(\Omega)^d\right)\}, \tag{4.20}$$

and for any function $\hat{w} \in \hat{X}$, we define $w \in X$ such that $w(\underline{x}) = \hat{w}(\mathcal{G}^{-1}(\mu;\underline{x}))$. Furthermore, we have

$$d\hat{\Omega} = \det R^{-1}(\mu)\ d\Omega\ , \tag{4.21}$$

$$d\hat{\Gamma} = \left|R^{-1}(\mu)\,\underline{e}^t\right|\,d\Gamma\ , \tag{4.22}$$

where $\underline{e}^t$ is a unit vector tangent to the boundary Γ, and

$$\left|R^{-1}(\mu)\,\underline{e}^t\,\right| = \left(\sum_{i=1}^{d}\left(R_{ij}e_j^t\right)^2\right)^{1/2}. \tag{4.23}$$

We can then map $(\hat{w},\hat{v}) \rightarrow (w,v)$, and express the term

$$\begin{aligned}
\int_{\hat{\Omega}} \frac{\partial \hat{w}}{\partial \hat{x}_j}\,\frac{\partial \overline{\hat{v}}}{\partial \hat{x}_l}\,d\hat{\Omega} &= \sum_{r=1}^{R}\int_{\hat{\Omega}^r} \frac{\partial \hat{w}}{\partial \hat{x}_j}\frac{\partial \overline{\hat{v}}}{\partial \hat{x}_l}d\hat{\Omega}^r \\
&= \sum_{r=1}^{R}\int_{\Omega^r}\left(R^r_{jj'}(\mu)\frac{\partial w}{\partial x_j}\right)\left(\overline{R^r_{ll'}(\mu)}\frac{\partial \overline{v}}{\partial x_l}\right)\,\det(\mathbf{R}^{-1}(\mu))\,d\Omega^r \\
&= \sum_{r=1}^{R}\int_{\Omega^r}\frac{\partial w}{\partial x_j}\left(R^r_{jj'}(\mu)\overline{R^r_{ll'}(\mu)}\,\det(\mathbf{R}^{-1}(\mu))\right)\frac{\partial \overline{v}}{\partial x_l}\,d\Omega^r \\
&= \sum_{r=1}^{R}\int_{\Omega^r}\frac{\partial w}{\partial x_j}C_{jl}(\mu)\frac{\partial \overline{v}}{\partial x_l}\,d\Omega,\ \forall w,v\in X,
\end{aligned} \tag{4.24}$$

where

$$C_{jl}(\mu) = R^r_{jj'}(\mu)\,\overline{R^r_{ll'}(\mu)}\,\det(\mathbf{R}^{-1}(\mu)). \tag{4.25}$$

Similarly,

$$\begin{aligned}
\int_{\hat{\Omega}} \hat{w}\,\overline{\hat{v}}\,d\hat{\Omega} &= \sum_{r=1}^{R}\int_{\hat{\Omega}^r}\hat{w}\,\overline{\hat{v}}\,d\hat{\Omega}^r \\
&= \sum_{r=1}^{R}\int_{\Omega^r} w\,\overline{v}\,\det(\mathbf{R}^{-1}(\mu))\,d\Omega^r,\ \forall w,v\in X;
\end{aligned}$$

We can apply the same decomposition techniques to the forcing term. However, for our examples, there is no change in the mapping of the forcing term from the original to the reference domain and we omit the decomposition for the same.

We present some mappings and the subsequent reduction to the abstract form (3.73).

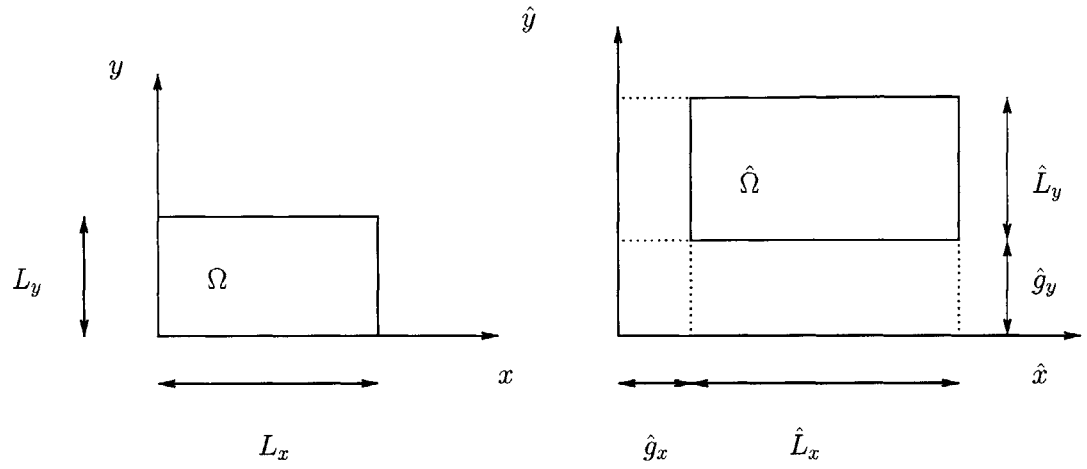


Figure 4-1: Affine transformation: Dilations in both x-direction and y-direction

4.2.3 Example 1: Dilations in both directions

We now consider the mapping from $\mathcal{G}(\hat{x};\mu) : \hat{\Omega}(\mu) \rightarrow \Omega(\mu)$ shown in Figure 4-1. The mapping involves dilations in both axis x and y and given by (4.17) where $\mathbf{g} = [\hat{g}_x \ \hat{g}_y]^T$ and

$$\mathbf{R} = \begin{bmatrix} t_x & 0 \\ 0 & t_y \end{bmatrix}, \tag{4.26}$$

where $t_x = \dfrac{\hat{L}_x}{L_x}$ and $t_y = \dfrac{\hat{L}_y}{L_y}$ are the scaling factor in the x and y directions, respectively.

We apply the mappings to (4.12) where

$$a(w,v) = \int_{\hat{\Omega}} \nabla w \, \nabla \overline{v} - k^2 \int_{\hat{\Omega}} w \, \overline{v}. \tag{4.27}$$

In the reference domain Ω, we obtain

$$a(w,v;\mu) = \frac{t_y}{t_x} \int_{\Omega} \frac{\partial w}{\partial x} \frac{\partial \overline{v}}{\partial x} + \frac{t_x}{t_y} \int_{\Omega} \frac{\partial w}{\partial y} \frac{\partial \overline{v}}{\partial y} - k^2 \, t_x \, t_y \int_{\Omega} w \, \overline{v}. \tag{4.28}$$

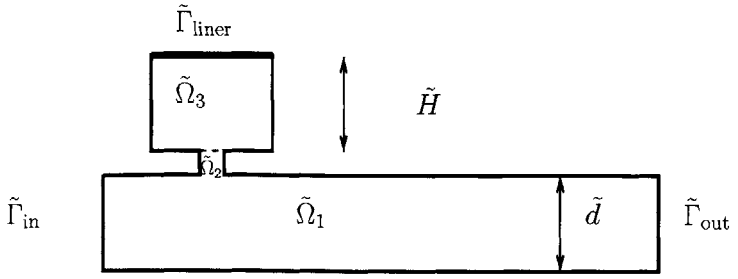


Figure 4-2: Band-Stop Filter: Original Domain (dimensional)

4.3 Band-Stop Filter

4.3.1 Problem Statement

In this example we consider the acoustic band-stop filter first introduced in Chapter 1. We consider a suitably regular (smooth) domain $\tilde{\Omega} \subset \mathbb{R}^2$ with Lipschitz-continuous boundary $\tilde{\Gamma}$. The waveguide $\tilde{\Omega}_1$ is coupled to a Helmholtz resonator, $\tilde{\Omega}_3$, by a thin neck $\tilde{\Omega}_2$. The waveguide is of depth $\tilde{d}$. The boundary $\tilde{\Gamma}_{\text{liner}}$ represents a backplate of fixed compliance — we shall refer to this as the "liner" throughout the thesis.

Dimensional Strong Form

We first pose our problem on $\tilde{\Omega}$ and present the actual boundary conditions specified on $\tilde{\Gamma}$. Let $\tilde{p}$ be the dimensional pressure, $\tilde{\rho}_0$ the density of the medium, $\tilde{c}_0$ the speed of sound in the medium and $\tilde{\omega}$ the driving frequency for the forcing system. We impose a fixed velocity $\tilde{U}_0$ at the inlet; $\tilde{n}$ denotes the outward normal to $\tilde{\Gamma}$.

The governing equation for $\tilde{p}$ is the Helmholtz acoustic equation

$$\nabla^2 \tilde{p} + \frac{\tilde{\omega}^2}{\tilde{c}_0^2} \tilde{p} = 0 \tag{4.29}$$

on the domain $\tilde{\Omega}$. We specify an inhomogeneous Neumann (imposed velocity) condition at the inlet,

$$\frac{\partial \tilde{p}}{\partial \tilde{n}}|_{\tilde{\Gamma}_{\text{in}}} = \mathrm{i}\, \tilde{\rho}_0\, \tilde{\omega}\, \tilde{U}_0. \tag{4.30}$$

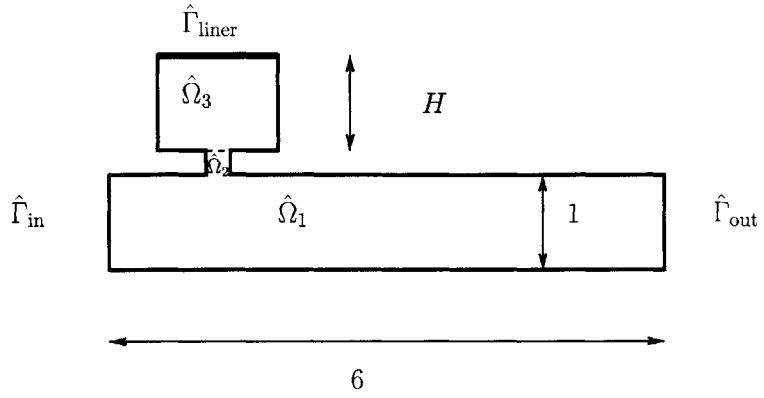


Figure 4-3: Band-Stop Filter: Non-dimensional Original Domain

We also have a resistive impedance condition on the compliant liner wall

$$\frac{\partial \tilde{p}}{\partial \tilde{n}}|_{\tilde{\Gamma}_{\text{liner}}} = -\mathrm{i}\,\tilde{\rho}_0\,\tilde{\omega}\,\frac{\tilde{p}}{\tilde{Z}} \tag{4.31}$$

where $\tilde{Z}$ is the (dimensional) acoustic impedance of the liner wall.

We impose a radiation boundary condition at the outflow; all other boundaries have homogeneous Neumann conditions.

Non-dimensional Strong Form

We show in Figure 4-3 the resulting (non-dimensional) original domain $\hat{\Omega}$. It consists of a long waveguide element $\hat{\Omega}_1$ coupled to a Helmholtz resonator element, $\hat{\Omega}_3$ by a thin neck $\hat{\Omega}_2$. We non-dimensionalize the problem with respect to the depth $\tilde{d}$ of the waveguide element: denote $k = \frac{\tilde{\omega}\,\tilde{d}}{\tilde{c}_0}$ as the non-dimensional wave-number.

We similarly non-dimensionalize the pressure and denote $p = \frac{\tilde{p}}{\tilde{\rho}_0\,\tilde{c}_0\,\tilde{U}_0}$ as the non-dimensional pressure per unit depth. The waveguide element has a non-dimensional depth of $L_y = 1$ and a (non-dimensional) length of $L_x = 6$.

We denote the non-dimensional original domain as $\hat{\Omega}$ with Lipschitz-continuous boundary $\hat{\Gamma}$, $\hat{n}$ denotes the outward normal to $\hat{\Gamma}$. We now pose the problem on $\hat{\Omega}$: we want to solve for p from $\nabla^2 p + k^2 p = 0$; the boundary condition at the inlet is specified as

$$\frac{\partial p}{\partial \hat{n}}|_{\hat{\Gamma}_{\text{in}}} = \mathrm{i}\,k. \tag{4.32}$$

We can also re-write the damping boundary condition as

$$\frac{\partial p}{\partial \hat{n}}|_{\hat{\Gamma}_{\text{liner}}} = -\mathrm{i}\,\frac{k}{Z_R}\,p \tag{4.33}$$

where $\tilde{Z} = \tilde{\rho}_0\,\tilde{c}_0\,Z_R$ and Z_R is the non-dimensional purely resistive acoustic impedance of the liner (i.e, the compliant resonator wall).

We next specify the radiation boundary condition as

$$\frac{\partial p}{\partial \hat{n}}|_{\hat{\Gamma}_{\text{out}}} = \sum_{j=1}^{N_{\text{prop}}} -\mathrm{i}\,\alpha_j\Big(\int_{\hat{\Gamma}_{\text{out}}} p\,\overline{\Xi}_j\Big)\Xi_j. \tag{4.34}$$

4.3.2 Parameters and Outputs

The parameters are the (non-dimensional) frequency squared, $\mu_1 \equiv k^2$, and the height of the Helmholtz resonator cavity (relative to the reference cavity height $H_{\text{ref}} = 1$), $\mu_2 \equiv H$; we consider the parameter domain $\mu \in \mathcal{D} \equiv [0.1, 5.0] \times [0.75, 1.25]$.

The non-dimensional frequency k corresponds to the dimensional frequency $\tilde{f} \equiv \frac{\tilde{c}_0\,k}{2\,\pi\,\tilde{d}}$. Typically, $\tilde{d} \approx 50$mm, for a sound speed $\tilde{c}_0 = 340$ m/s we are operating in a (dimensional) frequency range of 340 Hz to 2.4 KHz. These ranges are appropriate for the use of band-stop filters in industrial silencers and in acoustic liners for the suppression of aircraft engine noise [50, 105].

We shall assume that the liner (the compliant wall of the Helmholtz resonator) is of fixed compliance: we thus fix the non-dimensional acoustic impedance to $Z_R = 10$. The choice of the non-dimensional acoustic resistance, Z_R, was motivated by experimental studies [50, 105] of tunable electromechanical acoustic liners used for reduction of aircraft engine duct noise; other studies on MEMs-based acoustic liners [24] obtained lower ranges for acoustic impedance.

For our chosen range of k^2, we have $N_{\text{prop}} = 1$ and $\alpha_1 = k$. Thus, our radiation outflow condition is

$$\frac{\partial p}{\partial \hat{n}}|_{\hat{\Gamma}_{\text{out}}} = -\mathrm{i}\,k\Big(\int_{\hat{\Gamma}_{\text{out}}} p\,\overline{\Xi}_1\Big)\Xi_1. \tag{4.35}$$

The "output" of interest is the Transmission Coefficient (TC) given by

$$TC = -20\,log_{10}\Big(\frac{\sqrt{\int_{\hat{\Gamma}_{\text{out}}} \overline{p}\,p}}{\sqrt{\int_{\hat{\Gamma}_{\text{in}}} \overline{p}\,p}}\Big). = -10\,log_{10}\frac{\int_{\hat{\Gamma}_{\text{out}}} \overline{p}\,p}{\int_{\hat{\Gamma}_{\text{in}}} \overline{p}\,p} \tag{4.36}$$

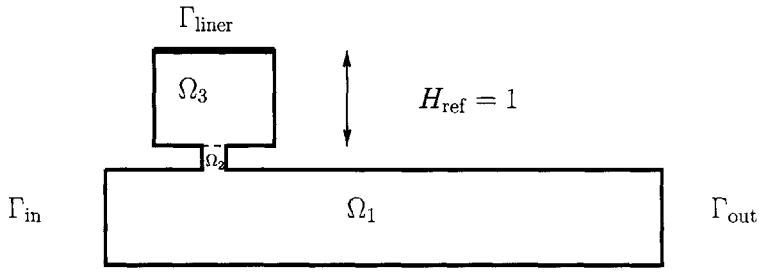


Figure 4-4: Band-Stop Filter: Reference Domain.

4.3.3 Weak Form on Original Domain

To derive the weak form of the governing equations, we introduce the complex vector space $\hat{X}^e$.

Next, multiplying by a test function $v \in \hat{X}^e$ and integrating by parts we obtain (substituting u^e for p)

$$\int_{\hat{\Omega}} \nabla u^e \nabla \overline{v} - k^2 \int_{\hat{\Omega}} u^e \overline{v} - \int_{\hat{\Gamma}_{\text{liner}}} \overline{v} \frac{\partial u^e}{\partial \hat{n}} - \int_{\hat{\Gamma}_{\text{out}}} \overline{v} \frac{\partial u^e}{\partial \hat{n}} = \int_{\hat{\Gamma}_{\text{in}}} \overline{v} \frac{\partial u^e}{\partial \hat{n}} \tag{4.37}$$

We then use our boundary conditions on inflow (4.32), liner (4.33) and outflow (4.35) to obtain the weak form on the non-dimensional original domain as

$$a(u^e, v) = f(v), \; \forall v \in \hat{X}^e, \tag{4.38}$$

where

$$a(u^e, v) = \int_{\hat{\Omega}} \nabla u^e \nabla \overline{v} - k^2 \int_{\hat{\Omega}} u^e \overline{v} - \frac{\mathrm{i}\, k}{Z_R} \int_{\hat{\Gamma}_{\text{liner}}} u^e \overline{v} + \mathrm{i}\, k \int_{\hat{\Gamma}_{\text{out}}} \Xi_1 \overline{v} \int_{\hat{\Gamma}_{\text{out}}} u^e \overline{\Xi}_1 \tag{4.39}$$

$$f(v) = \mathrm{i}\, k \int_{\hat{\Gamma}_{\text{in}}} \overline{v}. \tag{4.40}$$

4.3.4 Weak Form on Reference Domain

We obtain our reference domain Ω by fixing the height of the Helmholtz resonator $H_{\text{ref}} = 1$, as shown in Figure 4-4. Thus, while terms on $\hat{\Omega}_1$ and $\hat{\Omega}_2$ are mapped as is, we need to use the affine mappings for $\hat{\Omega}_3$. The transformation is piecewise affine: an identity for Ω_1 and Ω_2 and a dilation in the y-direction for Ω_3.

Thus, $u^e(\mu) \in X^e$ (X^e is the complex vector space on the reference domain Ω, $X^e(\Omega) = \hat{X}^e(\hat{\Omega})$)

q	Θ_a^q	$a^q(u^e, v)$
1	1	$\int_{\Omega_{1,2}} \nabla u^e \nabla \overline{v}$
2	H	$\int_{\Omega_3} \frac{\partial u^e}{\partial x} \frac{\partial \overline{v}}{\partial x}$
3	$\frac{1}{H}$	$\int_{\Omega_3} \frac{\partial u^e}{\partial y} \frac{\partial \overline{v}}{\partial y}$
4	$-k^2$	$\int_{\Omega_{1,2}} u^e \overline{v}$
5	$-k^2 H$	$\int_{\Omega_3} u^e \overline{v}$
6	$\frac{\mathrm{i}k}{Z_R}$	$\int_{\Gamma_{\text{liner}}} u^e \overline{v}$
7	$\mathrm{i}k$	$\int_{\Gamma_{\text{out}}} \Xi_1 \overline{v} \int_{\Gamma_{\text{out}}} u^e \overline{\Xi}_1$

Table 4.1: Affine decomposition of $a(u^e(\mu), v; \mu)$ in the reference domain for the Band-Stop Filter Problem.

satisfies the weak form of the μ-parametrized PDE

$$a(u^e(\mu), v; \mu) = f(v), \quad \forall\, v \in X^e. \tag{4.41}$$

The affine decomposition of $a(\cdot, \cdot; \mu) = \sum_{q=1}^{Q_a} \Theta_a^q(\mu)\, a^q(\cdot, \cdot)$ is shown in Table 4.1. The bilinear and linear forms admit affine separations for $Q_a = 7$ and $Q_f = 1$.

4.3.5 Expanded Formulation

To calculate our "actual" output (4.36) of interest, we note that we need to compute two quadratic outputs: (a)$\int_{\Gamma_{\text{in}}} \overline{p}\, p$, and (b)$\int_{\Gamma_{\text{out}}} \overline{p}\, p$. In Chapter 3, we introduced the expanded formulation to deal with a *single* quadratic output. To compute the transmission coefficient given by (4.36) we will need two separate expanded formulations for our two quadratic outputs $\int_{\Gamma_{\text{in}}} \overline{p}\, p$ and $\int_{\Gamma_{\text{out}}} \overline{p}\, p$.

We also introduce the complex "expanded" function space $\mathcal{X}^e \subset (X^e)^2$ that will be required for the expanded formulations. Thus, $\dim(\mathcal{X}^e) = 2\dim(X^e)$.

Pressure Intensity at Inflow: the Inlet System

We denote the pressure intensity over the inlet as our first quadratic output as

$$Q_1(u^e(\mu), u^e(\mu)) = \int_{\Gamma_{\text{in}}} \overline{u^e}\, u^e. \tag{4.42}$$

Given $\mu \in \mathcal{D} \in \mathbb{R}^P$, we find our "compliance" output of interest

$$\mathcal{S}_1^e(\mu) = \overline{\mathcal{F}(\mathcal{U}^e(\mu); \mu)}, \tag{4.43}$$

where $\mathcal{U}^e(\mu) = [U^{+e}, U^{-e}] \in \mathcal{X}^e$ ($\mathcal{X}^e \in (X^e)^2$) satisfies

$$\mathcal{A}(\mathcal{U}^e(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}; \mu), \quad \forall \mathcal{V} = [V^+, V^-]^T \in \mathcal{X}^e, \tag{4.44}$$

where

$$\begin{aligned}
\mathcal{A}(\mathcal{U}^e(\mu), \mathcal{V}; \mu) \quad &= a(U^{+e}, V^+; \mu) + \overline{a(V^+, U^{+e}; \mu)} \\
&+ a(U^{-e}, V^+; \mu) - \overline{a(V^+, U^{-e}; \mu)} \\
&+ \overline{a(V^-, U^{+e}; \mu)} - a(U^{+e}, V^-; \mu) \\
&- \overline{a(V^-, U^{-e}; \mu)} - a(U^{-e}, V^-; \mu) \\
&- Q_1(U^{+e}, V^+; \mu) - Q_1(U^{-e}, V^+; \mu) \\
&- Q_1(U^{+e}, V^-; \mu) - Q_1(U^{-e}, V^-; \mu),
\end{aligned} \tag{4.45}$$

and $\mathcal{F}$ as

$$\mathcal{F}(\mathcal{V}; \mu) = f(V^+; \mu) - f(V^-; \mu), \tag{4.46}$$

$\mathcal{A}(\cdot, \cdot; \mu) = \sum_{q=1}^{\mathcal{Q}_\mathcal{A}} \Theta_\mathcal{A}^q(\mu)\, \mathcal{A}(\cdot, \cdot)$ and $\mathcal{F} = \sum_{q=1}^{\mathcal{Q}_\mathcal{F}} \Theta_\mathcal{F}^q(\mu)\, \mathcal{F}^q(\cdot)$ are also "affine" in the parameter, we recover $\mathcal{Q}_\mathcal{A} = 7$ and $\mathcal{Q}_\mathcal{F} = 1$. We can relate each a^q term to its corresponding $\mathcal{A}^q$ term as shown in Table 4.2. In Table 4.2 we only show the terms for $q = 1$ and $q \neq 1$. We have the same number of terms in the expanded bilinear form $\mathcal{A}(\cdot, \cdot; \mu)$ as we do in $a(\cdot, \cdot; \mu)$: $Q_a = Q_\mathcal{A} = 7$ and $\Theta_\mathcal{A}^q(\mu) = \Theta_a^q(\mu)$. There is a one-to-one mapping of the expanded bilinear forms $\mathcal{A}^q(\cdot, \cdot)$, $q = 1, \ldots, Q_\mathcal{A}$ to $a^q(\cdot, \cdot)$, $q = 1, \ldots, Q_a$; however the mappings are slightly different for $q = 1$ relative to $q \neq 1$. We show the one-to-one mapping from $a^q(\cdot, \cdot)$ to $\mathcal{A}^q(\cdot, \cdot)$ in Table 4.2 — note that for $q = 1$, the $\mathcal{A}^1(\cdot, \cdot)$ has some additional terms related to the quadratic output $Q_1(\cdot, \cdot)$. The quadratic output $Q_1(\cdot, \cdot)$ is μ-independent, from Table 4.1 we see that $\Theta_a^1(\mu) \equiv \Theta_\mathcal{A}^1(\mu) = 1$; thus the quadratic output terms only enter into the bilinear form $\mathcal{A}^1(\cdot, \cdot)$.

q	$\mathcal{A}^q(\mathcal{U}^e, \mathcal{V})$
$q = 1$	$a^1(U^{+e}, V^+; \mu) + \overline{a^1(V^+, U^{+e}; \mu)}$ $+a^1(U^{-e}, V^+; \mu) - \overline{a^1(V^+, U^{-e}; \mu)}$ $+ \overline{a^1(V^-, U^{+e}; \mu)} - a^1(U^{+e}, V^-; \mu)$ $- \overline{a^1(V^-, U^{-e}; \mu)} - a^1(U^{-e}, V^-; \mu)$ $-Q_1(U^{+e}, V^+; \mu) - Q_1(U^{-e}, V^+; \mu)$ $- Q_1(U^{+e}, V^-; \mu) - Q_1(U^{-e}, V^-; \mu)$
$q \neq 1$	$a^q(U^{+e}, V^+; \mu) + \overline{a^q(V^+, U^{+e}; \mu)}$ $+a^q(U^{-e}, V^+; \mu) - \overline{a^q(V^+, U^{-e}; \mu)}$ $+ \overline{a^q(V^-, U^{+e}; \mu)} - a^q(U^{+e}, V^-; \mu)$ $- \overline{a^q(V^-, U^{-e}; \mu)} - a^q(U^{-e}, V^-; \mu)$

Table 4.2: Affine decomposition of $\mathcal{A}(\mathcal{U}^e(\mu), \mathcal{V}; \mu)$ for the Band-Stop Filter Inlet System

Pressure Intensity at Outflow: the Outlet System

We denote the pressure intensity over the outlet as our second quadratic output

$$Q_2(u^e(\mu), u^e(\mu)) = \int_{\Gamma_{\text{out}}} \overline{u^e}\, u^e. \tag{4.47}$$

Given $\mu \in \mathcal{D} \in \mathbb{R}^P$, we find our "compliance" output of interest in a similar manner to that shown for the inflow pressure intensity.

Given $\mu \in \mathcal{D} \in \mathbb{R}^P$, we find our "compliance" output of interest

$$\mathcal{S}_2^e(\mu) = \overline{\mathcal{F}(\mathcal{U}^e(\mu); \mu)}, \tag{4.48}$$

where $\mathcal{U}^e(\mu) = [U^{+e}, U^{-e}] \in \mathcal{X}^e$ satisfies

$$\mathcal{A}(\mathcal{U}^e(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}; \mu), \quad \forall \mathcal{V} = [V^+, V^-]^T \in \mathcal{X}^e, \tag{4.49}$$

where

$$\begin{aligned}
\mathcal{A}(\mathcal{U}^e(\mu),\mathcal{V};\mu) \quad &= a(U^{+e},V^+;\mu) + \overline{a(V^+,U^{+e};\mu)} \\
&+ a(U^{-e},V^+;\mu) - \overline{a(V^+,U^{-e};\mu)} \\
&+ \overline{a(V^-,U^{+e};\mu)} - a(U^{+e},V^-;\mu) \\
&- \overline{a(V^-,U^{-e};\mu)} - a(U^{-e},V^-;\mu) \\
&- Q_2(U^{+e},V^+;\mu) - Q_2(U^{-e},V^+;\mu) \\
&- Q_2(U^{+e},V^-;\mu) - Q_2(U^{-e},V^-;\mu),
\end{aligned} \tag{4.50}$$

and $\mathcal{F}$ as

$$\mathcal{F}(\mathcal{V};\mu) = f(V^+;\mu) - f(V^-;\mu), \tag{4.51}$$

$\mathcal{A}$ and $\mathcal{F}$ are also "affine" in the parameter, we recover $\mathcal{Q}_\mathcal{A} = 7$ and $\mathcal{Q}_\mathcal{F} = 1$, each a^q term can be easily corresponded with its associated $\mathcal{A}^q$ term as shown in Table 4.2 after replacing Q_1 with Q_2 for $q = 1$, the remaining terms are unaffected.

Note that the only difference in the inlet and outlet system formulations is on the left hand side: $\mathcal{A}^1(\cdot,\cdot;\mu)$ for the inlet system (4.45) has Q_1 while $\mathcal{A}^1(\cdot,\cdot;\mu)$ for the outlet system (4.50) has Q_2; the right hand side is the same for both formulations.

4.4 Low-Pass Filter

4.4.1 Problem Statement

We now consider a more complex parametrization with $P = 4$ parameters. In this example we consider the acoustic low-pass filter first introduced in Chapter 1. The low-pass filter problem is significantly different from the band-stop filter problem introduced earlier. The scaling up from $P = 2$ to $P = 4$ parameters requires the choice of smaller ranges of the parameter inputs primarily because the associated increase in complexity is fairly difficult to handle.

We consider a suitably regular (smooth) domain $\tilde{\Omega} \subset \mathbb{R}^2$ with Lipschitz-continuous boundary $\tilde{\Gamma}$. We show in Figure 4-5 the original domain $\tilde{\Omega}$. It consists of two waveguide sections $\tilde{\Omega}_1$ and $\tilde{\Omega}_3$

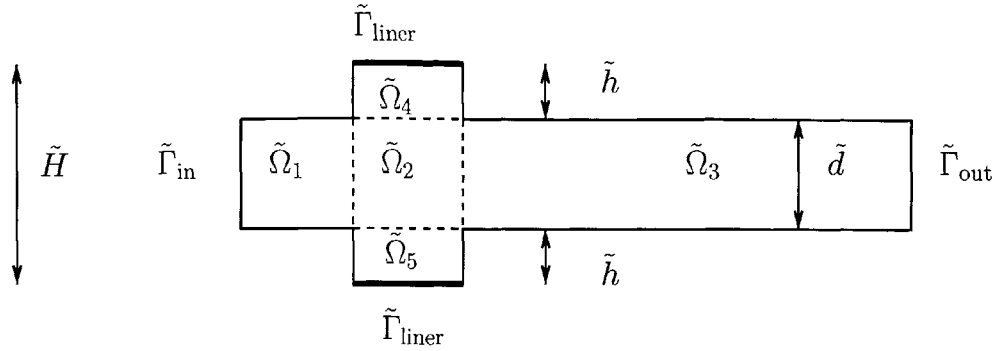


Figure 4-5: Low-Pass Filter: Original Domain (dimensional).

of fixed cross-section connected to an expansion chamber (represented by the sub-domains $\tilde{\Omega}_2$, $\tilde{\Omega}_4$, and $\tilde{\Omega}_5$) in the middle. The height of the expansion chamber, $\tilde{H}$, can be changed by changing $\tilde{h}$, the heights of the sub-domains $\tilde{\Omega}_4$ and $\tilde{\Omega}_5$ while the sub-domain $\tilde{\Omega}_2$ remains fixed. The waveguide is of fixed depth $\tilde{d}$. The top and bottom walls of the expansion chamber, denoted by $\tilde{\Gamma}_{\text{liner}}$ are compliant: we shall refer to them throughout this thesis as liner walls.

Dimensional Strong Form

We first pose our problem on $\tilde{\Omega}$ and present the actual boundary conditions specified on $\tilde{\Gamma}$. As in the band-stop filter problem, we denote $\tilde{p}$ as the dimensional pressure, $\tilde{\rho}_0$ as the density of the medium, $\tilde{c}_0$ as the speed of sound in the medium and $\tilde{\omega}$ as the driving frequency for the forcing system. We impose a fixed velocity $\tilde{U}_0$ at the inlet; $\tilde{n}$ denotes the outward normal to $\tilde{\Gamma}$.

The governing equation for $\tilde{p}$ is the Helmholtz acoustic equation

$$\nabla^2 \tilde{p} + \frac{\tilde{\omega}^2}{\tilde{c}_0^2} \tilde{p} = 0 \tag{4.52}$$

on the domain $\tilde{\Omega}$. We specify an inhomogeneous Neumann (imposed velocity) condition at the inlet,

$$\frac{\partial \tilde{p}}{\partial \tilde{n}}|_{\tilde{\Gamma}_{\text{in}}} = \mathrm{i}\, \tilde{\rho}_0\, \tilde{\omega}\, \tilde{U}_0. \tag{4.53}$$

We also have damping on the acoustic liner walls at the bottom and top of the expansion chamber

$$\frac{\partial \tilde{p}}{\partial \tilde{n}}|_{\tilde{\Gamma}_{\text{in}}} = -\mathrm{i}\, \tilde{\rho}_0\, \tilde{\omega}\, \frac{\tilde{p}}{\tilde{Z}} \tag{4.54}$$

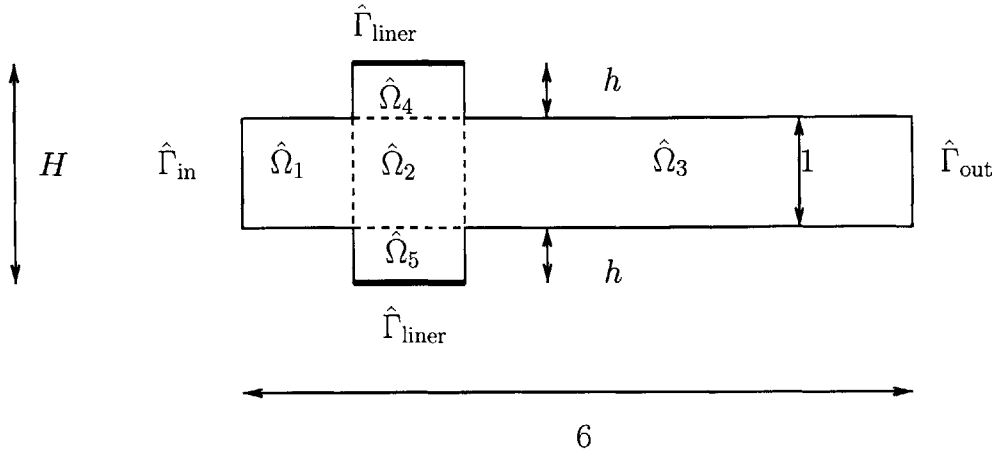


Figure 4-6: Low-Pass Filter: Non-dimensional Original Domain

where $\tilde{Z}$ is the acoustic impedance of the liner walls. The dimensional acoustic impedance $\tilde{Z} \in \mathbb{C}$ is not fixed.

We impose the radiation boundary condition at the outflow; all other walls have homogeneous Neumann conditions.

Non-dimensional Strong Form

We show in Figure 4-6 the resulting (non-dimensional) original domain $\hat{\Omega}$. It consists of two waveguide sections $\hat{\Omega}_1$ and $\hat{\Omega}_3$ of fixed cross-section connected to an expansion chamber (represented by the sub-domains $\hat{\Omega}_2$, $\hat{\Omega}_4$, and $\hat{\Omega}_5$) in the middle. The height of the expansion chamber, H, can be changed by changing h, the heights of the sub-domains $\hat{\Omega}_4$ and $\hat{\Omega}_5$ while the sub-domain $\hat{\Omega}_2$ remains fixed.

As before, we non-dimensionalize the problem with respect to the depth $\tilde{d}$ of the constant waveguide section: denote $k = \frac{\tilde{\omega}\tilde{d}}{\tilde{c}_0}$ as the non-dimensional wave-number, and $p = \frac{\tilde{p}}{\tilde{\rho}_0 \tilde{c} \tilde{U}_0}$ as the non-dimensional pressure per unit depth. The waveguide element has a non-dimensional depth of 1 and a (non-dimensional) length of 6.

The boundary condition at the inlet is then specified as

$$\frac{\partial p}{\partial \hat{n}}|_{\hat{\Gamma}_{\text{in}}} = \mathrm{i}\, k. \tag{4.55}$$

The boundary conditon at the acoustic liner walls is given as

$$\frac{\partial p}{\partial \hat{n}}|_{\hat{\Gamma}_{\text{liner}}} = -\mathrm{i}\,\frac{k}{Z_R + \mathrm{i}\,Z_I}\,p \tag{4.56}$$

where $\tilde{Z} = \tilde{\rho}_0\,\tilde{c}_0\,(Z_R + \mathrm{i}\,Z_I)$. Here Z_R and Z_I are the non-dimensional acoustic resistance and acoustic reactance of the liner wall.

We next specify the radiation boundary condition as

$$\frac{\partial p}{\partial \hat{n}}|_{\hat{\Gamma}_{\text{out}}} = \sum_{j=1}^{N_{\text{prop}}} -\mathrm{i}\,\alpha_j\Big(\int_{\hat{\Gamma}_{\text{out}}} p\,\overline{\Xi}_j\Big)\Xi_j. \tag{4.57}$$

4.4.2 Parameters and Outputs

The parameters are the (non-dimensional) frequency squared, $\mu_1 \equiv k^2$, the height of the expansion sections h (relative to the reference height $h_{\text{ref}} = 0.5$), $\mu_2 \equiv \frac{h}{h_{\text{ref}}}$, the non-dimensional acoustic resistance of the liner walls, $\mu_3 \equiv Z_R$, and the non-dimensional acoustic reactance of the liner walls, $\mu_4 \equiv Z_I$ respectively. We consider the parameter domain $\mu \in \mathcal{D} \equiv [0.1, 2.5] \times [0.75, 1.25] \times [4.8, 5.0] \times [1.8, 2.0]$.

The non-dimensional frequency k corresponds to the dimensional frequency $\tilde{f} \equiv \frac{\tilde{c}_0\,k}{2\,\pi\,\tilde{d}}$. Typically, $\tilde{d} \approx 50$mm, for a sound speed $\tilde{c}_0 = 340$ m/s we are operating in a (dimensional) frequency range of 340 Hz to 1.7 KHz. We choose fairly restrictive ranges for the non-dimensional acoustic resistance, Z_R, and the non-dimensional acoustic reactance, Z_I; this choice was motivated by experimental studies [50, 105] of tunable electromechanical acoustic liners.

For our chosen range of k^2, we have $N_{\text{prop}} = 1$ and $\alpha_1 = k$. Thus, our radiation outflow condition is

$$\frac{\partial p}{\partial \hat{n}}|_{\hat{\Gamma}_{\text{out}}} = -\mathrm{i}\,k\Big(\int_{\hat{\Gamma}_{\text{out}}} p\,\overline{\Xi}_1\Big)\,\Xi_1. \tag{4.58}$$

The "output" of interest is the transmission coefficient

$$TC = -10\,log_{10}\frac{\int_{\Gamma_{\text{out}}} \overline{p}\,p}{\int_{\Gamma_{\text{in}}} \overline{p}\,p}. \tag{4.59}$$

4.4.3 Weak Form on Original Domain

To derive the weak form of the governing equations, we introduce the complex vector space $\hat{X}^e$.

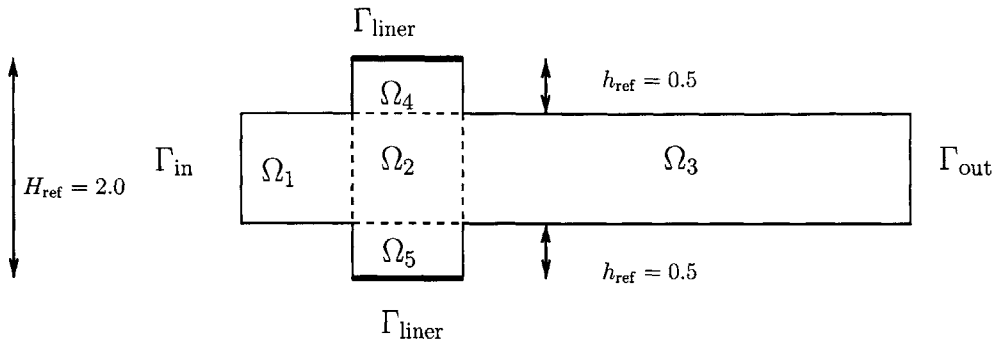


Figure 4-7: Low-Pass Filter: Reference Domain.

Next, multiplying by a test function $v \in \hat{X}^e$ and integrating by parts we obtain (substituting u^e for p)

$$\int_{\hat{\Omega}} \nabla u^e \nabla \overline{v} - k^2 \int_{\hat{\Omega}} u^e \overline{v} - \int_{\hat{\Gamma}_{\text{liner}}} \overline{v} \frac{\partial u^e}{\partial \hat{n}} - \int_{\hat{\Gamma}_{\text{out}}} \overline{v} \frac{\partial u^e}{\partial \hat{n}} = \int_{\hat{\Gamma}_{\text{in}}} \overline{v} \frac{\partial u^e}{\partial \hat{n}} \tag{4.60}$$

We then use our boundary conditions on inflow (4.55), liner (4.56) and outflow (4.58) to obtain the weak form as

$$a(u^e, v) = f(v), \; \forall v \in \hat{X}^e, \tag{4.61}$$

where

$$\begin{aligned} a(u^e, v) &= \int_{\hat{\Omega}} \nabla u^e \nabla \overline{v} - k^2 \int_{\hat{\Omega}} u^e \overline{v} - \frac{\mathrm{i}\, k}{Z_R + \mathrm{i} Z_I} \int_{\hat{\Gamma}_{\text{liner}}} u^e \overline{v} + \mathrm{i}\, k \int_{\hat{\Gamma}_{\text{out}}} \Xi_1 \overline{v} \int_{\hat{\Gamma}_{\text{out}}} u^e \overline{\Xi}_1 \\ f(v) &= \mathrm{i}\, k \int_{\hat{\Gamma}_{\text{in}}} \overline{v}. \end{aligned} \tag{4.62}$$

4.4.4 Weak Form on Reference Domain

We obtain our reference domain Ω by fixing the height of the expansion chamber $H_{\text{ref}} = 2.0$, ($h_{\text{ref}} = 0.5$) as shown in Figure 4-7. Thus, while terms on $\overline{\Omega}_1$-$\overline{\Omega}_3$ are mapped as is, we need to use the affine mappings for $\overline{\Omega}_4$ and $\overline{\Omega}_5$. The transformation is piecewise affine: an identity for Ω_1, Ω_2, and Ω_3 and a dilation in the y-direction for Ω_4 and Ω_5.

Thus, $u^e(\mu) \in X^e$ (X^e is the complex vector space on the reference domain Ω, $X^e(\Omega) = \hat{X}^e(\hat{\Omega})$) satisfies the weak form of the μ-parametrized PDE

$$a(u^e(\mu), v; \mu) = f(v), \quad \forall\, v \in X^e. \tag{4.63}$$

The affine decomposition of a is shown in Table 4.3. The bilinear and linear forms admit affine

q	Θ_a^q	$a^q(u^e, v)$
1	1	$\int_{\Omega_{1,2}} \nabla u^e \nabla \overline{v}$
2	H	$\int_{\Omega_3} \frac{\partial u^e}{\partial x} \frac{\partial \overline{v}}{\partial x}$
3	$\frac{1}{H}$	$\int_{\Omega_3} \frac{\partial u^e}{\partial y} \frac{\partial \overline{v}}{\partial y}$
4	$-k^2$	$\int_{\Omega_{1,2}} u^e \overline{v}$
5	$-k^2 H$	$\int_{\Omega_3} u^e \overline{v}$
6	$\frac{k Z_I}{Z_R^2+Z_I^2}$	$\int_{\Gamma_{\text{liner}}} u^e \overline{v}$
7	$\frac{\mathrm{i} k Z_R}{Z_R^2+Z_I^2}$	$\int_{\Gamma_{\text{liner}}} u^e \overline{v}$
8	$\mathrm{i} k$	$\int_{\Gamma_{\text{out}}} \Xi_1 \overline{v} \int_{\Gamma_{\text{out}}} u^e \overline{\Xi}_1$

Table 4.3: Affine decomposition of $a(u^e(\mu), v; \mu)$ in the reference domain for the Low-Pass Filter Problem.

separations for $Q_a = 8$ and $Q_f = 1$.

4.4.5 Expanded Formulation

To calculate our "actual" output of interest, we note that we need to compute two quadratic outputs: (a) $\int_{\Gamma_{\text{out}}} \overline{p}\, p$, and (b) $\int_{\Gamma_{\text{in}}} \overline{p}\, p$. We thus will need two expanded formulations to compute the "actual" output of interest.

As before, we introduce the complex "expanded" function space $\mathcal{X}^e \subset (X^e)^2$ that will be required for the expanded formulations. Thus, $\dim(\mathcal{X}^e) = 2\dim(X^e)$.

Pressure Intensity at Inflow: the Inlet System

We denote the pressure intensity over the inlet as our first quadratic output as

$$Q_1(u^e(\mu), u^e(\mu)) = \int_{\Gamma_{\text{in}}} \overline{u^e}\,(u^e). \tag{4.64}$$

Given $\mu \in \mathcal{D} \in \mathbb{R}^P$, we find our "compliance" output of interest

$$\mathcal{S}_1^e(\mu) = \overline{\mathcal{F}(\mathcal{U}^e(\mu); \mu)}, \tag{4.65}$$

where $\mathcal{U}^e(\mu) = [U^{+e}, U^{-e}] \in \mathcal{X}^e$ satisfies

$$\mathcal{A}(\mathcal{U}^e(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}; \mu), \quad \forall \mathcal{V} = [V^+, V^-]^T \in \mathcal{X}^e, \tag{4.66}$$

where

$$\begin{aligned}
\mathcal{A}(\mathcal{U}^e(\mu), \mathcal{V}; \mu) \quad &= a(U^{+e}, V^+; \mu) + \overline{a(V^+, U^{+e}; \mu)} \\
&+ a(U^{-e}, V^+; \mu) - \overline{a(V^+, U^{-e}; \mu)} \\
&+ \overline{a(V^-, U^{+e}; \mu)} - a(U^{+e}, V^-; \mu) \\
&- \overline{a(V^-, U^{-e}; \mu)} - a(U^{-e}, V^-; \mu) \\
&- Q_1(U^{+e}, V^+; \mu) - Q_1(U^{-e}, V^+; \mu) \\
&- Q_1(U^{+e}, V^-; \mu) - Q_1(U^{-e}, V^-; \mu),
\end{aligned} \tag{4.67}$$

and $\mathcal{F}$ as

$$\mathcal{F}(\mathcal{V}; \mu) = f(V^+; \mu) - f(V^-; \mu), \tag{4.68}$$

$\mathcal{A}$ and $\mathcal{F}$ are also "affine" in the parameter, we recover $\mathcal{Q}_\mathcal{A} = 8$ and $\mathcal{Q}_\mathcal{F} = 1$, each a^q term can be easily corresponded with its associated $\mathcal{A}^q$ term in a manner similar to Table 4.2.

Pressure Intensity at Outflow: the Outlet System

We denote the pressure intensity over the outlet as our second quadratic output

$$Q_2(u^e(\mu), u^e(\mu)) = \int_{\Gamma_{\text{out}}} \overline{u^e}\,(u^e). \tag{4.69}$$

Given $\mu \in \mathcal{D} \in \mathbb{R}^P$, we find our "compliance" output of interest in a similar manner to that shown for the inflow pressure intensity.

$\mathcal{A}$ and $\mathcal{F}$ are also "affine" in the parameter, we recover $\mathcal{Q}_\mathcal{A} = 8$ and $\mathcal{Q}_\mathcal{F} = 1$, each a^q term can be easily corresponded with its associated $\mathcal{A}^q$ term as shown in Table 4.2 after replacing Q_1 with Q_2 for $q = 1$, the remaining terms are unaffected.

Chapter 5

Finite Element Treatment of Waveguide Problems

The analytical solution of a partial differential equation is often difficult to compute. In such cases, numerical techniques are used to obtain approximate solutions as surrogates for the "exact" solution. The finite element method (FEM) is one of the most popular numerical methods available for constructing approximate solutions to the exact solution. In this chapter we shall review the standard finite element method relative to acoustic applications and then focus on the treatment of the radiation boundary condition in the finite element context. We then outline the "truth" formulation for acoustics waveguide problems with quadratic outputs.

For a detailed discussion on the finite element solution of the Helmholtz equation, please see [8, 47]. The finite element treatment of high wavenumbers requires the use of special finite element techniques; [54, 55] discuss these issues extensively. In the acoustic filter applications under consideration we are operating in low wavenumbers, this is made clear given that our radiation boundary condition only assumes that one mode is propagating. The accuracy of the radiation boundary condition however is not limited by the number of modes assumed to be propagating.

In what follows, we use ^{e} to refer to the exact solution to our problem on the reference domain for the prescribed mathematical model.

5.1 Finite Element Method

5.1.1 Weak formulation

The finite element method is not based on the strong form formulation, but rather an weighted-integral statement of it, which is the weak formulation. We described the weak form of the Helmholtz equation for our model acoustic problems in detail in Chapter 4. Here, we start with the abstract weak formulation for the exact statement: given $\mu \in \mathcal{D} \in \mathbb{R}^P$, find $s^e(\mu) = \ell(u^e(\mu))$, where $u^e(\mu) \in X^e$ is the solution of

$$a(u^e(\mu), v; \mu) = f(v), \quad \forall\, v \in X^e. \tag{5.1}$$

Here $a(\cdot, \cdot; \mu)$ is a μ-parametrized bilinear form, f is a linear functional, and X^e is an appropriate complex Hilbert space over the physical domain $\Omega \in \mathbb{R}^d$.

5.1.2 "Truth" Approximation Space and Basis

In the finite element method, we seek the approximate solution over a discretized domain known as a triangulation $\mathcal{T}_h$ of the physical domain Ω: $\overline{\Omega} = \bigcup_{T_h \in \mathcal{T}_h} \overline{T}_h$, where $T_h^k, k = 1, \ldots, K$, are the elements, $x_i, i = 1, \ldots, \mathcal{N}$, are the nodes, and subscript h denoting the diameter of the triangulation $\mathcal{T}_h$ is the maximum of the longest edges of all elements. The small elements that together constitute the domain are segments in $\mathbb{R}^1$, triangles/quadrilaterals in $\mathbb{R}^2$, and tetrahedras/hexagonals in $\mathbb{R}^3$. Simple elements such as triangular/tetrahedral elements work quite well for simple domains such as those with straight boundaries and planar surfaces. Complex domains involving curved boundaries or non-planar surfaces require the use of isoparametric elements that can represent the curvature of the boundaries/surfaces accurately.

We shall seek the approximation solution over $\mathcal{T}_h$. Our "exact" approximation space is complex

$$X^e = \{v = v_R + \mathrm{i} v_I \mid v_R \in H^1(\Omega),\ v_I \in H^1(\Omega)\}\ . \tag{5.2}$$

We next define a finite element "truth" approximation space $X \subset X^e$,

$$X = \{v = v_R + \mathrm{i} v_I \in X^e \mid v_R|_{T_h} \in \mathbb{P}_p(T_h),\ v_I|_{T_h} \in \mathbb{P}_p(T_h),\ \forall\, T_h \in \mathcal{T}_h\}\ , \tag{5.3}$$

where $\mathbb{P}_p(T_h)$ is the space of p^{th} degree polynomials over element T_h.

Our finite element "truth" approximation space is associated with an inner product

$$(w, v)_X = \int_\Omega \nabla w \nabla \bar{v} + w\bar{v} \ . \tag{5.4}$$

Recall that $\bar{v}$ and $|v|$ denote the complex conjugate and modulus of v, respectively. Note the notion of symmetry in the complex case, a bilinear form $a(w, v)$ is said to be symmetric if and only if $a(w, v) = \overline{a(v, w)}, \forall w, v \in X$. It is clear that $(\cdot, \cdot)_X$ defined above is symmetric.

To obtain the discrete equations of the weak form, we express the field variable $u(\mu) \in X$ in terms of the nodal basic functions $\varphi_i \in X$, $\varphi_i(x_j) = \delta_{ij}$, such that

$$X = \text{span}\,\{\varphi_1, \ldots, \varphi_{\mathcal{N}}\} \ , \tag{5.5}$$

$$u(\mu) = \sum_{j=1}^{\mathcal{N}} u_j(\mu)\varphi_j, \quad \forall\, v \in X \ ; \tag{5.6}$$

here $u_j(\mu), j = 1, \ldots, \mathcal{N}$, is the nodal value of $u(\mu)$ at node x_i and is complex.

5.1.3 Galerkin projection

We next introduce our Galerkin projection on the discrete space X to find the approximation $u(\mu) \in X$ to $u^e(\mu) \in X^e$. Given $\mu \in \mathcal{D} \in \mathbb{R}^P$, solve for $u(\mu) \in X$ from

$$a(u(\mu), v; \mu) = f(v), \quad \forall v \in X, \tag{5.7}$$

and then calculate the output $s(\mu)$ as

$$s(\mu) = f(u(\mu)). \tag{5.8}$$

We assume that the triangulation $\mathcal{T}_h$ represents the domain exactly and the approximation subspace X is chosen such that Galerkin projection directly inherits its properties from the exact formulation. In particular, the dual norm of f over $X \subset X^e$ is bounded by the dual norm of f over X^e; the Galerkin recipe preserves symmetry; a is continuous over X; and the affine expansions for

f and a remain valid for w, v restricted to X.

5.1.4 Discrete Equations

We next substitute the approximation $u(\mu) = \sum_{j=1}^{\mathcal{N}} u_j(\mu)\varphi_j$ into (5.7) and take v as the basis functions $\varphi_i, i = 1, \ldots, \mathcal{N}$, to obtain the desired linear system

$$\sum_{j=1}^{\mathcal{N}} a\left(\varphi_j, \varphi_i; \mu\right) u_j(\mu) = f\left(\varphi_i\right), \quad i = 1, \ldots, \mathcal{N} , \tag{5.9}$$

which can be written into matrix form

$$\underline{A}(\mu)\, \underline{u}(\mu) = \underline{F} . \tag{5.10}$$

Here $\underline{A}(\mu)$ is an $\mathcal{N} \times \mathcal{N}$ matrix with $A_{ij}(\mu) = a\left(\varphi_j, \varphi_i; \mu\right)$, $\underline{F}$ is an vector with $F_i = f\left(\varphi_i\right)$, and $\underline{u}(\mu)$ is an vector with $u_i(\mu) = u(x_i; \mu)$, where x_i is the coordinates of the node i. The matrix $\underline{A}$ and vector $\underline{F}$ depend on the finite element mesh and type of basis functions. They can be formed via assembling elemental matrices and vectors associated with each elements T_h of $\mathcal{T}_h$.

By solving the linear system, we obtain the nodal values $\underline{u}(\mu)$ and thus $u(\mu) = \sum_{i=1}^{\mathcal{N}} u_i(\mu)\varphi_i$. Finally, the output approximation $s(\mu)$ can be calculated as

$$s(\mu) = \ell(u(\mu)) . \tag{5.11}$$

A complete discussion and detailed implementation of the finite element procedure can be found in most finite element method textbooks (see, for example, [16]). The finite element treatment of Helmholtz problems in exterior domains can be found in [47]; special treatment for high wavenumbers can be found in [8, 54, 55].

5.1.5 *A Priori* Convergence

The finite element method seeks the approximate solution $u(\mu)$ (respectively, the approximate output $s(\mu)$) in the finite element "truth" approximation space X to the exact solution $u^e(\mu)$ (respectively, the exact output $s^e(\mu)$) of the underlying PDE. The *a priori* convergence analysis for the finite element approximation suggests that $\|u^e(\mu) - u(\mu)\|_X$ and $|s^e(\mu) - s(\mu)|$ will converge as

h^α and h^β, respectively; here α and β are positive constants whose value depend on the specific problem, the output functional, and the regularity of force functional and domain.

In general, we have $u(\mu) \to u^e(\mu)$ and $s(\mu) \to s^e(\mu)$, as $h \to 0$. For a particular case in which a is symmetric positive-definite, Ω and $f, (\ell = f)$ are sufficiently regular, $\|u^e(\mu) - u(\mu)\|_X$ and $|s^e(\mu) - s(\mu)|$ will vanish as $O(h)$ and $O(h^2)$, respectively, for $\mathbb{P}_1$ elements.

Increasing the required accuracy for a problem implies higher $\mathcal{N}$ to obtain accurate and reliable results; adequately converged truth approximations are thus achieved only for spaces X of very large dimension $\mathcal{N}$. For many medium or large-scale applications, $\mathcal{N}$ is typically in the order of $O(10^4)$ up to $O(10^5)$. Thus, the computational time for a particular "input" can be unacceptably large, rendering "real-time" and "many-query" contexts impossible.

5.1.6 Choice of $\mathcal{N}$

The reduced basis approach is built on the "truth" Galerkin solution of the μ-parametrized PDE. We measure the reduced basis approximation error relative to the "truth" solution. It is thus important that $\mathcal{N}$ is chosen large enough to satisfy achievable engineering accuracies. Furthermore, $\mathcal{N}$ should be chosen so that it can represent the "worst-case"/most complex solution over all $\mu \in \mathcal{D}$ accurately.

5.2 Radiation Boundary Condition (RBC)

In Chapter 3, we discussed the radiation boundary condition in detail. The radiation boundary condition is a Robin boundary condition specified at outflow, Γ_{out} that allows us to model semi-infinite waveguides with a truncated domain. The The domain truncation if done properly can reduce the size of the mesh $\mathcal{N}$ significantly. The domain truncation is important, since we will require adequate mesh resolution (a fine mesh) to ensure accuracy of the finite element solution for the chosen range of wavenumber; in the absence of the radiation boundary condition we would have conceivably required a longer domain length and therefore larger $\mathcal{N}$.

The development of the radiation boundary condition in Chapter 3 is based on domain truncation using the modal expansion on the boundary [21, 40]. We use the same modal expansion to obtain the eigen problem allowing easy calculation of the eigenvectors and associated eigenvalues for the different modes. Although, the eigen problem is dependent on the size of the truth mesh, the

reduced basis *offline-online* decomposition allows us to restrict the expensive truth-mesh dependent calculations to the *offline* stage.

We explore the effect of specifying the radiation boundary condition and truncating the domain on (a) the accuracy of the solution, and (b) the choice of the truncated domain length in more detail.

Consider the homogeneous acoustic waveguide element coupled with the Helmholtz resonator of fixed geometry shown in Figure 5-1. The governing (non-dimensional) equation is the acoustic Helmholtz equation $\nabla^2 p + k^2 p = 0$ for the (non-dimensional) pressure p; here k is the non-dimensional wavenumber. We impose homogeneous Neumann conditions $\frac{\partial p}{\partial n} = 0$ at Γ_{top}, Γ_{bot} and on the walls of the Helmholtz resonator Γ_{hr}. At the inflow, Γ_{in}, we have an inhomogeneous Neumann (velocity) condition $\frac{\partial p}{\partial n} = \mathrm{i}\, k$.

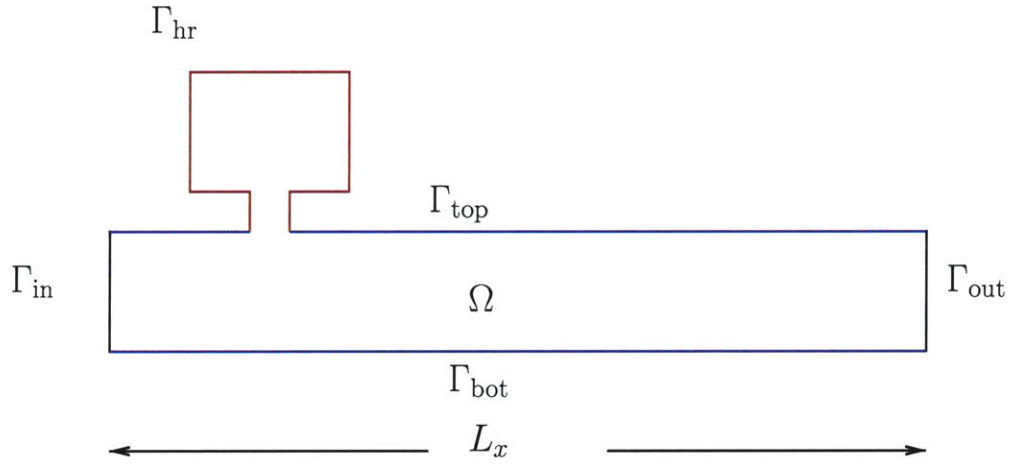


Figure 5-1: Homogeneous waveguide of length L_x; the waveguide is of unit depth.

5.2.1 Accuracy of RBC

We explore the correctness of the radiation boundary condition at outflow. We consider different choices for the outflow boundary condition and compare the different numerical solutions to demonstrate that the radiation boundary condition is indeed appropriate.

We fix the non-dimensional length of the waveguide $L_x = 6$ and consider three different choices for the boundary condition at outflow: (a) homogeneous Neumann $\frac{\partial p}{\partial n} = 0$; (b) the Sommerfeld radiation condition $\frac{\partial p}{\partial n} = -\mathrm{i}kp$; and (c) the radiation outflow condition $\frac{\partial p}{\partial n} = \sum_{j=1}^{N_{\text{prop}}} -\mathrm{i}\alpha_j (\int_{\Gamma_{\text{out}}} p\overline{\Xi}_j)\Xi_j$.

The Sommerfeld boundary condition is often used to truncate the domain when solving the exterior Helmholtz problem that represents the basic model for farfield propagation of waves in the frequency domain. The Sommerfeld condition [28]

$$\lim_{\tilde{r}\to\infty} \tilde{r}^{(n-1)/2} \left(\frac{\partial \tilde{u}}{\partial \tilde{r}} + \mathrm{i}k\tilde{u} \right) = 0, \tag{5.12}$$

can be seen as an expression of causality for the wave equation in that there can be no incoming waves at infinity. Thus, prescribing the Sommerfeld condition is akin to using a filter that only allows outgoing waves. The Sommerfeld sign convention depends on the sign convention used for the time-harmonic factor in the wave equation. Mathematically, the Sommerfeld condition ensures the well-posedness of the system; it approximates the replacement of an indefinite domain with an artifical closed boundary, denoted here by Γ_{out}. For our problem, we choose the first order approximation of the Sommerfeld radiation condition $\frac{\partial p}{\partial n} = -\mathrm{i}\, k\, p$; higher order approximation requires more sophisticated implementation. Other approximate radiation conditions have been developed, including the most accurate Bayliss-Turkel radiation condition [6, 17].

We choose to compare these three different formulations to a *reference* formulation. The reference formulation is of much longer extent, $L_x = 12$, and has the radiation boundary condition specified at outflow. We impose a constant velocity at the inlet

$$\frac{\partial p}{\partial n}|_{\Gamma_{\text{in}}} = \mathrm{i}\, k; \tag{5.13}$$

here n is the outward normal to Γ_{in}.

The homogeneous Neumann boundary conditions on Γ_{top} and Γ_{bot} help us identify $\Xi_j(y) = \cos(\lambda_j\, y)$, where $\lambda_j = \frac{j\,\pi}{L_y}, j = 0, 1, \ldots$. The two propagating modes have x-directional propagating coefficients $\alpha_1 = \sqrt{11}$ and $\alpha_2 = \sqrt{11 - \pi^2}$ respectively. Figure 5-2 compares the obtained solutions with the "reference" solution on $y = 0$ (the boundary Γ_{bot}) for the choice of $k^2 = 11$ when we have $N_{\text{prop}} = 2$. Note that the obtained solution for $L_x = 6$ (the curve (c) in Figure 5-2) lies over the "reference" solution (the curve (d) in Figure 5-2) in the range $0 \leq x \leq 6$. The homogeneous Neumann is clearly inappropriate because the propagating modes have variations in the pressure and are not constant. Also, we notice that the Sommerfeld radiation condition approximates the boundary condition at outflow quite well; had we chosen a longer domain length the Sommerfeld

condition would have been quite accurate.

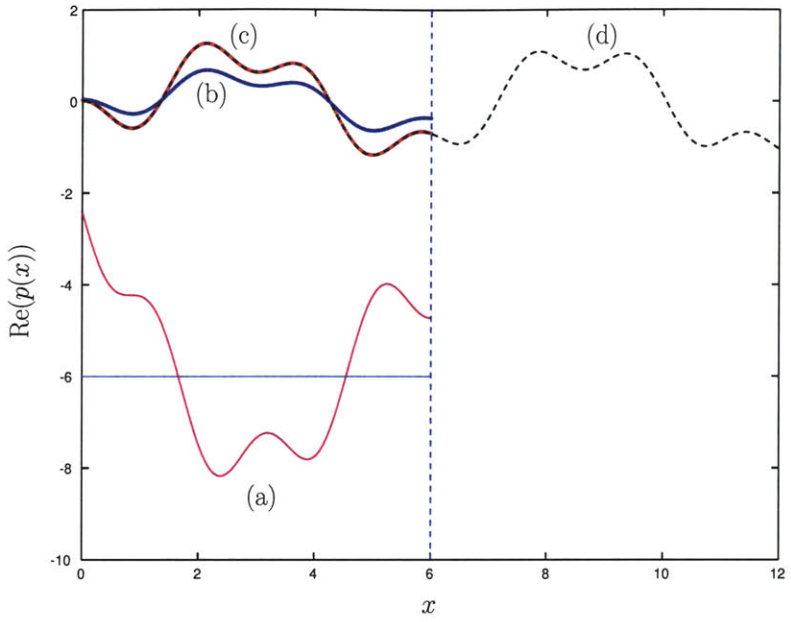


Figure 5-2: $\mathrm{Re}(p(x)))$ vs x at Γ_{bot} $(y = 0)$ for $k^2 = 11$ and $N_{\mathrm{prop}} = 2$. We show solutions for the four cases: (a) Homogeneous Neumann, (b) Sommerfeld, (c) Radiation B.C and (d) "Reference" Solution on waveguide of length $L_x = 12$ with radiation b.c prescribed at outflow.

5.2.2 Choice of Truncated Domain Length

We also need to specify an acceptable length for the truncated domain when we apply the radiation boundary condition at the outflow. The radiation boundary condition

$$\frac{\partial p}{\partial n} = \sum_{j=1}^{N_{\mathrm{prop}}} -\mathrm{i}\,\alpha_j \left(\int_{\Gamma_{\mathrm{out}}} p\,\overline{\Xi}_j \right) \Xi_j \tag{5.14}$$

propagates out *only* the N_{prop} propagating modes (of the form $e^{-\mathrm{i}\sqrt{|\gamma_j|}\,x}\,\Xi_j(y)$). But there is also an implicit assumption that the evanescent modes (of the form $e^{-:\sqrt{|\gamma_j|}\,x}\,\Xi_j(y)$) have all decayed and do *not* exist at the outflow.

Thus, the accuracy of the solution in the domain is affected by the length of the domain. If the chosen length is too small, the evanescent modes *assumed* to have decayed would still exist and the outflow boundary condition would be wrong — our "truth" solution will be unacceptably

inaccurate. On the other hand, if the chosen length is too long, our *offline* computational times related to the approximation of the reduced basis space and the calculation of the inf-sup lower bounds would be adversely affected.

Thus it is important to choose the length of the truncated domain accurately. We would like to choose the *minimum* length of domain required to ensure that the evanescent modes have decayed out.

Figure 5-3 shows the solution for $y = 0$ (i.e., on the boundary Γ_{bot}) for $k^2 = 11$ when we have 2 propagating modes ($N_{\text{prop}} = 2$) for different choices of L_x; we specify a constant velocity at the inlet (5.13) and impose the radiation boundary condition at outflow in all cases. It is clear that the radiation boundary condition works for relatively small L_x, in our case we can see that $L_x = 3$ is sufficient for our purposes.

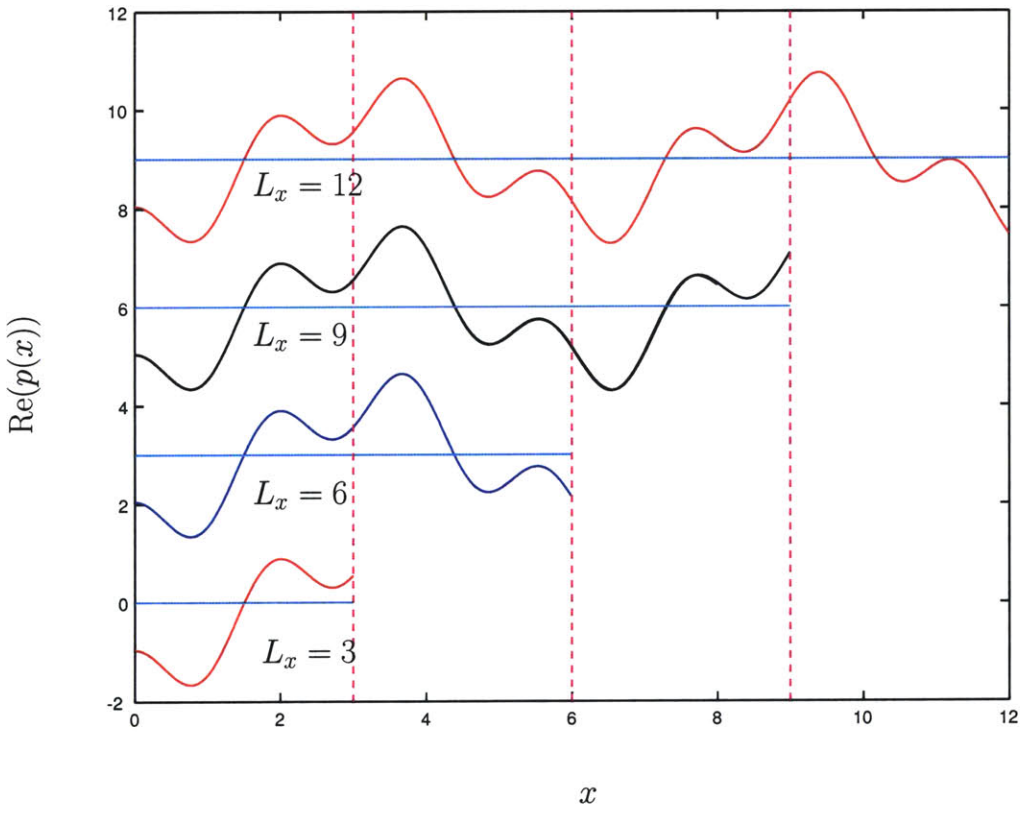


Figure 5-3: $\text{Re}(p(x))$ vs x on Γ_{bot} ($y = 0$) for different choice of L_x for $k^2 = 11$ and $N_{\text{prop}} = 2$. We use the radiation b.c at outflow in all the three cases.

However, it is not guaranteed that $L_x = 3$ will be sufficient for any choice of inflow boundary condition. The choice of truncated domain length can be incorrect in two ways: (i) if we have an evanescent mode that is prescribed in the inlet boundary condition and L_x is not sufficiently long to eliminate the effect of that mode; and (ii) an evanescent mode is generated because of the complex coupling of the waveguide with other acoustic elements (in this example, the Helmholtz resonator) and the the truncated length is not sufficiently long to let the mode decay. We illustrate the point by imposing the following boundary condition at inflow

$$\frac{\partial p}{\partial n}|_{\Gamma_{\text{in}}} = \frac{1}{4}\,\mathrm{i}\,|\alpha_1|\,\Xi_1(y) + 2\,\mathrm{i}\,|\alpha_2|\,\Xi_2(y). \tag{5.15}$$

The boundary condition described in (5.15) represents the imposition of a specific pressure at the inlet. We can derive this condition starting from the general modal expansion of the pressure

$$p = \sum_{j=1}^{\infty} c_n\, e^{-\mathrm{i}\,\alpha_n\, x}\,\Xi_n(y); \tag{5.16}$$

taking derivatives w.r.t x we obtain

$$\frac{\partial p}{\partial x} = \sum_{j=1}^{\infty} -\mathrm{i}\,\alpha_n\, c_n\, e^{-\mathrm{i}\,\alpha_n\, x}\,\Xi_n(y). \tag{5.17}$$

We obtain (5.15) by setting $x = 0$, $c_1 = \frac{1}{4}$, $c_2 = 2$, $c_j = 0, j = 3, \ldots$, and recognizing that the outward normal n points in the $-x$ direction on Γ_{in}.

We choose the wavenumber $k = 3$ for this problem. The propagation coefficient associated with the first mode, $\alpha_1 = 3.0$; this first mode is propagating and is of the form $e^{-\mathrm{i}\,\alpha_1\, x}$. We can similarly calculate the propagation coefficient associated with the second mode; we obtain $\alpha_2 = \sqrt{k^2 - \pi^2} \approx -0.93\,\mathrm{i}$ — this is the first evanescent mode and is of the form $e^{-|\alpha_2|\, x}$. We can see that this mode is going to evanesce fairly slowly: at $L_x = 3$, $e^{-|\alpha_2|\, L_x} = 0.06$; doubling the length of the truncated domain to $L_x = 6$, drops down the magnitude of the evanescent mode by one order, $e^{-|\alpha_2|\, 6} = 0.004$. We thus have $N_{\text{prop}} = 1$ propagating mode for this problem. Figure 5-4 shows the solution along Γ_{bot} for different choices of L_x. Note that a truncated domain length of $L_x = 1.5$ is clearly unacceptable; with increasing length L_x the solutions get more and more similar. Thus, the solution on a truncated domain length of $L_x = 3$ differs much more from the solution on the

truncated length $L_x = 4.5$ relative to the solution for $L_x = 6$.

This example demonstrates that one has to be careful about the choice of L_x; furthermore, the inflow boundary condition and the complex coupling of the acoustic waveguide with the other acoustic elements will determine which "evanescent" modes might be excited — while they will decay it is important to ensure that the truncated domain length is sufficiently long to eliminate all evanescent modes.

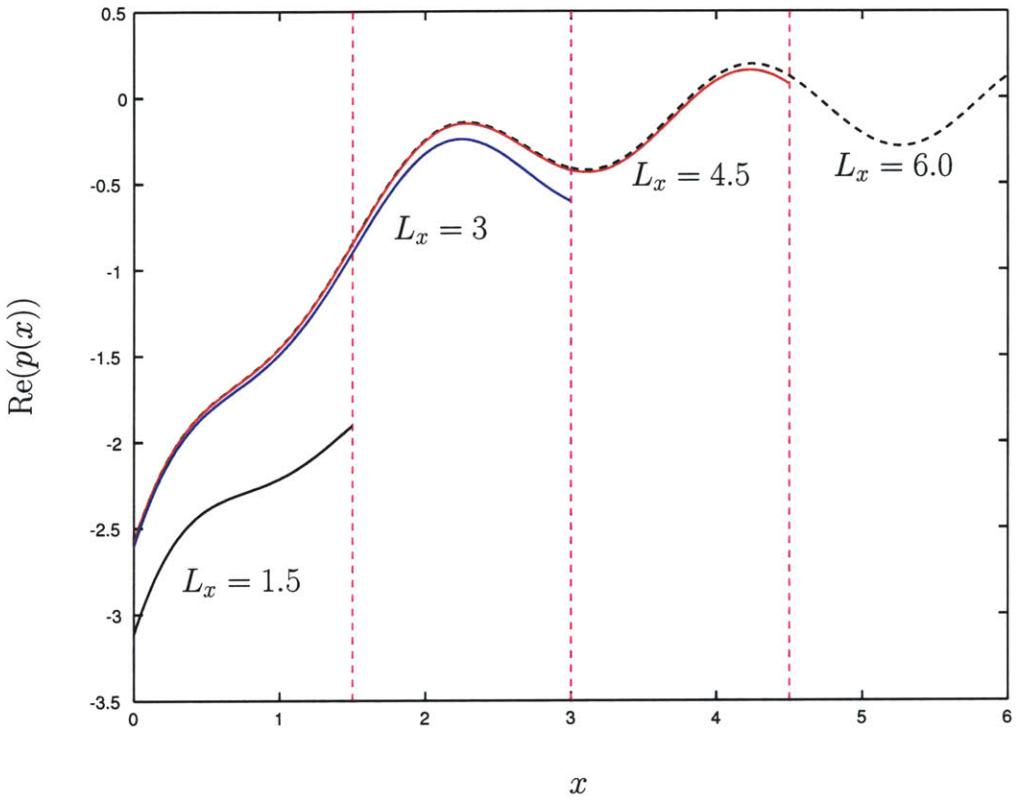


Figure 5-4: $\mathrm{Re}(p(x))$ vs x on Γ_{bot} $(y = 0)$ for different choice of L_x for $k^2 = 9$ and $N_{\mathrm{prop}} = 1$. We use the radiation b.c at outflow in all the three cases. The inflow boundary condition is given by (5.15).

5.2.3 Choice of N_{prop}

It is also very important to ensure that we choose N_{prop} correctly. For the sake of clarity, we introduce two new variables: $N_{\mathrm{prop}}^{\mathrm{actual}}$ and $N_{\mathrm{prop}}^{\mathrm{rbc}}$. $N_{\mathrm{prop}}^{\mathrm{actual}}$ denotes the actual (correct) number of propagating modes that will be present at outflow assuming we chose our L_x judiciously to

eliminate all evanescent modes. $N_{\text{prop}}^{\text{rbc}}$ denotes the number of propagating modes used in the radiation boundary condition.

We compare the effect of $N_{\text{prop}}^{\text{rbc}}$ on the power $\mathcal{P}$ at the inflow and outflow boundaries. In the absence of any dissipative forces in the system (i.e, damping, friction),the total power at the inflow boundary, $\mathcal{P}_{\text{in}}$, must equal the total power at the outflow boundary, $\mathcal{P}_{\text{out}}$, because there have been no losses. The expressions[1] for $\mathcal{P}_{\text{in}}$ and $\mathcal{P}_{\text{out}}$ are given by

$$\mathcal{P}_{\text{in}} = \frac{1}{2}\int_{\Gamma_{\text{in}}} \text{Re}(p), \tag{5.18}$$

$$\mathcal{P}_{\text{out}} = \frac{1}{2\,k}\sum_{n=1}^{N_{\text{prop}}^{\text{rbc}}} \alpha_n \,(\int_{\Gamma_{\text{out}}} p\,\overline{\Xi}_n)^2. \tag{5.19}$$

Note however that this relationship must hold true for *any* choice of $N_{\text{prop}}^{\text{rbc}}$ for the radiation boundary condition — however the obtained solutions and the power will be wrong if $N_{\text{prop}}^{\text{rbc}} \neq N_{\text{prop}}^{\text{actual}}$. We show in Figure 5-5 the solutions on Γ_{bot} obtained by specifying $N_{\text{prop}}^{\text{rbc}}$ propagating modes for the radiation boundary condition. We confirm that the solutions are different when we use $N_{\text{prop}}^{\text{rbc}} < N_{\text{prop}}^{\text{actual}}$.

The actual power at the outflow boundary, $\mathcal{P}_{\text{out}}$, will be distributed amongst the $N_{\text{prop}}^{\text{actual}}$ propagating modes. We denote by $\mathcal{P}_{\text{out}}^{\text{actual}}$ the output power that is obtained when we use $N_{\text{prop}}^{\text{actual}}$ number of propagating modes in *both* the radiation boundary condition and the expression for the output power:

$$\mathcal{P}_{\text{out}}^{\text{actual}} = \frac{1}{2\,k}\sum_{n=1}^{N_{\text{prop}}^{\text{actual}}} \alpha_n \,(\int_{\Gamma_{\text{out}}} p\,\overline{\Xi}_n)^2. \tag{5.20}$$

We denote by $\mathcal{P}_{\text{out}}^{\text{actual,rbc}}$ the output power obtained by choosing $N_{\text{prop}}^{\text{actual}}$ propagating modes in the radiation boundary condition and $N_{\text{prop}}^{\text{rbc}}$ modes in the calculation of the output power.

We calculate $\mathcal{P}_{\text{out}}^{\text{actual,rbc}}$ for different choices of $N_{\text{prop}}^{\text{rbc}}$ and compare them to the actual output power $\mathcal{P}_{\text{out}}^{\text{actual}}$. Table 5.1 shows $|\mathcal{P}_{\text{out}}^{\text{actual}} - \mathcal{P}_{\text{out}}^{\text{actual,rbc}}|$ for different choices of $N_{\text{prop}}^{\text{rbc}}$ and for $k^2 = 144$. We expect $N_{\text{prop}}^{\text{actual}} = 4$ propagating modes: we confirm $\mathcal{P}_{\text{out}}^{\text{actual,rbc}} = \mathcal{P}_{\text{out}}^{\text{actual}}$ for $N_{\text{prop}}^{\text{actual}} = N_{\text{prop}}^{\text{rbc}} = 4$. When $N_{\text{prop}}^{\text{rbc}} < N_{\text{prop}}^{\text{actual}}$, $\mathcal{P}_{\text{out}}^{\text{actual,rbc}} < \mathcal{P}_{\text{out}}^{\text{actual}}$; adding more than $N_{\text{prop}}^{\text{actual}}$ modes is not a problem because the evanescent modes decay and do not contribute to the output power as Table 5.1 clearly

[1] For a proof of the relationship $\mathcal{P}_{\text{in}} = \mathcal{P}_{\text{out}}$ and the derivation of the forms $\mathcal{P}_{\text{in}}$ and $\mathcal{P}_{\text{out}}$ for a non-dissipative acoustic waveguide system please see Appendix A.

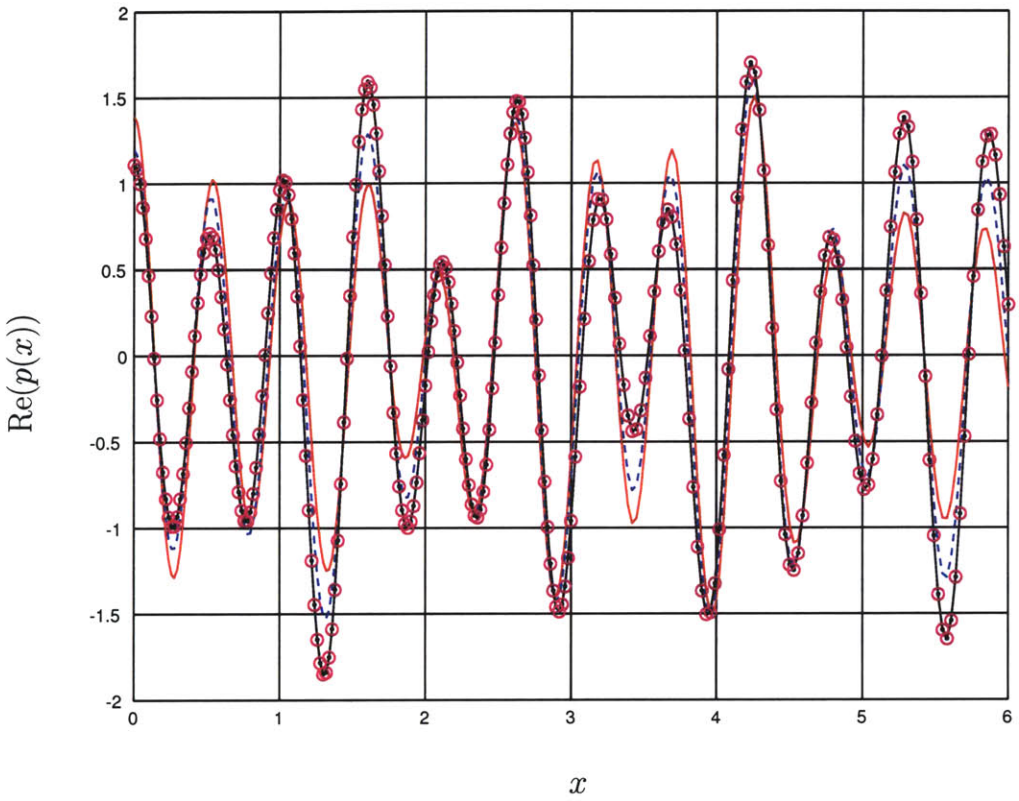


Figure 5-5: $\mathrm{Re}(p(x))$ vs x on Γ_{bot} $(y = 0)$ for different choice of N_{prop} for $k^2 = 9$ and $L_x = 6$. We specify $N_{\mathrm{prop}}^{\mathrm{rbc}} = 1 - -4$ propagating modes in the radiation boundary condition at outflow. We show the solutions on Γ_{bot} for $N_{\mathrm{prop}}^{\mathrm{rbc}} = 4$ (in red), $N_{\mathrm{prop}}^{\mathrm{rbc}} = 3$ (in blue), $N_{\mathrm{prop}}^{\mathrm{rbc}} = 2$ (in black), and $N_{\mathrm{prop}}^{\mathrm{rbc}} = 1$ (in magenta). Notice that the solutions progressively deviate from the correct solution (for $N_{\mathrm{prop}}^{\mathrm{actual}} = 4$) as we decrease the number of propagating modes used at outflow. We impose a constant velocity boundary condition at inflow.

$N_{\text{prop}}^{\text{rbc}}$	$\lvert \mathcal{P}_{\text{out}}^{\text{actual}} - \mathcal{P}_{\text{out}}^{\text{actual,rbc}} \rvert$
1	$4.25E-02$
2	$3.67E-02$
3	$1.7E-02$
≥ 4	$0.0E00$

Table 5.1: Effect of choice of $N_{\text{prop}}^{\text{rbc}}$ on the power for an acoustic waveguide system. Parameters are the non-dimensional wave-number square $k^2 = 144$, length of domain $L_x = 6$, and $N_{\text{prop}}^{\text{actual}} = 4$. We vary $N_{\text{prop}}^{\text{rbc}}$ and demonstrate that excluding propagating modes gives us wrong results for $\mathcal{P}_{\text{out}}^{\text{actual,rbc}}$ and $\mathcal{P}_{\text{out}}^{\text{actual,rbc}} < \mathcal{P}_{\text{out}}$. However, adding more than N_{prop} modes at the outflow does not change the output power since the evanescent modes have all decayed for the choice of $L_x = 6$.

demonstrates.

5.3 Expanded "Truth" Formulation

The starting point of the reduced basis approximation is the "truth" formulation. For our problems, we do not have closed-form analytical solutions — instead we compute a highly accurate "truth" numerical approximation (in our case, we use FEM) as a stand-in for the actual or exact solution. We use this "reference" numerical approximation both to construct our low-dimensional reduced basis space; we also measure the accuracy of the RB predictions against this truth approximation.

5.3.1 "Exact" Problem Statement

Consider the abstract form of the "exact" (superscript e) problem posed in the weak form: for any $\mu \in \mathcal{D} \subset \mathbb{R}^P$, calculate outputs

$$s_1^e(\mu) = \ell(u^e(\mu)), \tag{5.21}$$

and/or

$$s_2^e(\mu) = Q(u^e(\mu), u^e(\mu)), \tag{5.22}$$

where $u^e(\mu)$ satisfies the weak form of the μ-parametrized PDE

$$a(u^e(\mu), v; \mu) = f(v), \quad \forall\, v \in X^e. \tag{5.23}$$

Here μ and $\mathcal{D} \in \mathbb{R}^P$ are the input and (closed) input domain, respectively. Depending on the choice of problem, our "output" of interest can be linear ($s_1^e(\mu)$) or quadratic ($s_2^e(\mu)$). Here $u^e(x;\mu)$ is our

field variable and X^e is a complex vector space.

5.3.2 Expanded "Truth" Statement

The expanded truth statement is the starting point for the reduced basis approximation. In Chapter 3, we discussed some of the benefits of the expanded formulation in the reduced basis context — the expanded formulation enables the treatment of quadratic outputs as linear compliant outputs. By basing the reduced basis approximation upon the expanded weak form, we eliminate the need for the dual (adjoint) problem that would be required if we had chosen the primal-dual formulation, and we also obtain better *a posteriori* error estimators.

We now present the expanded finite element "truth" formulation: given $\mu \in \mathcal{D} \in \mathbb{R}^P$, calculate the "truth" output of interest

$$\mathcal{S}(\mu) = \overline{\mathcal{F}(\mathcal{U}(\mu); \mu)} \tag{5.24}$$

where $\mathcal{U}(\mu) = [U^+, U^-]^T \in \mathcal{X}$ satisfies

$$\mathcal{A}(\mathcal{U}(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}; \mu), \ \forall \mathcal{V} \in \mathcal{X}, \tag{5.25}$$

and $\mathcal{V} = [V^+, V^-]^T \in \mathcal{X}$, here $\mathcal{X} \in X^2 \subset (X^e)^2$.

$X \subset X^e$ is a complex "truth" finite element approximation space of dimension $\mathcal{N}$

$$X = \{v = v_R + \mathrm{i}\, v_I \mid v_R \in X_R,\ v_I \in X_R\}, \tag{5.26}$$

$$X_R = \{v \in (H^1(\Omega))^d \mid v = 0 \text{ on } \Gamma_D\}; \tag{5.27}$$

and associated inner product and norm

$$(w, v)_X = \int_\Omega \nabla w \, \nabla \overline{v} + w \, \overline{v}, \tag{5.28}$$

$$\|w\|_X = (w, w)_X. \tag{5.29}$$

.

The "expanded" complex function space $\mathcal{X} \in X^2$ is associated with an inner product

$$(\mathcal{W}, \mathcal{V})_{\mathcal{X}} = (W^+, V^+)_X + (W^-, V^-)_X \tag{5.30}$$

and norm

$$\|\mathcal{W}\|_{\mathcal{X}} = (\mathcal{W}, \mathcal{W})_{\mathcal{X}}. \tag{5.31}$$

Here $\mathcal{W} = [W^+, W^-]^T \in \mathcal{X}$, $\mathcal{V} = [V^+, V^-]^T$ and $W^+, W^-, V^+, V^- \in X$. Note that the dimension of $\mathcal{X}$, $|\mathcal{X}|$, is $2\mathcal{N}$.

We identify the "big" operator $\mathcal{A}$ as

$$\begin{aligned}
\mathcal{A}(\mathcal{U}(\mu), \mathcal{V}; \mu) \quad &= a(U^+, V^+; \mu) \; + \; \overline{a(V^+, U^+; \mu)} \\
&+ a(U^-, V^+; \mu) \; - \; \overline{a(V^+, U^-; \mu)} \\
&+ \overline{a(V^-, U^+; \mu)} \; - \; a(U^+, V^-; \mu) \\
&- \overline{a(V^-, U^-; \mu)} \; - \; a(U^-, V^-; \mu) \\
&- Q(U^+, V^+; \mu) \; - \; Q(U^-, V^+; \mu) \\
&- Q(U^+, V^-; \mu) \; - \; Q(U^-, V^-; \mu),
\end{aligned} \tag{5.32}$$

and $\mathcal{F}$ as

$$\mathcal{F}(\mathcal{V}; \mu) = f(V^+; \mu) - f(V^-; \mu) + \frac{1}{2} l(V^+; \mu) + \frac{1}{2} l(V^-; \mu), \tag{5.33}$$

where $\mathcal{U}(\mu) = [U^+, U^-]^T$, $\mathcal{V} = [V^+, V^-]^T$. Note that $\mathcal{A}(\mathcal{W}, \mathcal{V}; \mu) = \overline{\mathcal{A}(\mathcal{V}, \mathcal{W}; \mu)}$ — $\mathcal{A}$ is symmetric.

Purely Quadratic Output

If our output is purely quadratic (i.e. we do not have any linear output $\ell(v)$), we write $\mathcal{F}$ as

$$\mathcal{F}(\mathcal{V}; \mu) = f(V^+; \mu) - f(V^-; \mu), \tag{5.34}$$

while $\mathcal{A}$ (5.32) remains unchanged.

Linear *Non-Compliant* Output

If the problem has a linear *non-compliant* output where $\ell(v) \neq \overline{f(\overline{v})}$, we re-write $\mathcal{A}$ as

$$\begin{aligned}\mathcal{A}(\mathcal{U}(\mu),\mathcal{V};\mu) &= a(U^+,V^+;\mu) + \overline{a(V^+,U^+;\mu)} \\ &+ a(U^-,V^+;\mu) - \overline{a(V^+,U^-;\mu)} \\ &+ \overline{a(V^-,U^+;\mu)} - a(U^+,V^-;\mu) \\ &- \overline{a(V^-,U^-;\mu)} - a(U^-,V^-;\mu),\end{aligned} \tag{5.35}$$

while $\mathcal{F}$ (5.33) remains unchanged. This is identical to having a seperate adjoint (dual) problem and associated spaces — instead of writing the two systems (i) primal: $a(u,v;\mu) = f(v)$ and (ii) dual: $\overline{a(v,\psi;\mu)} = -\ell(v)$ seperately we have simply formulated it as one system with double the number of unknowns.

Linear *Compliant* Output

In the special case, where we have a *compliant* output where $\ell(v) = \overline{f(\overline{v})}$, we have no need for the expanded system and we return to the non-expanded system of equations with $\mathcal{N}$ unknowns

$$\begin{aligned}a(u,v;\mu) &= f(v), \forall v \in X \\ \ell(v) &= \overline{f(\overline{v})} \\ s(\mu) = \ell(\overline{u}) &= \overline{f(u)} = \overline{a(u,u;\mu)} = a(u,u;\mu).\end{aligned} \tag{5.36}$$

Chapter 6

Reduced Basis Approximation

6.1 Introduction

In this chapter, we discuss some of the key ingredients of the reduced basis output bound methods: dimension reduction effected by rapidly convergent global approximation spaces, offline-online decomposition enabled by affine parameter dependence and optimal sampling strategies enabling formation of efficient reduced basis spaces. In subsequent chapters, we shall develop the *a posteriori* error estimation for non-coercive elliptic problems more fully.

We start by stating the abstract problem for non-coercive elliptic partial differential equations with quadratic outputs.

6.2 Abstraction

6.2.1 "Exact" Problem Statement

We consider a suitably regular (smooth) domain $\Omega \subset \mathbb{R}^d, d = 1, 2,$ or 3[1] with Lipschitz-continuous boundary Γ. Our "exact" problem: for any $\mu \in \mathcal{D} \subset \mathbb{R}^P$, find

$$\mathcal{S}^e(\mu) = \overline{\mathcal{F}(\mathcal{U}^e(\mu))}, \tag{6.1}$$

[1] Note that Ω is a *reference domain* and hence does *not* depend on the parameter.

where $\mathcal{U}^e(\mu)$ satisfies the weak form of the μ-parametrized PDE

$$\mathcal{A}(\mathcal{U}^e(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}), \quad \forall\, \mathcal{V} \in \mathcal{X}^e. \tag{6.2}$$

Here μ and $\mathcal{D}$ are the input and (closed) input domain, respectively; $\mathcal{S}^e(\mu)$ is the "output of interest"; $\mathcal{U}^e(x;\mu)$ is our field variable; $\mathcal{X}^e$ is a complex Hilbert space defined over the physical domain $\Omega \subset \mathbb{R}^d$ with inner product $(\mathcal{W}, \mathcal{V})_{\mathcal{X}^e}$ and associated norm $\|\mathcal{W}\|_{\mathcal{X}^e} = \sqrt{(\mathcal{W}, \mathcal{W})_{\mathcal{X}^e}}$; and $\mathcal{A}(\cdot, \cdot; \mu)$ and $\mathcal{F}(\cdot)$ are $\mathcal{X}^e$-continuous bilinear and bounded linear functionals, respectively. Our function space $\mathcal{X}^e$ will thus satisfy $(H_0^1(\Omega))^\nu \subset \mathrm{Re}(\mathcal{X}^e), \mathrm{Im}(\mathcal{X}^e) \subset (H^1(\Omega))^\nu$, where $\nu = 1$ for a scalar field variable and $\nu = d$ for a vector field variable. Recall that $H^1(\Omega)$ (respectively, $H_0^1(\Omega)$) is the usual Hilbert space (respectively, the Hilbert space of functions that vanish on the domain boundary Γ).

6.2.2 "Truth" Finite Element Approximation

In actual practice, we replace $\mathcal{X}^e$ with $\mathcal{X} \subset \mathcal{X}^e$, a "truth" finite element approximation space of dimension $\mathcal{N}$. The inner product and norm associated with $\mathcal{X}$ are given by $(\cdot, \cdot)_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{X}} = (\cdot, \cdot)_{\mathcal{X}}^{1/2}$, respectively. A typical choice for $(\cdot, \cdot)_{\mathcal{X}}$ is

$$(\mathcal{W}, \mathcal{V})_{\mathcal{X}} = \int_\Omega \nabla \mathcal{W} \cdot \nabla \overline{\mathcal{V}} + \mathcal{W}\overline{\mathcal{V}}\ , \tag{6.3}$$

which is simply the standard $H^1(\Omega)$ inner product. We shall next denote by $\mathcal{X}'$ the dual space of $\mathcal{X}$. For a $h \in \mathcal{X}'$, the dual norm is given by

$$\|h\|_{\mathcal{X}'} \equiv \sup_{\mathcal{V} \in \mathcal{X}} \frac{h(\mathcal{V})}{\|\mathcal{V}\|_{\mathcal{X}}}\ . \tag{6.4}$$

Our "truth" finite element approximation $\mathcal{U}(\mu) \in \mathcal{X}$ to $\mathcal{U}^e(\mu)$ is then defined as the Galerkin projection of $\mathcal{U}^e(\mu)$ onto $\mathcal{X}$. The finite element approximation of the continuous problem can then be stated as: given $\mu \in \mathcal{D} \in \mathbb{R}^P$, find

$$\mathcal{S}(\mu) = \overline{\mathcal{F}(\mathcal{U}(\mu))} \tag{6.5}$$

where $\mathcal{U}(\mu) \in \mathcal{X}$ is the solution of the discretized weak form

$$\mathcal{A}(\mathcal{U}(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}), \quad \forall \mathcal{V} \in \mathcal{X} \ . \tag{6.6}$$

We shall assume — hence the appellation "truth" — that $\mathcal{X}$ is sufficiently rich that $\mathcal{U}$ (respectively, $\mathcal{S}$) is sufficiently close to $\mathcal{U}^e(\mu)$ (respectively, $\mathcal{S}^e(\mu)$) for all μ in the (closed) parameter domain $\mathcal{D}$. We must be certain that our formulation are *stable* and *efficient* as $\mathcal{N} \to \infty$. The reduced basis approximation will be built upon our "truth" finite element approximation, and the reduced basis error will thus be evaluated with respect to $\mathcal{U}(\mu) \in \mathcal{X}$ and $\mathcal{S}(\mu)$.

We now make several assumptions on the well-posedness and the nature of the parametric dependence of our problem.

6.2.3 Well-posedness

We shall assume that the bilinear form $\mathcal{A}$ is symmetric, $\mathcal{A}(\mathcal{W}, \mathcal{V}; \mu) = \overline{\mathcal{A}(\mathcal{V}, \mathcal{W}; \mu)}, \forall \mathcal{W}, \mathcal{V} \in \mathcal{X}, \forall \mu \in \mathcal{D}$ and non-coercive. The classical inf-sup and continuity parameters can then be introduced as

$$\beta(\mu) \equiv \inf_{\mathcal{W} \in \mathcal{X}} \sup_{\mathcal{V} \in \mathcal{X}} \frac{| \mathcal{A}(\mathcal{W}, \mathcal{V}; \mu) |}{\|\mathcal{W}\|_{\mathcal{X}} \|\mathcal{V}\|_{\mathcal{X}}}, \quad \forall \mu \in \mathcal{D} \tag{6.7}$$

and

$$\gamma(\mu) \equiv \sup_{\mathcal{W} \in \mathcal{X}} \sup_{\mathcal{V} \in \mathcal{X}} \frac{| \mathcal{A}(\mathcal{W}, \mathcal{V}; \mu) |}{\|\mathcal{W}\|_{\mathcal{X}} \|\mathcal{V}\|_{\mathcal{X}}}, \quad \forall \mu \in \mathcal{D}. \tag{6.8}$$

$\beta(\mu)$, the inf-sup stability constant, is the minimum (generalized) singular value associated with our differential operator; $\gamma(\mu)$ is the standard continuity constant; note that both these "constants" are functions of the parameter μ. We also assume that the linear functional $\mathcal{F} \in \mathcal{X}'$ is bounded.

We now suppose that $0 < \beta_0 \leq \beta(\mu)$ and $0 < \gamma(\mu) \leq \gamma_0 < \infty$, $\forall \mu \in \mathcal{D}$. It then follows that our problem is well-posed for all $\mu \in \mathcal{D}$.

We next introduce the parametrized linear supremizing operator $T^\mu : \mathcal{X} \to \mathcal{X}$ associated with our bilinear form a,

$$T^\mu \mathcal{W} = \arg \sup_{\mathcal{V} \in \mathcal{X}} \frac{a(\mathcal{W}, \mathcal{V}; \mu)}{\|\mathcal{X}\|_{\mathcal{X}}}. \tag{6.9}$$

We can explicitly represent T^μ : for any $\mathcal{W} \in \mathcal{X}$,

$$(T^\mu \mathcal{W}, \mathcal{V})_\mathcal{X} = \mathcal{A}(\mathcal{W}, \mathcal{V}; \mu), \quad \forall \mathcal{V} \in \mathcal{X}; \tag{6.10}$$

we note that T^μ is linear. The definition of $T^\mu \mathcal{W}$ in the complex case is identical to the real case — we do not require a complex modulus in the expression on the left hand side in (6.10). We can then rewrite out classical inf-sup and continuity parameters as

$$\beta(\mu) = \inf_{\mathcal{W} \in \mathcal{X}} \frac{\|T^\mu \mathcal{W}\|_\mathcal{X}}{\|\mathcal{W}\|_\mathcal{X}}, \quad \forall \mu \in \mathcal{D}, \tag{6.11}$$

and

$$\gamma(\mu) = \sup_{\mathcal{W} \in \mathcal{X}} \frac{\|T^\mu \mathcal{W}\|_\mathcal{X}}{\|\mathcal{W}\|_\mathcal{X}} \quad \forall \mu \in \mathcal{D}. \tag{6.12}$$

6.2.4 Affine Parameter Dependence

We also make certain assumptions on the nature of the parametric dependence for our problem. In particular, we shall suppose that, for some finite (preferably small) integer Q, $\mathcal{A}(\cdot, \cdot; \mu)$ may be expressed as an affine decomposition of the form

$$\mathcal{A}(\mathcal{W}, \mathcal{V}; \mu) = \sum_{q=1}^{Q} \Theta_\mathcal{A}^q(\mu) \mathcal{A}^q(\mathcal{W}, \mathcal{V}), \tag{6.13}$$

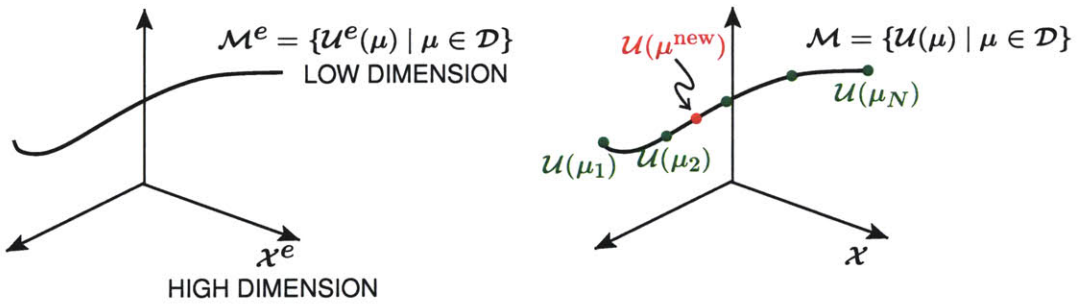
where for $1 \leq q \leq Q$, $\Theta_\mathcal{A}^q : \mathcal{D} \to \mathbb{C}$ are differentiable parameter-dependent coefficient functions and bilinear forms $\mathcal{A}^q : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ are parameter-independent. This assumption of affine parameter dependence is crucial for the computational efficiency of our method. To make matters simple we assume that the linear form $\mathcal{F}$ does not depend on the parameter; however affine (and even non-affine) parameter dependence can be treated similarly (see for example [108]).[2]

[2]Note that the assumption of affine parameter dependence can be relaxed; see [14, 109] for extensions to problems exhibiting non-affine parameter dependence or nonlinearities.

6.3 Reduced Basis Approximation

6.3.1 Manifold of Solutions

The reduced basis method recognizes that the field variable $\mathcal{U}^e(\mu)$ is not an arbitrary member of the infinite-dimensional solution space $\mathcal{X}^e$ associated with the underlying partial differential equation. All the possible values of $\mathcal{U}^e(\mu)$ do not "cover" the entire space, $\mathcal{X}^e$; the approximation space $\mathcal{X}^e$ is much too general and can approximate many functions that are not μ-dependent. Infact, $\mathcal{U}^e(\mu)$ resides on a very low-dimensional manifold $\mathcal{M}^e \equiv \{\mathcal{U}^e(\mu) \,|\, \mu \in \mathcal{D}\}$ induced by the parametric dependence. For example, for a single parameter, $\mu \in \mathcal{D} \subset \mathbb{R}^{P=1}$, $\mathcal{U}^e(\mu)$ will describe a one-dimensional filament that winds through $\mathcal{X}^e$ as depicted in Figure 6-1a. The manifold containing all possible solutions of the parametrized partial differential equation is much smaller than the function space $\mathcal{X}^e$.



(a) (b)

Figure 6-1: (a) Low-dimensional manifold in which the field variable resides; and (b) approximation of the solution at μ^{new} by a linear combination of pre-computed solutions $\mathcal{U}(\mu_i)$.

Even with the finite "truth" approximation, $\mathcal{X}$ is much too general — to approximate $\mathcal{U}(\mu)$, we need not represent every single function in $\mathcal{X}$, but rather only those which lie on $\mathcal{M}$. This critical observation presents a clear opportunity: we can effect significant, in many cases Draconian, dimension reduction in state space by restricting our attention to $\mathcal{M}$; $\mathcal{U}(\mu)$ can then be adequately approximated by a space of dimension $N \ll \mathcal{N}$.

Since all solutions of the parametrized partial differential equation live on the low-dimensional manifold $\mathcal{M}$, we need to construct an approximation space for the manifold. We choose to construct this approximation space by including the solutions $\mathcal{U}(\mu)$ at N selected points along the manifold

$\mathcal{M}$ as shown in Figure 6-1(b). We then approximate the solution, $\mathcal{U}(\mu^{\text{new}})$, at a newly chosen point μ^{new} by a projection onto this approximation space composed of solutions at pre-computed sample points $\mathcal{U}(\mu_n), n = 1, \ldots, N$. We now make these ideas more precise.

6.3.2 Reduced Basis Spaces

We now construct our reduced basis spaces W_N. Recall however that our solution for the "expanded" system consists of two sub-functions: $\mathcal{U}(\mu) = [U^+(\mu), U^-(\mu)]^T, \forall \mu \in \mathcal{D}$. We choose to construct our space in a way which lets us seperate the two sub-functions: given $\mu \in \mathcal{D}$, this is equivalent to introducing two basis functions $[\mathcal{U}^+(\mu), \mathbf{0}]^T$ and $[\mathbf{0}, \mathcal{U}^-(\mu)]^T$. We hope this approach will represent the sub-functions of $\mathcal{U}(\mu)$ more efficiently.

We now make these ideas more precise. We introduce the sample set in parameter space as

$$\begin{aligned} S_N &= \{\mu_1^+ \in \mathcal{D}, \ldots, \mu_{N/2}^+ \in \mathcal{D}, \mu_1^- \in \mathcal{D}, \ldots, \mu_{N/2}^- \in \mathcal{D}\}, \\ &= \left\{S_{N/2}^+ \equiv \{\mu_1^+ \in \mathcal{D}, \ldots, \mu_{N/2}^+ \in \mathcal{D}\}, S_{N/2}^- \equiv \{\mu_1^- \in \mathcal{D}, \ldots, \mu_{N/2}^- \in \mathcal{D}\}\right\} \end{aligned} \tag{6.14}$$

for $1 \leq N \leq N_{\max}$. Here $\mu_1^+ = \mu_1^- = \mu_1^\pm \in \mathcal{D}, \ldots, \mu_{N/2}^+ = \mu_{N/2}^- = \mu_{N/2}^\pm \in \mathcal{D}$ ($S_{N/2}^+ = S_{N/2}^- = S_{N/2}^\pm$) and specify the associated Lagrangian [81] reduced basis approximation space

$$\begin{aligned} W_N &= \text{span}\{[U^+(\mu_n^\pm), \mathbf{0}]^T, [\mathbf{0}, U^-(\mu_n^\pm)]^T, 1 \leq n \leq N/2\}, \quad 1 \leq N \leq N_{\max}, \\ &= \text{span}\{\zeta_n^+ = [U^+(\mu_n^\pm), \mathbf{0}]^T, \zeta_n^- = [\mathbf{0}, U^-(\mu_n^\pm)]^T, 1 \leq n \leq N/2\}, \quad 1 \leq N \leq N_{\max}, \\ &= \text{span}\{\zeta_n \equiv \{\zeta_n^+, \zeta_n^-\}, 1 \leq n \leq N\}, \quad 1 \leq N \leq N_{\max}, \end{aligned} \tag{6.15}$$

where $\mathcal{U}(\mu_n^\pm) = [U^+(\mu_n^\pm), U^-(\mu_n^\pm)]^T$ is the solution of (6.6) for $\mu = \mu_n^\pm$. We orthogonalize the basis with respect to the inner product (6.3). Note that, by construction, $W_1 \subset W_N \subset W_{N_{\max}}$.

We now present our reduced basis approximation: for any $\mu \in \mathcal{D}$, find

$$\mathcal{S}_N(\mu) = \overline{\mathcal{F}(\mathcal{U}_N(\mu))} \tag{6.16}$$

where $\mathcal{U}_N(\mu)$ is the Galerkin projection of $\mathcal{U}(\mu)$ onto W_N.

$$\mathcal{A}(\mathcal{U}_N(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}), \quad \forall\, \mathcal{V} \in W_N. \tag{6.17}$$

Orthogonalized Basis

In forming the reduced basis space W_N, the basis functions ζ_n must be selected such that they are linearly independent to make the algebraic system associated with (6.16) as well-conditioned as possible. Although the basis functions are the solutions of the parametrized partial differential equations for different μ, they are far from linearly independent. Consequently, the associated algebraic system will be very ill-conditioned for large N. To this end, using Gram-Scmidt orthogonalization we orthogonalize our basis with respect to the inner product associated with our space $\mathcal{X}$, $(\cdot,\cdot)_{\mathcal{X}}$ and thus obtain $(\zeta_i,\zeta_j)_{\mathcal{X}} = \delta_{ij}$, $1 \le i,j \le N$. Given $\mu_n^{\pm}$, we sketch the steps involved in incorporating $\mathcal{U}(\mu_n^{\pm}) = [\mathcal{U}^+(\mu_n^{\pm}, \mathcal{U}^-(\mu_n^{\pm})]^T$ into our basis. We assume that before the augmentation W_N comprises of $N = 2(n-1)$ basis functions. We first form $\zeta_n^+ = [\mathcal{U}^+(\mu_n^{\pm}), \mathbf{0}]^T$ and $\zeta_n^- = [\mathbf{0}, U^-(\mu_n^{\pm})]^T$ and collect them into $Z_{\text{new}} = [\zeta_n^+, \zeta_n^-]$. We then do Gram-Schmidt orthogonalization as indicated below

$$
\begin{aligned}
&\text{for } j = 1:2 \\
&\quad r = Z_{\text{new}}(j) \\
&\quad \text{for } k = 1:2\,(n-1) \\
&\qquad c_{jk} = (r, \zeta_k)_{\mathcal{X}} \\
&\qquad r = r - c_j k\, \zeta_k \\
&\quad \text{end for} \\
&\quad \zeta_{2\,(n-1)+j} = r/(r,r)_{\mathcal{X}} \\
&\quad N = N + 1 \\
&\text{end for}
\end{aligned}
\tag{6.18}
$$

Since $\mathcal{M}$ is low-dimensional and smooth, we thus anticipate that $\mathcal{U}_N(\mu) \to \mathcal{U}(\mu)$ very rapidly, and therefore we may choose $N \ll \mathcal{N}$. We now attempt to qualify our claim.

6.3.3 *A Priori* Convergence Theory

We consider here the convergence rate of $\mathcal{U}_N(\mu)$ and $\mathcal{S}_N(\mu)$ to $\mathcal{U}(\mu)$ and $\mathcal{S}(\mu)$, respectively. We first introduce the operator $T_N^{\mu} : W_N \to W_N$ such that, for any $\hat{\mathcal{W}}_N \in W_N$,

$$
(T_N^{\mu}\, \hat{\mathcal{W}}_N, \hat{\mathcal{V}}_N; \mu)_{\mathcal{X}} = \mathcal{A}(\hat{\mathcal{W}}_N, \hat{\mathcal{V}}_N; \mu), \quad \forall \hat{\mathcal{V}}_N \in W_N. \tag{6.19}
$$

We next define $\beta_N(\mu) \in \mathbb{R}$ as

$$\beta_N(\mu) \equiv \inf_{\hat{\mathcal{W}}_N \in W_N} \sup_{\hat{\mathcal{V}}_N \in W_N} \frac{|\mathcal{A}(\hat{\mathcal{W}}_N, \hat{\mathcal{V}}_N; \mu)|}{\|\hat{\mathcal{W}}_N\|_{\mathcal{X}} \|\hat{\mathcal{V}}_N\|_{\mathcal{X}}} = \inf_{\hat{\mathcal{W}}_N \in W_N} \frac{\|T_N^\mu \hat{\mathcal{W}}_N\|_{\mathcal{X}}}{\|\hat{\mathcal{W}}_N\|_{\mathcal{X}}}. \tag{6.20}$$

We shall now show that the reduced basis approximation $\mathcal{U}_N(\mu)$ is optimal in $\mathcal{X}$-norm,

$$\|\mathcal{U}(\mu) - \mathcal{U}_N(\mu)\|_{\mathcal{X}} \leq \left(1 + \frac{\gamma_0}{\beta_0}\right) \min_{\hat{\mathcal{W}}_N \in W_N} \|\mathcal{U}(\mu) - \hat{\mathcal{W}}_N(\mu)\|_{\mathcal{X}} \tag{6.21}$$

with the assumption that $\beta_N(\mu) \geq \beta_0 > 0$, and $\gamma(\mu) \leq \gamma_0 < \infty$, $\forall \mu \in \mathcal{D}$.

Proof. We first note from (6.6) and (6.17) that

$$\mathcal{A}(\mathcal{U}(\mu) - \mathcal{U}_N(\mu), \hat{\mathcal{V}}_N; \mu) = 0, \quad \forall \hat{\mathcal{V}}_N \in W_N. \tag{6.22}$$

It is follows from (6.20) that,

$$\beta_N(\mu) \, \|\hat{\mathcal{W}}_N\|_{\mathcal{X}} \, \|T_N^\mu \hat{\mathcal{W}}_N\|_{\mathcal{X}} \leq |\mathcal{A}(\hat{\mathcal{W}}_N, T_N^\mu \hat{\mathcal{W}}_N; \mu)|, \quad \forall \hat{\mathcal{W}}_N \in W_N. \tag{6.23}$$

It thus follows for any $\hat{\mathcal{W}}_N \in W_N$ that

$$\beta_N(\mu) \, \|\hat{\mathcal{W}}_N - \mathcal{U}_N\|_{\mathcal{X}} \, \|T_N^\mu (\hat{\mathcal{W}}_N - \mathcal{U}_N)\|_{\mathcal{X}} \leq |\mathcal{A}(\hat{\mathcal{W}}_N - \mathcal{U}_N, T_N^\mu (\hat{\mathcal{W}}_N - \mathcal{U}_N); \mu)|. \tag{6.24}$$

We write $\mathcal{U}_N$ as $\mathcal{U}_N = \mathcal{U} - (\mathcal{U} - \mathcal{U}_N)$ in the right hand expression to obtain

$$\begin{aligned} \mathcal{A}(\hat{\mathcal{W}}_N - \mathcal{U} + \mathcal{U} - \mathcal{U}_N, T_N^\mu(\hat{\mathcal{W}}_N - \mathcal{U}_N); \mu) &= \mathcal{A}(\hat{\mathcal{W}}_N - \mathcal{U}, T_N^\mu (\hat{\mathcal{W}}_N - \mathcal{U}_N); \mu) \\ &\quad + \mathcal{A}(\mathcal{U} - \mathcal{U}_N, T_N^\mu (\hat{\mathcal{W}}_N - \mathcal{U}_N); \mu). \end{aligned} \tag{6.25}$$

Since $\mathcal{A}(\mathcal{U} - \mathcal{U}_N, \hat{\mathcal{V}}_N) = 0, \forall \hat{\mathcal{V}}_N \in W_N$, we set $\hat{\mathcal{V}}_N = T_N^\mu (\hat{\mathcal{W}}_N - \mathcal{U}_N)$ to obtain

$$\mathcal{A}(\mathcal{U} - \mathcal{U}_N, T_N^\mu (\hat{\mathcal{W}}_N - \mathcal{U}_N); \mu) = 0. \tag{6.26}$$

We then combine (6.24), (6.25) and (6.26) to obtain

$$\begin{aligned}\beta_N(\mu)\,\|\hat{\mathcal{W}}_N-\mathcal{U}_N\|_{\mathcal{X}}\,\|T_N^{\mu}\,(\hat{\mathcal{W}}_N-\mathcal{U}_N)\|_{\mathcal{X}} &\leq |\mathcal{A}(\hat{\mathcal{W}}_N-\mathcal{U}_N,T_N^{\mu}\,(\hat{\mathcal{W}}_N-\mathcal{U}_N);\mu)|\\ &\leq |\mathcal{A}(\hat{\mathcal{W}}_N-\mathcal{U},T_N^{\mu}\,(\hat{\mathcal{W}}_N-\mathcal{U}_N);\mu)|\\ &\leq \gamma(\mu)\,\|\hat{\mathcal{W}}_N-\mathcal{U}\|_{\mathcal{X}}\,\|T_N^{\mu}\,(\hat{\mathcal{W}}_N-\mathcal{U}_N)\|_{\mathcal{X}}. \quad (6.27)\end{aligned}$$

We therefore obtain

$$\|\hat{\mathcal{W}}_N-\mathcal{U}_N\|_{\mathcal{X}}\leq\frac{\gamma(\mu)}{\beta_N(\mu)}\|\hat{\mathcal{W}}_N-\mathcal{U}\|_{\mathcal{X}}. \quad (6.28)$$

Finally, from triangle inequality, it follows that

$$\begin{aligned}\|\mathcal{U}(\mu)-\mathcal{U}_N(\mu)\|_{\mathcal{X}} &\leq \|\mathcal{U}(\mu)-\hat{\mathcal{W}}_N(\mu)\|_{\mathcal{X}}+\|\hat{\mathcal{W}}_N(\mu)-\mathcal{U}_N(\mu)\|_{\mathcal{X}}\\ &\leq \left(1+\frac{\gamma_0}{\beta_0}\right)\min_{\hat{\mathcal{W}}_N\in W_N}\|\mathcal{U}(\mu)-\hat{\mathcal{W}}_N(\mu)\|_{\mathcal{X}} \quad (6.29)\end{aligned}$$

which concludes our proof. □

We next show that, for the case where our output of interest $\mathcal{S}(\mu)$ is compliant, $\mathcal{S}_N(\mu)$ converges optimally to $\mathcal{S}(\mu)$ in $\mathcal{X}$-norm,

$$\mathcal{S}(\mu)-\mathcal{S}_N(\mu)\leq\gamma_0\left(1+\frac{\gamma_0}{\beta_0}\right)\min_{\hat{\mathcal{W}}_N\in W_N}\|\mathcal{U}-\hat{\mathcal{W}}_N\|_{\mathcal{X}}^2, \quad (6.30)$$

where $\gamma(\mu)\leq\gamma_0<\infty,\forall\mu\in\mathcal{D}$.

Proof. It is simply to show that,

$$\begin{aligned}\mathcal{S}(\mu)-\mathcal{S}_N(\mu) &= \overline{\mathcal{F}(\mathcal{U}-\mathcal{U}_N;\mu)}\\ &= \overline{\mathcal{A}(\mathcal{U},\mathcal{U}-\mathcal{U}_N;\mu)}\\ &= \mathcal{A}(\mathcal{U}-\mathcal{U}_N,\mathcal{U}-\hat{\mathcal{W}}_N+\hat{\mathcal{W}}_N;\mu)\\ &= \mathcal{A}(\mathcal{U}-\mathcal{U}_N,\mathcal{U}-\hat{\mathcal{W}}_N;\mu)\\ &\leq \gamma(\mu)\,\|\mathcal{U}-\mathcal{U}_N\|_{\mathcal{X}}\|\mathcal{U}-\hat{\mathcal{W}}_N\|_{\mathcal{X}}\\ &\leq \gamma_0\left(1+\frac{\gamma_0}{\beta_0}\right)\min_{\hat{\mathcal{W}}_N\in W_N}\|\mathcal{U}-\hat{\mathcal{W}}_N\|_{\mathcal{X}}^2, \quad (6.31)\end{aligned}$$

from the symmetry of $\mathcal{A}$, Galerkin orthogonality (6.22), continuity (6.12) and optimality of $\mathcal{U}_N(\mu)$ (6.21). □

The results (6.30) show that the output approximation, $\mathcal{S}_N(\mu)$, converges to $\mathcal{S}(\mu)$ as the square of the error in $\mathcal{U}_N(\mu)$.

6.3.4 Offline-Online Computational Procedure

Even though N is small, the elements of $\mathcal{W}_N$ are $\mathcal{N}$-dependent "truth" finite element basis functions $\zeta_n \equiv \{\zeta_n^+, \zeta_n^-\}, n = 1, \ldots, N$. To eliminate the $\mathcal{N}$ dependency we resort to offline-online computational procedures. To begin, we expand our reduced basis approximation as

$$\mathcal{U}_N(\mu) = \sum_{j=1}^{N} \mathcal{U}_{N\,j}(\mu)\, \zeta_j \ . \tag{6.32}$$

We then choose as test functions in (6.17) $v = \zeta_i$, $i = 1, \ldots, N$; it then follows from (6.17) that $\underline{\mathcal{U}}_N(\mu) \equiv [\mathcal{U}_{N\,1}(\mu), \ldots, \mathcal{U}_{N\,N}(\mu)]^T \in \mathbb{R}^N$ satisfies the $N \times N$ linear algebraic system

$$\mathcal{A}_N(\mu)\, \underline{\mathcal{U}}_N(\mu) = \mathcal{F}_N \tag{6.33}$$

where $\mathcal{A}_N(\mu) \in \mathbb{C}^{N\times N}$ and $F_N \in \mathbb{C}^N$ are given by $\mathcal{A}_{N\,i,j}(\mu) = \mathcal{A}(\zeta_j, \zeta_i; \mu)$, $1 \le i, j \le N$, and $\mathcal{F}_{N\,i} = \mathcal{F}(\zeta_i)$, $1 \le i \le N$, respectively.

We then invoke the affine decomposition (6.13) to obtain

$$\mathcal{A}_{N\,i,j}(\mu) = \sum_{q=1}^{Q_\mathcal{A}} \Theta_\mathcal{A}^q(\mu)\, \mathcal{A}^q(\zeta_i, \zeta_j); \tag{6.34}$$

which we can rewrite as

$$\mathcal{A}_N(\mu) = \sum_{q=1}^{Q_\mathcal{A}} \Theta_\mathcal{A}^q(\mu)\, \mathcal{A}_N^q. \tag{6.35}$$

Here the *parameter independent* quantities matrices $\mathcal{A}_N^q \in \mathbb{C}^{N\times N}$ are given by

$$\mathcal{A}_{N\,i,j}^q = \mathcal{A}^q(\zeta_i, \zeta_j), \quad 1 \le i, j \le N,\ 1 \le q \le Q_\mathcal{A}. \tag{6.36}$$

The reduced basis output can then be calculated as

$$\begin{aligned} \mathcal{S}_N(\mu) &= \overline{\mathcal{F}(\mathcal{U}_N(\mu))} \\ &= \sum_{j=1}^{N} \overline{\mathcal{U}_{N\,j}(\mu)}\,\overline{\mathcal{F}(\zeta_j)} = \sum_{j=1}^{N} \overline{\mathcal{U}_{N\,j}(\mu)}\,\overline{\mathcal{F}_{N,j}}. \end{aligned} \tag{6.37}$$

It is clear from (6.36) and (6.37) that we may pursue an offline-online computational strategy to economize the output evaluation.

In the *offline* stage — performed *once* — we first solve for the ζ_i, $1 \le i \le N_{\max}$: this requires $N_{\max}$ finite element solves; we then form *and store* $\mathcal{F}(\zeta_i)$, $1 \le i \le N_{\max}$, and $\mathcal{A}^q(\zeta_j, \zeta_i)$, $1 \le i, j \le N_{\max}$, $1 \le q \le Q_{\mathcal{A}}$: this requires $O(Q_{\mathcal{A}}\, N_{\max}^2\, \mathcal{N})$ operations and $O(Q_{\mathcal{A}}\, N_{\max}^2)$ storage.

In the *online stage* — performed *many times*, for each new value of μ — we first assemble and subsequently invert the (full) $N \times N$ "stiffness matrix" $\mathcal{A}_N(\mu)$ in (6.35) to solve for the reduced basis coefficients $\underline{u}_{N\,j}(\mu)$, $j \in \mathbb{N}$; we then evaluate the output approximation $\mathcal{S}_N(\mu)$ from (6.37). The operation count for the online stage is respectively $O(Q_{\mathcal{A}} N^2)$ to assemble (recall the $a^q(\zeta_j, \zeta_i)$, $1 \le i, j \le N$, $1 \le q \le Q_{\mathcal{A}}$, are *pre-stored*), $O(N^3)$ to invert the stiffness matrix, and $O(N)$ to evaluate the output inner product (recall the $\mathcal{F}(\zeta_j)$ are *pre-stored*).

The essential point is that the online complexity is dependent on $Q_{\mathcal{A}}$ and N but *independent of* $\mathcal{N}$, the dimension of the underlying truth finite element approximation space. Since $N \ll \mathcal{N}$, we expect — and often realize — significant, orders-of-magnitude computational economies relative to classical discretization approaches.

6.4 Sampling Strategy for the Construction of the Reduced Basis Space

We now discuss the construction of the nested sample set S_N and the associated reduced basis space W_N. For a given N, we need to find the optimal sample S_N that maximizes our reduced basis accuracy. We denote the maximum number of basis obtained from our sampling algorithm for some desired accuracy as $N_{\max}$. The key ingredient here is the rigorous, sharp and *inexpensive*

output bound $\Delta_N^{\mathcal{S}}(\mu)$ such that

$$|\mathcal{S}(\mu) - \mathcal{S}_N(\mu)| \leq \Delta_N^{\mathcal{S}}(\mu), \tag{6.38}$$

for all $\mu \in \mathcal{D}$ and for all N. Since we use $\Delta_N^{\mathcal{S}}(\mu)$ *in lieu* of the actual error $|\mathcal{S}(\mu) - \mathcal{S}_N(\mu)|$, we require the effectivity $\eta_N^{\mathcal{S}}(\mu)$ defined as

$$\eta_N^{\mathcal{S}}(\mu) \equiv \frac{\Delta_N^{\mathcal{S}}(\mu)}{|\mathcal{S}(\mu) - \mathcal{S}_N(\mu)|}, \tag{6.39}$$

to be as small as possible. *Rigor* dictates that $\eta_N^{\mathcal{S}}(\mu) \geq 1$ for all $\mu \in \mathcal{D}$ and for all N. Furthermore, we expect that we can compute $\Delta_N^{\mathcal{S}}(\mu)$ extremely fast. We discuss the construction of this error estimate in detail in Chapter 8.

We shall pursue a greedy algorithm [73, 110] to find the set S_N. We shall denote the particular "optimal" (nested) samples and (hierarchical) spaces selected by our greedy algorithm as

$$S_N^* \equiv \{\mu_1^*, \ldots, \mu_{N/2}^*\} \tag{6.40}$$

and

$$\mathcal{X}_N^*(\equiv W_N) = \text{span}\{[U^+(\mu_1^*), \mathbf{0}]^T, [\mathbf{0}, U^-(\mu_1^*)]^T \ldots, [U^+(\mu_{N/2}^*), \mathbf{0}]^T, [\mathbf{0}, U^-(\mu_{N/2}^*)]^T\}, \tag{6.41}$$

for $1 \leq N \leq N_{\max}$; here $\mathcal{U}(\mu_n^*) = [U^+(\mu_n^*), U^-(\mu_n^*)]^T, n = 1, \ldots, N/2$.

6.4.1 Standard Greedy Algorithm

We specify our training sample $\Xi_{\text{train}} \subset \mathcal{D}$ (a very fine sample over the parameter domain) and termination error tolerance $\epsilon_{\text{tol,min}}$. Choose $\mu_1^* \in \Xi_{\text{train}}$ at random as the first μ-sample to be added into the basis. Denote $S_1^* = \{\mu_1^*\}$ and $\mathcal{X}_1^* = \text{span}\{\zeta_1 \equiv [U^+(\mu_1^*), \mathbf{0}], [\mathbf{0}, U^-(\mu_1^*)]\}$ as the associated reduced basis space.

We introduce $\forall \mu \in \mathcal{D}$, the relative error bound $\epsilon_N^*(\mu)$ as the *rigorous* bound and "surrogate" for the actual relative error

$$\sqrt{\frac{|\mathcal{S}(\mu) - \mathcal{S}_N(\mu)|}{\mathcal{S}(\mu)}} \leq \sqrt{\frac{\Delta_N^{\mathcal{S}}(\mu)}{\mathcal{S}_N(\mu) - \Delta_N^s(\mu)}} \equiv \epsilon_N^*(\mu), \ \forall N. \tag{6.42}$$

The algorithm proceeds as follows (we provisionally set $N_{\max} = \overline{N}_{\max}$)

```
for N = 2: N̄_max
    μ*_N = arg max_{μ∈Ξ_train} ε*_{N-1}(μ);
    ε*_N = ε*_{N-1}(μ*_N);
    if ε*_N ≤ ε_{tol,min}
        N_max = N − 1;
        exit;
    end;
    S*_N = S*_{N-1} ∪ μ*_N;
    X*_N = X*_{N-1} + span{U⁺(μ*_N), 0], [0, U⁻(μ*_N)]};
end.
```

The most crucial point of this strategy is that the error estimator $\Delta_N^{\mathcal{S}}(\mu)$ can be computed "online-inexpensively" and is representative of the true error $|\mathcal{S}(\mu) - \mathcal{S}_N(\mu)|$. This permits us perform a very exhaustive ($n_{\text{train}} = |\Xi_{\text{train}}| \gg 1$) search for the best sample S_N and, hence, determine the sub-optimal number of basis N for which we achieve the desired accuracy. The greedy operation count is ($N_{\max}$ finite element solves + $Q\ N_{\max}$ Poisson-like problem solves + $Q^2\ N_{\max}^2\ (\cdot,\cdot)_{\mathcal{X}}$ inner-products + $n_{\text{train}}\ O(Q^2\ N_{\max}^2 + N_{\max}^3)$ reduced basis solutions).

This surrogate also permits us in the online stage, to determine the smallest N, and hence the most efficient approximation, for which we *rigorously* achieve the desired accuracy. This adaptive offline sampling procedure can thus be viewed as a (greedy, parameter space, "$L^\infty(\mathcal{D})$") variant of the POD economization procedure [98] in which — thanks to $\Delta_N^{\mathcal{S}}(\mu)$ — *we need never construct* the "rejected" snapshots.

6.5 Comparison of Primal-Dual and Adjoint Formulations for the Treatment of Quadratic Outputs

Finally, we comment on the some of the differences between the reduced basis formulation described in Section 6.3.2 and the typical treatment of quadratic outputs. We re-introduce the original

problem for the sake of convenience. For any $\mu \in \mathcal{D}$, we evaluate

$$s(\mu) \equiv Q(u(\mu), u(\mu), \mu) \tag{6.43}$$

where $u(\mu) \in X$ satisfies

$$a(u(\mu), v; \mu) = f(v), \ \forall v \in X. \tag{6.44}$$

We will assume that the bilinear form $a(\cdot, \cdot; \mu)$ has an affine seperation of the form $a(\cdot, \cdot; \mu) = \sum_{q=1}^{Q_a} \Theta_a^q(\mu)\, a^q(\cdot, \cdot)$.

We first present the primal-dual formulation and then comment on the differences between the reduced basis approximation of the expanded system and the primal-dual formulation for quadratic outputs.

6.5.1 Primal-Dual Formulation

Given $\mu \in \mathcal{D}$, our primal approximation $u_N(\mu) \in W_N^{\text{pr}}(\mu)$ satisfying

$$a(u_N(\mu), v; \mu) = f(v), \ \forall v \in W_N^{\text{pr}}; \tag{6.45}$$

we also introduce the primal residual, $r_N^{\text{pr}}(\cdot; \mu)$, associated with the primal approximation

$$r_N^{\text{pr}}(v; \mu) = f(v) - a(u_N(\mu), v; \mu), \ \forall v \in X. \tag{6.46}$$

The primal-dual formulation introduces a dual (adjoint) problem of the form

$$\overline{a(v, \psi(\mu); \mu)} = Q(u_N(\mu) + u(\mu), v; \mu), \ \forall v \in X. \tag{6.47}$$

Corresponding to the adjoint problem, we introduce the dual approximation $\psi_N(\mu) \in W_N^{\text{du}}(\mu)$ satisfying

$$\overline{a(v, \psi_N(\mu); \mu)} = Q(2\, u_N(\mu), v), \ \forall v \in W_N^{\text{du}}; \tag{6.48}$$

and associated dual residual, $r_N^{\text{du}}(\cdot; \mu)$ as

$$r_N^{\text{du}}(v; \mu) = Q(2\, u_N(\mu), v; \mu) - \overline{a(v, \psi_N(\mu); \mu)}, \ \forall v \in X; \tag{6.49}$$

note the dependence of the dual solution $\psi_N(\mu)$ on the primal approximation $u_N(\mu)$.

We introduce the reduced basis output $s_N(\mu)$ [53] as

$$s_N(\mu) = Q(u_N(\mu), u_N(\mu)) + r_N^{\rm pr}(\psi_N(\mu); \mu). \tag{6.50}$$

Subsequently, we bound the error in $|s(\mu) - s_N(\mu)|$ by the *a posteriori* error estimator $\Delta_N^s(\mu)$ [53, 96] given by

$$\Delta_N^s(\mu) = \frac{1}{\beta_{\rm LB}(\mu)} \|r_N^{\rm pr}(\mu)\|_{X'} \, \|r_N^{\rm du}\|_{X'} + \frac{\gamma_{\rm UB}^Q(\mu)}{\beta_{\rm LB}(\mu)} \|r_N^{\rm pr}(\mu)\|_{X'}^2 \tag{6.51}$$

where $\|r_N^{\rm pr}(\mu)\|_{X'}$ and $\|r_N^{\rm du}(\mu)\|_{X'}$ are the dual norm of the primal and dual residuals, $\beta_{\rm LB}(\mu)$ is a rigorous lower bound for the inf-sup stability constant, and $\gamma_{\rm UB}^Q(\mu)$ is the maximum eigenvalue of the generalized eigenproblem [53]

$$\gamma_{\rm UB}^Q(\mu) = \sup_{v \in X} \frac{Q(v, v; \mu)}{(v, v)_X}. \tag{6.52}$$

We discuss the construction of $\beta_{\rm LB}(\mu)$ in Chapter 7.

6.5.2 Expanded System for Reduced Basis Approximation

We now summarize the reduced basis approximation of the expanded system. The expanded formulation presented in Chapter 3 reformulates (6.43) and (6.44) into an equivalent problem: evaluate the linear *compliant* output

$$\mathcal{S}(\mu) = \overline{\mathcal{F}(\mathcal{U}(\mu); \mu)} \tag{6.53}$$

where $\mathcal{U}(\mu) \in \mathcal{X}$ satisfies

$$\mathcal{A}(\mathcal{U}(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}), \ \forall \mathcal{V} \in \mathcal{X}; \tag{6.54}$$

$\mathcal{A}(\cdot, \cdot; \mu)$ has an affine seperation of the form $\mathcal{A}(\cdot, \cdot; \mu) = \sum_{q=1}^{Q_\mathcal{A}} \Theta_\mathcal{A}^q(\mu) \, \mathcal{A}^q(\cdot, \cdot)$.

The reduced basis approximation $\mathcal{U}_N(\mu) \in W_N$ satisfies

$$\mathcal{A}(\mathcal{U}_N(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}; \mu), \ \forall \mathcal{V} \in W_N; \tag{6.55}$$

subsequently we calculate $\mathcal{S}_N(\mu)$ as

$$\mathcal{S}_N(\mu) = \overline{\mathcal{F}(\mathcal{U}_N(\mu); \mu)}. \tag{6.56}$$

We can rigorously bound the error in $|\mathcal{S}(\mu) - \mathcal{S}_N(\mu)|$ by the *a posteriori* error estimator $\Delta_N^{\mathcal{S}}(\mu)$ (please see Chapter 8) given by

$$\Delta_N^{\mathcal{S}}(\mu) = \frac{1}{\beta_{\text{LB}}(\mu)} \|r_N(\mu)\|_{\mathcal{X}'} \tag{6.57}$$

where the residual $r_N(\mu)$ is given by

$$r_N(\mathcal{V}; \mu) \quad = \quad \mathcal{F}(\mathcal{V}) - \mathcal{A}(\mathcal{U}_N(\mu), \mathcal{V}; \mu), \ \forall \mathcal{V} \in \mathcal{X}; \tag{6.58}$$

here $\|r_N(\mu)\|_{\mathcal{X}'}$ is the dual norm of the residual, and $\beta_{\text{LB}}(\mu)$ a rigorous lower bound for the inf-sup stability constant.

6.5.3 Comparison of the Primal-Dual and Expanded Reduced Basis Formulations

We now talk about some of the differences between the two approaches. If we denote by $\mathcal{N}$ the size of the algebraic system in (6.44) to solve for $u(\mu) \in X$, then in the expanded formulation (6.54), we solve an algebraic system of equations which is twice as big. However, for the primal-dual formulation, we still need to solve for $\psi(\mu)$ from (6.47). So we have replaced the solution of two sparse algebraic systems of size $\mathcal{N} \times \mathcal{N}$ for the primal-dual formulation with a single sparse algebraic system of size $2\mathcal{N} \times 2\mathcal{N}$ for the expanded formulation. We will need to solve the $\mathcal{N}$-dependent systems *offline* and store quantities related to the decomposition of the error residual equations as discussed in Section 6.3.4.

In the *online* stage, if we use the expanded reduced basis formulation, we first assemble the dense $N \times N$ reduced basis system in $O(Q_{\mathcal{A}} N^2)$ time and then solve for $\mathcal{U}_N(\mu)$ in $O(N^3)$ time, and the reduced basis output $\mathcal{S}_N(\mu)$ in $O(N)$ time. In the primal-dual formulation, the assembly of the primal and dual reduced basis systems takes $O(Q_a (N_{\text{pr}}^2 + N_{\text{du}}^2))$ time; we then solve the dense $N_{\text{pr}} \times N_{\text{pr}}$ for $u_N(\mu)$ in $O(N_{\text{pr}}^3)$ time; subsequently we solve the dense $N_{\text{du}} \times N_{\text{du}}$ for $\psi_N(\mu)$

in $O(N_{\text{du}}^3)$ time; the computation of the quadratic output takes $O(Q_a\, N_{\text{pr}} \times N_{\text{du}})$ time.

The computation of the *a posteriori* error estimator $\Delta_N^{\mathcal{S}}(\mu)$ for the expanded formulation takes $O(Q_{\mathcal{A}}^2 N^2)$ time (see Chapter 8 for the explanation of the computational cost). The primal-dual error estimator $\Delta_N^s(\mu)$ can be computed in $O(Q_a^2\,(N_{\text{pr}}^2 + N_{\text{du}}^2))$ time.

We also expect that for fixed $N \equiv (N_{\text{pr}}, N_{\text{du}})$, the *a posteriori* error estimator $\Delta_N^{\mathcal{S}}(\mu)$ will be *sharper* than $\Delta_N^s(\mu)$. We will discuss the *a posteriori* error estimators in more detail in Chapter 8.

While a direct comparison of the computational cost cannot be made for our problems, we still can make some observations. The size of the reduced basis space is *independent* of the dimension of $\mathcal{N}$: since the expanded weak statement of our quadratic formulation is approximating the same parametric dependence as the primal-dual formulation, we can expect that $N \approx (N_{\text{pr}}, N_{\text{du}})$. If N and $(N_{\text{pr}}, N_{\text{du}})$ are the same order, then clearly the online computational cost of the quadratic formulation will be significantly less compared to the primal-dual formulation. In practice, we might even have $N < N_{\text{pr}},\ N_{\text{du}}$ since the size of the reduced basis space depends on the sharpness of the *a posteriori* error estimator, and the error estimator is much more accurate for our quadratic formulation. We thus expect the quadratic formulation to be a better method relative to the primal-dual formulation with respect to the twin metrics of *online* computational cost (i.e, *faster* predictions) and *sharpness* of the *a posteriori* error estimator (i.e, *better* predictions).

Chapter 7

Lower Bound for the Inf-Sup Parameter

In the previous chapter, we have presented various aspects of the reduced basis method for general non-coercive elliptic problems with quadratic outputs. However, we have not addressed the calculation of the lower bound $\beta_{\text{LB}}(\mu)$ for the inf-sup stability parameter $\beta(\mu)$. The inf-sup lower bound is crucial to our error estimation since it enters in the denominator of the error bounds. *Upper* bounds for minimum eigenvalues are essentially "free"; however rigorous *lower* bounds are notoriously difficult to obtain. In earlier works [72, 83, 84, 108, 110, 111], a family of rigorous error estimators for reduced basis approximation of a wide class of partial differential equations have been introduced. These earlier approaches for the development of inf-sup lower bounds rely critically on the existence of *bound conditioners* that satisfy a spectral"bound" requirement, and fit into the crucial offline-online computational paradigm.

Furthermore, identifying appropriate *bound conditioners* is very difficult when the parameter domain $\mathcal{D}$ has significant variations in $\beta(\mu)$ or when the number of parameters P is large. The afore-mentioned approaches often do not scale well to accommodate the increased difficulty in tackling the parameter dependence and the *offline* effort can be unacceptably high. Furthermore, the *sharpness* of the *a posteriori* error estimators is compromised when the parameter domain $\mathcal{D}$ includes near-resonances ($\beta(\mu) \to 0$). More recently, the "natural norm" formulation [96] outlined a error estimation framework that (a) greatly simplifies and improves the inf-sup lower bound construction (offline) and evaluation (online), and (b) much better controls — significantly sharpens

— the output error bounds, in particular (through deflation) for parameter values corresponding to nearly singular solution behavior. Although an improvement over earlier efforts, it suffers from some of the same flaws — the offline effort can be prohibitive when the number of parameters or the parameter range is increased.

In this chapter, we introduce a much simpler "successive constraint" formulation for the construction of the lower bound to the inf-sup parameter. The successive constraint formulation as we shall see is very easy to implement as a black-box method to construct inf-sup lower bounds.

7.1 Preliminaries

We consider a suitably regular (smooth) domain $\Omega \subset \mathbb{R}^d, d = 1, 2$, or 3[1] with Lipschitz-continuous boundary Γ. Our exact output and field variable, $s^e(\mu) \in \mathbb{R}$ and $u^e(\mu) \in \mathcal{X}^e$, satisfy (6.1), (6.2). Here, for any $\mu \equiv (\mu_1 \cdots \mu_P) \in \mathcal{D} \subset \mathbb{R}^P$, $a(\cdot, \cdot; \mu)$: $\mathcal{X}^e \times \mathcal{X}^e \rightarrow \mathbb{C}$ is a bilinear form; f: $\mathcal{X}^e \rightarrow \mathbb{C}$ is a parameter-independent linear form. We shall consider second-order partial differential equations, and hence our "exact" complex vector space $\mathcal{X}^e$ satisfies $H_0^1(\Omega) \subset \mathrm{Re}(\mathcal{X}^e)$, $\mathrm{Im}(\mathcal{X}^e) \subset H^1(\Omega)$.

Our "truth" or "reference" finite element approximation to the exact output and field variable, $s(\mu)$ and $u(\mu) \equiv u(\mu) \in \mathcal{X}$, satisfies (6.5), (6.6). Given $\mu \in \mathcal{D}$,

$$s(\mu) = \overline{f(u(\mu))} \ , \tag{7.1}$$

where $u(\mu) \in \mathcal{X}$ satisfies

$$a(u(\mu), v; \mu) = f(v), \quad \forall \, v \in \mathcal{X} \ . \tag{7.2}$$

We shall suppose that our bilinear form is "affine" in the parameter: for some fixed integer Q — typically Q_a shall be larger than P, sometimes by a considerable factor — we require

$$a(w, v; \mu) = \sum_{q=1}^{Q_a} \Theta_a^q(\mu) a^q(w, v), \quad \forall \, w, v \in \mathcal{X}, \ \forall \, \mu \in \mathcal{D} \ , \tag{7.3}$$

where Θ_a^q: $\mathcal{D} \rightarrow \mathbb{C}$ and $a^q(w, v)$: $\mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$, $1 \le q \le Q_a$, are parameter-dependent functions and parameter-independent continuous bilinear forms, respectively. We shall further assume that $\Theta_a^q \in \mathcal{C}^1(\mathcal{D})$, $1 \le q \le Q_a$.

[1] Note that Ω is a *reference domain* and hence does *not* depend on the parameter.

$\mathcal{X}$, our "truth" approximation subspace of $\mathcal{X}^e$ is endowed with an appropriate $H^1(\Omega)$-equivalent inner product $(\cdot,\cdot)_{\mathcal{X}}$ (5.30), (6.3) and induced norm $\|\cdot\|_{\mathcal{X}}$ (5.31), (6.4).

7.2 Stability Parameters

We introduce [44, 65, 73, 110] the parametrized linear operator $T^\mu : \mathcal{X} \to \mathcal{X}$ such that, for any $\mu \in \mathcal{D}$ and any $\mathcal{W} \in \mathcal{X}$,

$$(T^\mu \mathcal{W}, \mathcal{V})_{\mathcal{X}} = \mathcal{A}(\mathcal{W}, \mathcal{V}; \mu), \quad \forall \mathcal{V} \in \mathcal{X}; \tag{7.4}$$

the classical inf-sup and continuity parameters are then given as

$$\beta(\mu) = \inf_{\mathcal{W}\in\mathcal{X}} \frac{\|T^\mu \mathcal{W}\|_{\mathcal{X}}}{\|\mathcal{W}\|_{\mathcal{X}}}, \quad \forall \mu \in \mathcal{D}, \tag{7.5}$$

and

$$\gamma(\mu) = \sup_{\mathcal{W}\in\mathcal{X}} \frac{\|T^\mu \mathcal{W}\|_{\mathcal{X}}}{\|\mathcal{W}\|_{\mathcal{X}}} \quad \forall \mu \in \mathcal{D}. \tag{7.6}$$

We assume that $\mathcal{A}(\cdot,\cdot;\mu)$ is inf-sup stable, $\beta(\mu) > 0$, $\forall \mu \in \mathcal{D}$. We wish to obtain a lower bound for $\beta(\mu)$, $\beta_{\rm LB}(\mu)$, such that

$$\beta(\mu) \geq \beta_{\rm LB}(\mu) \geq 0, \ \forall \mu \in \mathcal{D}. \tag{7.7}$$

We would also like $\beta_{\rm LB}(\mu)$ to be "close" to $\beta(\mu)$, such as $\beta_{\rm LB}(\mu)/\beta(\mu) \geq \frac{1}{5}$, $\forall \mu \in \mathcal{D}$.

We also introduce the linear operator $T^q : \mathcal{X} \to \mathcal{X}$ such that, for any $\mathcal{W} \in \mathcal{X}$,

$$(T^q \mathcal{W}, \mathcal{V})_{\mathcal{X}} = \mathcal{A}^q(\mathcal{W}, \mathcal{V}; \mu), \quad \forall \mathcal{V} \in \mathcal{X}. \tag{7.8}$$

Then, $T^\mu : \mathcal{X} \to \mathcal{X}$ can be expressed as

$$T^\mu \mathcal{W} = \sum_{q=1}^{Q_{\mathcal{A}}} \Theta^q_{\mathcal{A}}(\mu)\, T^q \mathcal{W} \tag{7.9}$$

from (7.3), (7.4), and (7.8).

We then conveniently define the square of the inf-sup stability constant

$$\sigma(\mu) = \beta^2(\mu), \forall \mu \in \mathcal{D}. \tag{7.10}$$

We now use the expression for $\beta(\mu)$ (7.5) to expand $\sigma(\mu)^2$ as

$$\sigma(\mu) = \inf_{\mathcal{W}\in\mathcal{X}} \sum_{q=1}^{Q_\mathcal{A}} \sum_{q'=1}^{Q_\mathcal{A}} \Theta^q_\mathcal{A}(\mu)\, \overline{\Theta^{q'}_\mathcal{A}(\mu)}\, \frac{(T^q\mathcal{W}, T^{q'}\mathcal{W})_\mathcal{X}}{\|\mathcal{W}\|^2_\mathcal{X}}. \tag{7.11}$$

For example, consider the case when $Q_\mathcal{A} = 2$. We have

$$T^\mu\mathcal{W} = \Theta^1_\mathcal{A}(\mu)\, T^1\mathcal{W} + \Theta^2_\mathcal{A}(\mu)\, T^2\mathcal{W}. \tag{7.12}$$

We note that $\Theta^q_\mathcal{A} \in \mathbb{C}, q = 1, \ldots, Q_\mathcal{A}$, and can be decomposed as

$$\Theta^q_\mathcal{A}(\mu) = \Theta^q_{\mathcal{A},R}(\mu) + \mathrm{i}\,\Theta^q_{\mathcal{A},I}(\mu),\ q = 1, \ldots, Q_\mathcal{A},\ \forall \mu \in \mathcal{D}, \tag{7.13}$$

where $\Theta^q_{\mathcal{A},R}(\mu), \Theta^q_{\mathcal{A},I}(\mu) \in \mathbb{R}$. Thus, $\Theta^1_\mathcal{A}(\mu) = \Theta^1_{\mathcal{A},R}(\mu) + \mathrm{i}\,\Theta^1_{\mathcal{A},I}$; $\Theta^2_\mathcal{A}$ can be similarly decomposed. Thus, $\sigma(\mu)$ is given by

$$\sigma(\mu) = \inf_{\mathcal{W}\in\mathcal{X}} \frac{(\Theta^1_\mathcal{A}(\mu)T^1\mathcal{W} + \Theta^2_\mathcal{A}(\mu)\mathcal{W}, \Theta^1_\mathcal{A}(\mu)T^1\mathcal{W} + \Theta^2_\mathcal{A}(\mu)\mathcal{W})_\mathcal{X}}{\|\mathcal{W}\|^2_\mathcal{X}}. \tag{7.14}$$

Recall that $\sigma(\mu)$ is real, thus the inner product $(T^\mu\mathcal{W}, T^\mu\mathcal{W})_\mathcal{X}$ has to be real since the denominator $\|\mathcal{W}\|^2_\mathcal{X}$ is real. We decompose $\Theta^q_\mathcal{A}(\mu)$ into their real and imaginary parts; and group the expanded numerator into sets of terms that are real

$$\begin{aligned}(T^\mu\mathcal{W}, T^\mu\mathcal{W})_\mathcal{X} \;=\; & \Theta^1_\mathcal{A}(\mu)^2\,(T^1\mathcal{W}, T^1\mathcal{W})_\mathcal{X} + \Theta^2_\mathcal{A}(\mu)^2\,(T^2\mathcal{W}, T^2\mathcal{W})_\mathcal{X} \\ & + \big(\Theta^1_{\mathcal{A},R}(\mu)\Theta^2_{\mathcal{A},R}(\mu) + \Theta^1_{\mathcal{A},I}(\mu)\Theta^2_{\mathcal{A},I}(\mu)\big)\,\big((T^1\mathcal{W}, T^2\mathcal{W})_\mathcal{X} + (T^2\mathcal{W}, T^1\mathcal{W})_\mathcal{X}\big) \\ & + \big(\Theta^1_{\mathcal{A},R}(\mu)\Theta^2_{\mathcal{A},R}(\mu) + \Theta^1_{\mathcal{A},I}(\mu)\Theta^2_{\mathcal{A},I}(\mu)\big)\,\mathrm{i}\big((T^1\mathcal{W}, T^2\mathcal{W})_\mathcal{X} - (T^2\mathcal{W}, T^1\mathcal{W})_\mathcal{X}\big).\end{aligned} \tag{7.15}$$

Note that all the terms in the expansion (7.15) are real: the first two terms are of the form $(\Theta^q_\mathcal{A})^2 = \Theta^q_\mathcal{A}\,\overline{\Theta^q_\mathcal{A}}$, $q = 1, 2$; the third term has a *real* coefficient of the form $\mathrm{Re}(\Theta^1_\mathcal{A}, \Theta^2_\mathcal{A})$ and is multiplied by the Hermitian (symmetric) bilinear form $(T^1\mathcal{W}, T^2\mathcal{W})_\mathcal{X} + (T^2\mathcal{W}, T^1\mathcal{W})_\mathcal{X}$; and the

[2]Please see Section 7.4 for a description of the Lanczos algorithm used to calculate the square of the inf-sup $\sigma(\mu)$.

fourth term has a *real* coefficient of the form $\mathrm{Im}(\Theta^1_{\mathcal{A}}, \Theta^2_{\mathcal{A}})$ and is again multiplied by a Hermitian bilinear form $\mathrm{i}\left((T^1\mathcal{W}, T^2\mathcal{W})_{\mathcal{X}} + (T^2\mathcal{W}, T^1\mathcal{W})_{\mathcal{X}}\right)$. The expansion of $\sigma(\mu)$ consists of real coefficients multiplied by Hermitian bilinear forms; since the eigenvalues of the Hermitian matrices associated with the Hermitian bilinear forms are *always* real, we are guaranteed that $\sigma(\mu)$ will be real.

We now focus on the general case. We expand $\sigma(\mu)$ in (7.11) as

$$\sigma(\mu) = \inf_{\mathcal{W}\in\mathcal{X}} \left\{ \sum_{q=1}^{Q_{\mathcal{A}}} \Theta^q_{\mathcal{A}}(\mu)\, \overline{\Theta^q_{\mathcal{A}}(\mu)}\, \frac{(T^q\mathcal{W}, T^q\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|^2_{\mathcal{X}}} + \right. \tag{7.16}$$

$$\sum_{q=1}^{Q_{\mathcal{A}}} \sum_{q'=q+1}^{Q_{\mathcal{A}}} \left(\mathrm{Re}(\Theta^q_{\mathcal{A}}(\mu)\, \overline{\Theta^{q'}_{\mathcal{A}}(\mu)}\,) \, \frac{(T^q\mathcal{W}, T^{q'}\mathcal{W})_{\mathcal{X}} + (T^{q'}\mathcal{W}, T^q\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|^2_{\mathcal{X}}} \right.$$

$$\left. \left. + \mathrm{Im}(\Theta^q_{\mathcal{A}}(\mu)\, \overline{\Theta^{q'}_{\mathcal{A}}(\mu)}\,) \, \frac{\mathrm{i}\left((T^q\mathcal{W}, T^{q'}\mathcal{W})_{\mathcal{X}} - (T^{q'}\mathcal{W}, T^q\mathcal{W})_{\mathcal{X}} \right)}{\|\mathcal{W}\|^2_{\mathcal{X}}} \right) \right\}. \tag{7.17}$$

We now introduce for $q = q'$ $(1 \le q = q' \le Q_{\mathcal{A}})$

$$\hat{\gamma}^-_{q,q',R} = \inf_{\mathcal{W}\in\mathcal{X}} \frac{(T^q\mathcal{W}, T^{q'}\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|^2_{\mathcal{X}}}, \tag{7.18}$$

$$\hat{\gamma}^+_{q,q',R} = \sup_{\mathcal{W}\in\mathcal{X}} \frac{(T^q\mathcal{W}, T^{q'}\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|^2_{\mathcal{X}}}, \tag{7.19}$$

$$\hat{\gamma}^-_{q,q',I} = 0, \tag{7.20}$$

$$\hat{\gamma}^+_{q,q',I} = 0. \tag{7.21}$$

We also introduce for $q \neq q'$ $(1 \le q < q' \le Q_{\mathcal{A}})$

$$\hat{\gamma}^-_{q,q',R} = \inf_{\mathcal{W}\in\mathcal{X}} \frac{(T^q\mathcal{W}, T^{q'}\mathcal{W})_{\mathcal{X}} + (T^{q'}\mathcal{W}, T^q\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|^2_{\mathcal{X}}}, \tag{7.22}$$

$$\hat{\gamma}^+_{q,q',R} = \sup_{\mathcal{W}\in\mathcal{X}} \frac{(T^q\mathcal{W}, T^{q'}\mathcal{W})_{\mathcal{X}} + (T^{q'}\mathcal{W}, T^q\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|^2_{\mathcal{X}}}, \tag{7.23}$$

$$\hat{\gamma}^-_{q,q',I} = \inf_{\mathcal{W}\in\mathcal{X}} \frac{\mathrm{i}\left((T^q\mathcal{W}, T^{q'}\mathcal{W})_{\mathcal{X}} - (T^{q'}\mathcal{W}, T^q\mathcal{W})_{\mathcal{X}} \right)}{\|\mathcal{W}\|^2_{\mathcal{X}}}, \tag{7.24}$$

$$\hat{\gamma}^+_{q,q',I} = \sup_{\mathcal{W}\in\mathcal{X}} \frac{\mathrm{i}\left((T^q\mathcal{W}, T^{q'}\mathcal{W})_{\mathcal{X}} - (T^{q'}\mathcal{W}, T^q\mathcal{W})_{\mathcal{X}} \right)}{\|\mathcal{W}\|^2_{\mathcal{X}}}. \tag{7.25}$$

Note that by continuity $\hat{\gamma}_{q,q',R}$ and $\hat{\gamma}_{q,q',I}$ are finite and real (since the associated eigen problems are all Hermitian), $1 \le q, q' \le Q_{\mathcal{A}}$.

We next introduce index mapping notations

$$\text{Ind}(q,q',R): \quad 1 \le q \le Q_{\mathcal{A}},\ 1 \le q' \le Q_{\mathcal{A}} \quad \longrightarrow r : 1,\ldots,\overline{Q}, \tag{7.26}$$

$$\text{Ind}(q,q',I): \quad 1 \le q \le Q_{\mathcal{A}},\ 1 \le q' \le Q_{\mathcal{A}} \quad \longrightarrow r : \overline{Q}+1,\ldots,2\,\overline{Q}, \tag{7.27}$$

where

$$\overline{Q} = \frac{Q_{\mathcal{A}}\,(Q_{\mathcal{A}}+1)}{2}. \tag{7.28}$$

For example, if $Q_{\mathcal{A}} = 2$, we can pick $\text{Ind}(1,1,R) = 1$, $\text{Ind}(1,2,R) = 2$, $\text{Ind}(2,2,R) = 3$, $\text{Ind}(1,1,I) = 4$, $\text{Ind}(1,2,I) = 5$, and $\text{Ind}(2,2,I) = 6$ where $\overline{Q} = 3$.

We then define for $q = q'$ $(1 \le q = q' \le Q_{\mathcal{A}})$

$$\begin{aligned} \hat{\Phi}_{\text{Ind}(q,q',R)}(\mu) &\equiv \Theta^q_{\mathcal{A}}(\mu)\,\overline{\Theta^{q'}_{\mathcal{A}}(\mu)} = |\Theta^q_{\mathcal{A}}|^2, \\ \hat{\Phi}_{\text{Ind}(q,q',I)}(\mu) &\equiv 0, \end{aligned} \tag{7.29}$$

and for $q \neq q'$ $(1 \le q < q' \le Q_{\mathcal{A}})$

$$\begin{aligned} \hat{\Phi}_{\text{Ind}(q,q',R)}(\mu) &\equiv \text{Re}(\Theta^q_{\mathcal{A}}(\mu)\,\overline{\Theta^{q'}_{\mathcal{A}}(\mu)}), \\ \hat{\Phi}_{\text{Ind}(q,q',I)}(\mu) &\equiv \text{Im}(\Theta^q_{\mathcal{A}}(\mu)\,\overline{\Theta^{q'}_{\mathcal{A}}(\mu)}), \end{aligned} \tag{7.30}$$

$\forall \mu \in \mathcal{D}$.

Similarly, we can define for $q = q'$ $(1 \le q = q' \le Q_{\mathcal{A}})$

$$\begin{aligned} \hat{\Psi}_{\text{Ind}(q,q',R)}(\mathcal{W}) &= \frac{(T^q\mathcal{W}, T^{q'}\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|^2_{\mathcal{X}}}, \\ \hat{\Psi}_{\text{Ind}(q,q',I)}(\mathcal{W}) &= 0, \end{aligned} \tag{7.31}$$

and for $q \neq q'$ $(1 \le q < q' \le Q_{\mathcal{A}})$

$$\begin{aligned} \hat{\Psi}_{\text{Ind}(q,q',R)}(\mathcal{W}) &= \frac{(T^q\mathcal{W}, T^{q'}\mathcal{W})_{\mathcal{X}} + (T^{q'}\mathcal{W}, T^q\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|^2_{\mathcal{X}}}, \\ \hat{\Psi}_{\text{Ind}(q,q',I)}(\mathcal{W}) &= \frac{\text{i}\,\big((T^q\mathcal{W}, T^{q'}\mathcal{W})_{\mathcal{X}} - (T^{q'}\mathcal{W}, T^q\mathcal{W})_{\mathcal{X}}\big)}{\|\mathcal{W}\|^2_{\mathcal{X}}}, \end{aligned} \tag{7.32}$$

$\forall \mathcal{W} \in \mathcal{X}$.

We effect a second mapping: for all $\mu \in \mathcal{D}$, we map $\hat{\Phi}_{\text{Ind}(\cdot,\cdot,R)} \rightarrow \Phi$.

$$\hat{\Phi}_{\text{Ind}(q,q',R)}(\mu), 1 \leq q \leq Q_{\mathcal{A}},\ q \leq q' \leq Q_{\mathcal{A}} \longmapsto \Phi_r(\mu), r = 1, \ldots, \overline{Q}, \tag{7.33}$$

and $\hat{\Phi}_{\text{Ind}(\cdot,\cdot,I)} \rightarrow \Phi$. as

$$\hat{\Phi}_{\text{Ind}(q,q',I)}(\mu), 1 \leq q \leq Q_{\mathcal{A}},\ q \leq q' \leq Q_{\mathcal{A}} \longmapsto \Phi_r(\mu), r = \overline{Q}+1, \ldots, 2\,\overline{Q}. \tag{7.34}$$

Similarly, we identify for all $\mathcal{W} \in \mathcal{X}$

$$\hat{\Psi}_{\text{Ind}(q,q',R)}(\mathcal{W}), 1 \leq q \leq Q_{\mathcal{A}},\ q \leq q' \leq Q_{\mathcal{A}} \longmapsto \Psi_r(\mathcal{W}), r = 1, \ldots, \overline{Q}, \tag{7.35}$$

and

$$\hat{\Psi}_{\text{Ind}(q,q',I)}(\mathcal{W}), 1 \leq q \leq Q_{\mathcal{A}},\ q \leq q' \leq Q_{\mathcal{A}} \longmapsto \Psi_r(\mathcal{W}), r = \overline{Q}+1, \ldots, 2\,\overline{Q}. \tag{7.36}$$

We can then write (substituting $\hat{\mathcal{Q}} = 2\,\overline{Q}$)

$$\sigma(\mu) = \beta^2(\mu) = \inf_{\mathcal{W} \in \mathcal{X}} \sum_{r=1}^{\hat{\mathcal{Q}}} \Phi_r(\mu)\ \Psi_r(\mathcal{W}) \tag{7.37}$$

from (7.17)– (7.36).

Although our formulation (7.37) is written as a single sum in terms of ($\Phi_r(\mu)$ and $\Psi_r(\mathcal{W})$), we can readily invert the mappings to correctly infer the expression as a double sum in ($\hat{\Phi}_{q,q',\cdot}(\mu)$ and $\hat{\Psi}_{q,q',\cdot}(\mathcal{W})$). We therefore choose to identify

$$\begin{aligned}
\hat{\gamma}^-_{q,q',R},\ 1 \leq q, q \leq q' \leq Q_{\mathcal{A}} &\longmapsto \gamma^-_r,\ 1 \leq r \leq \frac{\hat{\mathcal{Q}}}{2}, \\
\hat{\gamma}^-_{q,q',I},\ 1 \leq q, q \leq q' \leq Q_{\mathcal{A}} &\longmapsto \gamma^-_r,\ \frac{\hat{\mathcal{Q}}}{2}+1 \leq r \leq \hat{\mathcal{Q}}, \\
\hat{\gamma}^+_{q,q',I},\ 1 \leq q, q \leq q' \leq Q_{\mathcal{A}} &\longmapsto \gamma^+_r,\ 1 \leq r \leq \frac{\hat{\mathcal{Q}}}{2}, \\
\hat{\gamma}^+_{q,q',I},\ 1 \leq q, q \leq q' \leq Q_{\mathcal{A}} &\longmapsto \gamma^+_r,\ \frac{\hat{\mathcal{Q}}}{2}+1 \leq r \leq \hat{\mathcal{Q}}.
\end{aligned} \tag{7.38}$$

We have $\hat{Q}$ terms in the expansion but we should note that for $q = q'$, $\hat{\gamma}_{\text{Ind}(q,q',I)} = 0$: in reality

we have $\hat{Q} = Q_{\mathcal{A}}^2$ terms in our expansion. Further, the cross-terms $\hat{\gamma}_{\text{Ind}(q,q',R)}$ and $\hat{\gamma}_{\text{Ind}(q,q',I)}$ can often be very close to zero.

7.3 Successive Constraints Method (SCM)

By way of preliminaries we define the set $\mathcal{S}_0 \subset \mathbb{R}^{\hat{Q}}$,

$$\begin{aligned} \mathcal{S}_0 \quad \equiv \quad & \left\{ \underline{z} = (z_1, \ldots, z_{\hat{Q}}) \in \mathbb{R}^{\hat{Q}} \mid \exists\, \mathcal{W} \in \mathcal{X} \right. \\ & \left. \text{s.t } z_r = \Psi_r(\mathcal{W}),\ 1 \le r \le \hat{Q} \right\}. \end{aligned} \tag{7.39}$$

We further define the objective function $\mathcal{J} : \mathbb{R}^{\hat{Q}} \rightarrow \mathbb{R}$ as

$$\mathcal{J}(\mu; \underline{z}) = \sum_{r=1}^{\hat{Q}} \Phi_r(\mu)\, z_r; \tag{7.40}$$

we can then write $\sigma(\mu)$ as

$$\sigma(\mu) = \min_{\underline{z} \in \mathbb{R}^{\hat{Q}}} \mathcal{J}(\mu; \underline{z}). \tag{7.41}$$

We can also write $\sigma(\mu)$ in terms of the set $\mathcal{S}_0$ as

$$\begin{aligned} \sigma(\mu) \quad &= \quad \inf_{\mathcal{W} \in \mathcal{X}} \sum_{r=1}^{\hat{Q}} \Phi_r(\mu)\, \Psi_r(\mathcal{W}) \\ &= \quad \inf_{z_1, \ldots, z_{\hat{Q}} \in \mathcal{S}_0} \sum_{r=1}^{\hat{Q}} \Phi_r(\mu)\, z_r\ (z_r = \Psi_r(\mathcal{W})) \\ &= \quad \inf_{\underline{z} \in \mathcal{S}_0} \sum_{r=1}^{\hat{Q}} \Phi_r(\mu)\, z_r. \end{aligned} \tag{7.42}$$

We now introduce two parameter sets

$$\mathcal{C}_K \equiv \{\omega_1 \in \mathcal{D}, \ldots, \omega_K \in \mathcal{D}\} \tag{7.43}$$

and

$$\Xi_J \equiv \{\nu_1 \in \mathcal{D}, \ldots, \nu_J \in \mathcal{D}\}. \tag{7.44}$$

We associate $\mathcal{M}_\alpha \in \mathbb{N}$ and $\mathcal{M}_+ \in \mathbb{N}$ with $\mathcal{C}_K$ and Ξ_J respectively. Given $\mu \in \mathcal{D}$, we define $P_{\mathcal{M}_\alpha}(\mu; \mathcal{C}_K)$ as the set of $\mathcal{M}_\alpha$ points in $\mathcal{C}_K$ which are "closest" to μ in the Euclidean norm; similarly, $P_{\mathcal{M}_+}(\mu; \Xi_J)$ is the set of $\mathcal{M}_+$ points in Ξ_J which are "closest" to μ in the Euclidean norm.

Given the sets $\mathcal{C}_K$ and Ξ_J, we consider a relaxation

$$\begin{aligned}
\mathcal{S}_{\text{relax}}(\mu; \mathcal{C}_K, \Xi_J) \quad \equiv \quad & \Big\{ \underline{z} = (z_1, \ldots, z_{\hat{\mathcal{Q}}}) \in \mathbb{R}^{\hat{\mathcal{Q}}} \mid \\
& \gamma_r^- \leq z_r \leq \gamma_r^+, \; 1 \leq r \leq \hat{\mathcal{Q}}; \\
& \sum_{r=1}^{\hat{\mathcal{Q}}} \Phi_r(\omega') \, z_r \geq \sigma(\omega'), \; \forall \omega' \in P_{\mathcal{M}_\alpha}(\mu; \mathcal{C}_K); \\
& \sum_{r=1}^{\hat{\mathcal{Q}}} \Phi_r(\nu') \, z_r \geq 0, \; \forall \nu' \in P_{\mathcal{M}_+}(\mu; \Xi_J) \Big\}.
\end{aligned} \tag{7.45}$$

7.3.1 Lower Bound for $\sigma(\mu)$

We can then define our inf-sup lower bound, $\sigma_{\text{LB}}(\mu)$ as

$$\sigma_{\text{LB}}(\mu; \mathcal{C}_K, \Xi_J) = \min_{\underline{z} \in \mathcal{S}_{\text{relax}}(\mu; \mathcal{C}_K, \Xi_J)} \mathcal{J}(\mu; \underline{z}). \tag{7.46}$$

Proposition 2. *For given $\mathcal{C}_K$ (and associated $\mathcal{M}_\alpha \in \mathbb{N}$) and Ξ_J (and associated $\mathcal{M}_+ \in \mathbb{N}$),*

$$\sigma_{\text{LB}}(\mu; \mathcal{C}_K, \Xi_J) \leq \sigma(\mu), \forall \mu \in \mathcal{D}. \tag{7.47}$$

Proof. We only need to prove that for any $\mathcal{W}^z \in \mathcal{X}$, there exists a $\underline{z} = (z_1, \ldots, z_{\hat{\mathcal{Q}}}) \in \mathbb{R}^{\hat{\mathcal{Q}}}$ in $\mathcal{S}_{\text{relax}}$ for which $z_1 = \Psi_1(\mathcal{W}^z), \ldots, z_{\hat{\mathcal{Q}}} = \Psi_{\hat{\mathcal{Q}}}(\mathcal{W}^z)$.

But since

$$\gamma_r^- = \inf_{\mathcal{W} \in \mathcal{X}} \Psi_r(\mathcal{W}), \; 1 \leq r \leq \hat{\mathcal{Q}} \tag{7.48}$$

and

$$\gamma_r^+ = \sup_{\mathcal{W} \in \mathcal{X}} \Psi_r(\mathcal{W}), \; 1 \leq r \leq \hat{\mathcal{Q}}; \tag{7.49}$$

it is clear that $\gamma_r^- \le \Psi_r(\mathcal{W}^z) \le \gamma_r^+$, $1 \le r \le \hat{\mathcal{Q}}$.

Furthermore, since

$$\sum_{r=1}^{\hat{\mathcal{Q}}} \Phi_r(\omega') \, \Psi_r(\mathcal{W}) \ge \sigma(\omega'), \forall \omega' \in P_{\mathcal{M}_\alpha}(\mu; \mathcal{C}_K) \tag{7.50}$$

applies to any $\mathcal{W} \in \mathcal{X}$, it applies to $\mathcal{W}^z$.

Also, the positivity criteria $\sum_{r=1}^{\hat{\mathcal{Q}}} \Phi_r(\nu') \, z_r \ge 0$, $\forall \nu' \in P_{\mathcal{M}_+}(\mu; \Xi_J)$ are trivially satisfied for any $\mathcal{W} \in \mathcal{X}$ since by assumption $\sigma(\nu') > 0, \forall \nu' \in \mathcal{D}$, it then applies to $\mathcal{W}^z$.

Thus, each constraint in $\mathcal{S}_{\text{relax}}$ is satisfied by all members of $\mathcal{S}_0$. The desired result follows directly.

□

We note that (7.45), (7.46) is in fact a linear optimization problem or Linear Program (LP); indeed (7.45), (7.46) resembles a discretized linear semi-infinite program [39]. We formally write out the LP

$$\begin{array}{rcll} & \text{Minimize } \sum_{r=1}^{\hat{\mathcal{Q}}} \Phi_r(\mu) \, z_r & & \\ & \text{subject to} & & \\ \gamma_r^- & \le & z_r, \quad 1 \le r \le \hat{\mathcal{Q}} & \\ \gamma_r^+ & \le & z_r, \quad 1 \le r \le \hat{\mathcal{Q}} & \\ \sum_{r=1}^{\hat{\mathcal{Q}}} \phi_r(\omega_k) \, z_r & \ge & \sigma(\omega_k), \quad 1 \le k \le M_\alpha & \\ \sum_{r=1}^{\hat{\mathcal{Q}}} \phi_r(\nu_j) \, z_r & \ge & 0, \quad 1 \le j \le M_+. & \end{array} \tag{7.51}$$

Our LP (7.51) contains $\hat{\mathcal{Q}}$ design variables and $2\hat{\mathcal{Q}} + M_\alpha + M_+$ (one-sided) inequality constraints: the operation count for the Online stage $\mu \rightarrow \sigma_{LB}(\mu)$ is *independent* of $\mathcal{N}$.

7.3.2 Upper Bound for $\sigma(\mu)$

Let us introduce the set

$$\begin{aligned} \mathcal{S}_{\text{UB}}(\mu; \mathcal{C}_K) \quad \equiv \quad & \left\{ \underline{z}^*(\omega_k) = (z_1^*(\omega_k), \ldots, z_{\hat{\mathcal{Q}}}^*(\omega_k)) \in \mathbb{R}^{\hat{\mathcal{Q}}}, 1 \le k \le K \mid \right. \\ & \left. \underline{z}^*(\omega_k) = \arg \min_{\underline{z} \in \mathbb{R}^{\hat{\mathcal{Q}}}} \mathcal{J}(\omega_k; \underline{z}), 1 \le k \le K. \right. \end{aligned} \tag{7.52}$$

We can also define a *rigorous* upper bound to $\sigma(\mu)$, $\sigma_{\text{UB}}(\mu; \mathcal{C}_K)$ as

$$\sigma_{\text{UB}}(\mu; \mathcal{C}_K) = \min_{\underline{z} \in \mathcal{S}_{\text{UB}}(\mu; \mathcal{C}_K)} \mathcal{J}(\mu; \underline{z}) \tag{7.53}$$

Proposition 3. *For given* $\mathcal{C}_K$,

$$\sigma_{\text{UB}}(\mu; \mathcal{C}_K) \geq \sigma(\mu), \forall \mu \in \mathcal{D}. \tag{7.54}$$

Proof. Since the idea

$$\sum_{r=1}^{\hat{Q}} \Phi_r(\mu)\, \Psi_r(\mathcal{W}) \geq \sigma(\mu) \tag{7.55}$$

applies to any $\mathcal{W} \in \mathcal{X}$, it applies to $\mathcal{W}_k^* \in \mathcal{X}, 1 \leq k \leq K$, the inf-sup minimizers associated with $\omega_k \in \mathcal{C}_K$ as well.

Thus, $\mathcal{S}_{\text{UB}} \subset \mathcal{S}_0$, and we can take the minimum upper bound obtained in the set $\mathcal{S}_{\text{UB}}$ as our "best" upper bound. □

Modified Upper Bound

We expect that if there is some $\omega_k \in \mathcal{C}_K$ "close" to μ; then the infimizer $\mathcal{W}_k^*$ associated with ω_k will be close to the infimizer associated with μ and our upper bound will be *sharp*. Thus, we choose to construct our upper bound, $\sigma_{\text{UB}}(\mu)$, from sample points in the set $\mathcal{P}_{\mathcal{M}_\alpha}(\mu; \mathcal{C}_K)$ — the set of $\mathcal{M}_\alpha$ points "closest" to μ in the Euclidean norm. This construction also gives us a rigorous upper bound, however we look for the upper bound in a smaller sub-set of $\mathcal{C}_K$. Thus,

$$\sigma_{\text{UB}}(\mu) = \min_{\underline{z} \in \mathcal{S}_{\text{UB}}(\mu; \mathcal{P}_{\mathcal{M}_\alpha}(\mu; \mathcal{C}_K))} \mathcal{J}(\mu; \underline{z}). \tag{7.56}$$

Looking for the upper bounds only over the closest points is helpful particularly when K is very large; also the modified upper bound will not degrade in its sharpness since the closest point is usually optimal.

7.3.3 Selection of $\mathcal{C}_K$

We must first determine $\mathcal{C}_K$ by an offline "greedy" algorithm before we can compute our inf-sup lower and upper bounds.

We specify our training sample $\Xi_{\text{train}} \subset \mathcal{D}$ (a very fine sample over the parameter domain) and termination error tolerance $\epsilon_{\text{tol,min}}$. For each $\mu \in \Xi_{\text{train}}$, define the set $P_{\mathcal{M}_+}(\mu; \Xi_{\text{train}})$ as the $\mathcal{M}_+$ "closest" points to μ in the Euclidean norm in the set Ξ_{train}.

Choose $\mu_1^* \in \Xi_{\text{train}}$ at random as the first μ-sample to be added into the set $\mathcal{C}_K$. Denote

$$C_1^* = \{\mu_1^*\} \tag{7.57}$$

and associated

$$\mathcal{S}_{\text{UB},1}^* = \left\{ \underline{z}(\mu_1^*) = \arg \min_{\underline{z} \in \mathbb{R}^{\hat{\mathcal{Q}}}} \mathcal{J}(\mu_1^*; \underline{z}) \right\}; \tag{7.58}$$

and relative error bound

$$\epsilon_k^*(\mu) = \frac{\sigma_{\text{UB}}(\mu; \mathcal{C}_k^*) - \sigma_{\text{LB}}(\mu; \mathcal{C}_k^*, \Xi_{\text{train}})}{\sigma_{\text{UB}}(\mu; \mathcal{C}_k^*)}, \ \forall k. \tag{7.59}$$

The algorithm proceeds as follows (we provisionally define an upper bound on the size of K, $\overline{K}$)

```
while k = 2: K̄
    μ_k^* = arg max_{μ∈Ξ_train} ε_{k-1}^*(μ);
    ε_k^* = ε_{k-1}^*(μ_k^*);
    if ε_k^* ≤ ε_tol,min
        K = k − 1;
        exit;
    end;
    C_k^* = C_{k-1}^* ∪ μ_k^*;
    S_{UB,k}^* = { S_{UB,k-1}^* ; z(μ_k^*) = arg min_{z∈R^Q̂} J(μ_k^*; z) };
end.
```

7.3.4 Offline-Online Decomposition

In the *offline* stage, we construct our constraint set $\mathcal{C}_K$ using the greedy generation algorithm outlined above. The notable offline computations are (i) $2\,\hat{\mathcal{Q}}$ eigenproblems to compute the $\gamma_r^{\pm}, 1 \leq r \leq \hat{\mathcal{Q}}$ *once*; (ii) K eigen problems to solve for $\sigma(\omega_k)$ and $\mathcal{W}(\omega_k) = \inf_{\mathcal{W}\in\mathcal{X}} \sigma(\mu;\mathcal{W})$, $k = 1,\ldots,K$. The $\sigma(\omega_k)$, $1 \leq k \leq K$ can be very efficiently calculated using a Lanczos algorithm (see Section 7.4); (iii) $O(\mathcal{N}\,\hat{\mathcal{Q}}\,K$ operations to form $\mathcal{S}_{UB}$; and (iv) $|\Xi_{\text{train}}|\,K$ Linear Programs (LPs) of size $O(2\,\hat{\mathcal{Q}} + \mathcal{M}_\alpha + \mathcal{M}_+)$.

In the *online* stage, for a given $\mu \in \mathcal{D}$, we first compute the $\Psi_r, 1 \leq r \leq \hat{\mathcal{Q}}$ which requires $\hat{\mathcal{Q}}$ operations; then identify $P_{\mathcal{M}_\alpha}(\mu;\mathcal{C}_K)$ in $P\,M_\alpha$ operations; and solve the LP of size $O(2\,\hat{\mathcal{Q}} + \mathcal{M}_\alpha + \mathcal{M}_+)$ to calculate $\sigma_{\text{LB}}(\mu)$. We obtain the inf-sup lower bound as $\beta_{\text{LB}}(\mu) = \sqrt{\sigma_{\text{LB}}(\mu)}$.

We shall discuss actual computational costs for the construction of $\mathcal{C}_K$ *offline*, and the lower bounds *online* in Section 7.5 for the low-pass filter example.

7.4 Inf-Sup Calculation using the Lanczos Algorithm

We now address the calculation of the classical inf-sup parameter given as

$$\beta(\mu) = \inf_{\mathcal{W}\in\mathcal{X}} \frac{\|T^\mu\mathcal{W}\|_{\mathcal{X}}}{\|\mathcal{W}\|_{\mathcal{X}}}, \quad \forall \mu \in \mathcal{D}; \tag{7.60}$$

where for any $\mathcal{W}$ in $\mathcal{X}$, the linear operator $T^\mu\colon \mathcal{W} \to T^\mu\mathcal{W}$ is given as

$$(T^\mu\mathcal{W},\mathcal{V})_{\mathcal{X}} = \mathcal{A}(\mathcal{W},\mathcal{V};\mu), \quad \forall\, \mathcal{V} \in \mathcal{X}\ . \tag{7.61}$$

We can then define a symmetric positive-semidefinite eigenproblem related to the (square of the) singular values of our partial differential operator: given $\mu \in \mathcal{D}$, $(\Phi_i(\mu) \in \mathcal{X},\ \sigma_i(\mu) \in \mathbb{R})$, $i = 1,\ldots,\mathcal{N} \equiv dim(\mathcal{X})$, satisfies

$$(T^\mu\Phi_i(\mu), T^\mu\mathcal{V})_{\mathcal{X}} = \sigma_i(\mu)\,(\Phi_i(\mu),\mathcal{V})_{\mathcal{X}}, \quad \forall\, \mathcal{V}\ \in \mathcal{X}\ ; \tag{7.62}$$

the eigenvalues are ordered such that $0 \leq \sigma_1 \leq \sigma_2 \leq \cdots \sigma_{\mathcal{N}}$. We normalize our eigenfunctions as $\|\Phi_i(\mu)\|_{\mathcal{X}} = 1$, $i = 1, \ldots, \mathcal{N}$, and hence orthogonality reads

$$(T^\mu \Phi_i(\mu), T^\mu \Phi_j(\mu))_{\mathcal{X}} = \sigma_i \, (\Phi_i(\mu), \Phi_j(\mu))_{\mathcal{X}} = \sigma_i \delta_{ij}, \quad 1 \leq i, j \leq \mathcal{N}, \tag{7.63}$$

where δ_{ij} is the Kronecker-delta symbol.

We may then identify $\beta(\mu) \equiv \sqrt{\sigma_1(\mu)}$, where $\beta(\mu)$ is the classical inf-sup parameter. Note that $\mathcal{V} = T^\mu \mathcal{W}$ is the supremizer in (7.60).

We now consider the algebraic system of the eigenproblem. We first identify the matrix representations of the different bilinear forms. Let $\underline{X} \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$ be the finite element matrix for the $(\cdot, \cdot)_{\mathcal{X}}$ norm — note that the $\mathcal{X}$-norm is μ-independent. We denote $\underline{B} \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$ the Cholesky factorizer of $\mathcal{X}$; thus

$$\underline{B}^H \, \underline{B} = \underline{X}, \tag{7.64}$$

note that the μ-independence of $\mathcal{X}$ implies we only need to do the Cholesky factorization *once* and store the matrix $\underline{B}$.

We denote $\underline{A}(\mu) \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$ as the μ-dependent finite element matrix $A_{ij}(\mu) = \mathcal{A}(\zeta_j, \zeta_i; \mu)$, $\zeta_i, \zeta_j \in \mathcal{X}$. We also introduce $\underline{T^\mu}$ as the supremizing matrix associated with the supremizing operator $T^\mu : \mathcal{W} \rightarrow T^\mu \mathcal{W}$. The algebraic system corresponding to (7.61) is

$$\underline{X} \, \underline{T^\mu} \, \underline{\mathcal{W}} = \underline{A}(\mu) \, \underline{\mathcal{W}}; \tag{7.65}$$

here $\underline{\mathcal{W}}$ is the vector corresponding to $\mathcal{W} \in \mathcal{X}$.

We can then write down the algebraic system corresponding to the eigenproblem as

$$\underline{T^\mu}^H \, \underline{X} \, \underline{T^\mu} \, \underline{\phi_i}(\mu) = \sigma_i(\mu) \, \underline{X} \, \underline{\phi_i}(\mu), \; i = 1, \ldots, \mathcal{N}; \tag{7.66}$$

were $\underline{\phi_i}$, $i = 1, \ldots, \mathcal{N}$ are the eigenvectors and $\sigma_i(\mu)$, $i = 1, \ldots, \mathcal{N}$ the eigenvalues of the eigenproblem. We now make use of (7.65) to re-write (7.66) as

$$\underline{A}^H(\mu) \, \underline{X}^{-1} \, \underline{A}(\mu) \, \underline{\phi_i} = \sigma_i \, \underline{X} \, \underline{\phi_i}. \tag{7.67}$$

We now expand $\underline{X}$ as $\underline{X} = \underline{B}^H \underline{B}$, and substitute $\underline{B}\,\underline{\phi_i} = \underline{\hat{\phi}_i}$ to obtain the system

$$\underline{B}^{-H}\,\underline{A}^H(\mu)\,\underline{B}^{-1}\,\underline{B}^{-H}\,\underline{A}(\mu)\underline{B}^{-1}\underline{\hat{\phi}_i} = \sigma_i\,\underline{\hat{\phi}_i}. \tag{7.68}$$

We denote $\hat{A} = \underline{B}^{-H}\,\underline{A}^H(\mu)\,\underline{B}^{-1}\,\underline{B}^{-H}\,\underline{A}(\mu)\underline{B}^{-1}$ and note that $\hat{A}$ is Hermitian (i.e., $\hat{A}^H = \hat{A}$). We now consider the Hermitian eigenvalue problem (HEP) associated with (7.60)

$$\hat{A}x = \lambda x; \tag{7.69}$$

here $x \equiv \underline{\hat{\phi}_i}$ and $\lambda \equiv \sigma_i(\mu)$ for $i = 1, \ldots, \mathcal{N}$.

We are interested in only the first eigenvalue of the Hermitian eigenproblem in (7.69); the first eigenvalue is the square of the inf-sup stability constant. For the Helmholtz problem: (i) the spectrum for the lowest eigenvalues is "well-spaced" — thus we do not require any spectral transformation for rapid convergence and therefore we do not require any decomposition or inversion of the parameter-dependent and typically non-coercive matrix $\underline{A}(\mu)$ (the Helmholtz operator); and (ii) $\underline{B}$ is parameter independent, and hence the expensive Cholesky can be formed once; subsequently we need to effect only back/forward substitutions.

The Lanczos method [9, 106] is well-suited to the task of finding the *minimum* eigenvalue $\lambda_{\min}$ of the eigenproblem (7.69). The key to rapid Lanczos convergence is a good "relative" spacing between the first few eigenvalues (in our case, singular values squared); fortunately for the Helmholtz problem this occurs quite naturally.

We present in Figure 7-1 the Lanczos algorithm [106]. In Steps 4a–4c, we compute the residual vector and the mutual orthogonalized bases $Q_n = \{q_1, \ldots, q_n\}$. In step 5a. we update the tridiagonal matrix T_n from T_{n-1}. In steps 5b. and 5c. we compute the first N_e eigenvalues Λ_n and approximate eigenvectors X_n. In step 6. we check for convergence. We see that the Lanczos iteration works by replacing the eigenproblem (7.69) with a much simpler eigenproblem (associated with T_n) that approximates well (certain part of) the spectrum.

The Lanczos iteration is the Arnoldi iteration [9] specialized to the case when $\hat{A}$ is Hermitian. The algorithm builds up an orthogonal basis Q_n of the Krylov subspace $\text{span}\{q, \hat{A}q, \ldots, \hat{A}^{n-1}q\}$. In the new orthogonal basis Q_n, the operator $\hat{A}$ is replaced by the real symmetric tridiagonal matrix T_n. At the "n"-th step, we compute the first N_e eigensolutions of $T_n\,S_n = S_n\,\lambda_n$; the Ritz value λ_j

and its associated Ritz vector $X_j = Q_n(:, 1 : N_e)S_j$ will be a good approximation to an eigenpair of $\hat{A}$ if the residual has small norm. The residual for the Ritz pair is given by

$$r_j = \hat{A}X_j - \lambda_j X_j \tag{7.70}$$

$$\|r_j\| = \beta_n |S_{n,j}|; \tag{7.71}$$

$\|r_j\|, j = 1, \ldots, N_e$ is a *rigorous* bound on the error between the exact and approximate eigenvalues [9] only if the nearest truth eigenvalues correspond to the first N_e eigenvalues. We can also define rigorous lower bounds $\lambda_{j,\mathrm{LB}}$ based on the Ritz value and the error bound as

$$\lambda_{j,\mathrm{LB}} = \lambda_j - \|r_j\|; \tag{7.72}$$

these lower bounds could be used in the *a posteriori* error estimators. Finally, we compute a ratio ϵ_j, $j = 1, \ldots, N_e$, which if less than unity, assures a *relative* accuracy of ϵ_{tol}; we stop the iteration when $\max_{j=1,\ldots,N_e} \epsilon_j \leq 1$.

7.5 Examples

7.5.1 Model Problem: 1-D Helmholtz

We consider the weak statement for a model problem $\mathcal{A}(\mathcal{U}, \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V})$, $\forall \mathcal{V} \in \mathcal{X}$ on the reference domain $\Omega =]0, 1[$. We choose our function space $\mathcal{X} = \mathcal{X}^e = H_0^1(\Omega)$, with associated inner product $(\mathcal{W}, \mathcal{V})_{\mathcal{X}} = \int_\Omega \mathcal{W}_x \mathcal{V}_x$ and norm $\|\mathcal{W}\|_{\mathcal{X}} = |\mathcal{W}|_{H^1(\Omega)}$.

Bilinear Form and Affine Decomposition

Our bilinear form on the reference domain is given by

$$\mathcal{A}(\mathcal{W}, \mathcal{V}; \mu) = \int_0^1 \mathcal{W}_x \, \mathcal{V}_x \; - \; \mu \, \mathcal{W} \, \mathcal{V}, \; \forall \mathcal{W}, \mathcal{V} \in \mathcal{X}. \tag{7.73}$$

We thus identify $Q_{\mathcal{A}} = 2$, $\Theta^1_{\mathcal{A}}(\mu) = 1$, $\mathcal{A}^1(\mathcal{W}, \mathcal{V}) = \int_0^1 \mathcal{W}_x \mathcal{V}_x$, $\Theta^2_{\mathcal{A}}(\mu) = -\mu$ and $\mathcal{A}^2(\mathcal{W}, \mathcal{V}) = \int_0^1 \mathcal{W} \, \mathcal{V}$.

1. Set $\beta_0 = 0, q_0$
1. Set b = arbitrary, $q_1 = \frac{b}{\|b\|}, Q(:,1) = q_1$
3. **for** $n = 1, 2, \ldots,$ **until** convergence
4a. $\quad v = \hat{A}\, q_n - \beta_{n-1}\, q_{n-1}$
4b. $\quad \alpha_n = q_n^H\, v$
4c. $\quad v = v - \alpha_n\, q_n$
4d. $\quad \beta_n = \sqrt{v^H\, v}$
4d. $\quad q_{n+1} = v/\beta_n$
4e. $\quad Q(:, n+1) = q_{n+1}$
5a. $\quad T_n =$
$$\begin{pmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{n-1} \\ & & & \beta_{n-1} & \alpha_n \end{pmatrix}$$
5b. $\quad$ Compute N_e eigenvalues Λ_n from $T_n = S_n\, \Lambda_n\, S_n^H$
5c. $\quad$ Compute N_e eigenvectors $X_n = Q(:, 1 : N_e)\, S_n$
6. $\quad$ Test for convergence
6a. $\quad$ **for** $j = 1, \ldots, N_e$
6b. $\quad\quad \|r_j\| = \beta_n\, \|S_{n,j}\|$
6c. $\quad\quad \epsilon_j = \frac{\|r_j\|\,(1+\epsilon_{\text{tol}})}{\epsilon_{\text{tol}}\,\lambda_j}$
6d. $\quad$ **end for**
6e. $\quad$ **if** $\max \epsilon_j \leq 1$
6f. $\quad\quad$ **exit**
6g. $\quad$ **end if**
7. **end for**

Figure 7-1: Lanczos Algorithm for HEP [106].

Eigenfunctions and Eigen values

We describe the operators by means of their eigenfunctions and eigenvalues. We expand $\mathcal{U}$ as

$$\mathcal{U}_m = \alpha_m \sin(m\,\pi\,x),\ m = 1, 2, \ldots \tag{7.74}$$

where the scalar α_m is chosen such that

$$\|\mathcal{U}_m\|_X \equiv |\mathcal{U}_m|_{H^1(\Omega)} = 1. \tag{7.75}$$

We denote δ_{mn} as the Dirac delta function. We have that

$$\begin{aligned}
\mathcal{A}^1(\mathcal{U}_m, \mathcal{U}_n) &= \alpha_m^2\, m^2\, \frac{\pi^2}{2}\, \delta_{mn} && (7.76)\\
\mathcal{A}^2(\mathcal{U}_m, \mathcal{U}_n) &= \alpha_m^2\, \frac{1}{2}\, \delta_{mn} && (7.77)\\
\mathcal{A}(\mathcal{U}_m, \mathcal{U}_n; \mu) &= \frac{\alpha_m^2}{2}\, (m^2\,\pi^2 - \mu)\, \delta_{mn}. && (7.78)
\end{aligned}$$

Thus, $T^\mu \mathcal{U}_m$ is given by

$$T^\mu \mathcal{U}_m = \frac{\alpha_m^2}{2}\, (m^2\,\pi^2 - \mu)\, \mathcal{U}_m \tag{7.79}$$

from $(T^\mu \mathcal{U}, \mathcal{V})_\mathcal{X} = \mathcal{A}(\mathcal{U}, \mathcal{V}; \mu), \forall \mathcal{V} \in \mathcal{X}$. Similarly, we identify

$$\begin{aligned}
T^1 \mathcal{U}_m &= \frac{\alpha_m^2}{2}\, m^2\, \pi^2\, \mathcal{U}_m, \\
T^2 \mathcal{U}_m &= \frac{\alpha_m^2}{2}\, \mathcal{U}_m. && (7.80)
\end{aligned}$$

We can then write $\sigma(\mu)$ as

$$\sigma(\mu) = \min_{m=1,2,\ldots} \frac{|m^2\,\pi^2 - \mu|^2}{m^4\,\pi^4} \tag{7.81}$$

Parametrization

We choose $\mu \in \mathcal{D} \equiv [\mu_{\min}, \mu_{\max}]$ with $\pi^2 < \mu_{\min} < \mu_{\max} < (2\,\pi)^2$ to avoid the resonances. We specify $\mu_{\min} = \pi^2 + 2$ and $\mu_{\max} = 4\,\pi^2 - 2$.

We identify $\sigma(\mu)$ exactly as

$$\sigma(\mu) = \min\left(\frac{|\pi^2 - \mu|^2}{\pi^4}, \frac{|4\,\pi^2 - \mu|^2}{256\,\pi^4}\right). \tag{7.82}$$

For this problem, we have $\overline{Q} = Q_{\mathcal{A}}(Q_{\mathcal{A}} + 1)/2 = 3$; since we have no complex terms in our bilinear form we also have $\hat{\mathcal{Q}} = 3$. We choose the mappings

$$\begin{aligned}
\mathrm{Ind}(q = 1, q' = 1) &= 1, \\
\mathrm{Ind}(q = 1, q' = 2) &= 2 \\
\mathrm{Ind}(q = 2, q' = 1) &= 2 \\
\mathrm{Ind}(q = 2, q' = 2) &= 3.
\end{aligned} \tag{7.83}$$

We can thus calculate (in this case exactly)

$$\begin{aligned}
\gamma_1^- &= \inf_{w\in X} \frac{(T^1\mathcal{W}, T^1\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|_{\mathcal{X}}} &= 1 \\
\gamma_1^+ &= \sup_{w\in X} \frac{(T^1\mathcal{W}, T^1\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|_{\mathcal{X}}} &= 1 \\
\gamma_2^- &= \inf_{w\in X} \frac{(T^1\mathcal{W}, T^2\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|_{\mathcal{X}}} &= 0 \\
\gamma_2^+ &= \sup_{w\in X} \frac{(T^1\mathcal{W}, T^2\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|_{\mathcal{X}}} &= \tfrac{1}{\pi^2} \\
\gamma_3^- &= \inf_{w\in X} \frac{(T^2\mathcal{W}, T^2\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|_{\mathcal{X}}} &= 0 \\
\gamma_3^+ &= \sup_{w\in X} \frac{(T^2\mathcal{W}, T^2\mathcal{W})_{\mathcal{X}}}{\|\mathcal{W}\|_{\mathcal{X}}} &= \tfrac{1}{\pi^4};
\end{aligned}$$

by dropping the Θ_I^q related-terms in the expansions (7.17)– (7.37).

We also obtain $\Phi_1(\mu) = 1$, $\Phi_2(\mu) = -2\,\mu$ and $\Phi_3(\mu) = \mu^2$. Thus, we define our lower bound

$$\sigma_{\mathrm{LB}}(\mu) = \inf_{\underline{z}\in\mathcal{S}_{\mathrm{relax}}} \sum_{q=1}^{3} \Phi_q(\mu) z_q, \tag{7.84}$$

where the set $\mathcal{S}_{\text{relax}}$ is given by

$$\begin{aligned}
\mathcal{S}_{\text{relax}}(\mu; \mathcal{C}_K, \Xi_J) \quad \equiv \quad & \Big\{ \underline{z} = (z_1, z_2, z_3) \in \mathbb{R}^3 \mid \\
& z_1 = 1 \\
& 0 \le z_2 \le \frac{1}{\pi^2}, \\
& 0 \le z_3 \le \frac{1}{\pi^4}, \\
& \sum_{r=1}^{3} \Phi_r(\omega')\, z_r \ge \sigma(\omega'),\ \forall \omega' \in P_{\mathcal{M}_\alpha}(\mu; \mathcal{C}_K), \\
& \sum_{r=1}^{3} \Phi_r(\nu')\, z_r \ge 0,\ \forall \nu' \in P_{\mathcal{M}_+}(\mu; \Xi_J) \Big\}. \qquad (7.85)
\end{aligned}$$

Thus, (7.84) constitutes a Linear Program (LP) in $\hat{\mathcal{Q}} = \overline{\mathcal{Q}} = 3$ variables with $2\,\hat{\mathcal{Q}} + \mathcal{M}_\alpha + \mathcal{M}_+$ constraints.

Spectrum of Inf-Sup Eigenproblem

We now present the first three singular values of the eigenproblem for $\sigma(\mu)$ in Figure 7-2. Note that $\sigma_1(\mu)$ and $\sigma_2(\mu)$ have the same value at $\mu \approx 15.8$; similarly $\sigma_2(\mu)$ and $\sigma_3(\mu)$ have the same value at $\mu \approx 17.76$. These two points represent "mode crossing" points — for example, $\mu \approx 15.8$ is the point where the first and second modes become close, then identical (multiplicity two), and then cross over such that the first mode becomes the second mode and vice versa.

Normally, we would have to consider the block Lanczos [41] algorithm (with block size 2) at sample points which are near "mode crossings". However, the regular Lanczos algorithm works well even near mode crossings. In Table 7.1, we show the rapid convergence of the Lanczos algorithm to the first two singular values across the range of μ. For a wide range of μ, we show the first two singular values, the number of Lanczos iterations, and the time required for the Lanczos algorithm. Note that even for samples near the mode crossing at $\mu \approx 15.8$, the Lanczos algorithm requires roughly the same number of iterations; thus, even when the relative gap between the two singular values is quite small the convergence still remains rapid.

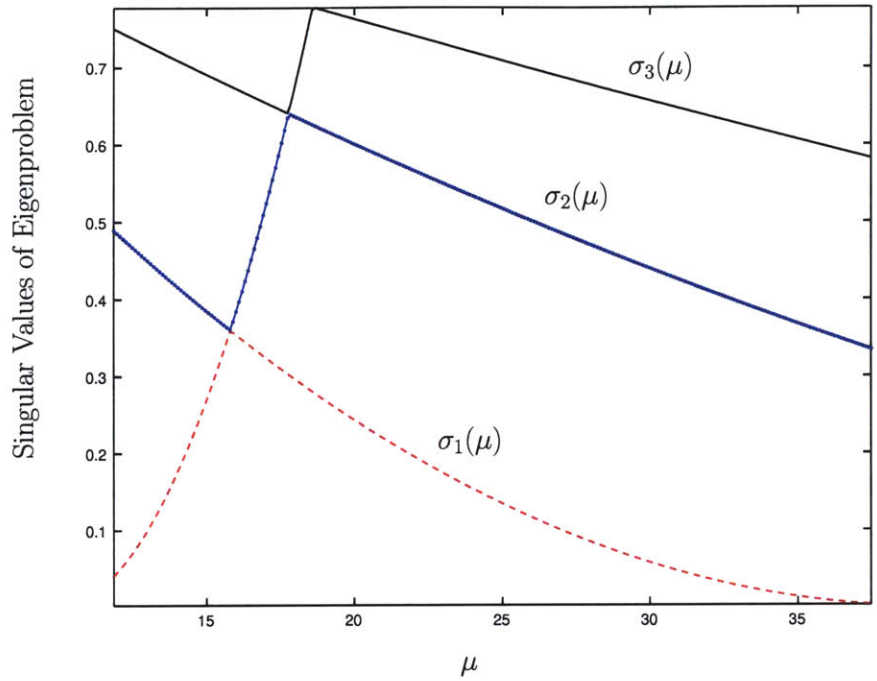


Figure 7-2: The first three singular values of the eigenproblem. Note the "mode crossings" near $\mu = 15.8$ and $\mu = 17.76$.

μ	$\sigma_1(\mu)$	$\sigma_2(\mu)$	n_{iter}	T_{lanc}
$\pi^2 + 2$	0.0411	0.4891	7	0.05
12.7952	0.0879	0.4568	7	0.05
15.5721	0.3338	0.3667	7	0.04
15.6749	0.346	0.3636	7	0.05
15.7778	0.3583	0.3604	8	0.04
15.8806	0.3573	0.3709	7	0.05
15.9835	0.3542	0.3837	7	0.04
16.9091	0.3268	0.5087	7	0.04
22.0514	0.1949	0.5651	8	0.05
27.1938	0.0968	0.4814	8	0.06
32.3361	0.0327	0.4045	9	0.06
37.4784	0.0026	0.3342	11	0.08

Table 7.1: The number of Lanczos iterations required to satisfy a convergence criteria of $\epsilon_{\text{tol}} = 0.005$. n_{iter} is the number of Lanczos iterations, and T_{lanc} is the time required for the Lanczos algorithm in seconds. The size of the problem is $\mathcal{N} = 500$.

Results

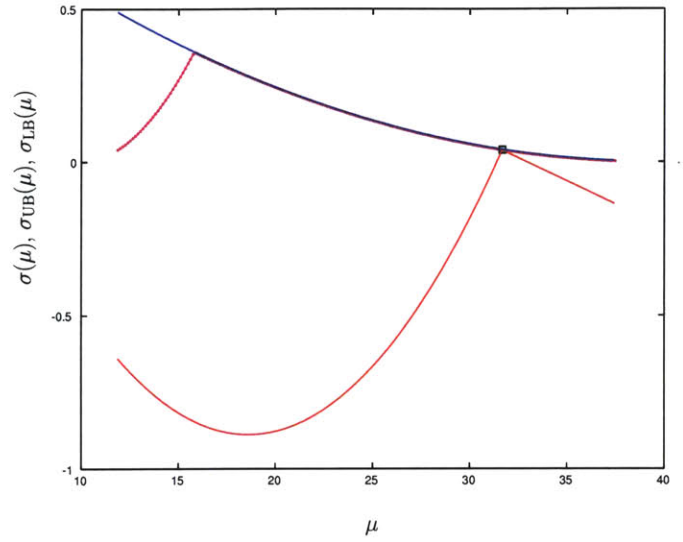
We now apply the SCM greedy algorithm over a fine sample-set Ξ_{train}. We choose $\mathcal{M}_\alpha = 5$ and $\mathcal{M}_+ = 0$ and run the greedy algorithm with a termination tolerance of $\epsilon_{\text{tol,min}} = 0.75$. We need $K = 5$ to satisfy our required tolerance; for the samples in Ξ_{train} we obtain $\beta_{\text{LB}}(\mu) \leq \beta(\mu) \leq 1.41\,\beta_{\text{LB}}(\mu)$.

Since K is very small, we choose to show the upper and lower bounds obtained at each iteration, and compare with the exact $\sigma(\mu)$. Figure 7-3 shows the SCM quantitites for $k = 1$ and $k = 2$, Figure 7-4 for $k = 3$ and $k = 4$, and Figure 7-5 for $k = K = 5$.

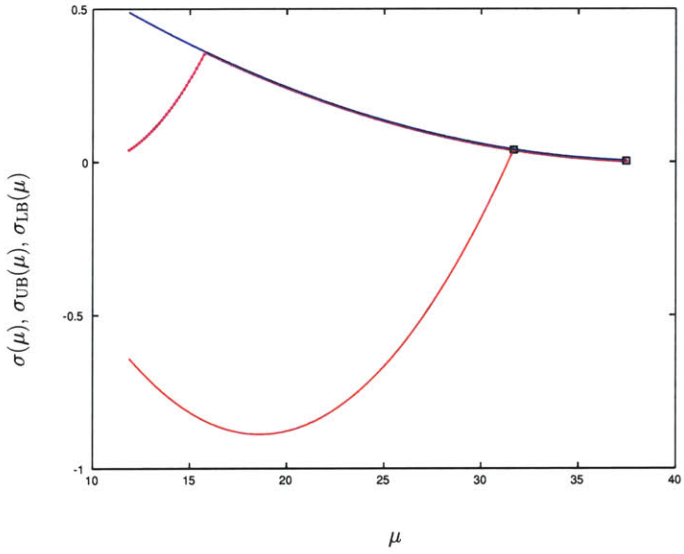
We specify $\mathcal{M}_\alpha = 5$ and $\mathcal{M}_+ = 0$. Figures 7-3, 7-4, and 7-5 show the rapid convergence of the lower bounds $\sigma_{\text{LB}}(\mu)$ to the exact $\sigma(\mu)$. Also note that the upper bound converges extremely fast (infact much faster relative to the lower bound) to the exact $\sigma(\mu)$ — even for $k = 1$, for the most part $\sigma_{\text{UB}}(\mu) = \sigma(\mu)$. Thus, the upper bound is an extremely sharp surrogate for $\sigma(\mu)$.

We need $K = 5$ to satisfy our required tolerance of $\epsilon_{\text{tol,min}} = 0.75$; we then have $\beta_{\text{LB}}(\mu) \leq \beta(\mu) \leq 1.68\,\beta_{\text{LB}}(\mu)$. From Figure 7-5 we can see that our lower bound is extremely "close" to the inf-sup throughout the domain. For more difficult problems, we will have to relax $\epsilon_{\text{tol,min}}$ and increase $\mathcal{M}_\alpha$ and $\mathcal{M}_+$: we discuss these issues in greater detail for our next example.

However, it is clear that the method works in a "naive" setting; although we still need to see how the method scales with increased number of parameters, P, and larger number of affine terms, Q. While it is difficult to say how the algorithm will perform in higher parameter dimensions, there is one aspect that gives us hope: the greedy algorithm is essetially trying to close the "gap" between the upper and lower bounds – since the upper bound is a good approximation for the exact $\sigma(\mu)$, we expect that when the "gap" is closed the lower bound also has to be very good. Furthermore, relative to a SCM constraint point $\omega_k \in \mathcal{C}_K$, we expect the upper bounds to be good in the local neighborhood; if between two samples, one sample has a much worse lower bound then the gap will also be much worse for that point since the upper bounds are well approximated in the local neighborhood and roughly will be of the same order. We thus hope that the SCM algorithm will perform well in higher dimensions also.

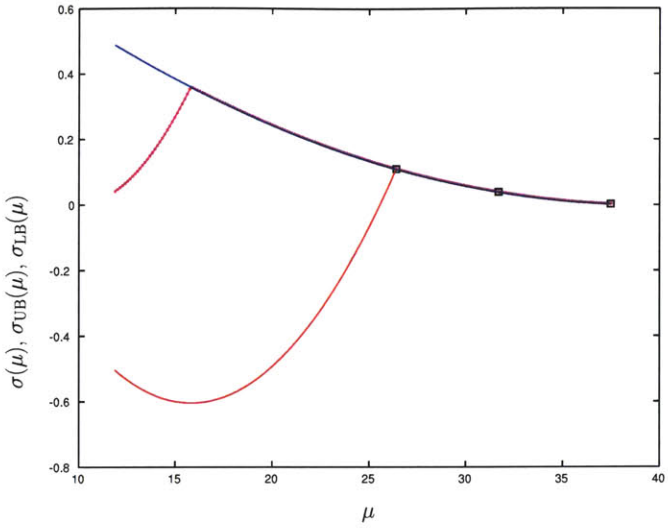


(a)

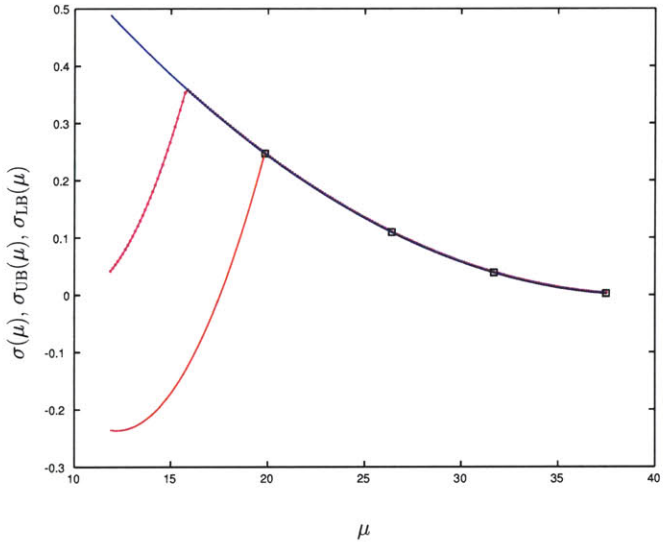


(b)

Figure 7-3: Construction of $\mathcal{C}_K$: we show the exact $\sigma(\mu)$ in magenta, the upper bound $\sigma_{\mathrm{UB}}(\mu)$ in blue, (a) the lower bound $\sigma_{\mathrm{LB}}(\mu)$ in red, and the SCM constraint points $\omega_k \in \mathcal{C}_K$ as black squares for (a) $k = 1$ and (b) $k = 2$. The maximum ratio, $\max_{\mu \in \Xi_{\mathrm{train}}} \frac{\sigma_{\mathrm{UB}}(\mu) - \sigma_{\mathrm{LB}}(\mu)}{\sigma_{\mathrm{UB}}(\mu)}$, is 54.23 with $k = 1$ constraint, and 6.1 with $k = 2$ constraints.



(a)



(b)

Figure 7-4: Construction of $\mathcal{C}_K$: we show the exact $\sigma(\mu)$ in magenta, the upper bound $\sigma_{\mathrm{UB}}(\mu)$ in blue, (a) the lower bound $\sigma_{\mathrm{LB}}(\mu)$ in red, and the SCM constraint points $\omega_k \in \mathcal{C}_K$ as black squares for (a) $k = 3$ and (b) $k = 4$. The maximum ratio, $\max_{\mu \in \Xi_{\mathrm{train}}} \frac{\sigma_{\mathrm{UB}}(\mu) - \sigma_{\mathrm{LB}}(\mu)}{\sigma_{\mathrm{UB}}(\mu)}$, is 3.03 with $k = 3$ constraints, and 1.52 with $k = 4$ constraints.

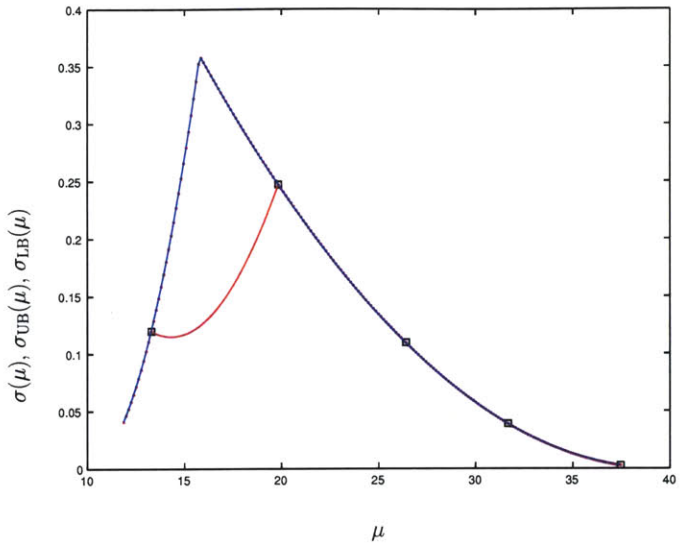


Figure 7-5: Construction of $\mathcal{C}_K$: we show the exact $\sigma(\mu)$ in magenta, the upper bound $\sigma_{\rm UB}(\mu)$ in blue, (a) the lower bound $\sigma_{\rm LB}(\mu)$ in red, and the SCM constraint points $\omega_k \in \mathcal{C}_K$ as black squares for $k = 5$. We have $\max_{\mu \in \Xi_{\rm train}} \frac{\sigma_{\rm UB}(\mu) - \sigma_{\rm LB}(\mu)}{\sigma_{\rm UB}(\mu)} = 0.65$ for $k = 5$; thus $\dim(\mathcal{C}_K) = K = 5$.

7.5.2 Low-pass Filter ($P = 2$)

Problem Statement

We consider the expanded outlet system for the low-pass filter described extensively in 4.4. The weak statement for the problem on the reference domain Ω is $a(u(\mu), v; \mu) = f(v), \forall v \in X, \forall \mu \in \mathcal{D}$; X is a complex Hilbert space with associated inner product $(\cdot, \cdot)_X$ (5.4) and norm $\|\cdot\|_X$. Here $u(\mu)$ is the abstract field variable representing in this case the pressure p ($u(\mu) \equiv p$).

Bilinear Form and Affine Decomposition

We fix $Z_R = 2$ and $Z_I = 0$; $\mu \equiv (k^2, H)$ and a is separable as shown in Table 7.2. The bilinear and linear forms admit affine separations for $Q_a = 7$ and $Q_f = 1$.

We denote the pressure intensity over the outlet as our quadratic output of interest

$$Q_{\rm out}(u(\mu), u(\mu)) = \int_{\Gamma_{\rm out}} \overline{u}\, u. \tag{7.86}$$

The expanded weak system for the outlet pressure intensity can then be posed: given $\mu \in \mathcal{D} \in$

q	Θ_a^q	$a^q(u,v)$
1	1	$\int_{\Omega_{1,2}} \nabla u \nabla \overline{v}$
2	H	$\int_{\Omega_3} \frac{\partial u}{\partial x} \frac{\partial \overline{v}}{\partial x}$
3	$\frac{1}{H}$	$\int_{\Omega_3} \frac{\partial u}{\partial y} \frac{\partial \overline{v}}{\partial y}$
4	$-k^2$	$\int_{\Omega_{1,2}} u \overline{v}$
5	$-k^2 H$	$\int_{\Omega_3} u \overline{v}$
6	$\frac{\mathrm{i}k}{Z_R}$	$\int_{\Gamma_{\text{liner}}} u \overline{v}$
7	$\mathrm{i}k$	$\int_{\Gamma_{\text{out}}} \Xi_1 \overline{v} \int_{\Gamma_{\text{out}}} u \overline{\Xi}_1$

Table 7.2: Affine decomposition of $a(u(\mu), v; \mu) = \sum_{q=1}^{Q_a} \Theta_a^q(\mu)\, a^q(w,v)$ for the two-parameter ($\mu \equiv (k^2, H)$) Low-Pass Filter Problem.

$\mathbb{R}^2$, we find our "compliance" output of interest

$$\mathcal{S}(\mu) = \overline{\mathcal{F}(\mathcal{U}(\mu); \mu)}, \tag{7.87}$$

where $\mathcal{U}(\mu) = [U^+, U^-] \in \mathcal{X}$ satisfies

$$\mathcal{A}_{\text{out}}(\mathcal{U}(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}; \mu), \quad \forall \mathcal{V} = [V^+, V^-]^T \in \mathcal{X}, \tag{7.88}$$

where

$$\begin{aligned}
\mathcal{A}_{\text{out}}(\mathcal{U}(\mu), \mathcal{V}; \mu) \;=\; & a(U^+, V^+; \mu) + \overline{a(V^+, U^+; \mu)} \\
& + a(U^-, V^+; \mu) - \overline{a(V^+, U^-; \mu)} \\
& + \overline{a(V^-, U^+; \mu)} - a(U^+, V^-; \mu) \\
& - \overline{a(V^-, U^-; \mu)} - a(U^-, V^-; \mu) \\
& - Q_{\text{out}}(U^+, V^+; \mu) - Q_{\text{out}}(U^-, V^+; \mu) \\
& - Q_{\text{out}}(U^+, V^-; \mu) - Q_{\text{out}}(U^-, V^-; \mu),
\end{aligned} \tag{7.89}$$

and $\mathcal{F}$ as

$$\mathcal{F}(\mathcal{V}; \mu) \;=\; f(V^+; \mu) - f(V^-; \mu), \tag{7.90}$$

q	$\Theta^q_{\mathcal{A}}$	$\mathcal{A}^q_{\text{out}}(\mathcal{U},\mathcal{V})$
$q=1$	1	$a^1(U^+,V^+;\mu)+\overline{a^1(V^+,U^+;\mu)}$ $+a^1(U^-,V^+;\mu)-\overline{a^1(V^+,U^-;\mu)}$ $+\overline{a^1(V^-,U^+;\mu)}-a^1(U^+,V^-;\mu)$ $-\overline{a^1(V^-,U^-;\mu)}-a^1(U^-,V^-;\mu)$ $-Q_{\text{out}}(U^+,V^+;\mu)-Q_{\text{out}}(U^-,V^+;\mu)$ $-Q_{\text{out}}(U^+,V^-;\mu)-Q_{\text{out}}(U^-,V^-;\mu)$
$q\neq 1$	Θ^q	$a^q(U^+,V^+;\mu)+\overline{a^q(V^+,U^+;\mu)}$ $+a^q(U^-,V^+;\mu)-\overline{a^q(V^+,U^-;\mu)}$ $+\overline{a^q(V^-,U^+;\mu)}-a^q(U^+,V^-;\mu)$ $-\overline{a^q(V^-,U^-;\mu)}-a^q(U^-,V^-;\mu)$

Table 7.3: Affine decomposition of $\mathcal{A}_{\text{out}}(\mathcal{U}(\mu),\mathcal{V};\mu) = \sum_{q=1}^{Q_{\mathcal{A}}}\Theta^q_{\mathcal{A}}(\mu)\,\mathcal{A}^q_{\text{out}}(\mathcal{U}(\mu),\mathcal{V})$ for the 2-parameter Low-Pass Filter Outlet System

$\mathcal{A}_{\text{out}}$ and $\mathcal{F}$ are also "affine" in the parameter, we recover $Q_{\mathcal{A}}=7$ and $Q_{\mathcal{F}}=1$, each a^q term can be easily corresponded with its associated $\mathcal{A}^q_{\text{out}}$ term as shown in Table 7.3

Parametrization

We choose $\mu\in\mathcal{D}\equiv[0.5,5]\times[0.75,1.25]$. For this problem, $\Theta^q_{\mathcal{A},I}=0, 1\le q\le 5$ and $\Theta^q_{\mathcal{A},R}=0, q=6,7$. We identify $\hat{\mathcal{Q}}=28$ and we choose the mappings as shown in Table 7.4.

We also obtain $\Phi_r(\mu), r=1,\ldots,\hat{\mathcal{Q}}$ using the indexing scheme in Table 7.4. We then define our lower bound

$$\sigma_{\text{LB}}(\mu)=\inf_{z\in\mathcal{S}_{\text{relax}}}\sum_{q=1}^{28}\Phi_q(\mu)z_q, \tag{7.91}$$

where the set $\mathcal{S}_{\text{relax}}$ is given in (7.45). Thus, (7.91) constitutes a Linear Program (LP) in $\hat{\mathcal{Q}}=28$ variables with $2\,\hat{\mathcal{Q}}+\mathcal{M}_\alpha+\mathcal{M}_+$ constraints.

Ind(q, q', R)							
	1	2	3	4	5	6	7
1	1	2	3	4	5		
2		6	7	8	9		
3			10	11	12		
4				13	14		
5					15		
6						16	17
7							18
Ind(q, q', I)							
1						19	20
2						21	22
3						23	24
4						25	26
5						27	28

Table 7.4: Index Mappings for the 2-parameter Low-Pass Filter Outlet System: Ind(q, q', R), $1 \le q \le 7$, $q \le q' \le 7 \longmapsto r = 1, \ldots, 18$; and Ind$(q, q', I)$, $1 \le q \le 7$, $q \le q' \le 7 \longmapsto r = 19, \ldots, 28$.

Results

For this problem,we use the ratio $\frac{\sigma_{\text{UB}}(\mu)-\sigma_{\text{LB}}(\mu)}{\sigma_{\text{LB}}(\mu)}$ in the construction algorithm for $\mathcal{C}_K$. We specify $\mathcal{M}_\alpha = 20$, $\mathcal{M}_+ = 10$ and $\epsilon_{\text{tol,min}} = 0.75$ and aim to find the set $\mathcal{C}_K$ from our exhaustive sampling set $\Xi_{\text{train}} \subset \mathcal{D}$ such that $\max_{\mu \in \Xi_{\text{train}}} \frac{\sigma_{\text{UB}}(\mu;\mathcal{C}_K)-\sigma_{\text{LB}}(\mu;\mathcal{C}_K)}{\sigma_{\text{UB}}(\mu;\mathcal{C}_K)} \le \epsilon_{\text{tol,min}}$. Figures 7-6 and 7-7 show slices of the obtained lower and upper bounds at different stages in the offline algorithm for different values of k. We reach our desired tolerance for $K = 210$. Note that our upper bounds $\sigma_{\text{UB}}(\mu)$ is almost indistinguishable from the exact inf-sup $\sigma(\mu)$ for a large range of μ — this observation reflects our assumption that for μ near ω_k, the infimizer of ω_k will be a very good approximation for the infimizer associated with μ.

7.6 Efficiency of the Successive Constraint Method

It is essential that we understand the relative importance of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ in obtaining the *rigorous* inf-sup lower bound $\sigma_{\text{LB}}(\mu)$. Modifying $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ affects the size of K obtained in the *offline* stage, and the computational cost of the LP in the *online* stage.

We discuss the efficiency of the successive constraint method relative to the low-pass filter problem discussed in the earlier section. We investigate the effects of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ on the offline and online stages separately.

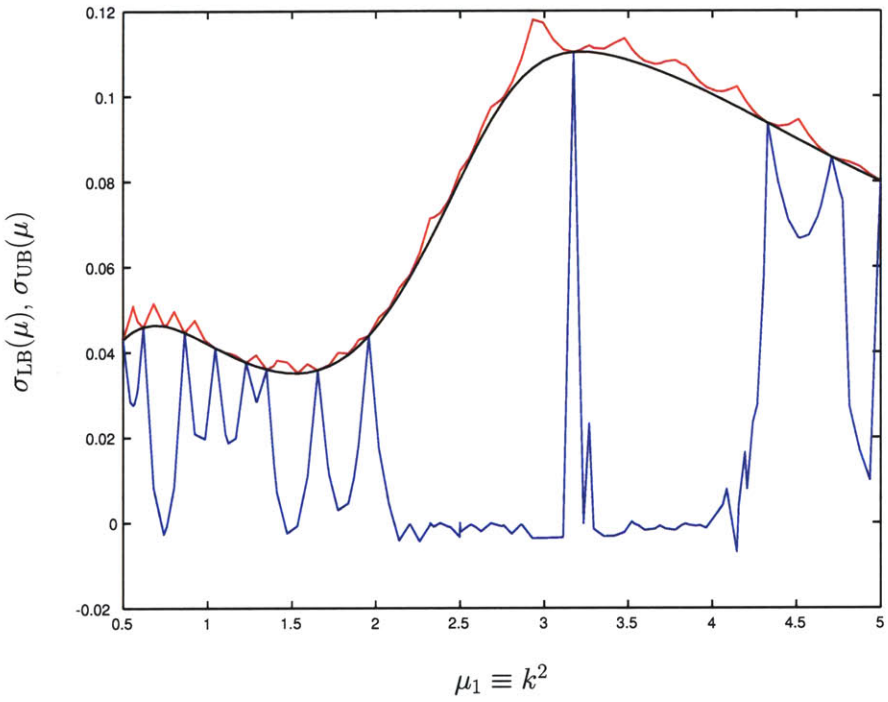


Figure 7-6: Construction of $\mathcal{C}_K$: with $k = 150$ constraints we have a maximum ratio $\max_{\mu \in \mathcal{D}} \frac{\sigma_{\mathrm{UB}}(\mu) - \sigma_{\mathrm{LB}}(\mu)}{\sigma_{\mathrm{UB}}(\mu)} = 1.082$. We show the slice for $H = 0.75$. The lower bound, $\sigma_{\mathrm{LB}}(\mu)$ is shown in blue, the exact inf-sup $\sigma(\mu)$ in black, and the upper bound $\sigma_{\mathrm{UB}}(\mu)$ in red. The points where $\sigma_{\mathrm{LB}}(\mu)$ touches $\sigma(\mu)$ represent our successive constraint points $\omega_k \in \mathcal{C}_K$.

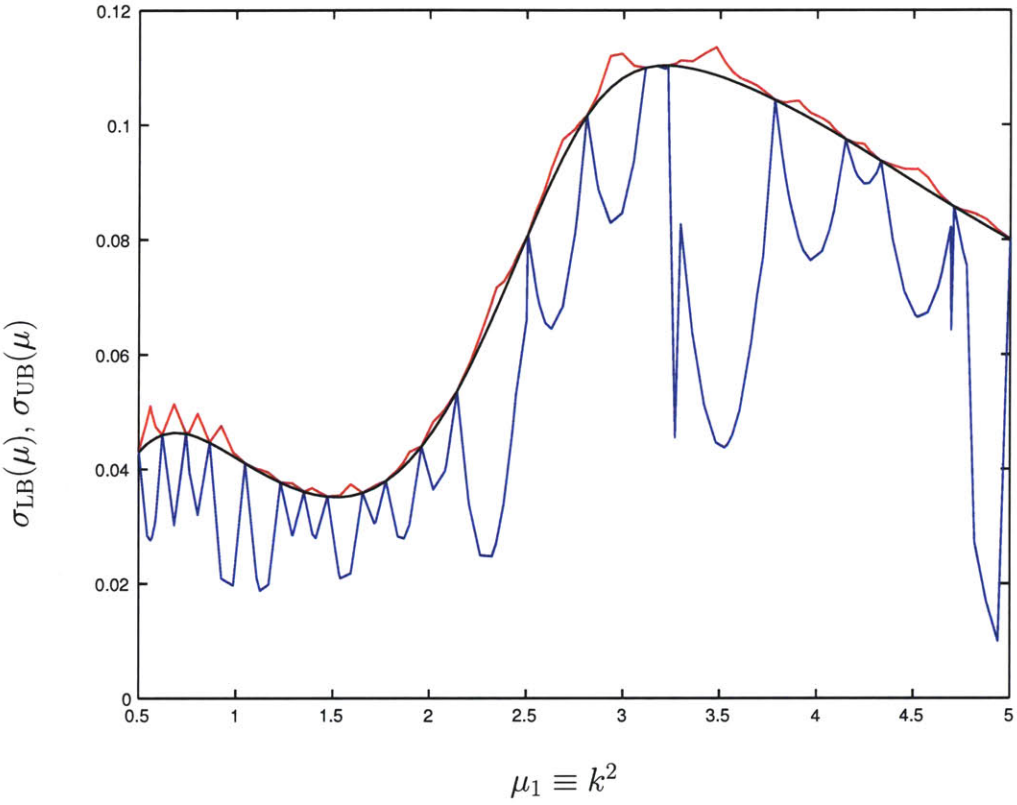


Figure 7-7: Construction of $\mathcal{C}_K$: with $K = 210$ constraints we have a maximum ratio $\max_{\mu \in \mathcal{D}} \frac{\sigma_{\mathrm{UB}}(\mu) - \sigma_{\mathrm{LB}}(\mu)}{\sigma_{\mathrm{UB}}(\mu)} = 0.73$. We show the slice for $H = 0.75$. The lower bound, $\sigma_{\mathrm{LB}}(\mu)$ is shown in blue, the exact inf-sup $\sigma(\mu)$ in black, and the upper bound $\sigma_{\mathrm{UB}}(\mu)$ in red. The points where $\sigma_{\mathrm{LB}}(\mu)$ touches $\sigma(\mu)$ represent our successive constraint points $\omega_k \in \mathcal{C}_K$.

7.6.1 Offline Greedy Algorithm

Intuitively, when we increase $\mathcal{M}_\alpha$ and $\mathcal{M}_+$, we are considering a larger set of constraints and we should obtain better lower bounds that are now computationally more expensive since the associated Linear Program to be solved is of larger size. At the same time, the lower bounds are better and conceivably we will require a smaller constraint set $\mathcal{C}_K$. Thus, we are really trying to optimize the tradeoff of increased cost for running the Linear Programs versus the benefit of a reduced constraint set — the question really is whether increasing $\mathcal{M}_\alpha$ and/or $\mathcal{M}_+$ is really worthwhile, and if so, to what extent. It is not clear what the tradeoffs of M_α and $\mathcal{M}_+$ are on (i) the *offline* cost of computing the set $\mathcal{C}_K$; and (ii) the *online* cost of obtaining lower bounds.

In Table 7.5, we present the values of K, the size of the constraint set $\mathcal{C}_K$, obtained using different values of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$. We consider the reduction of the set $\mathcal{C}_K$ in two ways: first, we fix $\mathcal{M}_+$ and consider the effect of varying $\mathcal{M}_\alpha$ on k; and second, we fix $\mathcal{M}_\alpha$ and consider the effect of variations in $\mathcal{M}_+$ on K. For the sake of consistency, in each of these distinctly different runs of the greedy algorithm, we use the same training set, Ξ_{train} ($\dim(\Xi_{\text{train}}) = 2000$), and the same termination criteria, $\epsilon_{\text{tol,min}} = 0.75$.

$\mathcal{M}_\alpha$	$\mathcal{M}_+$	K
10	5	272
10	10	223
10	15	217
15	5	250
15	10	214
15	15	210
20	5	230
20	10	210
20	15	208

Table 7.5: Low Pass Filter Results: Effect of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ on the size of the successive constraint set, C_K.

Note that for relatively low $\mathcal{M}_\alpha$ (say 10), increasing $\mathcal{M}_+$ is very useful, but that utility quickly tails off as the effect of positivity constraints diminishes. Thus, for $\mathcal{M}_\alpha = 10$, we get a reduction of 49 constraints when we go from 5 to 10 constraints; but a further increase of another 5 positivity constraints only decreases the number of constraints by 6. We observe a similar pattern when we increase $\mathcal{M}_\alpha$ and then attempt to vary $\mathcal{M}_+$ — however the impact of increasing $\mathcal{M}_+$ decreases rapidly with increasing $\mathcal{M}_\alpha$. Thus, for $\mathcal{M}_\alpha = 20$, moving from 5 to 10 positivity constraints

only reduces the constraint set by 20 units, a further increase of 5 positivity constraints only produces a reduction of 2 constraints. Thus, $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ both follow in some sense the law of diminishing returns — since K is roughly the same for the choice of $(\mathcal{M}_\alpha = 10, \mathcal{M}_+ = 15)$, $(\mathcal{M}_\alpha = 15, \mathcal{M}_+ = 10)$, or $(\mathcal{M}_\alpha = 20, \mathcal{M}_+ = 10)$, the choice of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ should depend on the computational cost of the greedy algorithm.

Table 7.6 presents the breakdown of the *average* computational cost per iteration for the different activities involved in running the greedy algorithm. We identify five separate *average* costs: (i) the calculation of $\sigma(\omega_k)$ and $\mathcal{W}^*(\omega_k)$, the associated infimizer for the inf-sup at the "k"-th constraint point using the Lanczos algorithm (denoted by T_{lanc}) ; (ii) the cost of computing $\underline{z}^*(\omega_k) = (z_1^*(\omega_k), \ldots, z_{\hat{\mathcal{Q}}}^*(\omega_k))$ where $z_r^*(\omega_k) = \psi_r(\omega_k), 1 \leq r \leq \hat{\mathcal{Q}}$ (denoted by $T_{\text{infim coeffs}}$); (c) the cost of running Ξ_{train} LPs of size $2\hat{\mathcal{Q}} + \mathcal{M}_\alpha + \mathcal{M}_+$ (denoted by T_{LP}); (d) the cost of calculating the upper bound for the k-th constraint (denoted by T_{UB}); and (e) updating the list of the $\mathcal{M}_\alpha$ closest points in $\mathcal{C}_K$ (denoted by T_{sort}).

$\mathcal{M}_\alpha$	$\mathcal{M}_+$	T_{lanc}	$T_{\text{infim coeffs}}$	T_{LP}	T_{UB}	T_{sort}
10	5	4.34	14.52	22.81	0.02	0.41
10	15	4.19	14.48	23.1	0.02	0.39
15	10	4.08	14.51	23.41	0.02	0.83
20	10	3.97	14.66	23.74	0.02	1.16

Table 7.6: Low Pass Filter Results: Breakdown of the *average* (or per iteration) cost (in seconds) of the greedy algorithm to compute $\mathcal{C}_K$ for different choices of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$. Note that T_{LP} is the cost of solving 2000 LPs using Tomlab's linprog function.

We note that we have roughly the same timings for the offline for different $\mathcal{M}_\alpha$ and $\mathcal{M}_+$. We have slightly increased cost of obtaining upper bounds as we increase $\mathcal{M}_\alpha$. The bulk of the per iteration cost is in calculating the infimizer coefficients and to run the Linear Programs. For $\mathcal{M}_\alpha = 10$ and $\mathcal{M}_+ = 5$ we obtain $K = 272$ and a per iteration cost of ≈ 42.5 seconds; for $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$ we obtain $K = 210$ and the algorithm requires ≈ 43.36 seconds per iteration. The tradeoff between $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ is now clear — $\mathcal{M}_\alpha^{\text{crit}}$ and $\mathcal{M}_+^{\text{crit}}$ are critical choices of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ for which K is relatively insensitive to further increases in both $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ *and* the *average* computational cost is relatively insensitive to decreases in both $\mathcal{M}_\alpha$ and $\mathcal{M}_+$. A choice of $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$ (or $\mathcal{M}_\alpha = 15$ and $\mathcal{M}_+ = 10$) certainly seem like good compromises, and in all possibility are close to the critical values $\mathcal{M}_\alpha^{\text{crit}}$ and $\mathcal{M}_+^{\text{crit}}$.

Post-Processing of $\mathcal{C}_K$

The *offline* algorithm and the optimization of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ are intimately related to the success of the *online* algorithm. However, there is a second issue we would like to talk about. At the k-th step of the greedy algorithm to identify the set $\mathcal{C}_K$, the sample point with the worst $\frac{\sigma_{\text{UB}}(\mu)-\sigma_{\text{LB}}(\mu)}{\sigma_{\text{LB}}(\mu)}$ ratio is picked. However, once we have identified the optimal sample points and constructed the set $\mathcal{C}_K$, it is conceivable that much smaller values of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ might suffice in the *online* stage if we relax our (ultra-conservative) offline termination criteria by a little. For example, the offline termination criteria $\frac{\sigma_{\text{UB}}(\mu)-\sigma_{\text{LB}}(\mu)}{\sigma_{\text{LB}}(\mu)} \leq 0.75$ translates into

$$0.25 \leq \frac{\sigma_{\text{LB}}(\mu)}{\sigma_{\text{UB}}(\mu)} \Rightarrow 0.5 \leq \frac{\beta_{\text{LB}}(\mu)}{\beta_{\text{UB}}(\mu)}. \tag{7.92}$$

Since the upper bound is usually very accurate, $\beta_{\text{UB}}(\mu) \approx \beta(\mu)$; hence $\frac{\beta_{\text{LB}}(\mu)}{\beta(\mu)} \approx 0.5$. Thus, relaxing our stringent offline termination criteria can still guarantee very *good* lower bounds.

We now verify the claim that we can obtain good lower bounds with smaller $(\mathcal{M}_\alpha, \mathcal{M}_+)$. We construct an exhaustive verification sample set $\Xi_{\text{post-process}}$ of size 1500, and compare the "exact" inf-sup with the lower bounds obtained from the SCM approach for different choices of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ in Table 7.7. In table 7.7 we present (a) the *best* inf-sup to inf-sup lower bound ratio $\frac{\beta(\mu)}{\beta_{\text{LB}}(\mu)}$, (d) the *worst* inf-sup to inf-sup lower bound ratio $\frac{\beta(\mu)}{\beta_{\text{LB}}(\mu)}$, and (e) the value of the *offline* termination criterion max $\frac{\sigma_{\text{UB}}(\mu)-\sigma_{\text{LB}}(\mu)}{\sigma_{\text{UB}}(\mu)}$.

We confirm that for the choice of $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$ we satisfy the required termination condition of the offline greedy algorithm, $\epsilon_{\text{tol,min}} = 0.75$. Notice that the SCM inf-sup lower bound is extremely sharp: $1.0002 \leq \frac{\beta(\mu)}{\beta_{\text{LB}}(\mu)} \leq 1.92$.

We note that we can reduce $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ for *online* computations *if* we relax the stringent tolerance of the offline termination criteria while using the same SCM constraint set $\mathcal{C}_K$. In that case, we can choose $\mathcal{M}_\alpha = 14$ and $\mathcal{M}_+ = 8$, and still obtain extremely sharp lower bounds, $1.0005 \leq \frac{\beta(\mu)}{\beta_{\text{LB}}(\mu)} \leq 2.29$. Also note that we cannot arbitrarily reduce $\mathcal{M}_+$, even for $\mathcal{M}_\alpha = 14$, choosing $\mathcal{M}_+ = 6$ instead of $\mathcal{M}_+ = 8$ increases the ratio $\frac{\sigma_{\text{UB}}(\mu)-\sigma_{\text{LB}}(\mu)}{\sigma_{\text{LB}}(\mu)}$ from 0.81 to 2.098 in the worst case. In the *offline* stage, after the completion of the greedy algorithm a simple test would help identify the appropriate values for use in the *online* stage.

$\mathcal{M}_\alpha$	$\mathcal{M}_+$	min $\frac{\beta(\mu)}{\beta_{\text{LB}}(\mu)}$	max $\frac{\beta(\mu)}{\beta_{\text{LB}}(\mu)}$	max $\frac{\sigma_{\text{UB}}(\mu)-\sigma_{\text{LB}}(\mu)}{\sigma_{\text{UB}}(\mu)}$
12	10	-	5.596	1.04
14	6	-	5.254	2.098
14	8	1.0005	2.291	8.11E-01
14	10	1.0002	2.278	8.09E-01
16	8	1.0002	2.291	8.11E-01
16	10	1.0002	2.278	8.09E-01
18	6	1.0002	2.291	8.11E-01
18	8	1.0002	2.291	8.11E-01
18	10	1.0002	2.278	8.09E-01
20	6	1.0002	2.095	7.73E-01
20	8	1.0002	2.092	7.63E-01
20	10	1.0002	1.924	7.43E-01

Table 7.7: Offline post-processing: the SCM constraint set C_K of size $K = 210$ is generated for $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$. We calculate the sharpness of inf-sup lower bounds for smaller choices of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ over a large set $\Xi_{\text{post-process}}$ of size 1500 using the same constraint set.

7.6.2 Online LP Evaluation

In the *offline* stage we generated the constraint sets C_K for different choices of $(\mathcal{M}_\alpha, \mathcal{M}_+)$. The choice of $(\mathcal{M}_\alpha, \mathcal{M}_+)$ and associated K presents us with at tradeoff: we would like to evaluate if by reducing $(\mathcal{M}_\alpha, \mathcal{M}_+)$ *offline* and obtaining a larger set K we gain any computational advantages *online*. More precisely, if the online LP cost for lower $(\mathcal{M}_\alpha, \mathcal{M}_+)$ is not significantly smaller relative to LPs with larger $(\mathcal{M}_\alpha, \mathcal{M}_+)$, then we are better off reducing *offline* cost by choosing more (conservative) values for $\mathcal{M}_\alpha$ and $\mathcal{M}_+$.

We construct an exhaustive verification sample set Ξ_{verif} of size 2000, and compare the "exact" inf-sup with the lower bounds obtained from the SCM approach for different choices of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ in Table 7.8. In table 7.8 we present (a) the *average* cost for the inf-sup lower calculation $T_{\text{LB}}^{\text{av}}$, (b) the *average* computational savings relative to the true inf-sup evaluation $T_{\text{B}}^{\text{av}}/T_{\text{LB}}^{\text{av}}$, (c) the *average* cost for the inf-sup upper bound, (d) the *best* inf-sup to inf-sup lower bound ratio $\frac{\beta(\mu)}{\beta_{\text{LB}}(\mu)}$, (e) the *worst* inf-sup to inf-sup lower bound ratio $\frac{\beta(\mu)}{\beta_{\text{LB}}(\mu)}$, and (f) the value of the *offline* termination criterion max $\frac{\sigma_{\text{UB}}(\mu)-\sigma_{\text{LB}}(\mu)}{\sigma_{\text{UB}}(\mu)}$.

We confirm that the required offline termination condition of the offline greedy algorithm, $\epsilon_{\text{tol,min}} = 0.75$ is satisfied in all cases. Notice that the LP (with 86 constraints, $\hat{Q} = 28$, $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$) is extremely fast: the average time to compute the inf-sup lower bound is ≈ 12 milli-seconds (ms) and is ≈ 346 times faster than the inf-sup calculation. Moreover, the SCM

inf-sup lower bound is extremely sharp: $1.0002 \leq \frac{\beta(\mu)}{\beta_{\rm LB}(\mu)} \leq 1.89$.

Note that there is no significant savings in computational cost for reduced $(\mathcal{M}_\alpha, \mathcal{M}_+)$. For example, the average computational cost decreased only by ≈ 3.5 percent when we go from $(\mathcal{M}_\alpha = 10, \mathcal{M}_+ = 5)$ to $(\mathcal{M}_\alpha = 20, \mathcal{M}_+ = 10)$. There are two reasons for that: (i) the LP size $2\hat{\mathcal{Q}}+M_\alpha+M_+$ is dominated by $\hat{\mathcal{Q}} = 28$ for all the choices of $\mathcal{M}_\alpha$ (max. 20) and $\mathcal{M}_+$ (max 10); and (ii) the LP solver (Tomlab's linprog function) is very efficient for LPs with $\approx O(85)$ constraints as our numerical experiments demonstrate.

We conclude by stating that for this problem, using larger $(\mathcal{M}_\alpha, \mathcal{M}_+)$ to reduce the *offline* constraint set size $\mathcal{C}_K$ is much more beneficial than trying to calculate the *online* LPs with lower $(\mathcal{M}_\alpha, \mathcal{M}_+)$. Relative to the problems attempted in this thesis, the size of the LP probably mandates more emphasis on reducing offline cost than attempting to reduce online cost since the LP is insensitive to the changes in $(\mathcal{M}_\alpha, \mathcal{M}_+)$ as we have shown here.

$\mathcal{M}_\alpha$	$\mathcal{M}_+$	K	$T_{\rm LB}^{\rm av}$	$T_{\rm B}^{\rm av}/T_{\rm LB}^{\rm av}$	min $\frac{\beta(\mu)}{\beta_{\rm LB}(\mu)}$	max $\frac{\beta(\mu)}{\beta_{\rm LB}(\mu)}$	max $\frac{\sigma_{\rm UB}(\mu)-\sigma_{\rm LB}(\mu)}{\sigma_{\rm UB}(\mu)}$
10	5	272	1.145E-02	361.88	1.00024	1.981	7.48E-01
10	10	223	1.155E-02	358.72	1.00024	1.968	7.46E-01
10	15	217	1.161E-02	356.89	1.00024	1.931	7.35E-01
15	5	250	1.141E-02	363.32	1.00014	1.975	7.46E-01
15	10	214	1.17E-02	354.06	1.00024	1.895	7.40E-01
15	15	210	1.187E-02	349.00	1.00024	1.894	7.32E-01
20	5	230	1.159E-02	357.62	1.00024	1.899	7.26E-01
20	10	210	1.187E-02	349.07	1.00024	1.924	7.43E-01
20	15	208	1.198E-02	345.86	1.00024	1.893	7.21E-01

Table 7.8: Low Pass Filter Results: Timings for different choices of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ over a verification set $\Xi_{\rm verif}$ of size 2000: $T_{\rm LB}^{\rm av}$ is the average time to compute the lower bound in milliseconds; $T_{\rm B}^{\rm av}/T_{\rm LB}^{\rm av}$ is the ratio of the average times to compute the inf-sup and the lower bound respectively. We use TOMLAB's linprog solver to calculate the LP associated with the SCM inf-sup lower bound approximation.

Chapter 8

A Posteriori Error Estimation

From Chapter 6, we know that in theory, we can obtain $\mathcal{S}_N(\mu)$ very inexpensively: the online computational effort is independent of $\mathcal{N}$; and the number of basis N can be chosen quite small. But we do not know how small N should be chosen in order for the reduced basis method to produce desired accuracy for all $\mu \in \mathcal{D}$: if N is too small, then our reduced basis approximation is unacceptably inaccurate — if N is too large, our reduced basis approximation unnecessary expensive. We thus need *a posteriori* error estimator for $\mathcal{S}_N(\mu)$: not only the *a posteriori* error estimator helps us to construct the reduced basis nested parameter set, it also tells us how good we are doing: our reduced basis solution is *certified* and *reliable*. Furthermore, the *rigorous* error bounds we obtain curtesy of the *a posteriori* error estimation procedures play a critical role in the construction of the reduced basis space offline.

Here we discuss the development of *a posteriori* error estimators for general non-coercive elliptic partial differential equations with quadratic outputs: in particular, we focus on the "expanded" form of the weak statement with compliant outputs.

8.1 Abstraction

8.1.1 Preliminaries

We consider the "exact" problem posed on the reference domain Ω: for any $\mu \in \mathcal{D} \subset \mathbb{R}^P$, find

$$\mathcal{S}^e(\mu) = \overline{\mathcal{F}(\mathcal{U}^e(\mu))}, \tag{8.1}$$

where $\mathcal{U}^e(\mu)$ satisfies the weak form of the μ-parametrized PDE

$$\mathcal{A}(\mathcal{U}^e(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}), \quad \forall\, \mathcal{V} \in \mathcal{X}^e. \tag{8.2}$$

Here μ and $\mathcal{D}$ are the input and (closed) input domain, respectively; $\mathcal{S}^e(\mu)$ is the "output of interest"; $\mathcal{U}^e(x;\mu)$ is our field variable; $\mathcal{X}^e$ is a complex Hilbert space defined over the physical domain $\Omega \subset \mathbb{R}^d$ with inner product $(\mathcal{W}, \mathcal{V})_{\mathcal{X}^e}$ and associated norm $\|\mathcal{W}\|_{\mathcal{X}^e} = \sqrt{(\mathcal{W}, \mathcal{W})_{\mathcal{X}^e}}$; and $\mathcal{A}(\cdot, \cdot; \mu)\colon \mathcal{X}^e \times \mathcal{X}^e \to \mathbb{C}$ and $\mathcal{F}\colon \mathcal{X}^e \to \mathbb{C}$ are $\mathcal{X}^e$-continuous bilinear and bounded linear functionals, respectively.

Our function space $\mathcal{X}^e$ satisfies $(H_0^1(\Omega))^\nu \subset \mathrm{Re}(\mathcal{X}^e)$, $\mathrm{Im}(\mathcal{X}^e) \subset (H^1(\Omega))^\nu$, where $\nu = 1$ for a scalar field variable and $\nu = d$ for a vector field variable. Recall that $H^1(\Omega)$ (respectively, $H_0^1(\Omega)$) is the usual Hilbert space (respectively, the Hilbert space of functions that vanish on the domain boundary Γ).

8.1.2 Problem Statement

Our "truth" or "reference" finite element approximation to the exact output and field variable, $\mathcal{S}(\mu)$ and $\mathcal{U}(\mu) \equiv \mathcal{U}(\mu) \in \mathcal{X}$, satisfies (6.5), (6.6). Given $\mu \in \mathcal{D}$,

$$\mathcal{S}(\mu) = \overline{\mathcal{F}(\mathcal{U}(\mu))}\ , \tag{8.3}$$

where $\mathcal{U}(\mu) \in \mathcal{X}$ satisfies

$$\mathcal{A}(\mathcal{U}(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}), \quad \forall\, \mathcal{V} \in \mathcal{X}\ . \tag{8.4}$$

We shall suppose that our bilinear form is symmetric

$$\mathcal{A}(\mathcal{W}, \mathcal{V}; \mu) = \overline{\mathcal{A}(\mathcal{V}, \mathcal{W}; \mu)}, \forall \mathcal{W}, \mathcal{V} \in X, \tag{8.5}$$

and "affine" in the parameter: for some fixed integer $Q_\mathcal{A}$ — typically $Q_\mathcal{A}$ shall be larger than P, sometimes by a considerable factor — we require

$$\mathcal{A}(\mathcal{W}, \mathcal{V}; \mu) = \sum_{q=1}^{Q} \Theta_\mathcal{A}^q(\mu) \mathcal{A}^q(\mathcal{W}, \mathcal{V}), \quad \forall\, \mathcal{W}, \mathcal{V} \in \mathcal{X},\ \forall\, \mu \in \mathcal{D}\ , \tag{8.6}$$

where $\Theta_{\mathcal{A}}^q$: $\mathcal{D} \to \mathbb{R}$ and $\mathcal{A}^q(\mathcal{W},\mathcal{V})$: $\mathcal{X} \times \mathcal{X} \to \mathbb{C}$, $1 \le q \le Q$, are parameter-dependent functions and parameter-independent continuous bilinear forms, respectively. We shall further assume that $\Theta_{\mathcal{A}}^q \in \mathcal{C}^1(\mathcal{D})$, $1 \le q \le Q$. $\mathcal{X}$, our "truth" approximation subspace of $\mathcal{X}^e$ is endowed with an appropriate $H^1(\Omega)$-equivalent inner product $(\cdot,\cdot)_{\mathcal{X}}$ (6.3) and induced norm $\|\cdot\|_{\mathcal{X}}$ (6.4).

In essence, $\mathcal{U}(\mu) \in \mathcal{X}$ is a calculable surrogate for $\mathcal{U}^e(\mu)$ upon which we will build our reduced basis approximation and with respect to which we will evaluate the reduced basis error; $\mathcal{U}(\mu)$ shall also serve as the "classical alternative" relative to which we will assess the efficiency of our approach. We assume that $\|\mathcal{U}^e(\mu) - \mathcal{U}(\mu)\|$ is suitably small and hence that $\mathcal{N}$ is typically very large: our formulation must be both *stable* and *efficient* as $\mathcal{N} \to \infty$.

The classical inf-sup, $\beta(\mu)$, and continuity, $\gamma(\mu)$, parameters are given by

$$\beta(\mu) \equiv \inf_{\mathcal{W}\in\mathcal{X}} \sup_{\mathcal{V}\in\mathcal{X}} \frac{|\,\mathcal{A}(\mathcal{W},\mathcal{V};\mu)\,|}{\|\mathcal{W}\|_{\mathcal{X}}\,\|\mathcal{V}\|_{\mathcal{X}}}, \quad \forall\,\mu \in \mathcal{D} \tag{8.7}$$

and

$$\gamma(\mu) \equiv \sup_{\mathcal{W}\in\mathcal{X}} \sup_{\mathcal{V}\in\mathcal{X}} \frac{|\,\mathcal{A}(\mathcal{W},\mathcal{V};\mu)\,|}{\|\mathcal{W}\|_{\mathcal{X}}\,\|\mathcal{V}\|_{\mathcal{X}}}, \quad \forall\,\mu \in \mathcal{D}. \tag{8.8}$$

We now suppose that our problem is well-posed for $\mu \in \mathcal{D}$: $0 < \beta_0 \le \beta(\mu)$ and $0 < \gamma(\mu) \le \gamma_0 < \infty$.

We shall find it convenient to state our error bound proofs by introducing a supremizing operator $T^\mu : \mathcal{X} \to \mathcal{X}$ such that, for any $\mathcal{W}$ in $\mathcal{X}$

$$(T^\mu\mathcal{W},\mathcal{V})_{\mathcal{X}} = \mathcal{A}(\mathcal{W},\mathcal{V};\mu), \quad \forall\, v \in \mathcal{X}\ . \tag{8.9}$$

We then define

$$\sigma(\mathcal{W};\mu) \equiv \frac{\|T^\mu\mathcal{W}\|_{\mathcal{X}}}{\|\mathcal{W}\|_{\mathcal{X}}}, \tag{8.10}$$

and note that

$$\beta(\mu) \equiv \inf_{\mathcal{W}\in\mathcal{X}} \sup_{\mathcal{V}\in\mathcal{X}} \frac{|\mathcal{A}(\mathcal{W},\mathcal{V};\mu)|}{\|\mathcal{W}\|_{\mathcal{X}}\|\mathcal{V}\|_{\mathcal{X}}} = \inf_{\mathcal{W}\in\mathcal{X}} \frac{\|T^\mu\mathcal{W}\|_{\mathcal{X}}}{\|\mathcal{W}\|_{\mathcal{X}}} = \inf_{\mathcal{W}\in\mathcal{X}} \sigma(\mathcal{W};\mu) \tag{8.11}$$

and

$$\gamma(\mu) \equiv \sup_{\mathcal{W}\in\mathcal{X}} \sup_{\mathcal{V}\in\mathcal{X}} \frac{|\mathcal{A}(\mathcal{W},\mathcal{V};\mu)|}{\|\mathcal{W}\|_{\mathcal{X}}\|\mathcal{V}\|_{\mathcal{X}}} = \sup_{\mathcal{W}\in\mathcal{X}} \frac{\|T^\mu\mathcal{W}\|_{\mathcal{X}}}{\|\mathcal{W}\|_{\mathcal{X}}} = \sup_{\mathcal{W}\in\mathcal{X}} \sigma(\mathcal{W};\mu). \tag{8.12}$$

8.2 Reduced Basis Approximation

In this section we review briefly the reduced basis approximation since many details has been already discussed in Chapter 6. We shall not discuss approximation approaches other than Galerkin projection, in particular the Petrov-Galerkin projection.

The Petrov-Galerkin[66, 89] approximation is advantageous for non-coercive problem with non-symmetric $\mathcal{A}$ and where $\mathcal{L}(\mathcal{V}) \neq \overline{\mathcal{F}(\overline{\mathcal{V}})}$. In addition to the actual (primal) problem, a dual (adjoint) problem is introduced to deal with the non-symmetric $\mathcal{A}$ or the non-compliant output; we retain the quadratic convergence in the output.

In Section 6.5, we outlined the primal-dual formulation [53] relevant to the solution of quadratic outputs of non-coercive elliptic problems. The quadratic output complicates the primal-dual formulation, the dual problem is not separated from the primal problem. The construction of the dual reduced basis space can proceed only after the primal reduced basis space is constructed. Furthermore, the form of the *a posteriori* error estimators (6.51) includes a term, $\frac{\gamma^Q_{\rm UB}(\mu)}{\beta_{\rm LB}(\mu)}$, that pollutes the effect of the quadratic convergence in the output. Here $\gamma^Q_{\rm UB}(\mu)$ is the maximum eigenvalue of the generalized eigen valued problem associated with the quadratic output functional and is given in (6.52).

The expanded formulation for the quadratic outputs has two benefits: (*i*) our bilinear form $\mathcal{A}$ is symmetric $\mathcal{A}(\mathcal{W}, \mathcal{V}; \mu) = \overline{\mathcal{A}(\mathcal{V}, \mathcal{W}; \mu)}$, $\forall \mathcal{W}, \mathcal{V} \in \mathcal{X}$ and $\forall \mu \in \mathcal{D}$; and (*ii*) the output is compliant $\mathcal{L}(\mathcal{V}) = \overline{\mathcal{F}(\overline{\mathcal{V}})}, \forall \mathcal{V} \in \mathcal{X}$. The expanded system essentially allows us to handle the quadratic output as the "compliant" output for a slightly different problem with twice the number of unknowns (please see Section 6.5 for a comparison with the primal-dual formulation). The *compliance* of the output makes the output bound much simpler to obtain, furthermore we eliminate the polluting effect of $\gamma^Q_{\rm UB}(\mu)$ present in the error estimators obtained using primal-dual Petrov-Galerkin formulations while retaining the effect of the quadratic convergence in the output. Hence, we need only consider Galerkin approximations for the quadratic outputs treated as linear compliant outputs. We will return to the efficiency of this new approach once we introduce our error estimators.

8.2.1 Galerkin Approximation

In the "Lagrangian" [81] reduced basis approach, the field variable $\mathcal{U}(\mu)$ is approximated by (typically) Galerkin projection onto a space spanned by solutions of the governing PDE at N selected

points in parameter space. We introduce nested parameter samples

$$\begin{aligned} S_N &= \{\mu_1^+ \in \mathcal{D}, \ldots, \mu_{N/2}^+ \in \mathcal{D}, \mu_1^- \in \mathcal{D}, \ldots, \mu_{N/2}^- \in \mathcal{D}\}, \\ &= \left\{S_{N/2}^+ \equiv \{\mu_1^+ \in \mathcal{D}, \ldots, \mu_{N/2}^+ \in \mathcal{D}\}, S_{N/2}^- \equiv \{\mu_1^- \in \mathcal{D}, \ldots, \mu_{N/2}^- \in \mathcal{D}\}\right\}, \end{aligned} \tag{8.13}$$

for $1 \leq N \leq N_{\max}$. In practice, we choose $\mu_1^+ = \mu_1^- = \mu_1^\pm \in \mathcal{D}, \ldots, \mu_{N/2}^+ = \mu_{N/2}^- = \mu_{N/2}^\pm \in \mathcal{D}$ ($S_{N/2}^+ = S_{N/2}^- = S_{N/2}^\pm$) and specify the associated reduced basis space

$$\begin{aligned} W_N &= \text{span}\{[U^+(\mu_n^\pm), \mathbf{0}]^T, [\mathbf{0}, U^-(\mu_n^\pm)]^T, 1 \leq n \leq N/2\}, \quad 1 \leq N \leq N_{\max}, \\ &= \text{span}\{\zeta_n^+ = [U^+(\mu_n^\pm), \mathbf{0}]^T, \zeta_n^- = [\mathbf{0}, U^-(\mu_n^\pm)]^T, 1 \leq n \leq N/2\}, \quad 1 \leq N \leq N_{\max}, \\ &= \text{span}\{\zeta_n \equiv \{\zeta_n^+, \zeta_n^-\}, 1 \leq n \leq N\}, \quad 1 \leq N \leq N_{\max}, \end{aligned} \tag{8.14}$$

where $\mathcal{U}(\mu_n^\pm) = [U^+(\mu_n^\pm), U^-(\mu_n^\pm)]^T$ is the solution of (8.4) for $\mu = \mu_n^\pm$. We orthogonalize the basis with respect to the inner product (6.3). We next apply Galerkin projection onto W_N to obtain $\mathcal{U}_N(\mu) \in W_N$ from

$$\mathcal{A}(\mathcal{U}_N(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}), \quad \forall\, \mathcal{V} \in W_N, \tag{8.15}$$

in terms of which the reduced basis approximation to $\mathcal{S}(\mu)$ is then calculated as

$$\mathcal{S}_N(\mu) = \overline{\mathcal{F}(\mathcal{U}_N(\mu))}. \tag{8.16}$$

8.3 Lower Bound for Inf-Sup Parameter

Our output error bound requires an inexpensive (online) and reasonably accurate lower bound for the inf-sup stability parameter. In Chapter 7 we discussed the development and calculation of the inf-sup lower bound using the successive constraint method. We introduced a *rigorous* lower bound for the square of the inf-sup parameter, $\sigma(\mu)$, as

$$\sigma_{\text{LB}}(\mu; \mathcal{C}_K, \Xi_J) \leq \sigma(\mu), \forall \mu; \in \mathcal{D}. \tag{8.17}$$

and described the *offline* construction of the constraint set $\mathcal{C}_K$ (and associated $\mathcal{M}_\alpha \in \mathbb{N}$). The constraint set Ξ_J (and associated $\mathcal{M}_+ \in \mathbb{N}$) is used in the *offline* stage to enforce "positivity"

constraints at $\mathcal{M}_+$ "nearby" points in parameter space.

In the *online* stage, for a given $\mu \in \mathcal{D}$, we solve a Linear Program of size $O(2 * \hat{Q} + \mathcal{M}_\alpha + \mathcal{M}_+)$ to calculate $\sigma_{\text{LB}}(\mu)$. Here $\mathcal{M}_+$ can be set to zero (i.e., no positivity constraints), or we can use an exhaustive set of sample points as the set Ξ_J from which we pick the $\mathcal{M}_+$ positivity constraints. We obtain the inf-sup lower bound as

$$\beta_{\text{LB}}(\mu) = \sqrt{\sigma_{\text{LB}}(\mu)}; \tag{8.18}$$

such that

$$0 < \beta_{\text{LB}}(\mu) \leq \beta(\mu), \forall \mu \in \mathcal{D}. \tag{8.19}$$

8.4 *A Posteriori* Error Estimation

8.4.1 Objective

We wish to develop *a posteriori* error bounds $\Delta_N(\mu)$ and $\Delta_N^{\mathcal{S}}(\mu)$ such that

$$\|\mathcal{U}(\mu) - \mathcal{U}_N(\mu)\|_{\mathcal{X}} \leq \Delta_N(\mu) \ , \tag{8.20}$$

and

$$|\mathcal{S}(\mu) - \mathcal{S}_N(\mu)| \leq \Delta_N^{\mathcal{S}}(\mu) \ . \tag{8.21}$$

It shall prove convenient to introduce the notion of effectivity, defined (here) as

$$\eta_N(\mu) \equiv \frac{\Delta_N(\mu)}{\|u(\mu) - u_N(\mu)\|_X}, \tag{8.22}$$

$$\eta_N^s(\mu) \equiv \frac{\Delta_N^{\mathcal{S}}(\mu)}{|s(\mu) - s_N(\mu)|} \ . \tag{8.23}$$

We can re-state our requirements (8.20) and (8.21) as

$$\eta_N(\mu) \geq 1, \ \forall \mu \in \mathcal{D} \tag{8.24}$$

$$\eta_N^s(\mu) \geq 1, \forall \ \mu \in \mathcal{D}. \tag{8.25}$$

However, for efficiency, we must also require $\eta_N(\mu) \leq \eta_0$ and $\eta_N^s(\mu) \leq \eta_0$; here $\eta_0 \geq 1$ is a

constant independent of N and μ; preferably, η_0 is close to unity, thus ensuring that we choose the *smallest* N — and hence most economical — reduced basis approximation consistent with the specified error tolerance.

8.4.2 Error Bound Formulation

We assume that we may calculate μ-dependent lower bound $\beta_{\rm LB}(\mu)$ for the inf-sup parameter $\beta(\mu)$: $\beta(\mu) \geq \beta_{\rm LB}(\mu) \geq \beta_0 > 0, \forall \mu \in \mathcal{D}$. The calculation of $\beta_{\rm LB}(\mu)$ has been extensively studied in the previous chapter. We next introduce the dual norm of the residual

$$\varepsilon_N(\mu) = \sup_{\mathcal{V} \in \mathcal{X}} \frac{\mathcal{R}(\mathcal{V};\mu)}{\|\mathcal{V}\|_{\mathcal{X}}} \ , \tag{8.26}$$

where

$$\mathcal{R}(\mathcal{V};\mu) = \mathcal{F}(\mathcal{V}) - \mathcal{A}(\mathcal{U}_N(\mu), \mathcal{V};\mu), \quad \forall\, \mathcal{V} \in \mathcal{X} \tag{8.27}$$

is the residual associated with $\mathcal{U}_N(\mu)$.

We can now define our energy error bound

$$\Delta_N(\mu) \equiv \frac{\varepsilon_N(\mu)}{\beta_{\rm LB}(\mu)} \ , \tag{8.28}$$

and output error bound

$$\Delta_N^{\mathcal{S}}(\mu) \equiv \varepsilon_N^2(\mu)/\beta_{\rm LB}(\mu) \ . \tag{8.29}$$

We shall prove that $\Delta_N(\mu)$ and $\Delta_N^{\mathcal{S}}(\mu)$ are rigorous and sharp bounds for $\|\mathcal{U}(\mu) - \mathcal{U}_N(\mu)\|_{\mathcal{X}}$ and $|\mathcal{S}(\mu) - \mathcal{S}_N(\mu)|$, respectively.

8.4.3 Bound Properties

Proposition 4. *For the error bounds $\Delta_N(\mu)$ of (8.28) and $\Delta_N^{\mathcal{S}}(\mu)$ of (8.29), the corresponding effectivities satisfy*

$$1 \leq \eta_N(\mu) \leq \frac{\gamma(\mu)}{\beta_{\rm LB}(\mu)}, \quad \forall\, \mu \in \mathcal{D} \ , \tag{8.30}$$

$$1 \leq \eta_N^{\mathcal{S}}(\mu), \quad \forall \mu \ \in \mathcal{D} \ . \tag{8.31}$$

Proof. We first note from (8.4) and (8.27) that the error $e(\mu) \equiv \mathcal{U}(\mu) - \mathcal{U}_N(\mu)$ satisfies

$$\mathcal{A}(e(\mu), \mathcal{V}; \mu) = \mathcal{R}(\mathcal{V}; \mu), \quad \forall\, \mathcal{V} \in \mathcal{X}, \tag{8.32}$$

Furthermore, from standard duality argument we have

$$\varepsilon_N(\mu) = \|\hat{e}(\mu)\|_{\mathcal{X}} \ , \tag{8.33}$$

where

$$(\hat{e}(\mu), \mathcal{V})_{\mathcal{X}} = \mathcal{R}(\mathcal{V}; \mu), \quad \forall\, \mathcal{V} \in \mathcal{X} \ . \tag{8.34}$$

It then follows from (8.9), (8.32) and (8.34) that

$$\|\hat{e}(\mu)\|_{\mathcal{X}} = \|T^{\mu} e(\mu)\|_{\mathcal{X}}. \tag{8.35}$$

In addition from (8.10) we know that

$$\|e(\mu)\|_{\mathcal{X}} = \frac{\|T^{\mu} e(\mu)\|_{\mathcal{X}}}{\sigma(e(\mu); \mu)}. \tag{8.36}$$

It then follows from (8.22), (8.28), (8.32), (8.35) and (8.36) that

$$\eta_N(\mu) = \frac{\sigma(e(\mu); \mu)}{\beta_{\mathrm{LB}}(\mu)}; \tag{8.37}$$

this proves the desired result (8.30) since $\gamma(\mu) \geq \sigma(e(\mu); \mu) \geq \beta(\mu) \geq \beta_{\mathrm{LB}}(\mu)$.

Furthermore, from (8.36), (8.35) we note that

$$\|e(\mu)\|_{\mathcal{X}} = \frac{\|T^{\mu} e(\mu)\|_{\mathcal{X}}}{\sigma(e(\mu); \mu)} \leq \frac{\|\hat{e}(\mu)\|_{\mathcal{X}}}{\beta_{\mathrm{LB}}(\mu)} \equiv \frac{\varepsilon_N(\mu)}{\beta_{\mathrm{LB}}(\mu)} = \Delta_N(\mu); \tag{8.38}$$

using $\beta_{\mathrm{LB}}(\mu) \leq \beta(\mu) \leq \sigma(e(\mu); \mu)$.

Finally, we use the symmetry of $\mathcal{A}$ (8.5), compliance of the output, Galerkin orthogonality (8.32),

and the result (8.38) to prove that

$$
\begin{aligned}
|\mathcal{S}(\mu) - \mathcal{S}_N(\mu)| &= |\overline{\mathcal{F}(e(\mu))}| \\
&= |\overline{\mathcal{A}(\mathcal{U}(\mu), e(\mu); \mu)}| \\
&= |\mathcal{A}(e(\mu), \mathcal{U}(\mu); \mu)| \\
&= |\mathcal{A}(e(\mu), e(\mu); \mu)| \\
&= |\mathcal{R}(e(\mu); \mu)| \\
&\leq \|\mathcal{R}\|_{\mathcal{X}'} \|e(\mu)\|_{\mathcal{X}} \\
&\leq \frac{\|\hat{e}(\mu)\|_{\mathcal{X}}^2}{\beta_{\mathrm{LB}}(\mu)};
\end{aligned}
$$

which concludes our proof. □

While we know that $\eta_N(\mu)$ is greater than 1 by construction; the upper bound on $\eta_N(\mu)$ serves an important purpose. Note, that we can develop an upper bound on η_N only because the problem is well-posed and $\gamma(\mu)$ is bounded, $\gamma(\mu) \leq \gamma_0$, $\forall \mu \in \mathcal{D}$. The development of the upper bound on $\eta_N(\mu)$ is important because it tells us something about the *sharpness* of our *a posteriori* error estimators in the worst case. We would ideally like the ratio $\frac{\gamma(\mu)}{\beta_{\mathrm{LB}}(\mu)}$ to be as small as possible — while we cannot control $\gamma(\mu)$, a good inf-sup lower bound construction is essential. Furthermore, in systems where the eigen spectrum is small (the seperation between $\gamma(\mu)$ and $\beta(\mu)$ is not too large), we are guaranteed very good effectivities, sometimes $\eta_N(\mu) \approx O(1)$.

We also briefly touch upon the issue of the efficiency of the error estimators obtained using the expanded formulation relative to primal-dual formulations for quadratic outputs. In Chapter 6 we introduced the *a posteriori* error estimators (6.51) for the quadratic output of interest obtained from mixed primal-dual [53] formulations. We note that both approaches, the expanded formulation and the primal-dual formulation, guarantee quadratic convergence in the output of interest. However, a conservative term $\gamma_{\mathrm{UB}}^{Q}(\mu)$ (6.52), the maximum eigenvalue of a generalized eigenproblem associated with the quadratic output functional, compromises the *sharpness* of the primal-dual *a posteriori* error estimator. This observation has been empirically validated for real-valued elliptic coercive problems with quadratic outputs [52, 53]: the effectivities from the calculation of the *a posteriori* error estimators for the stress intensity factor using the primal-dual formulation [53] are significantly

higher compared to those computed using the expanded formulation [52].

8.4.4 Offline-Online Decomposition

It remains to develop associated offline-online computational procedure for the evaluation of $\Delta_N^{\mathcal{S}}(\mu)$. We begin from our reduced basis approximation $\mathcal{U}_N(\mu) = \sum_{n=1}^N \mathcal{U}_{Nn}(\mu)\zeta_n$ and the definition of $\mathcal{R}$

$$\mathcal{R}(\mathcal{V};\mu) = \mathcal{F}(\mathcal{V}) - \sum_{q=1}^{Q_\mathcal{A}} \sum_{n=1}^{N} \Theta_\mathcal{A}^q(\mu)\, \mathcal{U}_{Nn}(\mu)\, \mathcal{A}^q(\zeta_n, \mathcal{V}), \quad \forall \mathcal{V} \in \mathcal{X}. \tag{8.39}$$

It is clear from linear superposition that we can express $\hat{e}(\mu) \in \mathcal{X}$ as

$$(\hat{e}(\mu), \mathcal{V})_\mathcal{X} = \mathcal{C} + \sum_{q=1}^{Q_\mathcal{A}} \sum_{n=1}^{N} \Theta_\mathcal{A}^q(\mu)\, \mathcal{U}_{Nn}(\mu)\, \mathcal{L}_n^q, \tag{8.40}$$

where $(\mathcal{C}, \mathcal{V})_\mathcal{X} = \mathcal{F}(\mathcal{V})$, $\forall \mathcal{V} \in \mathcal{X}$, and $\mathcal{L}_n^q = -\mathcal{A}^q(\zeta_n, \mathcal{V})$, $\forall \mathcal{V} \in \mathcal{X}$, $1 \le n \le N$, $1 \le q \le Q_\mathcal{A}$; note that all of the above are simple parameter-independent. It thus directly follows that

$$\begin{aligned} \|\hat{e}(\mu)\|_\mathcal{X}^2 \;=\; & (\mathcal{C}, \mathcal{C})_\mathcal{X} \\ & + 2\,\mathrm{Re}\Big(\sum_{q'=1}^{Q_\mathcal{A}} \sum_{n=1}^{N} \Theta_\mathcal{A}^{q'}(\mu) \mathcal{U}_{Nn}(\mathcal{C}, \mathcal{L}_n^{q'})_\mathcal{X}\Big) \\ & + \sum_{q=1}^{Q} \sum_{q'=1}^{Q_\mathcal{A}} \sum_{n=1}^{N} \sum_{n'=1}^{N} \Theta_\mathcal{A}^q(\mu) \overline{\Theta_\mathcal{A}^{q'}(\mu)}\, \mathcal{U}_{Nn}\, \overline{\mathcal{U}_{Nn'}}\, (\mathcal{L}_n^q, \mathcal{L}_n^{q'})_\mathcal{X}. \end{aligned} \tag{8.41}$$

The expression (8.41) is simply a summation of parameter-dependent functions and parameter-independent inner products. The offline-online decomposition is now clear.

In the offline stage, we first solve for the quantities $\mathcal{C}$ and $\mathcal{L}_n^q$, $1 \le n \le N$, $1 \le q \le Q_\mathcal{A}$ and then perform and store the parameter-independent inner products, $(\mathcal{C}, \mathcal{C})_\mathcal{Y}$, ; $(\mathcal{C}, \mathcal{L}_n^{q'})_\mathcal{X}$, $1 \le n \le N$, $1 \le q' \le Q_\mathcal{A}$ and $(\mathcal{L}_n^q, \mathcal{L}_n^{q'})_\mathcal{X}$, $1 \le n, n' \le N$, $1 \le q, q' \le Q_\mathcal{A}$. This requires $O(NQ_\mathcal{A})$ "truth" finite element solutions and $O(N_\mathcal{A}^2\, Q_\mathcal{A}^2 + N\, Q_\mathcal{A} + 1)$ inner products.

In the online stage, given a new parameter value μ, we simply evaluate the sum (8.41) in terms of $\Theta_\mathcal{A}^q(\mu)$, and $\mathcal{U}_{Nn}(\mu)$ and the pre-computed inner products. The operation count for this stage is only $O(N^2\, Q_\mathcal{A}^2 + N\, Q_\mathcal{A} + 1)$ — totally independent of $\mathcal{N}$.

We also comment on the effect of the expanded (in effect, doubled) system. The original

algebraic system of the problem, from which the expanded system is derived, is given by

$$a(u(\mu), v; \mu) = f(v; \mu), \forall v \in X, \ \forall \mu \in \mathcal{D}; \tag{8.42}$$

and is of dimension $\hat{\mathcal{N}}$; we assume a compliant output $s(\mu) = \overline{f(u(\mu))}$. We compare the standard (8.42) problem to our reformulated expanded problem: evaluate $\mathcal{S}(\mu) = \overline{\mathcal{F}(\mathcal{U}(\mu))}$ where $\mathcal{U}(\mu)$ satisfies $\mathcal{A}(\mathcal{U}(\mu); v) = \mathcal{F}(\mu), \forall \mathcal{V} \in \mathcal{X}, \ \forall \mu \in \mathcal{D}$; here the algebraic system is of dimension $\mathcal{N} = 2\hat{\mathcal{N}}$. We assume a and $\mathcal{A}$ have affine decompositions of the form $a(\cdot, \cdot; \mu) = \sum_{q=1}^{Q_a} \Theta_a^q(\mu)\, a^q(\cdot, \cdot)$ and $\mathcal{A}(\cdot, \cdot; \mu) = \sum_{q=1}^{Q_\mathcal{A}} \Theta_\mathcal{A}^q\, \mathcal{A}^q(\cdot, \cdot)$.

We now compare the computational cost of the expanded formulation relative to the standard problem (8.42). We assume for convenience that both problems require the same number of basis $N_{\max}$. In the *offline* stage, we first solve for the reduced basis functions ζ_i, $1 \le i \le N_{\max}$: we need to solve a sparse algebraic system of size $\mathcal{N} \times \mathcal{N}$ (respectively, $\hat{\mathcal{N}} \times \hat{\mathcal{N}}$) $N_{\max}$ times. Furthermore, we need to calculate $\mathcal{F}(\zeta_i)$, $1 \le i \le N_{\max}$ (respectively, $f(\zeta_i)$, $1 \le i \le N_{\max}$), and $\mathcal{A}^q(\zeta_j, \zeta_i)$, $1 \le i, j \le N_{\max}$, $1 \le q \le Q_\mathcal{A}$ (respectively, $a^q(\zeta_j, \zeta_i)$, $1 \le i, j \le N_{\max}$, $1 \le q \le Q_a$) requiring $O(Q_\mathcal{A}\, N_{\max}^2\, \mathcal{N})$ (respectively, $O(Q_a\, N_{\max}^2\, \hat{\mathcal{N}})$) operations.

The offline storage costs are not impacted by the expanded system. Similarly, the *online* costs for the solution of the reduced basis system remains approximately identical for the standard and expanded problem because of our assumption that we need roughly the same N.

Chapter 9

Applications in Acoustics

9.1 Introduction

In this chapter we bring together the different methodological pieces: the reduced basis approximation of the expanded system to handle quadratic outputs, the in-sup lower bound formulation and the *a posteriori* error estimation to tackle the two acoustics applications which have been repeatedly used through out the thesis.

9.2 Band-Stop Filter

9.2.1 Problem Statement

We consider the acoustic band-stop filter first introduced in Chapter 1 and elaborated on in Chapter 4. The band-stop filter waveguide element under consideration consists of a rectangular acoustic waveguide coupled with a Helmholtz resonator [93, 94] with an acoustic liner of fixed impedance on the top wall. This example is representative of some acoustic filters used in particular industrial contexts: for example, typical passive acoustic liners used in turbofan engine nacelles for suppression of engine noise are essentially Helmholtz resonators with porous face-sheets and rigid backing sheets on the resonator cavity wall [105]. Helmholtz resonators themselves are used widely for sound attenuation [20].

Figure 9-1 shows the (non-dimensional) original domain $\hat{\Omega}$ of our band-stop filter problem. It consists of a long waveguide element $\hat{\Omega}_1$ coupled to a Helmholtz resonator element, $\hat{\Omega}_3$ by a thin

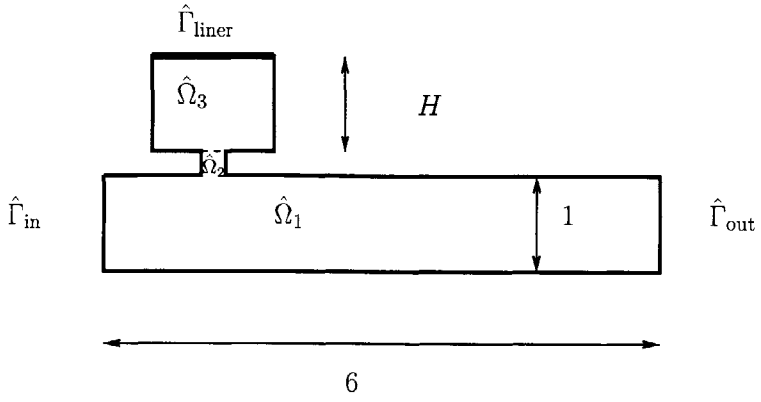


Figure 9-1: Band-Stop Filter: Non-dimensional Original Domain

neck $\hat{\Omega}_2$.

We summarize the governing equations on the (non-dimensional) original domain $\hat{\Omega}$, and associated boundary conditions on $\hat{\Gamma}$. Given $\mu \in \mathbb{R}^{P=2} \subset \mathcal{D}$, we solve the following set of equations (please see Section 4.3.1 for a fuller description of the formulation) for $u(\mu)$

$$\begin{aligned}
\nabla^2 u(\mu) + k^2 u(\mu) &= 0, \text{ in } \hat{\Omega}, && (9.1)\\
\frac{\partial u(\mu)}{\partial \hat{n}} &= \mathrm{i}\, k, \text{ on } \hat{\Gamma}_{\text{in}}, && (9.2)\\
\frac{\partial u(\mu)}{\partial \hat{n}} &= -\mathrm{i}\, \frac{k}{Z_R}\, u, \text{ on } \hat{\Gamma}_{\text{liner}}, && (9.3)\\
\frac{\partial u(\mu)}{\partial \hat{n}} &= -\mathrm{i}\, \alpha_1 \left(\int_{\hat{\Gamma}_{\text{out}}} u(\mu)\, \overline{\Xi}_1 \right) \Xi_1, \text{ on } \hat{\Gamma}_{\text{out}}, && (9.4)\\
\frac{\partial u(\mu)}{\partial \hat{n}} &= 0, \text{ on } \hat{\Gamma} \backslash \left(\hat{\Gamma}_{\text{in}} \cup \hat{\Gamma}_{\text{liner}} \cup \hat{\Gamma}_{\text{out}} \right); && (9.5)
\end{aligned}$$

and calculate our output of interest

$$s(\mu) = -20\, \log_{10} \frac{\sqrt{\int_{\hat{\Gamma}_{\text{out}}} \overline{u(\mu)}\, u(\mu)}}{\sqrt{\int_{\hat{\Gamma}_{\text{in}}} \overline{u(\mu)}\, u(\mu)}}. \qquad (9.6)$$

9.2.2 Parameters

The parameters are the (non-dimensional) frequency squared, $\mu_1 \equiv k^2$, and the height of the Helmholtz resonator cavity (relative to the reference cavity height $H_{\text{ref}} = 1$), $\mu_2 \equiv H$; we consider the parameter domain $\mu \in \mathcal{D} \equiv [0.1, 5.0] \times [0.75, 1.25]$. We also fix the non-dimensional acoustic

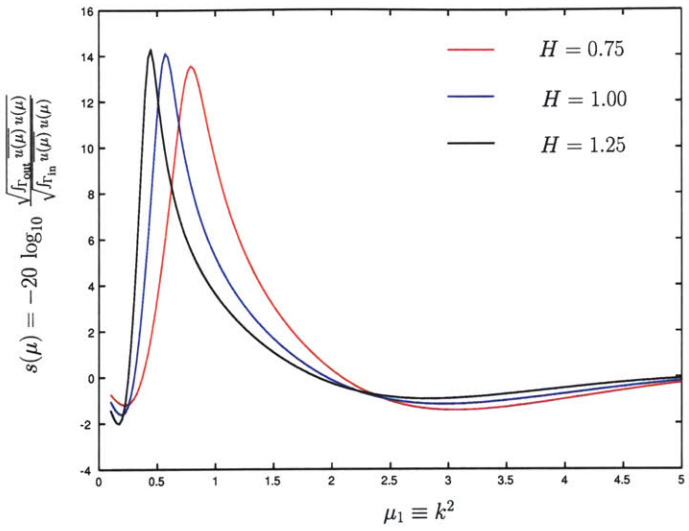


Figure 9-2: Transmission Coefficient curves for the Acoustic Band-Stop Filter Waveguide.

impedance of the liner wall as $Z_R = 10$. The non-dimensional frequency range of k corresponds to a dimensional frequency range of 340 Hz to 2.4 KHz — these ranges are appropriate for the use of band-stop filters in industrial silencers and acoustic liners for the suppression of aircraft engine noise. The fixed non-dimensional impedance $Z_R = 10$ is within the range of normalized acoustic impedance obtained from experiments conducted on tunable electromechanical acoustic liners designed to suppress aircraft engine noise [50, 105].

Figure 9-2 shows the variation of $s(\mu)$ with $\mu_1 \equiv k^2$ for three choices of μ_2. The peaks roughly correspond to the resonant frequency of the Helmholtz resonator. The TC curves are obtained using the reduced basis method for the choice of $N = 35$.

9.2.3 Resonant Frequency of the Helmholtz Resonator

The Transmission Coefficient (TC) peaks we obtain for our band-stop filter are obtained at the resonant frequencies of the Helmholtz resonator for the different geometries. At the natural frequency of the Helmholtz resonator, the incident wave propagating in the waveguide is blocked (and effectively reflected) because of the high impedance (recall, a sound-hard surface where you have perfect reflectivity corresponds to an infinite impedance); thus the outlet pressure intensity is orders of magnitudes lower than the inlet pressure intensity and we obtain a TC peak. So, we can indirectly infer the resonant frequency of the Helmholtz resonator for any particular geometry

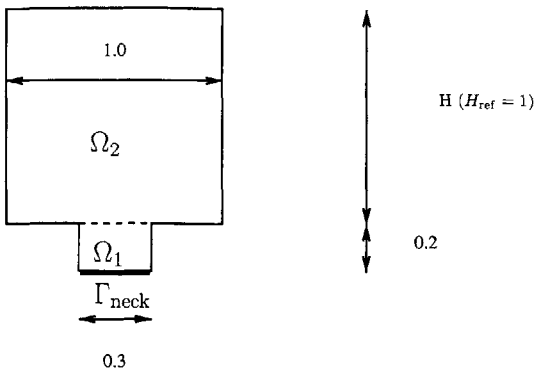


Figure 9-3: Picture of the Helmholtz resonator used in the band-stop filter. We also show the (non-dimensional) dimensions of the resonator.

by obtaining the associated TC curve. Here, we compare the inverse prediction of the resonant frequency of the Helmholtz resonator (corresponding to the TC peak) to those obtained from (i) a simple lumped-element (analytical) model, and (ii) a eigensolution of the Helmholtz equation for just the Helmholtz resonator.

Analytical Resonances

The analytical expression [22] for the natural frequency of a Helmholtz resonator with a circular neck and a circular cavity is given by

$$f = \frac{c}{2\pi}\sqrt{\frac{S}{V\ell_{\text{eff}}}}; \tag{9.7}$$

where S is the cross-sectional area of the neck, V is the volume of the cavity, and ℓ_{eff} is given by

$$\ell_{\text{eff}} = \ell_{\text{neck}} + 1.7\,r;$$

for a neck with flanged ends on both sides (on one side the flanged side of the resonator, and on the other the flanged sides of the waveguide). Here all the units of S, V and ℓ_{eff} are dimensional.

The non-dimensional units of the Helmholtz resonator are given in Figure 9-3. The non-dimensionalization was performed with respect to the actual dimensional depth of the waveguide $\tilde{d}$. We thus calculate the following terms as

$$
\begin{aligned}
S &= 0.3\tilde{d}\tilde{D} \\
V &= H\tilde{d}^2\tilde{D} \\
\ell_{\text{eff}} &= (0.2 + 0.85 * 0.3)\,\tilde{d};
\end{aligned}
$$

where $\tilde{D}$ is the dimensional depth in the third dimension that we effectively ignore; thus we obtain the resonant frequency as

$$
\begin{aligned}
f &= \frac{c}{2\,\pi}\sqrt{\frac{0.3\tilde{d}\tilde{D}}{H\tilde{d}^2\tilde{D}\,(0.2 + 0.85 * 0.3)\,\tilde{d}}} \\
f &= \frac{c}{2\pi\tilde{d}}\sqrt{\frac{0.3}{H\,(0.2 + 0.85 * 0.3)}}.
\end{aligned}
$$

Now, we can calculate $k = \frac{2\pi f\tilde{d}}{c}$, we obtain

$$
k = \sqrt{\frac{0.3}{H\,(0.2 + 0.85 * 0.3)}}
$$

or

$$
k^2|_{\text{helm,res}} = \frac{0.3}{H\,(0.2 + 0.85 * 0.3)}. \tag{9.8}
$$

We note that the resonant frequency calculation is for a circular resonator, our actual Helmholtz resonator is rectangular (where the depth $\tilde{D}$ in the third dimension is much larger than the other dimensions).

Resonances from Eigenproblem

We can also calculate the resonances from the following eigen problem

$$
\int_{\Omega_1} \nabla w \cdot \nabla v + H\int_{\Omega_2} w_x v_x + \frac{1}{H}\int_{\Omega_2} w_y v_y + \frac{3\pi}{8\,r}\int_{\Gamma_{\text{neck}}} wv = k^2_{\text{helm,res}}\Big(\int_{\Omega_1} wv + H\int_{\Omega_2} wv\Big); \quad w, v \in X; \tag{9.9}
$$

where we prescribed the radiation boundary condition at Γ_{neck} as $\frac{\partial p}{\partial n} = -\frac{3\pi}{8\,r}p$ [22], where r is half the width of the neck ($r = 0.15$). The radiation boundary condition used corresponds to a circular cavity and circular neck; in our case, although we do not have a circular neck and cavity, we shall see that the condition yields close agreement with obtained solutions.

Comparison of Obtained Resonances

We denote by $k^2_{\text{anal,res}}$ the analytical prediction of the resonance, $k^2_{\text{HR,res}}$ the eigenproblem prediction (obtained using the radiation boundary condition for a Helmholtz resonator with circular cross-section), and $k^2_{\text{TCpeak,res}}$ as the frequency obtained corresponding to the TC peak. Table 9.1 shows the $k^2_{\text{TCpeak,res}}$, $k^2_{\text{anal,res}}$, and $k^2_{\text{HR,res}}$ as a function of the Helmholtz resonator height H. We note that $k^2_{\text{HR,res}}$ is very close to $k^2_{\text{TCpeak,res}}$; as expected, the analytical prediction is most erroneous.

H	$k^2_{\text{anal,res}}$	$k^2_{\text{HR,res}}$	$k^2_{\text{TLpeak,res}}$
0.75	0.8791	0.80846	0.7798
0.8125	0.8115	0.73773	0.7156
0.875	0.7535	0.67707	0.6547
0.9375	0.7033	0.6245	0.605
1.0	0.6593	0.57857	0.5678
1.0625	0.6206	0.5381	0.5186
1.125	0.5861	0.50218	0.494
1.1875	0.5552	0.47012	0.4574
1.25	0.5275	0.44134	0.4201

Table 9.1: Helmholtz resonator resonances calculated using different methods for different H. The most accurate should be $k^2_{\text{TLpeak,res}}$, then $k^2_{\text{HR,res}}$ and then $k^2_{\text{anal,res}}$. Furthermore, the resonances from the HR resonances are fairly close to the TC peak resonances.

9.2.4 Weak Formulation

Weak Form on Reference Domain

To derive the weak form of the problem, we first introduce a complex function space

$$X = \{v = v_R + \mathrm{i}\, v_I \mid v_R \in H_1(\Omega), v_I \in H_1(\Omega)\}, \tag{9.10}$$

and associated inner product

$$(w, v)_X = \int_\Omega \nabla w \, \nabla \overline{v} + w \, \overline{v}. \tag{9.11}$$

We then summarize the original weak form on the μ-independent reference domain Ω. Given $\mu \in \mathcal{D}$, find $u(\mu) \in X$ such that

$$a(u, v; \mu) = f(v; \mu), \; \forall v \in X; \tag{9.12}$$

here the forms are given by

$$\begin{aligned} a(w,v;\mu) \;=\; & \int_{\Omega_{1,2}} \nabla w \, \nabla \overline{v} + \mu_2 \int_{\Omega_3} \frac{\partial w}{\partial x}\frac{\partial \overline{v}}{\partial x} + \frac{1}{\mu_2}\int_{\Omega_3} \frac{\partial w}{\partial y}\frac{\partial \overline{v}}{\partial y} \\ & - \mu_1 \int_{\Omega_{1,2}} w\,\overline{v} - \mu_1\,\mu_2 \int_{\Omega_3} w\;\overline{v} \\ & + \mathrm{i}\,\frac{\sqrt{\mu}_1}{Z_R}\int_{\Gamma_{\text{liner}}} w\,\overline{v} + \mathrm{i}\,\sqrt{\mu}_1 \int_{\Gamma_{\text{out}}} \Xi_1\,\overline{v} \int_{\Gamma_{\text{out}}} w\,\overline{\Xi}_1; \end{aligned} \tag{9.13}$$

and

$$f(v;\mu) = \mathrm{i}\,\sqrt{\mu}_1 \int_{\Gamma_{\text{in}}} \overline{v}. \tag{9.14}$$

We identify that a is affinely separable, $a(\cdot,\cdot;\mu) = \sum_{q=1}^{Q_a} \Theta_a^q(\mu)\, a^q(\cdot,\cdot)$, $\forall \mu \in \mathcal{D}$, for $Q_a = 7$. Similarly, f is affinely separable, $f(\cdot;\mu) = \sum_{q=1}^{Q_f} \Theta_f^q(\mu) f^q(\cdot)$, $\forall \mu \in \mathcal{D}$, for $Q_f = 1$.

We re-write our true output of interest $s(\mu)$ in terms of two intermediate quadratic outputs $s_1(\mu)$ and $s_2(\mu)$

$$s(\mu) = -10\,\log_{10}\frac{s_{\text{out}}(\mu)}{s_{\text{in}}(\mu)}; \tag{9.15}$$

where $s_1(\mu)$ and $s_2(\mu)$ are quadratic outputs given by

$$s_{\text{in}}(\mu) \equiv Q_{\text{in}}(u(\mu),u(\mu);\mu) \;=\; \int_{\Gamma_{\text{in}}} u(\mu)\,\overline{u(\mu)},\ \forall \mu \in \mathcal{D}, \tag{9.16}$$

$$s_{\text{out}}(\mu) \equiv Q_{\text{out}}(u(\mu),u(\mu);\mu) \;=\; \int_{\Gamma_{\text{out}}} u(\mu)\,\overline{u(\mu)},\ \forall \mu \in \mathcal{D}. \tag{9.17}$$

Expanded Weak Form

We can then identify two systems of equations to calculate our output of interest. In Chapter 3, we introduced the expanded formulation to deal with a *single* quadratic output. To compute the transmission coefficient given by (9.15), we will need two separate expanded formulations for our two quadratic outputs $s_{\text{in}}(\mu)$ and $s_{\text{out}}(\mu)$. The first expanded formulation is related to $s_{\text{in}}(\mu)$ — we denote $\mathcal{S}_{\text{in}}(\mu)$ as the "compliant" output associated with the expanded inlet formulation. The second expanded formulation is to deal with $s_{\text{out}}(\mu)$ — we denote $\mathcal{S}_{\text{out}}(\mu)$ as the linear compliant output associated with the expanded outlet formulation.

We also introduce the complex "expanded" function space $\mathcal{X} \subset (X)^2$ that will be required for the expanded formulations. Thus, $\dim(\mathcal{X}) = 2\dim(X)$.

Inlet System

Given $\mu \in \mathcal{D}$, calculate the "truth" output of interest

$$\mathcal{S}_{\text{in}}(\mu) = \overline{\mathcal{F}(\mathcal{U}_{\text{in}}(\mu); \mu)} \tag{9.18}$$

where $\mathcal{U}_{\text{in}}(\mu) = [U_{\text{in}}^+, U_{\text{in}}^-]^T \in \mathcal{X}$ satisfies

$$\mathcal{A}_{\text{in}}(\mathcal{U}_{\text{in}}(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}; \mu), \ \forall \mathcal{V} \in \mathcal{X}; \tag{9.19}$$

here $\mathcal{X} \in X^2$ and $\mathcal{V} = [V^+, V^-]^T \in \mathcal{X}$.

Outlet System

Given $\mu \in \mathcal{D}$, calculate the "truth" output of interest

$$\mathcal{S}_{\text{out}}(\mu) = \overline{\mathcal{F}(\mathcal{U}_{\text{out}}(\mu); \mu)} \tag{9.20}$$

where $\mathcal{U}_{\text{out}}(\mu) = [U_{\text{out}}^+, U_{\text{out}}^-]^T \in \mathcal{X}$ satisfies

$$\mathcal{A}_{\text{out}}(\mathcal{U}_{\text{out}}(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}; \mu), \ \forall \mathcal{V} \in \mathcal{X}; \tag{9.21}$$

here $\mathcal{X} \in X^2$ and $\mathcal{V} = [V^+, V^-]^T \in \mathcal{X}$.

The "expanded" complex function space $\mathcal{X} \in X^2$ is associated with an inner product

$$(\mathcal{W}, \mathcal{V})_{\mathcal{X}} = (W^+, V^+)_X + (W^-, V^-)_X \tag{9.22}$$

and norm

$$\|\mathcal{W}\|_{\mathcal{X}} = (\mathcal{W}, \mathcal{W})_{\mathcal{X}}; \tag{9.23}$$

where $\mathcal{W} = [W^+, W^-]^T \in \mathcal{X}$, $\mathcal{V} = [V^+, V^-]^T$ and $W^+, W^-, V^+, V^- \in X$. Note that the dimension

of $\mathcal{X}$ is twice the dimension of X.

9.2.5 Bilinear Forms and Affine Decomposition

Given $\mu \in \mathcal{D}$ and $\mathcal{W}, \mathcal{V} \in \mathcal{X}$, $\mathcal{A}_{\text{in}}$ can be affinely separated as

$$\mathcal{A}_{\text{in}}(\mathcal{W}, \mathcal{V}; \mu) = \sum_{q=1}^{Q_{\mathcal{A},\text{in}}} \Theta^q_{\mathcal{A},\text{in}}(\mu)\, \mathcal{A}^q_{\text{in}}(\mathcal{U}, \mathcal{V}); \tag{9.24}$$

$\mathcal{A}_{\text{out}}$ can be similarly decomposed as

$$\mathcal{A}_{\text{out}}(\mathcal{W}, \mathcal{V}; \mu) = \sum_{q=1}^{Q_{\mathcal{A},\text{out}}} \Theta^q_{\mathcal{A},\text{in}}(\mu)\, \mathcal{A}^q_{\text{out}}(\mathcal{W}, \mathcal{V}). \tag{9.25}$$

We identify the μ-independent bilinear forms $\mathcal{A}^q_{\text{in}}(\mathcal{U}_{\text{in}}, \mathcal{V})$ and $\mathcal{A}^q_{\text{out}}(\mathcal{U}_{\text{out}}, \mathcal{V})$ in Table 9.2. Note that the number of terms in the affine decomposition of the bilinear forms associated with the inlet and outlet systems are the same, $Q_{\mathcal{A},\text{in}} = Q_{\mathcal{A},\text{out}} \equiv Q_{\mathcal{A}}$; similarly the parameter-dependent coefficients are identical for the inlet and outlet systems, $\Theta^q_{\mathcal{A},\text{in}}(\mu) = \Theta^q_{\mathcal{A},\text{out}}(\mu) \equiv \Theta^q_{\mathcal{A}}$ for $q = 1, \ldots, Q_{\mathcal{A}}$. We present the $\Theta^q_{\mathcal{A}}(\mu)$ in Table 9.3.

We can obtain the field solution for the original problem, $u(\mu)$, from both the expanded formulations. Given $\mu \in \mathcal{D}$, we solve the expanded inlet system to obtain $\mathcal{U}_{\text{in}}(\mu) = [U^+_{\text{in}}(\mu), U^-_{\text{in}}(\mu)]^T$; similarly, we solve the expanded outlet system to obtain $\mathcal{U}_{\text{out}}(\mu) = [U^+_{\text{out}}(\mu), U^-_{\text{out}}(\mu)]^T$. We have

$$\begin{aligned} u(\mu) &= U^+_{\text{in}}(\mu) + U^-_{\text{in}}(\mu) \\ &= U^+_{\text{out}}(\mu) + U^-_{\text{out}}(\mu). \end{aligned} \tag{9.26}$$

We should bear in mind that the actual field solution for the original problem, $u(\mu)$, is identical whether we use the inlet or outlet systems. We only need two separate formulations because we want to solve for two *separate* quadratic outputs $\mathcal{S}_{\text{in}}(\mu)$ and $\mathcal{S}_{\text{out}}(\mu)$ respectively.

9.2.6 Truth Solutions

We next show the "truth" mesh for the band-stop filter problem in Figure 9-4; note that our reduced basis solution error will be measured against the solution on the truth mesh.

$\mathcal{A}^q_{\text{in}}(\mathcal{U}_{\text{in}}, \mathcal{V})$	
$q = 1$	$a^1(U^+_{\text{in}}, V^+; \mu) + \overline{a^1(V^+, U^+_{\text{in}}; \mu)}$ $+a^1(U^-_{\text{in}}, V^+; \mu) - \overline{a^1(V^+, U^-_{\text{in}}; \mu)}$ $+ \overline{a^1(V^-, U^+_{\text{in}}; \mu)} - a^1(U^+_{\text{in}}, V^-; \mu)$ $- \overline{a^1(V^-, U^-_{\text{in}}; \mu)} - a^1(U^-_{\text{in}}, V^-; \mu)$ $-Q_{\text{in}}(U^+_{\text{in}}, V^+; \mu) - Q_{\text{in}}(U^-_{\text{in}}, V^+; \mu)$ $- Q_{\text{in}}(U^+_{\text{in}}, V^-; \mu) - Q_{\text{in}}(U^-_{\text{in}}, V^-; \mu)$
$q \neq 1$	$a^q(U^+_{\text{in}}, V^+; \mu) + \overline{a^q(V^+, U^+_{\text{in}}; \mu)}$ $+a^q(U^-_{\text{in}}, V^+; \mu) - \overline{a^q(V^+, U^-_{\text{in}}; \mu)}$ $+ \overline{a^q(V^-, U^+_{\text{in}}; \mu)} - a^q(U^+_{\text{in}}, V^-; \mu)$ $- \overline{a^q(V^-, U^-_{\text{in}}; \mu)} - a^q(U^-_{\text{in}}, V^-; \mu)$
$\mathcal{A}^q_{\text{out}}(\mathcal{U}_{\text{out}}, \mathcal{V})$	
$q = 1$	$a^1(U^+_{\text{out}}, V^+; \mu) + \overline{a^1(V^+, U^+_{\text{out}}; \mu)}$ $+a^1(U^-_{\text{out}}, V^+; \mu) - \overline{a^1(V^+, U^-_{\text{out}}; \mu)}$ $+ \overline{a^1(V^-, U^+_{\text{out}}; \mu)} - a^1(U^+_{\text{out}}, V^-; \mu)$ $- \overline{a^1(V^-, U^-_{\text{out}}; \mu)} - a^1(U^-_{\text{out}}, V^-; \mu)$ $-Q_{\text{out}}(U^+_{\text{out}}, V^+; \mu) - Q_{\text{out}}(U^-_{\text{out}}, V^+; \mu)$ $- Q_{\text{out}}(U^+_{\text{out}}, V^-; \mu) - Q_{\text{out}}(U^-_{\text{out}}, V^-; \mu)$
$q \neq 1$	$a^q(U^+_{\text{out}}, V^+; \mu) + \overline{a^q(V^+, U^+_{\text{out}}; \mu)}$ $+a^q(U^-_{\text{out}}, V^+; \mu) - \overline{a^q(V^+, U^-_{\text{out}}; \mu)}$ $+ \overline{a^q(V^-, U^+_{\text{out}}; \mu)} - a^q(U^+_{\text{out}}, V^-; \mu)$ $- \overline{a^q(V^-, U^-_{\text{out}}; \mu)} - a^q(U^-_{\text{out}}, V^-; \mu)$

Table 9.2: Affine decomposition of $\mathcal{A}_{\text{in}}(\mathcal{U}_{\text{in}}(\mu), \mathcal{V}; \mu)$ and $\mathcal{A}_{\text{out}}(\mathcal{U}_{\text{out}}(\mu), \mathcal{V}; \mu)$ for the Band-Stop Filter. Here $\mathcal{U}_{\text{in}}(\mu)$, $\mathcal{U}_{\text{out}}(\mu)$, and $\mathcal{V}$ belong to the complex vector space $\mathcal{X}$. The first bilinear form, $q = 1$, is different for the inlet and outlet expanded formulations. The bilinear forms $\mathcal{A}^q_{\text{in}}(\cdot, \cdot)$ and $\mathcal{A}^q_{\text{out}}(\cdot, \cdot)$ are identical for $q = 2, \ldots, Q_{\mathcal{A}}$ ($Q_{\mathcal{A},\text{in}} = Q_{\mathcal{A},\text{out}} \equiv Q_{\mathcal{A}}$).

q	1	2	3	4	5	6	7
$\Theta^q_{\mathcal{A}}(\mu)$	1	μ_2	$\frac{1}{\mu_2}$	$-\mu_1$	$-\mu_1\,\mu_2$	$\frac{\mathrm{i}\sqrt{\mu_1}}{Z_R}$	$\mathrm{i}\sqrt{\mu_1}$

Table 9.3: $\Theta^q_{\mathcal{A}}(\mu)$ for the expanded inlet and outlet systems of the Band-Stop Filter. Note that the parameter-dependent coefficients are identical for both inlet and outlet systems.

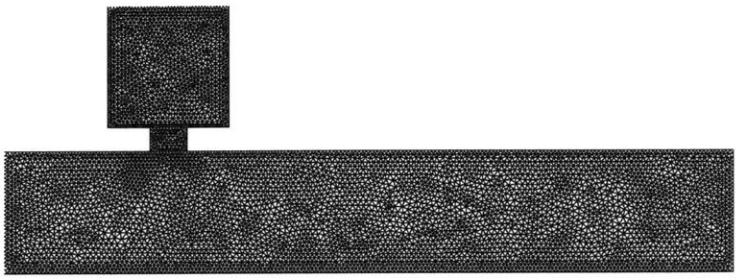


Figure 9-4: Truth mesh for the band-stop filter problem: note here that the solutions are plotted on the original mesh with $\mathcal{N} = 5815$; the expanded system has $\mathcal{N} = 11630$ unknowns.

Figure 9-5 shows a few FEM solutions for slightly different parameters. We observe that the parameter dependence induces rich variations in solution structure and magnitude. Approximating the solution will be quite difficult, and N may be quite large to achieve sufficient accuracy. Recall that in our case, the specification of the radiation boundary condition required only one propagating mode ($N_{\text{prop}} = 1$) at outflow for our chosen frequency range. We show solutions for different k^2 — at low frequencies 9-5(a) the incident wave propagates slowly, as k^2 increases, the wavelength of the propagating wave gets shorter and the domain length includes more than one wavelength as Figures 9-5(c) and 9-5(d) demonstrate. In Figure 9-5(b) we see the effect of the Helmholtz resonator where the incident wave is reflected back and the pressure intensity at outflow is significantly smaller compared to the pressure intensity at inflow; this corresponds to the resonant frequency of the Helmholtz resonator. For frequencies that are further away from the resonant frequency of the Helmholtz resonator, the Helmholtz resonator does effect the solution but the reflection is significantly lower.

Recall that the reduced basis approximation and associated *a posteriori* error estimators are

developed not just for the output of interest but also for the actual solution; for related numerical results including the convergence and effectivities, rigorousness of our error bounds, as well as computational savings relative to the "truth" solution please look in Sections 6.3.3 and 8.4.

9.2.7 Reduced Basis Approximation

Offline Convergence

The reduced basis for the inlet and outlet systems is constructed by running the greedy algorithm offline. In each case, our intent is to reduce the maximum relative error in the output over a large sample in $\mathcal{D}$ below a specified tolerance. Since we *never* calculate the exact error in the output, we use the *a posteriori* error estimators to guide us in the selection of the sample points; the reduced basis space is the span of the solutions at these sample points.

We denote $\mathcal{S}_{\text{in},N}(\mu)$ and $\mathcal{S}_{\text{out},N}(\mu)$ as the reduced basis outputs for the inlet and outlet systems respectively. We also introduce the outputs bounds (8.29) of the inlet and outlet systems (please see Section 8.4 for a proof)

$$\Delta_N^{\mathcal{S}_{\text{in}}}(\mu) = \frac{\epsilon_N^{\mathcal{S}_{\text{in}}}(\mu)^2}{\beta_{\text{LB,in}}(\mu)}, \tag{9.27}$$

$$\Delta_N^{\mathcal{S}_{\text{out}}}(\mu) = \frac{\epsilon_N^{\mathcal{S}_{\text{out}}}(\mu)^2}{\beta_{\text{LB,out}}(\mu)} \tag{9.28}$$

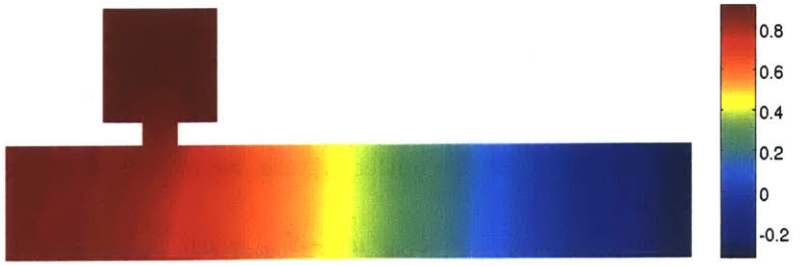
such that

$$\begin{aligned} |\mathcal{S}_{\text{in}}(\mu) - \mathcal{S}_{\text{in},N}(\mu)| &\leq \Delta_N^{\mathcal{S}_{\text{in}}}(\mu) \\ |\mathcal{S}_{\text{out}}(\mu) - \mathcal{S}_{\text{out},N}(\mu)| &\leq \Delta_N^{\mathcal{S}_{\text{out}}}(\mu). \end{aligned}$$

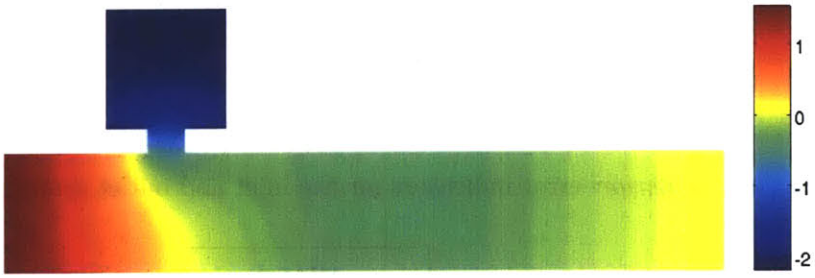
Here $\epsilon_N^{\mathcal{S}_{\text{in}}}(\mu)$ (respectively, $\epsilon_N^{\mathcal{S}_{\text{out}}}(\mu)$) is the dual norm of the residual for the inlet (respectively, outlet) system. $\beta_{\text{LB,in}}(\mu)$ and $\beta_{\text{LB,out}}(\mu)$ are the inf-sup lower bound approximations for the inlet and outlet system respectively. The construction of the inf-sup lower bound approximation is discussed in Section 9.2.8.

Given $\mu \in \mathcal{D}$, for any choice of N, we denote the relative output error for the inlet system by

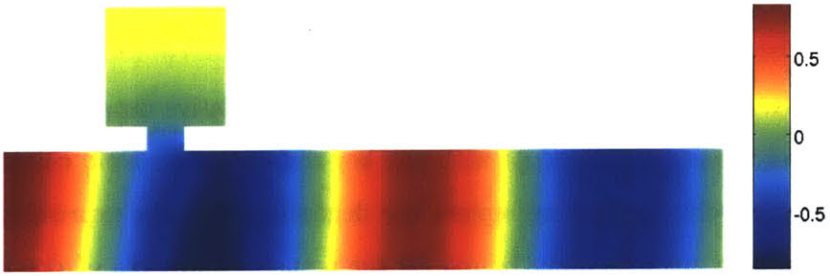
$$E_{\text{in},N}(\mu) = \frac{|\mathcal{S}_{\text{in}}(\mu) - \mathcal{S}_{\text{in},N}(\mu)|}{\mathcal{S}_{\text{in}}(\mu)}, \tag{9.29}$$



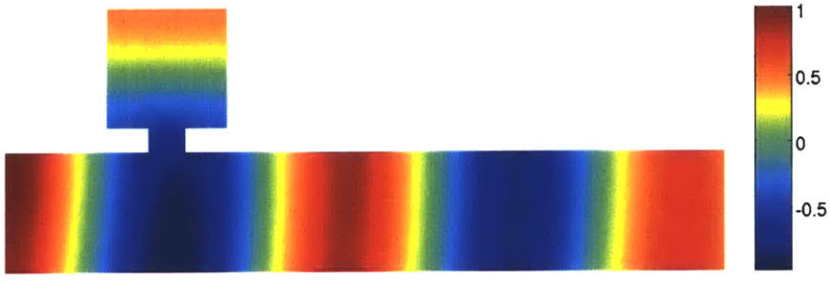
(a)



(b)



(c)



(d)

Figure 9-5: Band-Stop Filter FEM solutions for (a) $\mu = [0.1, 0.75]$; (b) $\mu = [0.56, 1.0]$; (c) $\mu = [2.5, 0.75]$; and (d) $\mu = [5.0, 1.25]$.

and the relative output error bound by

$$\mathcal{E}_{\text{in},N}(\mu) = \frac{\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)}{\mathcal{S}_{\text{in}}(\mu)}; \tag{9.30}$$

thus,

$$E_{\text{in},N}(\mu) \leq \mathcal{E}_{\text{in},N}(\mu), \tag{9.31}$$

for all N and for all $\mu \in \mathcal{D}$. Similarly, for the outlet system, we introduce the relative output error

$$E_{\text{out},N}(\mu) = \frac{|\mathcal{S}_{\text{out}}(\mu) - \mathcal{S}_{\text{out},N}(\mu)|}{\mathcal{S}_{\text{out}}(\mu)}, \tag{9.32}$$

and the relative output error bound by

$$\mathcal{E}_{\text{out},N}(\mu) = \frac{\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}{\mathcal{S}_{\text{out}}(\mu)}. \tag{9.33}$$

We then introduce the *a posteriori* error indicators for the inlet and outlet systems as

$$\Delta_N^{\text{in}}(\mu) = \max_{\mu \in \mathcal{D}} \sqrt{\frac{\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)}{\mathcal{S}_{\text{in},N}(\mu) - \Delta_N^{\mathcal{S}_{\text{in}}}(\mu)}}, \tag{9.34}$$

$$\Delta_N^{\text{out}}(\mu) = \max_{\mu \in \mathcal{D}} \sqrt{\frac{\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}{\mathcal{S}_{\text{out},N}(\mu) - \Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}}. \tag{9.35}$$

$\Delta_N^{\text{in}}(\mu)$ (respectively, $\Delta_N^{\text{out}}(\mu)$) is a measure of the maximum relative output error bound $\max_{\mu \in \mathcal{D}} \mathcal{E}_{\text{in},N}(\mu)$ (respectively, $\max_{\mu \in \mathcal{D}} \mathcal{E}_{\text{out},N}(\mu)$) with one key difference: to avoid calculating the actual output $\mathcal{S}_{\text{in}}(\mu)$ (respectively, $\mathcal{S}_{\text{out}}(\mu)$) we replace the denominator in the expression for the relative output error bound by a *rigorous* lower bound for the actual output: $\mathcal{S}_{\text{in},N}(\mu) - \Delta_N^{\mathcal{S}_{\text{in}}}(\mu) \leq \mathcal{S}_{\text{in}}(\mu)$ (respectively, $\mathcal{S}_{\text{out},N}(\mu) - \Delta_N^{\mathcal{S}_{\text{out}}}(\mu) \leq \mathcal{S}_{\text{out}}(\mu)$).

For the inlet system, we obtain

$$E_{\text{in},N}(\mu) = \frac{|\mathcal{S}_{\text{in}}(\mu) - \mathcal{S}_{\text{in},N}(\mu)|}{\mathcal{S}_{\text{in}}(\mu)} \tag{9.36}$$

$$\leq \frac{|\mathcal{S}_{\text{in}}(\mu) - \mathcal{S}_{\text{in},N}(\mu)|}{\mathcal{S}_{\text{in},N}(\mu) - \Delta_N^{\mathcal{S}_{\text{in}}}(\mu)} \tag{9.37}$$

$$\leq \frac{\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)}{\mathcal{S}_{\text{in},N}(\mu) - \Delta_N^{\mathcal{S}_{\text{in}}}(\mu)} \tag{9.38}$$

$$\leq \sqrt{\Delta_N^{\text{in}}(\mu)}, \qquad \forall \mu \in \mathcal{D},\ \forall N; \tag{9.39}$$

similarly, we obtain for the outlet system

$$E_{\text{out},N}(\mu) \leq \sqrt{\Delta_N^{\text{out}}(\mu)}, \qquad \forall \mu \in \mathcal{D},\ \forall N. \tag{9.40}$$

The greedy algorithm is run over a large training sample set $\Xi_{\text{train}} \subset \mathcal{D}$ of size 2500; we specify a termination error tolerance $\epsilon_{\text{tol,min}} = 1\text{E}-05$. We then pursue our offline greedy algorithm to construct our reduced basis space over the exhaustive sample set Ξ_{train}. Note that for small N, the output bounds although rigorous will be very large: thus, $\mathcal{S}_{\text{in},N}(\mu) \leq \Delta_N^{\mathcal{S}_{\text{in}}}(\mu)$ for some $\mu \in \mathcal{D}$; in such cases, we replace $\mathcal{S}_{\text{in},N}(\mu) - \Delta_N^{\mathcal{S}_{\text{in}}}(\mu)$ in the denominator for Δ_N^{in} by $\mathcal{S}_{\text{in},N}(\mu)$. With increasing N, the output bounds will get tighter and we can use the exact expression for Δ_N^{in}. This approach is followed in the construction of the reduced basis space for the outlet system as well.

We present the offline convergence of the maximum relative error bound ($\Delta_N^{\text{in}}(\mu)$ and $\Delta_N^{\text{out}}(\mu)$) used to construct our reduced basis spaces W_N^{in} and W_N^{out} for the inlet and outlet systems respectively. Figure 9.2.7 plots Δ_N^{in} and Δ_N^{out} as a function of the number of offline iterations (or the size of the reduced basis space) for the inlet (respectively, outlet) system. We obtain $N_{\text{max,in}} = 80$ and $N_{\text{max,out}} = 74$ respectively for $\epsilon_{\text{tol,min}} = 1\text{E}-05$. Note that the actual relative error is less than $1\text{E}-10$ for $N_{\text{max}} = 80$ for the inlet system, and $N_{\text{max}} = 74$ for the outlet system. From Figure 9.2.7, we note that there is some minimum number of basis required before the reduced basis converges uniformly in the error. The rapid convergence also confirms the efficiency of the offline sampling procedure and the utility of the error bound as a "good" surrogate for the exact error (please see Section 9.2.9 for the results for the online error estimation of the outputs of interest).

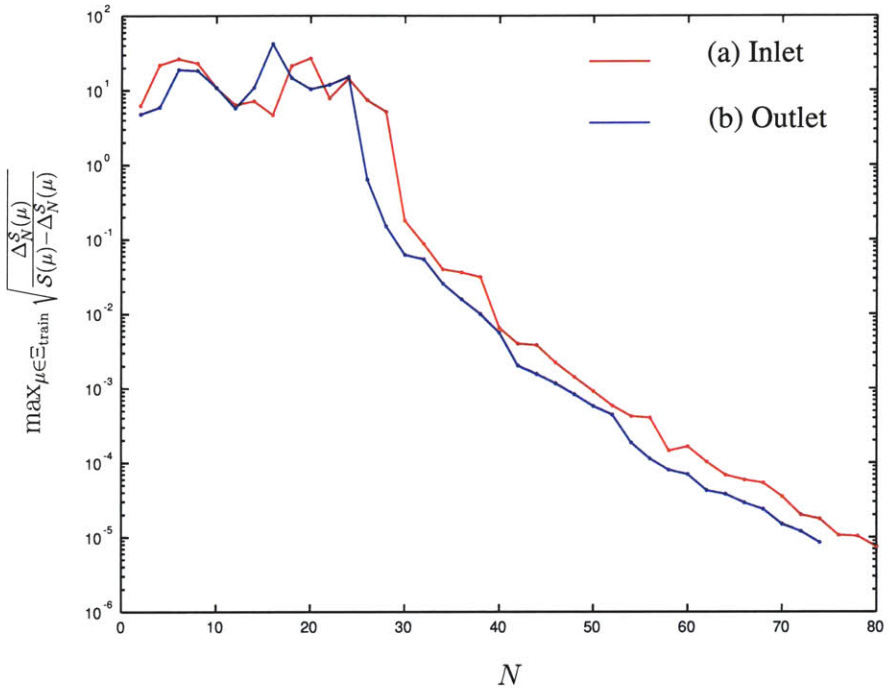


Figure 9-6: Greedy algorithm to generate reduced basis for the band-stop filter - maximum offline relative error bound for (a) inlet system: $\Delta_N^{\text{in}} = \max_{\mu\in\Xi_{\text{train}}} \sqrt{\frac{\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)}{\mathcal{S}_{\text{in},N}(\mu)-\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)}}$ and (b) outlet system: $\Delta_N^{\text{out}} = \max_{\mu\in\Xi_{\text{train}}} \sqrt{\frac{\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}{\mathcal{S}_{\text{out},N}(\mu)-\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}}$ shown as a function of N. We specify $\varepsilon_{\text{tol,min}} = 1\text{E}-05$, $\dim(\Xi_{\text{train}}) = 2500$.

Online Convergence

We test the accuracy of the obtained reduced basis approximations over an exhaustive sample set Ξ_{verif} of size 3000. For the verification sample, we compute the "truth" solutions and present the maximum relative error for different N. Recall that $E_{\text{in},N}(\mu)$ (respectively, $E_{\text{out},N}(\mu)$) denotes the relative error in the output for the inlet (respectively, outlet) system for any choice of N and for any $\mu \in \mathcal{D}$.

We present in Table 9.4 the convergence of the reduced basis output of the inlet system, $\mathcal{S}_{\text{in},N}(\mu)$, to its "truth" counterpart $\mathcal{S}_{\text{in}}(\mu)$ given by (9.18) ($\equiv s_{\text{in}}(\mu)$): we show the maximum relative error $E_{\text{in},N}(\mu)$ as a function of N.

N	$\max_{\mu\in\Xi_{\text{verif}}} E_{\text{in},N}(\mu)$	Online Time $\mathcal{S}_{\text{in},N}$
25	1.55E-01	1.58E-03
30	3.67E-04	1.67E-03
35	3.05E-05	2.56E-03
40	6.98E-07	2.78E-03
45	4.82E-07	3.41E-03

Table 9.4: Band-Stop Filter: Reduced basis convergence result and the online time to evaluate $\mathcal{S}_{\text{in},N}(\mu)$ as a function of N; the timing results are normalized with respect to the "average" time to calculate the "exact" output $s(\mu)$ directly.

We present in Table 9.5 the convergence of the reduced basis output for the outlet system, $\mathcal{S}_{\text{out},N}(\mu)$, to $\mathcal{S}_{\text{out}}(\mu)$ given by (9.20) ($\equiv s_{\text{out}}(\mu)$) over our verification sample set Ξ_{verif}): we show the maximum relative error $E_{\text{out},N}(\mu)$ as a function of N.

N	$\max_{\mu\in\Xi_{\text{verif}}} E_{\text{out},N}(\mu)$	Online Time $\mathcal{S}_{\text{out},N}$
25	5.37E-02	1.69E-03
30	1.56E-04	1.72E-03
35	1.63E-05	2.28E-03
40	4.04E-07	2.69E-03
45	1.38E-07	3.38E-03

Table 9.5: Band Pass Filter: Reduced basis convergence result and the online time to evaluate $\mathcal{S}_{\text{out},N}(\mu)$ as a function of N; the timing results are normalized with respect to the "average" time to calculate the "exact" output $s(\mu)$ directly.

We also present in Table 9.4 (respectively, Table 9.5) the online reduced basis computational cost to evaluate $\mathcal{S}_{\text{in},N}(\mu)$ (respectively, $\mathcal{S}_{\text{out},N}(\mu)$) compared to the finite element cost to evaluate

$s(\mu)$ for any given μ. We can see that the reduced basis solution $\mathcal{S}_{\text{in},N}(\mu)$ (respectively, $\mathcal{S}_{\text{out},N}(\mu)$) and the "truth" solution $\mathcal{S}_{\text{in}}(\mu)$ (respectively, $\mathcal{S}_{\text{out}}(\mu)$) are indistinguishable for $N \geq 35$. We defer the discussion on computational cost to Section 9.2.9 where we compute the true output $s(\mu)$.

9.2.8 Inf-Sup Lower Bound Approximation

A crucial ingredient of our *a posteriori* error estimators is the inf-sup lower bound. Here we discuss the inf-sup lower bound approximation using the SCM approach for the inlet and outlet systems respectively. We construct the constraint set $\mathcal{C}_K^{\text{in}}$ (respectively, $\mathcal{C}_K^{\text{out}}$) for the inlet (respectively, outlet) system *offline*; and verify the accuracy of the inf-sup lower bounds *online*. We denote $\sigma(\mu)$ as the square of the inf-sup parameter, $\sigma(\mu) = \beta^2(\mu)$. Given $\mu \in \mathcal{D}$, we denote $\sigma_{\text{LB}}(\mu)$ and $\sigma_{\text{UB}}(\mu)$ as the SCM lower and upper bounds to $\sigma(\mu)$; we have $\sigma_{\text{LB}}(\mu) \leq \sigma(\mu) \leq \sigma_{\text{UB}}(\mu), \forall \mu \in \mathcal{D}$.

Given $\mu \in \mathcal{D}$, we denote the inf-sup lower bound approximation for the inlet (respectively, outlet) system as $\beta_{\text{LB,in}}(\mu)$ (respectively, $\beta_{\text{LB,out}}(\mu)$). We also identify for all $\mu \in \mathcal{D}$, $\sigma_{\text{UB,in}}(\mu)$ (respectively, $\sigma_{\text{LB,out}}(\mu)$) and $\sigma_{\text{LB,in}}(\mu)$ (respectively, $\sigma_{\text{LB,out}}(\mu)$) as the upper and lower bounds of $\sigma(\mu)$ for the inlet (respectively, outlet) system.

Construction of $\mathcal{C}_K$

We first introduce an exhaustive sample-set $\Xi_{\text{train}} \subset \mathcal{D}$ of size 2000. For the SCM algorithm, we specify $\mathcal{M}_\alpha = 20$, $\mathcal{M}_+ = 10$ and a termination criteria of $\epsilon_{\text{tol,min}} = 0.85$. We aim to find the set $\mathcal{C}_K \in \Xi_{\text{train}} \subset \mathcal{D}$ such that

$$\max_{\mu \in \Xi_{\text{train}}} \frac{\sigma_{\text{UB}}(\mu; \mathcal{C}_K) - \sigma_{\text{LB}}(\mu; \mathcal{C}_K)}{\sigma_{\text{UB}}(\mu; \mathcal{C}_K)} \leq \epsilon_{\text{tol,min}}. \tag{9.41}$$

Given $\mu \in \mathcal{D}$, and a SCM constraint set $\mathcal{C}_k$ of dimension $k \in \mathbb{R}_+$, we denote the relative gap between $\sigma_{\text{UB}}(\mu)$ and $\sigma_{\text{LB}}(\mu)$ by $\epsilon_k(\mu)$ shown below

$$\epsilon_k(\mu) = \frac{\sigma_{\text{UB}}(\mu; \mathcal{C}_k) - \sigma_{\text{LB}}(\mu; \mathcal{C}_k)}{\sigma_{\text{UB}}(\mu; \mathcal{C}_k)}. \tag{9.42}$$

For given $k \in \mathbb{R}_+$, we denote $\mathcal{C}_k^{\text{in}}$ and $\mathcal{C}_k^{\text{out}}$ as the SCM constraint sets associated with the inf-sup lower bound approximation for the inlet and outlet systems respectively. For any $\mu \in \Xi_{\text{train}}$, we

identify $\epsilon_k^{\text{in}}(\mu)$ and $\epsilon_k^{\text{out}}(\mu)$ as

$$\epsilon_k^{\text{in}}(\mu) = \frac{\sigma_{\text{UB,in}}(\mu; \mathcal{C}_k^{\text{in}}) - \sigma_{\text{LB,in}}(\mu; \mathcal{C}_k^{\text{in}})}{\sigma_{\text{UB,in}}(\mu; \mathcal{C}_k^{\text{in}})}, \tag{9.43}$$

$$\epsilon_k^{\text{out}}(\mu) = \frac{\sigma_{\text{UB,out}}(\mu; \mathcal{C}_k^{\text{out}}) - \sigma_{\text{LB,out}}(\mu; \mathcal{C}_k^{\text{out}})}{\sigma_{\text{UB,out}}(\mu; \mathcal{C}_k^{\text{out}})}, \tag{9.44}$$

for the inlet and outlet system respectively.

We run the SCM algorithm for the inlet and outlet systems specifying $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$: we obtain $\dim(\mathcal{C}_K^{\text{in}}) = 970$ and $\dim(\mathcal{C}_K^{\text{out}}) = 425$ respectively.

Given the training set Ξ_{train} of 2000 sample points, we define $\epsilon_k^{\text{max,in}}$, $\epsilon_k^{\text{mean,in}}$, and $\epsilon_k^{\text{med,in}}$ as the maximum of $\epsilon_k^{\text{in}}(\mu)$ in Ξ_{train}, the mean of $\epsilon_k^{\text{in}}(\mu)$ in Ξ_{train}, and the median of $\epsilon_k^{\text{in}}(\mu)$ in Ξ_{train} for fixed $k \in \mathbb{R}^+$. In Table 9.6, we present the maximum, mean, and median values of $\epsilon_k^{\text{in}}(\mu)$ obtained from the SCM algorithm for the inlet system.

k	$\epsilon_k^{\text{max,in}}$	$\epsilon_k^{\text{mean,in}}$	$\epsilon_k^{\text{med,in}}$
250	1.1697	0.7917	1.0026
500	1.1503	0.5989	0.6842
750	1.0164	0.4374	0.3407
K=970	0.8476	0.259	0.2653

Table 9.6: Band-Stop Filter: offline construction of $\mathcal{C}_K^{\text{in}}$, the SCM constraint set for the inlet system. Here $\dim(\Xi_{\text{train}}) = 2000$, $\epsilon_{\text{tol,min}} = 0.85$, $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$. We obtain $\dim(\mathcal{C}_K^{\text{in}}) = 970$ for $\mu \in \mathcal{D} = [0.1, 5] \times [0.75, 1.25]$.

For fixed k, we denote $\epsilon_k^{\text{max,out}}$, $\epsilon_k^{\text{mean,out}}$, and $\epsilon_k^{\text{med,out}}$, as the maximum, mean, and median values of $\epsilon_k^{\text{out}}(\mu)$ for all samples in Ξ_{train}. In Table 9.7, we present the maximum, mean, and median values of $\epsilon_k^{\text{out}}(\mu)$ obtained from the SCM algorithm for the outlet system.

k	$\epsilon_k^{\text{max,out}}$	$\epsilon_k^{\text{mean,out}}$	$\epsilon_k^{\text{med,out}}$
100	1.2306	0.9532	1.0086
200	1.1323	0.8152	1.0032
300	1.2191	0.6807	0.7884
K=425	0.8491	0.3453	0.3484

Table 9.7: Band-Stop Filter: offline construction of $\mathcal{C}_K^{\text{out}}$, the SCM constraint set for the outlet system. Here $\dim(\Xi_{\text{train}}) = 2000$, $\epsilon_{\text{tol,min}} = 0.85$, $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$. We obtain $\dim(\mathcal{C}_K^{\text{out}}) = 425$ for $\mu \in \mathcal{D} = [0.1, 5] \times [0.75, 1.25]$.

Note that it is not necessary that the convergence of $\epsilon_k^{\max}$ with k be smooth: for some sample

points it is quite possible that the the obtained lower bound is better with a smaller number of constraints — we observe this effect for the outlet system where $\epsilon_k^{\max,\text{out}}$ is larger for $k = 300$ compared to $k = 200$. There are always some sample points where the lower bound improves *only* when a particular constraint is activated; in a similar manner, at certain sample points, the lower bound is worse when we activate a particular constraint — however, on *average* the lower bound approximation improves with increased size of k.

We also observe that the set $\mathcal{C}_K^{\text{in}}$ is much larger than $\mathcal{C}_K^{\text{out}}$. Figure 9.2.8 helps explain the large difference in the size of $\mathcal{C}_K$ for the inlet and outlet systems. To understand why we need such a comparatively larger constraint set $\mathcal{C}_K^{\text{in}}$ we plot $\sigma(\mu)$ as a function of μ_1 (for fixed $\mu_2 = 0.75$) for both the inlet and outlet systems. Notice that $\sigma(\mu)$ for the inlet system shown as a solid line in Figure 9.2.8 lies below the $\sigma(\mu)$ for the outlet system shown as a dashed line. Also notice that there is a sharp dip as $k^2 \rightarrow 1$ — this near-singular eigen value roughly corresponds to the resonant frequency of the Helmholtz resonator. The TC peak for $\mu_2 = 0.75$ is obtained around $k^2 = 0.8$ (per our earlier discussion this corresponds to the resonant frequency of the Helmholtz resonator); the addition of the quadratic functionals Q_{in} and Q_{out} into our bilinear forms $\mathcal{A}_{\text{in}}$ and $\mathcal{A}_{\text{out}}$ shifts the frequency where we obtain our lowest eigenvalue to the right. Also note, that relative to the inlet system, the outlet system is not as badly affected by the presence of the Helmholtz resonator as Figure 9.2.8 shows: the outlet system is a considerable distance away from the resonator and therefore its effect is also reduced. As we change the height of the Helmholtz resonator, the frequency where we obtain the minimum eigenvalue also shifts but the region where our near-singular eigenvalues lie remains concentrated in a small band around $k^2 = 1$ for all H. Accordingly, we need a large number of constraints in that region to make $\sigma_{\text{LB}}(\mu) > 0$. Hence, obtaining inf-sup lower bounds for the inlet system is much more difficult compared to the outlet system.

In Table 9.8, we present the values of K, the size of the constraint set $\mathcal{C}_K$, that we obtain for the inlet system using different values of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ for the offline greedy algorithm. For the sake of consistency, in each of these distinctly different runs of the greedy algorithm, we use the same training set, Ξ_{train} ($\dim(\Xi_{\text{train}}) = 2000$), and the same termination criteria, $\epsilon_{\text{tol,min}} = 0.85$. Table 9.8 also presents the breakdown of the *average* computational cost per iteration for the different activities involved in running the greedy algorithm. We identify five separate *average*

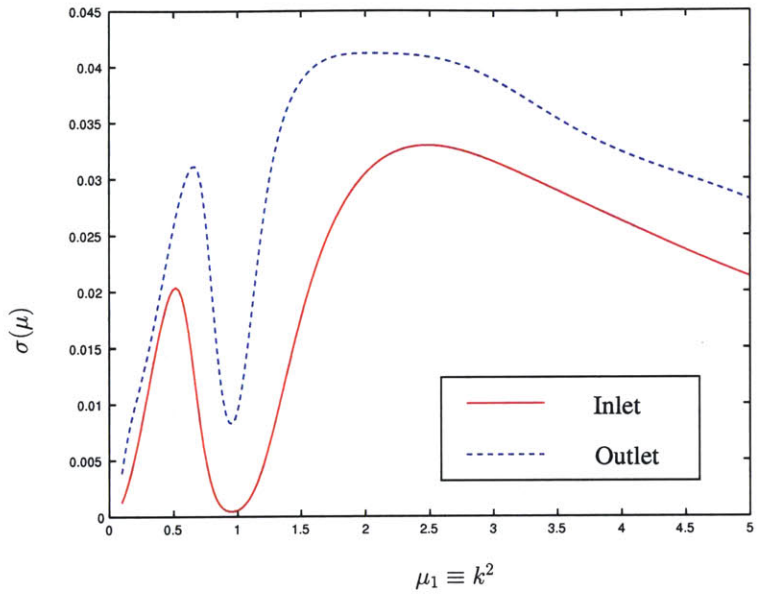


Figure 9-7: Band-stop filter: $\sigma(\mu)$ for inlet and outlet systems: we show $\sigma_{in}(\mu)$ in red and $\sigma_{\text{out}}(\mu)$ in blue. Here $\mu \equiv (\mu_1 \in [0.1, 5], \mu_2 = 0.75)$; we obtain $\min_\mu \sigma_{\text{in}}(\mu) = 4.64E-04$ and $\min_\mu \sigma_{\text{out}}(\mu) = 3.9E-03$.

costs: (i) the calculation of $\sigma(\omega_k)$ and $\mathcal{W}^*(\omega_k)$, the associated infimizer for the inf-sup at the "k"-th constraint point using the Lanczos algorithm (denoted by T_{lanc}) ; (ii) the cost of computing $\underline{z}^*(\omega_k) = (z_1^*(\omega_k), \ldots, z_{\hat{\mathcal{Q}}}^*(\omega_k))$ where $z_r^*(\omega_k) = \psi_r(\omega_k), 1 \le r \le \hat{\mathcal{Q}}$ (denoted by $T_{\text{infim coeffs}}$); (c) the cost of running Ξ_{train} LPs of size $2\hat{\mathcal{Q}} + \mathcal{M}_\alpha + \mathcal{M}_+$ (denoted by T_{LP}); (d) the cost of calculating the upper bound for the k-th constraint (denoted by T_{UB}); and (e) the cost of sorting the set $\mathcal{C}_K$ to identify the closest $\omega_k \in \mathcal{C}_K$ to use in the LP (denoted by T_{sort}). We note that we have roughly the same timings for the offline for different $\mathcal{M}_\alpha$ and $\mathcal{M}_+$. We have slightly increased cost of obtaining upper bounds as we increase $\mathcal{M}_\alpha$. The bulk of the per iteration cost is in calculating the infimizer coefficients and to run the Linear Programs. The choice of $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$ is quite justified — increasing $\mathcal{M}_+$ decreases K only mildly while reducing $\mathcal{M}_\alpha$ increases K without any significant reduction in computational cost.

In Table 9.9, we present the values of K, the size of the constraint set $\mathcal{C}_K$, that we obtain for the outlet system using different values of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ for the offline greedy algorithm. For the sake of consistency, in each of these distinctly different runs of the greedy algorithm, we use the same training set, Ξ_{train} ($\dim(\Xi_{\text{train}}) = 2000$), and the same termination criteria, $\epsilon_{\text{tol,min}} = 0.85$. Table 9.9 also presents the breakdown of the *average* computational cost per iteration for the

$\mathcal{M}_\alpha$	$\mathcal{M}_+$	K	T_{lanc}	$T_{\text{infim coeffs}}$	T_{LP}	T_{UB}	T_{sort}
10	10	983	3.48	11.39	19 91	0.04	0.3754
10	15	980	3.50	11.71	20.13	0.04	0.4164
15	10	977	3.51	11.70	20.43	0.04	0.7439
15	15	970	3.50	11.72	20.41	0.04	0.7359
20	10	970	3.51	11.67	20.55	0.04	1.07
20	15	969	3.51	11.76	20.63	0.04	1.13

Table 9.8: Results for the band-stop filter inlet system: Effect of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ on the size of the successive constraint set, C_K^{in}. We also present the breakdown of the *average* (or per iteration) cost (in seconds) of the greedy algorithm.

different activities involved in running the greedy algorithm. Note that the choice of $\mathcal{M}_\alpha = 15$ and $\mathcal{M}_+ = 15$ is sufficient. However, the initial convergence results (for the RB approximation and the inf-sup lower bounds) were obtained with the choice of $\mathcal{M}_+ = 20$ and $\mathcal{M}_+ = 10$; the results for different $(\mathcal{M}_\alpha, \mathcal{M}_+)$ were obtained very recently to better explore their effect on K on the outlet system. Also, note that the timings obtained for the outlet system in Table 9.9 are slightly higher relative to Table 9.8 — we actually expect the computational costs to be of the same per iteration for the different activities involved in running the greedy algorithm.

$\mathcal{M}_\alpha$	$\mathcal{M}_+$	K	T_{lanc}	$T_{\text{infim coeffs}}$	T_{LP}	T_{UB}	T_{sort}
10	10	442	3.55	11.08	21.87	0.03	0.3763
10	15	431	3.55	11.14	22.36	0.03	0.3788
15	10	430	3.73	11.77	22.33	0.03	0.7777
15	15	428	3.75	11.82	22.89	0.03	0.8045
20	10	425	3.76	11.70	22.12	0.03	1.1323
20	15	425	3.75	11.76	23.03	0.03	1.2214

Table 9.9: Results for the band-stop filter outlet system: Effect of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ on the size of the successive constraint set, C_K^{out}. We also present the breakdown of the *average* (or per iteration) cost (in seconds) of the greedy algorithm.

Post-processing of $\mathcal{C}_K$

In the construction of $\mathcal{C}_K$ (C_K^{in} or C_K^{out}), we specified $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$ respectively. Now for fixed K (for $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$), we experiment with different choices of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ to check for smaller values of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ that might suffice in the *online* stage if we relax our (ultra-conservative) offline termination criteria by a little. We want to reduce $(\mathcal{M}_\alpha, \mathcal{M}_+)$ but want to maintain approximately the same sharpness of the lower bound.

We present the results for the inlet (respectively, outlet) system in Table 9.10 (respectively, Table 9.11). We confirm that our post-processing sample set Ξ_{pp} satisfies the offline termination criteria for $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$.

For the inlet system, we arrive at two *minimum* viable settings: $(\mathcal{M}_\alpha = 20, \mathcal{M}_+ = 10)$ and $(\mathcal{M}_\alpha = 15, \mathcal{M}_+ = 15)$ for *online* computations. Both of these settings produce the same maximum $\frac{\beta}{\beta_{\text{LB}}}$ ratio. Furthermore, both settings require approximately the same amount of time for the LP.

$\mathcal{M}_\alpha$	$\mathcal{M}_+$	$\min_{\mu\in\Xi_{\text{pp}}} \frac{\sigma(\mu)}{\sigma_{\text{LB,in}}(\mu)}$	$\max_{\mu\in\Xi_{\text{pp}}} \frac{\sigma(\mu)}{\sigma_{\text{LB,in}}(\mu)}$	$\max_{\mu\in\Xi_{\text{pp}}} \epsilon_K^{\text{in}}(\mu)$
10	5	–	87.97	4.68
10	10	–	29.10	3.093
10	15	–	29.10	3.093
15	5	–	17.32	4.68
15	10	–	17.32	2.84
15	15	1.0038	5.98	8.41E-01
20	5	–	5.98	3.58
20	10	1.0038	5.98	8.47E-01
20	15	1.0038	5.98	8.41E-01

Table 9.10: Post-processing of $\mathcal{C}_K^{\text{in}}$ $(K = 970)$: approximation of the inf-sup lower bound for the band-stop filter inlet system for different choices of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ over Ξ_{pp} $(\dim(\Xi_{\text{pp}}) = 2000)$. Here $\sigma_{\text{UB,in}}(\mu)$ and $\sigma_{\text{LB,in}}(\mu)$ are the SCM upper and lower bounds for $\sigma(\mu)$ of the expanded inlet system.

For the outlet system, we may choose $\mathcal{M}_\alpha = 15$ and $\mathcal{M}_+ = 10$ for online computations when we use the SCM constraint set generated with $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$. Increasing $\mathcal{M}_\alpha$ from 15 to 20 did not improve the maximum $\frac{\beta}{\beta_{\text{LB}}}$ ratio. Furthermore, reducing $\mathcal{M}_\alpha$ from 20 to 15 while fixing $\mathcal{M}_+$, reduces the average inf-sup lower bound time from 12.52 milliseconds to 12.34 milliseconds — a decrease of only 1 percent.

We thus conclude, that our choice of $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$ represent fair choices for the *online* stage also.

Online Efficiency

We also compare the efficiency of the inf-sup lower bounds *and* the computational costs of the online LP evaluation amongst the different constraint sets obtained by specifying different $(\mathcal{M}_\alpha, \mathcal{M}_+)$ in the offline stage. We construct an exhaustive verification sample set Ξ_{verif} of size 2000, and compare the "exact" inf-sup with the lower bounds obtained from the SCM approach for different choices

$\mathcal{M}_\alpha$	$\mathcal{M}_+$	$\min_{\mu\in\Xi_{\rm pp}} \frac{\sigma(\mu)}{\sigma_{\rm LB,out}(\mu)}$	$\max_{\mu\in\Xi_{\rm pp}} \frac{\sigma(\mu)}{\sigma_{\rm LB,out}(\mu)}$	$\max_{\mu\in\Xi_{\rm pp}} \epsilon_k^{\rm out}(\mu)$
10	5	–	85.17	5.01
10	10	1.0012	85.17	1.18
10	15	1.0012	85.17	9.88E-01
15	5	–	7.92	3.23
15	10	1.0012	5.56	8.22E-01
15	15	1.0012	5.56	8.22E-01
20	5	1.0012	7.17	1.126
20	10	1.0012	5.56	8.22E-01
20	15	1.0012	5.56	8.22E-01

Table 9.11: Post-processing of $C_K^{\rm out}$ ($K = 425$): approximation of the inf-sup lower bound for the band-stop filter outlet system for different choices of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ over $\Xi_{\rm pp}$ ($\dim(\Xi_{\rm pp}) = 2000$). Here $\sigma_{\rm UB,out}(\mu)$ and $\sigma_{\rm LB,out}(\mu)$ are the SCM upper and lower bounds for $\sigma(\mu)$ of the expanded outlet system.

of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$. We present (a) the *average* cost for the inf-sup lower calculation $T_{\rm LB}^{\rm av}$, (b) the *average* computational savings relative to the true inf-sup evaluation $T_{\rm B}^{\rm av}/T_{\rm LB}^{\rm av}$, (c) the *average* cost for the inf-sup upper bound, (d) the *best* inf-sup to inf-sup lower bound ratio $\frac{\beta(\mu)}{\beta_{\rm LB}(\mu)}$, (e) the *worst* inf-sup to inf-sup lower bound ratio $\frac{\beta(\mu)}{\beta_{\rm LB}(\mu)}$, and (f) the maximum relative gap $\epsilon_k(\mu)$ for the inlet (respectively, outlet) system in Table 9.12 (respectively, Table 9.13. We note that the *average* time to compute the inf-sup lower bound over the verification set is slightly larger than those obtained over the training sample, however the timings are approximately of the same order.

We confirm that the required offline termination condition of the offline greedy algorithm, $\epsilon_{\rm tol,min} = 0.85$ is satisfied by all the SCM constraint sets. Notice that the LP (with 86 constraints, $\hat{\mathcal{Q}} = 28$, $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$) is extremely fast: the average time to compute the inf-sup lower bound for both the systems is $\approx$ 13 milli-seconds (ms). For $(\mathcal{M}_\alpha = 20, \mathcal{M}_+ = 10)$, the inlet inf-sup lower bound is $\approx$ 248 times faster than the inf-sup calculation, while the outlet inf-sup lower bound is $\approx$ 291 times faster. Moreover, both the SCM inf-sup lower bounds are extremely sharp: $1.001 \leq \frac{\beta_{\rm in}(\mu)}{\beta_{\rm LB,in}(\mu)} \leq 1.966$ for the inlet system, and $1.0003 \leq \frac{\beta_{\rm out}(\mu)}{\beta_{\rm LB,out}(\mu)} \leq 1.895$ for the outlet system.

We now discuss the offline-online computational cost tradeoff for the inlet system. From Table 9.12 we note that the LP time with $(\mathcal{M}_\alpha = 10, \mathcal{M}_+ = 10)$ setting is only 0.5 milliseconds faster relative to the $(\mathcal{M}_\alpha = 20, \mathcal{M}_+ = 10)$ setting but the offline cost is substantially higher. For the inlet system, we cannot make any improvement and choose the SCM constraint set of size $K = 970$ (for $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$) for *online* computations.

$\mathcal{M}_\alpha$	$\mathcal{M}_+$	K	T_{B}^{av}	$T_{\text{B}}^{\text{av}}/T_{\text{LB}}^{\text{av}}$	$\min_{\mu\in\Xi_{\text{verif}}} \frac{\beta(\mu)}{\beta_{\text{LB,in}}(\mu)}$	$\max_{\mu\in\Xi_{\text{verif}}} \frac{\beta(\mu)}{\beta_{\text{LB,in}}(\mu)}$	$\max_{\mu\in\Xi_{\text{verif}}} \epsilon_k^{\text{in}}(\mu)$
10	10	983	1.247E-02	259.16	1.003	1.962	8.491E-01
10	15	980	1.26E-02	256.52	1.0003	1.962	8.491E-01
15	10	977	1.267E-02	255.09	1.002	1.971	8.498E-01
15	15	970	1.296E-02	249.37	1.003	1.967	8.492E-01
20	10	970	1.303E-02	248.05	1.001	1.966	8.491E-01
20	15	969	1.324E-02	244.12	1.001	1.966	8.491E-01

Table 9.12: Results for the low pass filter inlet system: inf-sup lower bounds $\sigma_{\text{LB,in}}(\mu)$, and *averaged* timings for different choices of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ over a verification set Ξ_{verif} of size 2000. $T_{\text{LB}}^{\text{av}}$ is the average time to compute the lower bound in seconds; $T_{\text{B}}^{\text{av}}/T_{\text{LB}}^{\text{av}}$ is the ratio of the average times to compute the inf-sup and the lower bound respectively.

We now discuss the offline-online computational cost tradeoff for the outlet system. From Table 9.13 we note that the LP time with $(\mathcal{M}_\alpha = 10, \mathcal{M}_+ = 10)$ setting is again only 0.5 milliseconds faster relative to the $(\mathcal{M}_\alpha = 20, \mathcal{M}_+ = 10)$ setting. However, the constraint sets are of approximately the same size for the different $(\mathcal{M}_\alpha, \mathcal{M}_+)$ settings. For the outlet system, we are indifferent between the SCM constraint sets corresponding to $(\mathcal{M}_\alpha = 15, \mathcal{M}_+ = 10)$ and $(\mathcal{M}_\alpha = 20, \mathcal{M}_+ = 10)$ for *online* computations.

$\mathcal{M}_\alpha$	$\mathcal{M}_+$	K	T_{B}^{av}	$T_{\text{B}}^{\text{av}}/T_{\text{LB}}^{\text{av}}$	$\min_{\mu\in\Xi_{\text{verif}}} \frac{\beta(\mu)}{\beta_{\text{LB,out}}(\mu)}$	$\max_{\mu\in\Xi_{\text{verif}}} \frac{\beta(\mu)}{\beta_{\text{LB,out}}(\mu)}$	$\max_{\mu\in\Xi_{\text{verif}}} \epsilon_k^{\text{out}}(\mu)$
10	10	442	1.252E-02	300.61	1.0003	1.986	8.497E-01
10	15	431	1.266E-02	299.54	1.0003	1.986	8.493E-01
15	10	430	1.274E-02	297.66	1.0003	1.986	8.497E-01
15	15	428	1.306E-02	290.37	1.0003	1.986	8.493E-01
20	10	425	1.302E-02	291.26	1.0003	1.895	8.485E-01
20	15	425	1.131E-02	289.48	1.0003	1.895	8.485E-01

Table 9.13: Results for the low pass filter outlet system: inf-sup lower bounds $\sigma_{\text{LB,out}}(\mu)$, and *averaged* timings for different choices of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ over a verification set Ξ_{verif} of size 2000. $T_{\text{LB}}^{\text{av}}$ is the average time to compute the lower bound in seconds; $T_{\text{B}}^{\text{av}}/T_{\text{LB}}^{\text{av}}$ is the ratio of the average times to compute the inf-sup and the lower bound respectively.

9.2.9 A Posteriori Error Estimation

Output Bound Formulation

While we have obtained output bounds separately for the inlet and outlet systems we have not yet identified a *rigorous* output bound for the actual output $s(\mu)$ given by (9.15). Here we identify $\Delta_N^s(\mu)$ and prove that it is a rigorous bound.

Proposition 5. *Given* $\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)$ *(9.27) and* $\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)$ *(9.28) such that*

$$|\mathcal{S}_{\text{in}}(\mu) - \mathcal{S}_{\text{in},N}(\mu)| \leq \Delta_N^{\mathcal{S}_{\text{in}}}(\mu),\ \forall \mu \in \mathcal{D},\ \forall N \tag{9.45}$$

$$|\mathcal{S}_{\text{out}}(\mu) - \mathcal{S}_{\text{out},N}(\mu)| \leq \Delta_N^{\mathcal{S}_{\text{out}}}(\mu),\ \forall \mu \in \mathcal{D},\ \forall N \tag{9.46}$$

we can define a simple error bound for the output of interest in

$$|s(\mu) - s_N(\mu)| \ \leq\ \Delta_N^s(\mu) \equiv \frac{1}{2}\left(\mathcal{M}_N(\mu) - \mathcal{P}_N(\mu)\right),\ \forall\, \mu \in \mathcal{D},\ \forall N; \tag{9.47}$$

where

$$\mathcal{P}_N(\mu) \ = \ -10\ \log_{10} \frac{\mathcal{S}_{\text{out}}(\mu) + \Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}{\mathcal{S}_{\text{in}}(\mu) - \Delta_N^{\mathcal{S}_{\text{in}}}(\mu)}, \tag{9.48}$$

$$\mathcal{M}_N(\mu) \ = \ -10\ \log_{10} \frac{\mathcal{S}_{\text{out}}(\mu) - \Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}{\mathcal{S}_{\text{in}}(\mu) + \Delta_N^{\mathcal{S}_{\text{in}}}(\mu)} \tag{9.49}$$

and

$$s_N(\mu) = \frac{1}{2}\left(\mathcal{P}_N(\mu) + \mathcal{M}_N(\mu)\right). \tag{9.50}$$

Proof. To prove (9.47) we need to show that

$$s_N(\mu) - \Delta_N^s(\mu) \leq s(\mu) \tag{9.51}$$

and

$$s(\mu) \leq s_N(\mu) + \Delta_N^s(\mu). \tag{9.52}$$

We first prove (9.51). We insert the expressions of $s_N(\mu)$ (9.50) and $\Delta_N^s(\mu)$ (9.47) to obtain

$$s_N(\mu) - \Delta_N^s(\mu) = \frac{1}{2}\left(\mathcal{P}_N(\mu) + \mathcal{M}_N(\mu)\right) - \frac{1}{2}\left(\mathcal{M}_N(\mu) - \mathcal{P}_N(\mu)\right) = \mathcal{P}_N(\mu). \tag{9.53}$$

Hence, $s_N(\mu) - \Delta_N^s(\mu) \leq s(\mu)$ if $\mathcal{P}_N(\mu) \leq s(\mu)$, for any $\mu \in \mathcal{D}$. We know that

$$\frac{\mathcal{S}_{\text{in},N}(\mu) - \Delta_N^{\mathcal{S}_{\text{in}}}(\mu)}{\mathcal{S}_{\text{in}}(\mu)} \leq \frac{\mathcal{S}_{\text{out},N}(\mu) + \Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}{\mathcal{S}_{\text{out}}(\mu)} \tag{9.54}$$

since $\mathcal{S}_{\text{in},N}(\mu) - \Delta_N^{\mathcal{S}_{\text{in}}}(\mu) \leq \mathcal{S}_{\text{in}}(\mu)$ and $\mathcal{S}_{\text{out}}(\mu) \leq \mathcal{S}_{\text{out},N}(\mu) - \Delta_N^{\mathcal{S}_{\text{out}}}(\mu)$ from (9.45) and (9.46).

We re-arrange (9.54) to obtain

$$\begin{aligned} \frac{\mathcal{S}_{\text{out}}(\mu)}{\mathcal{S}_{\text{in}}(\mu)} &\leq \frac{\mathcal{S}_{\text{out},N}(\mu)+\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}{\mathcal{S}_{\text{in},N}(\mu)-\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)} \\ \Rightarrow -\log_{10} \frac{\mathcal{S}_{\text{out},N}(\mu)+\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}{\mathcal{S}_{\text{in},N}(\mu)-\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)} &\leq -\log_{10} \frac{\mathcal{S}_{\text{out}}(\mu)}{\mathcal{S}_{\text{in}}(\mu)} \\ \Rightarrow \mathcal{P}_N(\mu) &\leq s(\mu). \end{aligned} \tag{9.55}$$

The proof (9.51) follows directly.

We now prove (9.52). We insert the expressions of $s_N(\mu)$ (9.50) and $\Delta_N^s(\mu)$ (9.47) to obtain

$$s_N(\mu) + \Delta_N^s(\mu) = \frac{1}{2}\left(\mathcal{P}_N(\mu) + \mathcal{M}_N(\mu)\right) + \frac{1}{2}\left(\mathcal{M}_N(\mu) - \mathcal{P}_N(\mu)\right) = \mathcal{M}_N(\mu). \tag{9.56}$$

Hence, $s(\mu) \leq s_N(\mu) + \Delta_N^s(\mu)$ if $s(\mu) \leq \mathcal{M}_N(\mu)$, for any $\mu \in \mathcal{D}$. We know that

$$\frac{\mathcal{S}_{\text{in}}(\mu)}{\mathcal{S}_{\text{in},N}(\mu) + \Delta_N^{\mathcal{S}_{\text{in}}}(\mu)} \leq \frac{\mathcal{S}_{\text{out}}(\mu)}{\mathcal{S}_{\text{out},N}(\mu) - \Delta_N^{\mathcal{S}_{\text{out}}}(\mu)} \tag{9.57}$$

since $\mathcal{S}_{\text{in}}(\mu) \leq \mathcal{S}_{\text{in},N}(\mu) + \Delta_N^{\mathcal{S}_{\text{in}}}(\mu)$ and $\mathcal{S}_{\text{out},N}(\mu) - \Delta_N^{\mathcal{S}_{\text{out}}}(\mu) \leq \mathcal{S}_{\text{out}}(\mu)$ from (9.45) and (9.46).

We re-arrange (9.57) to obtain

$$\begin{aligned} \frac{\mathcal{S}_{\text{out},N}(\mu)-\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}{\mathcal{S}_{\text{in},N}(\mu)+\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)} &\leq \frac{\mathcal{S}_{\text{out}}(\mu)}{\mathcal{S}_{\text{in}}(\mu)} \\ \Rightarrow -\log_{10} \frac{\mathcal{S}_{\text{out}}(\mu)}{\mathcal{S}_{\text{in}}(\mu)} &\leq -\log_{10} \frac{\mathcal{S}_{\text{out},N}(\mu)-\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}{\mathcal{S}_{\text{in},N}(\mu)+\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)} \\ \Rightarrow s(\mu) &\leq \mathcal{M}_N(\mu). \end{aligned} \tag{9.58}$$

The proof (9.52) follows directly.

We thus obtain the desired result. □

Output Bounds: Online Convergence

We first define the effectivities associated with $\mathcal{S}_{\text{in}}(\mu)$, $\mathcal{S}_{\text{out}}(\mu)$ and $s(\mu)$ as

$$\eta_N^{\mathcal{S}_{\text{in}}}(\mu) = \frac{\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)}{|\mathcal{S}_{\text{in}}(\mu) - \mathcal{S}_{\text{in},N}(\mu)|}, \tag{9.59}$$

$$\eta_N^{\mathcal{S}_{\text{out}}}(\mu) = \frac{\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}{|\mathcal{S}_{\text{out}}(\mu) - \mathcal{S}_{\text{out},N}(\mu)|}, \tag{9.60}$$

$$\eta_N^{s}(\mu) = \frac{\Delta_N^{s}(\mu)}{|s(\mu) - s_N(\mu)|}, \tag{9.61}$$

respectively; these effectivities represent the usual measure of the efficiency of the *a posteriori* error estimators.

Given the verification set Ξ_{verif} of 3000 sample points, we define $\eta_{\text{in},N}^{\max}$, $\eta_{\text{in},N}^{\text{mean}}$, and $\eta_{\text{in},N}^{\text{med}}$ as the maximum of $\eta_N^{\mathcal{S}_{\text{in}}}(\mu)$ over all points in Ξ_{verif}, the mean of $\eta_N^{\mathcal{S}_{\text{in}}}(\mu)$ in Ξ_{verif}, and the median of $\eta_N^{\mathcal{S}_{\text{in}}}(\mu)$ over Ξ_{verif}, respectively. We also introduce another measure, a ratio of maxima,

$$\rho_N^{\text{err,in}}(\mu) = \frac{\max_{\mu\in\Xi_{\text{verif}}} \Delta_N^{\mathcal{S}_{\text{in}}}(\mu)}{\max_{\mu\in\Xi_{\text{verif}}} |\mathcal{S}_{\text{in}}(\mu) - \mathcal{S}_{\text{in},N}(\mu)|}, \tag{9.62}$$

$$\rho_N^{\text{err,out}}(\mu) = \frac{\max_{\mu\in\Xi_{\text{verif}}} \Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}{\max_{\mu\in\Xi_{\text{verif}}} |\mathcal{S}_{\text{out}}(\mu) - \mathcal{S}_{\text{out},N}(\mu)|}, \tag{9.63}$$

$$\rho_N^{\text{err,s}}(\mu) = \frac{\max_{\mu\in\Xi_{\text{verif}}} \Delta_N^{s}(\mu)}{\max_{\mu\in\Xi_{\text{verif}}} |s(\mu) - s_N(\mu)|} \tag{9.64}$$

We present in Figure 9-8 $\eta_{\text{in},N}^{\max}$, $\eta_{\text{in},N}^{\text{mean}}$, and $\eta_{\text{in},N}^{\text{med}}$ as a function of N. We observe that effectivity does not depend strongly on N, hence, we can expect stability and boundedness as N increases. Figure 9-8 that there are some samples in Ξ_{verif} (and generally, in $\mathcal{D}$) where the effectivities are very large. To understand these effectivities, we present in Figure 9-9, a scatter plot of $\eta_N^{\mathcal{S}_{\text{in}}}(\mu)$ vs $|\mathcal{S}_{\text{in}}(\mu) - \mathcal{S}_{\text{in},N}(\mu)|$ for all $\mu \in \Xi_{\text{verif}}$ for the representative case of $N = 35$. From the scatter plot in Figure 9-9, it is clear that the poor effectivities are uniformly associated with points in $\Xi_{\text{verif}} \subset \mathcal{D}$ for which the actual error is quite or very small (relative to, say, the average error over Ξ_{verif}). Since the large effectivities multiply very small actual errors they do not present a problem; moreover these smaller errors and associated error bounds are typically well below the admissible/acceptable tolerance. We can make similar observations regarding the effectivities $\eta_N^{\mathcal{S}_{\text{out}}}(\mu)$ and $\eta_N^{s}(\mu)$.

We then present in Tables 9.14, 9.15, and 9.16 the error bounds and effectivities for $\mathcal{S}_{\text{in}}(\mu)$, $\mathcal{S}_{\text{out}}(\mu)$, and the true output $s(\mu)$ respectively. As before, we present the results over the large

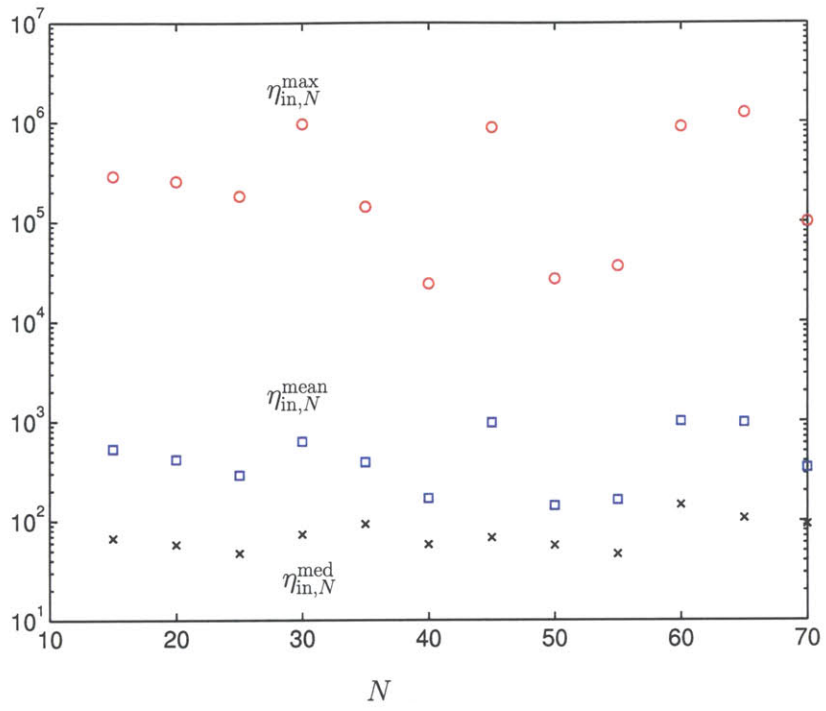


Figure 9-8: Band-Stop Filter: A plot depicting $\eta_{\text{in},N}^{\max}$, $\eta_{\text{in},N}^{\text{mean}}$, and $\eta_{\text{in},N}^{\text{med}}$ as a function of the dimension of the reduced basis space, N.

verification set Ξ_{verif} ($\dim(\Xi_{\text{verif}}) = 3000$) used for the reduced basis approximation in Tables 9.4 and 9.5. Note that we reject from our sample Ξ_{verif} those parameter points for which the actual error (i.e, $|\mathcal{S}_{\text{in}}(\mu) - \mathcal{S}_{\text{in},N}(\mu)|$, $|\mathcal{S}_{\text{out}}(\mu) - \mathcal{S}_{\text{out},N}(\mu)|$, and $|s(\mu) - s_N(\mu)|$) are of the order of machine precision, as for these parameter points the calculation is contaminated by round-off - this is akin to throwing out the isolated points where the actual error is very very small so our effectivity results are more representative of the entire sample rather than be tainted by these points with extremely small error.

We report the maximum relative error in the output $E_{\text{in},N}(\mu)$ given by (9.29), the maximum relative output bound $\mathcal{E}_{\text{in},N}(\mu)$ given by (9.30), and the ratio $\rho_N^{\text{err,in}}(\mu)$ as a function of N for the inlet system in Table 9.14. We also present in Table 9.14 the computational time to compute the error bound $\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)$ given by (9.27) in terms of the dual norm of the residual for the inlet system, $\epsilon_N^{\mathcal{S}_{\text{in}}}(\mu)$, and the associated inf-sup lower bound, $\beta_{\text{LB,in}}(\mu)$. We obtain relative accuracies of $O(10^{-5})$ for $N = 40$, the computational cost to evaluate $\mathcal{S}_{\text{in},N}(\mu)$ and $\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)$ is ≈ 24 times faster than the evaluation of the true output $s(\mu)$. Also, note that the computational time to calculate the a

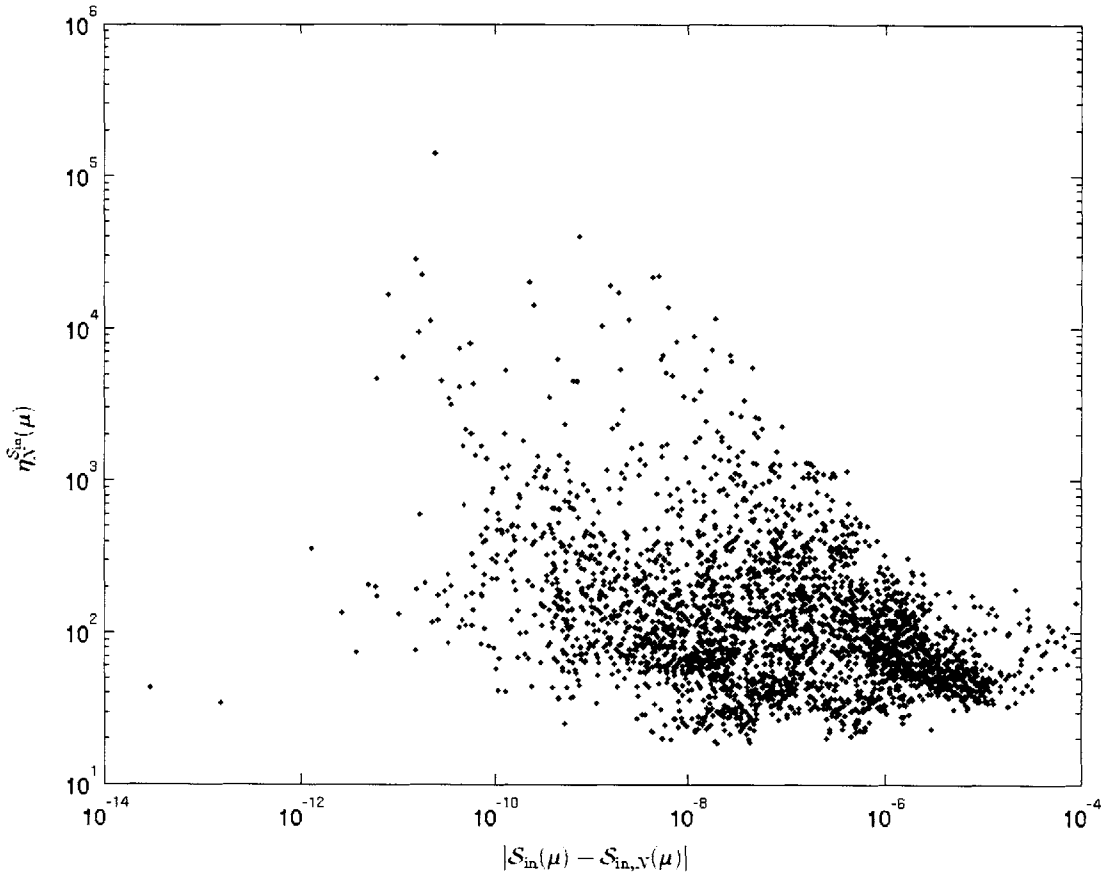


Figure 9-9: Band-Stop Filter: A scatter plot of $\eta_N^{\mathcal{S}_{\text{in}}}(\mu)$ vs $|\mathcal{S}_{\text{in}}(\mu) - \mathcal{S}_{\text{in},N}(\mu)|$ for all $\mu \in \Xi_{\text{verif}}$ for the representative case of $N = 35$.

N	$\max_{\mu\in\Xi_{\text{verif}}} E_{\text{in},N}(\mu)$	$\max_{\mu\in\Xi_{\text{verif}}} \mathcal{E}_{\text{in},N}(\mu)$	$\rho_N^{\text{err,in}}$	Online Time		
				$\mathcal{S}_{\text{in},N}$	$\Delta_N^{\mathcal{S}_{\text{in}}}$	
					$\epsilon_N^{\mathcal{S}_{\text{in}}}$	$\beta_{\text{LB,in}}$
25	1.55E-01	2.68E+00	17.29	1.58E-03	6.40E-03	
30	3.67E-04	2.82E-02	76.77	1.67E-03	7.36E-03	
35	3.05E-05	1.29E-03	42.48	2.56E-03	9.03E-03	2.83E-02
40	6.98E-07	5.32E-05	76.28	2.78E-03	9.37E-03	
45	4.82E-07	6.94E-06	14.39	3.41E-03	1.02E-02	

Table 9.14: Band-Stop Filter: Reduced basis convergence result and the online time to evaluate $\mathcal{S}_{\text{in},N}$, $\Delta_N^{\mathcal{S}_{\text{in}}}$ as a function of N; the timing results are normalized with respect to the time to calculate the "truth" output $s(\mu)$. $\Delta_N^{\mathcal{S}_{\text{in}}}(\mu) \equiv \frac{(\epsilon_N^{\mathcal{S}_{\text{in}}}(\mu))^2}{\beta_{\text{LB,in}}(\mu)}$, where $\epsilon_N^{\mathcal{S}_{\text{in}}}(\mu)$ and $\beta_{\text{LB,in}}(\mu)$ are the dual norm of the residual and the inf-sup lower bound for the inlet system.

posteriori error estimator[1] is of the same order as the time required for the reduced basis output — this is important if we are to ensure rapid exploration of the parameter space to construct our basis offline, and real-time evaluation of the actual output $s(\mu)$ online.

We now present the error estimation results for our outlet system: we report the maximum relative error in the output $E_{\text{out},N}(\mu)$ given by (9.32), the maximum relative output bound $\mathcal{E}_{\text{out},N}(\mu)$ given by (9.33), and the ratio $\rho_N^{\text{err,out}}(\mu)$ as a function of N in Table 9.15. We also present in Table 9.15 the computational time to compute the error bound $\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)$ given by (9.28) in terms of the dual norm of the residual for the outlet system, $\epsilon_N^{\mathcal{S}_{\text{out}}}(\mu)$, and the associated inf-sup lower bound, $\beta_{\text{LB,out}}(\mu)$. We obtain relative accuracies of $O(10^{-5})$ for $N = 40$, the computational cost to evaluate $\mathcal{S}_{\text{out},N}(\mu)$ and $\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)$ is ≈ 24 times faster than the evaluation of the true output $s(\mu)$.

Finally, in Table 9.16, we present the results for the actual output of interest, the transmission coefficient. We use the reduced basis solutions $\mathcal{S}_{\text{in},N}(\mu)$ (respectively, $\mathcal{S}_{\text{out},N}(\mu)$) and associated error bounds $\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)$ (respectively, $\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)$) to construct the reduced basis output $s_N(\mu)$ (9.50) and associated error bound $\Delta_N^s(\mu)$ (9.47). Note that $s_N(\mu)$ is a derived output, similarly $\Delta_N^s(\mu)$ is a derived output bound — we cannot obtain *realistic* solutions and *rigorous* bounds for any choice of N. The expressions for $s_N(\mu)$ and $\Delta_N^s(\mu)$ are only valid if

$$\Delta_N^{\mathcal{S}_{\text{in}}}(\mu) \leq \mathcal{S}_{\text{in},N}(\mu), \tag{9.65}$$

$$\Delta_N^{\mathcal{S}_{\text{out}}}(\mu) \leq \mathcal{S}_{\text{out},N}(\mu); \tag{9.66}$$

[1] Recall that the calculation of the dual norm of the residual requires $O(Q_{\mathcal{A}}^2 N^2)$ operations; althought we include the $O(N^3)$ computational cost of computing the reduced basis coefficients in the computational cost of the output, the dual norm calculation is more expensive.

N	$\max_{\mu\in\Xi_{\text{verif}}} E_{\text{out},N}(\mu)$	$\max_{\mu\in\Xi_{\text{verif}}} \mathcal{E}_{\text{out},N}(\mu)$	$\rho_N^{\text{err,out}}$	Online Time		
				$\mathcal{S}_{\text{out},N}$	$\Delta_N^{\mathcal{S}_{\text{out}}}$	
					$\epsilon_N^{\mathcal{S}_{\text{out}}}$	$\beta_{\text{LB,out}}$
25	5.37E-02	8.76E-01	16.39	1.69E-03	6.44E-03	
30	1.56E-04	3.96E-03	25.44	1.72E-03	7.16E-03	
35	1.63E-05	3.53E-04	21.71	2.28E-03	7.49E-03	2.78E-02
40	4.04E-07	4.45E-05	110.11	2.69E-03	9.18E-03	
45	1.38E-07	2.03E-06	14.69	3.38E-03	1.02E-02	

Table 9.15: Band-Stop Filter: Reduced basis convergence result and the online time to evaluate $\mathcal{S}_{\text{out},N}$, $\Delta_N^{\mathcal{S}_{\text{out}}}$ as a function of N; the timing results are normalized with respect to the time to calculate the "truth" output $s(\mu)$. $\Delta_N^{\mathcal{S}_{\text{out}}}(\mu) \equiv \frac{(\epsilon_N^{\mathcal{S}_{\text{out}}}(\mu))^2}{\beta_{\text{LB,out}}(\mu)}$, where $\epsilon_N^{\mathcal{S}_{\text{out}}}(\mu)$ and $\beta_{\text{LB,out}}(\mu)$ are the dual norm of the residual and the inf-sup lower bound for the outlet system.

N	$\max_{\mu\in\Xi_{\text{verif}}} \frac{\lvert s(\mu)-s_N(\mu)\rvert}{s(\mu)}$	$\max_{\mu\in\Xi_{\text{verif}}} \frac{\Delta_N^s(\mu)}{s(\mu)}$	$\rho_N^{\text{err,s}}(\mu)$
30	7.23E-03	1.18E-01	16.32
35	1.14E-04	4.09E-03	35.85
40	1.89E-05	9.13E-04	48.31
45	2.81E-06	7.16E-05	25.48

Table 9.16: Band-Stop Filter: Reduced basis convergence of $s_N(\mu)$ and $\Delta_N^s(\mu)$ as a function of N.

thus, we need to choose N such that the *a posteriori* output estimators $\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)$ and $\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)$ are *small* enough.

We report the maximum relative error, the maximum relative error bound, $\Delta_N^s(\mu)/|s(\mu)|$, and the ratio of the maxima $\rho_N^{\text{err,s}}(\mu)$ as a function of N over the same verification sample set Ξ_{verif} used for the inlet and outlet systems. We need atleast $N = 28$ basis functions so that we may calculate $s_N(\mu)$ and $\Delta_N^s(\mu)$ — in general, lower N also suffices for most samples, but there are some isolated sample points where we cannot calculate $s(\mu)$. We observe relatively good effectivity, on *average* our effectivities are $O(20)$. We obtain relative accuracies of $O(10^{-5})$ for $N = 40$, the computational cost to evaluate $s_N(\mu)$ and $\Delta_N^s(\mu)$ is ≈ 12 times faster than the evaluation of the true output $s(\mu)$. The computational savings are approximately halved for the true output because we need to integrate the solutions and associated error bounds for the inlet and outlet system. Note that $s_N(\mu)$ is essentially indistinguishable from $s(\mu)$ for $N \geq 35$. Furthermore, the "compliant" nature of $\mathcal{S}_{\text{in}}(\mu)$ (respectively, $\mathcal{S}_{\text{out}}(\mu)$) and expected quadratic convergence in $|\mathcal{S}_{\text{in}}(\mu) - \mathcal{S}_{\text{in},N}(\mu)|$ (respectively, $|\mathcal{S}_{\text{out}}(\mu) - \mathcal{S}_{\text{out},N}(\mu)|$) contributes to the rapid convergence of $s_N(\mu)$ to $s(\mu)$ with increasing N.

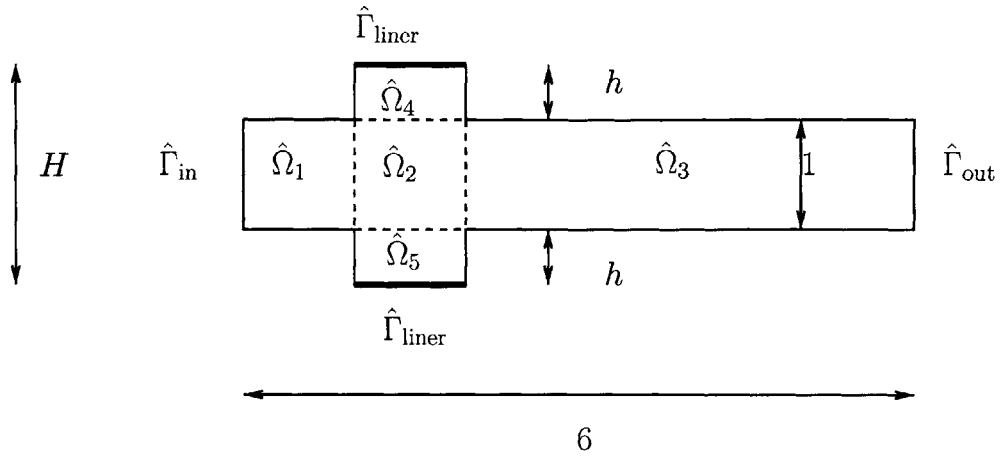


Figure 9-10: Low-Pass Filter: Non-dimensional Original Domain

9.3 Low-Pass Filter

9.3.1 Problem Statement

We return to the treatment of the $P = 4$ parameter low-pass filter system described extensively in Section 4.4. The low-pass filter element consists of an expansion chamber in the middle of a long acoustic waveguide; the top and bottom walls of the expansion chamber are compliant and represent general impedance terminations. We consider variations in the height of the expansion chamber, and also allow small variations in the impedance characteristics of the compliant walls of the expansion chamber (we refer to them here as liners). This example is representative of reactive mufflers used in internal combustion engines for low-frequency broadband noise attenuation. The idea of varying the height of the expansion chamber finds practical relevance in the design of reactive silencers with flexible panels inside expansion chambers to obtain the widest stopband at a particular frequency; these silencers find use in reducing low-frequency ducted fan noise in HVAC systems [51, 61].

Figure 9-10 shows the (non-dimensional) original domain $\hat{\Omega}$. It consists of two waveguide sections $\hat{\Omega}_1$ and $\hat{\Omega}_3$ of fixed cross-section connected to an expansion chamber (represented by the sub-domains $\hat{\Omega}_2$, $\hat{\Omega}_4$, and $\hat{\Omega}_5$) in the middle. The height of the expansion chamber, H, can be changed by changing h, the heights of the sub-domains $\hat{\Omega}_4$ and $\hat{\Omega}_5$ while the sub-domain $\hat{\Omega}_2$ remains fixed.

We first summarize the problem on the (non-dimensional) original domain $\hat{\Omega} \subset \mathbb{R}^2$ with

Lipschitz-continuous boundary $\hat{\Gamma}$. Given $\mu \in \mathcal{D} \subset \mathbb{R}^4$, we solve the μ-dependent acoustic transmission problem (please see Section 4.4 for a more detailed description of the problem) for $u(\mu)$

$$\begin{aligned}
\nabla^2 u(\mu) + k^2 u(\mu) &= 0, \text{ in } \hat{\Omega}, && (9.67)\\
\frac{\partial u(\mu)}{\partial n} &= \mathrm{i}\, k, \text{ on } \hat{\Gamma}_{\text{in}}, && (9.68)\\
\frac{\partial u(\mu)}{\partial n} &= -\mathrm{i}\, \frac{k}{Z_R + \mathrm{i}\, Z_I}\, u(\mu), \text{ on } \hat{\Gamma}_{\text{liner}}, && (9.69)\\
\frac{\partial u(\mu)}{\partial n} &= -\mathrm{i}\, \alpha_1 \left(\int_{\Gamma_{\text{out}}} u(\mu)\, \overline{\Xi}_1\right) \Xi_1, \text{ on } \hat{\Gamma}_{\text{out}}, && (9.70)\\
\frac{\partial u(\mu)}{\partial n} &= 0, \text{ on } \hat{\Gamma} \backslash (\hat{\Gamma}_{\text{in}} \cup \hat{\Gamma}_{\text{liner}} \cup \hat{\Gamma}_{\text{out}}); && (9.71)
\end{aligned}$$

and calculate our output of interest

$$s(\mu) = -20\, \log_{10} \frac{\sqrt{\int_{\hat{\Gamma}_{\text{out}}} \overline{u(\mu)}\, u(\mu)}}{\sqrt{\int_{\hat{\Gamma}_{\text{in}}} \overline{u}(\mu)\, u(\mu)}}. \qquad (9.72)$$

9.3.2 Parameters

The parameters are the (non-dimensional) frequency squared, $\mu_1 \equiv k^2$, the height of the expansion sections h (relative to the reference height $h_{\text{ref}} = 0.5$), $\mu_2 \equiv \frac{h}{h_{\text{ref}}}$, the non-dimensional acoustic resistance of the liner walls, $\mu_3 \equiv Z_R$, and the non-dimensional acoustic reactance of the liner walls, $\mu_4 \equiv Z_I$ respectively. We consider the parameter domain $\mu \in \mathcal{D} \equiv [0.1, 2.5] \times [0.75, 1.25] \times [4.8, 5.0] \times [1.8, 2.0]$. The low-pass filter problem is significantly different from the band-stop filter problem discussed earlier. The scaling up from $P = 2$ to $P = 4$ parameters requires the choice of *smaller* input parameter ranges primarily because the associated increase in *offline* complexity imposes restrictions on what is practically feasible.

The non-dimensional frequency range of k corresponds to a dimensional frequency range of 340 Hz to 1.7 KHz: exhaust mufflers typically try to reduce the sound level for even smaller frequency ranges. We vary the non-dimensional acoustic resistance, Z_R, and reactance, Z_I, well within the range of normalized acoustic impedance obtained from experiments conducted on tunable electromechanical acoustic liners [50, 105].

The parametric variation of the height of the expansion chamber, and the acoustic impedance characteristics of the liners considered in this problem serve two purposes: first, we can explore the

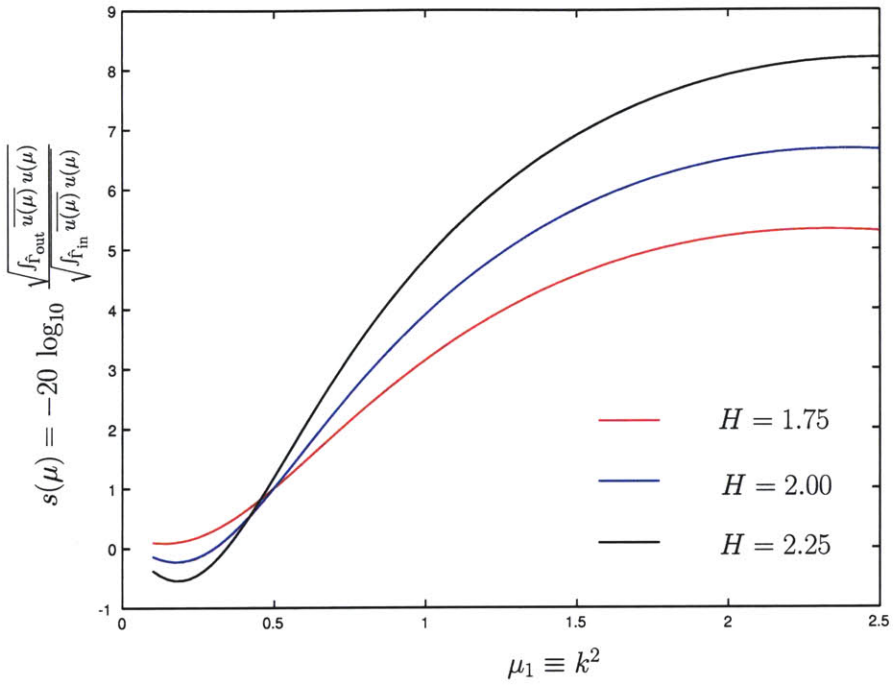


Figure 9-11: Low-Pass filter transmission coefficient as a function of k^2 for different H; the non-dimensional acoustic resistance $Z_R = 4.8$ and reactance $Z_I = 1.8$.

sensitivity of the output to changes in the parameter to identify optimal configurations; and second, the reduced-order model can be used to estimate unknown parameters (e.g., the liner impedances) by inverse procedures.

In Figure 9-11 we present the variation of the transmission coefficient, $s(\mu)$, as a function of the non-dimensional wave-number squared, $\mu_1 \equiv k^2$, for different choice of $\mu_2 \equiv H$.

9.3.3 Weak Formulation

Weak Form on Reference Domain

The development of the weak form for the low-pass filter problem is similar to the treatment of the band-stop filter in Section 9.2.4. We summarize the original weak form on the μ-independent reference domain Ω. Given $\mu \in \mathcal{D}$, find $u(\mu) \in X^2$ from

$$a(u(\mu), v; \mu) = f(v; \mu), \ \forall v \in X; \tag{9.73}$$

[2]Here X is a complex function space equipped with an inner product $(\cdot,\cdot)_X$ (9.10) and norm $\|\cdot\|_X$ (9.11).

where

$$
\begin{aligned}
a(w,v;\mu) \;=\; & \int_{\Omega_{1,2}} \nabla w \, \nabla \overline{v} + \mu_2 \int_{\Omega_3} \frac{\partial w}{\partial x}\frac{\partial \overline{v}}{\partial x} + \frac{1}{\mu_2}\int_{\Omega_3} \frac{\partial w}{\partial y}\frac{\partial \overline{v}}{\partial y} \\
& - \mu_1 \int_{\Omega_{1,2}} w\,\overline{v} - \mu_1\,\mu_2 \int_{\Omega_3} w\;\overline{v} \\
& + \frac{\sqrt{\mu}_1\,\mu_4}{\mu_3^2+\mu_4^2}\int_{\Gamma_{\text{liner}}} w\,\overline{v} + \mathrm{i}\,\frac{\sqrt{\mu}_1\,\mu_3}{\mu_3^2+\mu_4^2}\int_{\Gamma_{\text{liner}}} w\,\overline{v} \\
& + \mathrm{i}\,\sqrt{\mu}_1 \int_{\Gamma_{\text{out}}} \Xi_1\,\overline{v} \int_{\Gamma_{\text{out}}} w\,\overline{\Xi}_1;
\end{aligned}
\tag{9.74}
$$

and

$$
f(v;\mu) = \mathrm{i}\,\sqrt{\mu}_1 \int_{\Gamma_{\text{in}}} \overline{v}. \tag{9.75}
$$

We identify that a is affinely separable[3] for $Q_a = 8$ and f is affinely separable[4] for $Q_f = 1$. We also identify our output of interest can be written as

$$
s(\mu) = -10\,\log_{10}\frac{s_{\text{out}}(\mu)}{s_{\text{in}}(\mu)}; \tag{9.76}
$$

where $s_{\text{in}}(\mu)$ and $s_{\text{out}}(\mu)$ are quadratic outputs described in (9.16) and (9.17) respectively.

Expanded Weak Form

As in Section 9.2.4, we identify two distinct systems of equations to calculate our output of interest $s(\mu)$. The first expanded formulation is related to $s_{\text{in}}(\mu)$ — we denote $\mathcal{S}_{\text{in}}(\mu)$ as the "compliant" output associated with the expanded inlet formulation. The second expanded formulation is to deal with $s_{\text{out}}(\mu)$ — we denote $\mathcal{S}_{\text{out}}(\mu)$ as the linear compliant output associated with the expanded outlet formulation.

We will also require the complex "expanded" function space $\mathcal{X} \subset (X)^2$; here $\dim(\mathcal{X}) = 2\,\dim(X)$. The complex function spaces $\mathcal{X}$, described in Section 9.2.4, is associated with the inner product$(\cdot,\cdot)_{\mathcal{X}}$ (9.22) and norm $\|\cdot\|_{\mathcal{X}}$ (9.23) respectively.

Inlet System

[3] $a(w,v;\mu) = \sum_{q=1}^{Q_a} \Theta_a^q(\mu)\,a^q(w,v),\ \forall \mu \in \mathcal{D}$.
[4] $f(v;\mu) = \sum_{q=1}^{Q_f} \Theta_f^q(\mu)\,f^q(v),\ \forall \mu \in \mathcal{D}$.

Given $\mu \in \mathcal{D}$, calculate the "truth" output of interest

$$\mathcal{S}_{\text{in}}(\mu) = \overline{\mathcal{F}(\mathcal{U}_{\text{in}}(\mu); \mu)} \tag{9.77}$$

where $\mathcal{U}_{\text{in}}(\mu) = [U_{\text{in}}^+, U_{\text{in}}^-]^T \in \mathcal{X}$ satisfies

$$\mathcal{A}_{\text{in}}(\mathcal{U}_{\text{in}}(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}; \mu), \ \forall \mathcal{V} \in \mathcal{X}; \tag{9.78}$$

here $\mathcal{X} \in X^2$ and $\mathcal{V} = [V^+, V^-]^T \in \mathcal{X}$.

Outlet System

Given $\mu \in \mathcal{D}$, calculate the "truth" output of interest

$$\mathcal{S}_{\text{out}}(\mu) = \overline{\mathcal{F}(\mathcal{U}_{\text{out}}(\mu); \mu)} \tag{9.79}$$

where $\mathcal{U}_{\text{out}}(\mu) = [U_{\text{out}}^+, U_{\text{out}}^-]^T \in \mathcal{X}$ satisfies

$$\mathcal{A}_{\text{out}}(\mathcal{U}_{\text{out}}(\mu), \mathcal{V}; \mu) = \mathcal{F}(\mathcal{V}; \mu), \ \forall \mathcal{V} \in \mathcal{X}; \tag{9.80}$$

here $\mathcal{X} \in X^2$ and $\mathcal{V} = [V^+, V^-]^T \in \mathcal{X}$.

9.3.4 Bilinear Forms and Affine Decomposition

The bilinear forms $\mathcal{A}_{\text{in}}$ and $\mathcal{A}_{\text{out}}$ are affinely separable as

$$\mathcal{A}_{\text{in}}(\mathcal{W}, \mathcal{V}; \mu) = \sum_{q=1}^{Q_{\mathcal{A},\text{in}}} \Theta_{\mathcal{A},\text{in}}^q(\mu)\, \mathcal{A}_{\text{in}}^q(\mathcal{W}, \mathcal{V}), \tag{9.81}$$

$$\mathcal{A}_{\text{out}}(\mathcal{W}, \mathcal{V}; \mu) = \sum_{q=1}^{Q_{\mathcal{A},\text{out}}} \Theta_{\mathcal{A},\text{out}}^q(\mu)\, \mathcal{A}_{\text{out}}^q(\mathcal{W}, \mathcal{V}), \tag{9.82}$$

q	1	2	3	4	5	6	7	8
$\Theta^q_{\mathcal{A}}(\mu)$	1	μ_2	$\frac{1}{\mu_2}$	$-\mu_1$	$-\mu_1\,\mu_2$	$\frac{\sqrt{\mu}_1\,\mu_4}{\mu_3^2+\mu_4^2}$	$\frac{\mathrm{i}\,\sqrt{\mu}_1\,\mu_4}{\mu_3^2+\mu_4^2}$	$\mathrm{i}\,\sqrt{\mu}_1$

Table 9.17: $\Theta^q_{\mathcal{A}}(\mu)$ ($\equiv \Theta^q_{\mathcal{A},\mathrm{in}}(\mu) = \Theta^q_{\mathcal{A},\mathrm{out}}(\mu)$) for the expanded inlet and outlet systems of the Low-Pass Filter.

for all $\mathcal{W}, \mathcal{V} \in \mathcal{X}$, $\mu \in \mathcal{D}$. The bilinear forms $\mathcal{A}^q_{\mathrm{in}}(\mathcal{U}_{\mathrm{in}}, \mathcal{V})$ and $\mathcal{A}^q_{\mathrm{out}}(\mathcal{U}_{\mathrm{out}}, \mathcal{V})$ are identical to the bilinear forms $\mathcal{A}^q_{\mathrm{in}}(\mathcal{U}_{\mathrm{in}}, \mathcal{V})$ and $\mathcal{A}^q_{\mathrm{out}}(\mathcal{U}_{\mathrm{out}}, \mathcal{V})$ described for the band-stop filter in Table 9.2. We have the same number of affine terms in the affine separation of the inlet and outlet bilinear forms, thus $Q_{\mathcal{A},\mathrm{in}} = Q_{\mathcal{A},\mathrm{out}} \equiv Q_{\mathcal{A}}$; furthermore, the parameter-dependent coefficients associated with the inlet and outlet bilinear forms are also identical; thus $\Theta^q_{\mathcal{A},\mathrm{in}}(\mu) = \Theta^q_{\mathcal{A},\mathrm{out}}(\mu) \equiv \Theta^q_{\mathcal{A}}(\mu), q = 1, \ldots, Q_{\mathcal{A}}$. We summarize the $\Theta^q_{\mathcal{A}}(\mu)$ in Table 9.17.

9.3.5 Truth Solutions

We next show the "truth" mesh for the low-pass filter problem in Figure 9-12 — the accuracy of our reduced basis solution will be measured against the solution on the truth mesh.

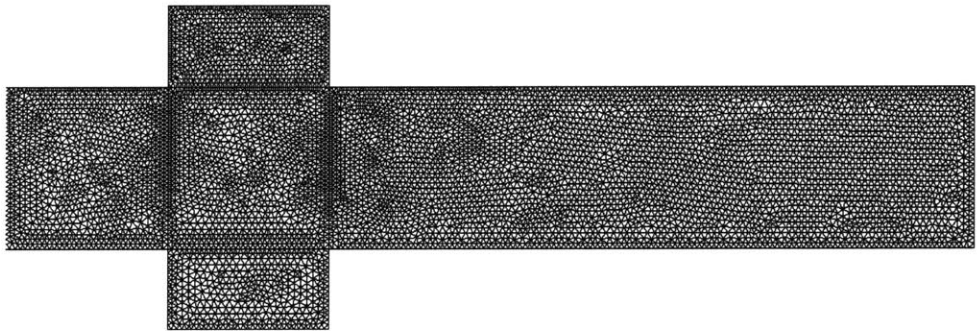


Figure 9-12: Truth mesh for the low-pass filter problem: note here that the solutions are plotted on the original mesh with $\mathcal{N} = 6616$; the expanded system has $\mathcal{N} = 13332$ unknowns.

Figure 9-13 shows a few of the FEM solutions for slightly different parameters. The parameter dependence induces significant variations in solution structure and magnitude. This will create

approximation difficulty in the reduced basis method, and thus it may require large N to achieve sufficient accuracy.

9.3.6 Reduced Basis Approximation

Offline Convergence

We pursue our offline greedy algorithm to construct our reduced basis space over the exhaustive sample set Ξ_{train} of size 5000. We specify the termination error tolerance $\epsilon_{\text{tol,min}} = 1\text{E} - 05$.

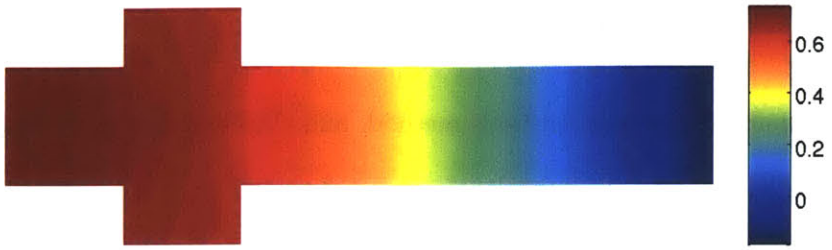
We present the offline convergence of the maximum relative error used to construct our reduced basis spaces W_N^{in} and W_N^{out} for the inlet and outlet systems respectively. The treatment here is similar to Section 9.2.7. We introduce the output bound $\Delta_N^{\mathcal{S}_{\text{in}}}(\mu) \equiv \frac{\epsilon_N^{\mathcal{S}_{\text{in}}}(\mu)^2}{\beta_{\text{LB,in}}(\mu)}$ (respectively, $\Delta_N^{\mathcal{S}_{\text{out}}}(\mu) \equiv \frac{\epsilon_N^{\mathcal{S}_{\text{out}}}(\mu)^2}{\beta_{\text{LB,out}}(\mu)}$) given by (9.27) (respectively, (9.28)) for the inlet (respectively, outlet) system. Here $\epsilon_N^{\mathcal{S}_{\text{in}}}(\mu)$ (respectively, $\epsilon_N^{\mathcal{S}_{\text{out}}}(\mu)$) is the dual norm of the residual for the inlet (respectively, outlet) system. $\beta_{\text{LB,in}}(\mu)$ and $\beta_{\text{LB,out}}(\mu)$ are the inf-sup lower bound approximations for the inlet and outlet system respectively. We discuss the inf-sup lower bound approximation in Section 9.3.7.

Figure 9.3.6 plots the maximum offline *a posteriori* error indicators Δ_N^{in} given by (9.34) (respectively, Δ_N^{out} given by (9.35)) for the inlet (respectively, outlet) system as a function of the number of offline iterations (or the size of the reduced basis space). We obtain $N_{\text{max,in}} = 80$ and $N_{\text{max,out}} = 74$ respectively for $\epsilon_{\text{tol,min}} = 1\text{E} - 05$. The actual relative error is less than $1\text{E} - 10$ for $N_{\text{max}} = 80$ for the inlet system, and $N_{\text{max}} = 74$ for the outlet system. The reduced basis converges uniformly after adding $\approx N = 20$ solutions into the basis; prior to that we note that the maximum relative error fluctuates without showing any signs of decreasing. Inspite of the higher parameter dimensionality ($P = 4$) compared to the band-stop filter problem, we obtain rapid convergence with increasing N.

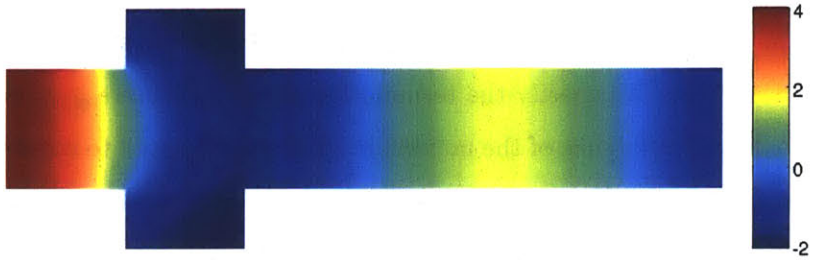
Online Convergence

We test the accuracy of the obtained reduced basis approximations over an exhaustive sample set Ξ_{verif} of size 5000. We present in Table 9.18 the convergence of the reduced basis output of the inlet system, $\mathcal{S}_{\text{in},N}(\mu)$, to its "truth" counterpart $\mathcal{S}_{\text{in}}(\mu)$ given by (9.18) ($\equiv s_{\text{in}}(\mu)$): we present the maximum relative error as a function of N.

We present in Table 9.19 the convergence of the reduced basis output for the outlet system, $\mathcal{S}_{\text{out},N}(\mu)$, to $\mathcal{S}_{\text{out}}(\mu)$ given by (9.20) ($\equiv s_{\text{out}}(\mu)$) over our verification sample set Ξ_{verif}.



(a)



(b)

(c)

(d)

Figure 9-13: Low Pass Filter FEM solutions for (a) $\mu = [0.1, 0.75, 4.8, 1.8]$; (b) $\mu = [2.0, 1.25, 4.8, 1.8]$; (c) $\mu = [1.0, 1.0, 5.0, 2.0]$; and (d) $\mu = [0.5, 1.0, 4.9, 1.9]$.

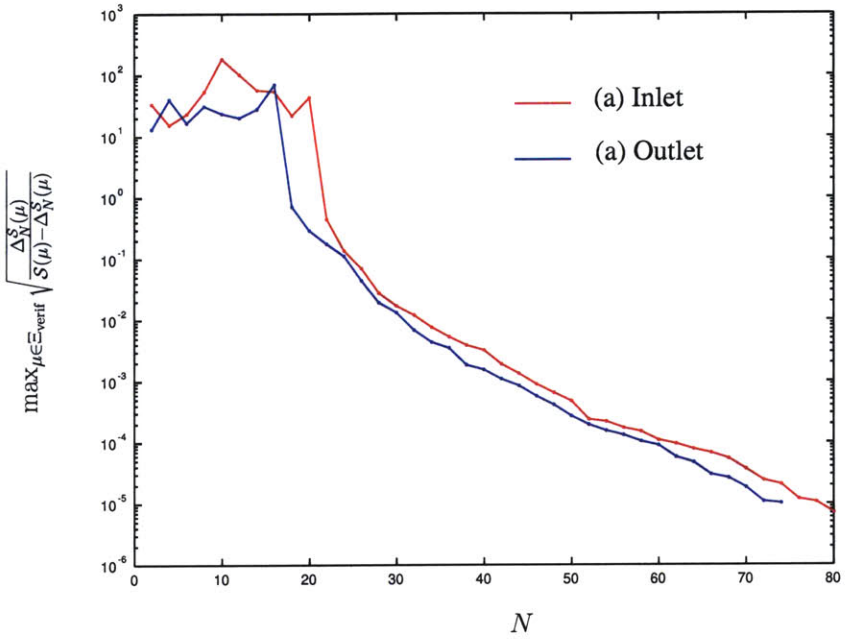


Figure 9-14: Greedy algorithm to generate reduced basis for the low-pass filter - maximum offline relative error bound for (a) inlet system: $\Delta_N^{\text{in}} = \max_{\mu\in\Xi_{\text{train}}} \sqrt{\frac{\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)}{\mathcal{S}_{\text{in},N}(\mu)-\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)}}$ and (b) outlet system: $\Delta_N^{\text{out}} = \max_{\mu\in\Xi_{\text{train}}} \sqrt{\frac{\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}{\mathcal{S}_{\text{out,N}}(\mu)-\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)}}$ shown as a function of N. We specify $\varepsilon_{\text{tol,min}} = 1\text{E}-05$, $\dim(\Xi_{\text{train}}) = 5000$.

N	$\max_{\mu\in\Xi_{\text{verif}}} E_{\text{in},N}(\mu)$	Online Time
		$\mathcal{S}_{\text{in},N}$
20	5.45E-02	1.3E-03
25	8.04E-05	2.3E-03
30	8.45E-07	2.56E-03
35	5.13E-07	2.7E-03
40	9.16E-08	4.03E-03

Table 9.18: Reduced basis convergence result and the online time to evaluate $\mathcal{S}_{\text{in},N}(\mu)$ as a function of N; the timing results are normalized with respect to the "average" time to calculate the "exact" output $s(\mu)$ directly.

N	$\max_{\mu \in \Xi_{\text{verif}}} E_{\text{out},N}(\mu)$	Online Time
		$\mathcal{S}_{\text{out},N}$
20	2.00E-03	1.3E-03
25	3.72E-04	2.3E-03
30	1.32E-05	2.6E-03
35	6.84E-07	2.75E-03
40	1.03E-07	4.00E-03

Table 9.19: Reduced basis convergence result and the online time to evaluate $\mathcal{S}_{\text{out},N}(\mu)$ as a function of N; the timing results are normalized with respect to the "average" time to calculate the "exact" output $s(\mu)$ directly.

We also present in Table 9.18 (respectively, Table 9.19) the online reduced basis computational cost to evaluate $\mathcal{S}_{\text{in},N}(\mu)$ (respectively, $\mathcal{S}_{\text{out},N}(\mu)$) compared to the finite element cost to evaluate $s(\mu)$ for any given μ. We can see that the reduced basis solution $\mathcal{S}_{\text{in,N}}(\mu)$ (respectively, $\mathcal{S}_{\text{out},N}(\mu)$) and the "truth" solution $\mathcal{S}_{\text{in}}(\mu)$ (respectively, $\mathcal{S}_{\text{out}}(\mu)$) are indistinguishable for $N \geq 25$.

9.3.7 Inf-Sup Lower Bound Approximation

We now discuss the inf-sup lower bound approximation using the SCM approach for the inlet and outlet systems respectively. We construct the constraint set $\mathcal{C}_K^{\text{in}}$ (respectively, $\mathcal{C}_K^{\text{out}}$) for the inlet (respectively, outlet) system *offline*; and verify the accuracy of the inf-sup lower bounds *online*. Given $\mu \in \mathcal{D}$, we denote $\sigma_{\text{LB}}(\mu)$ and $\sigma_{\text{UB}}(\mu)$ as the SCM lower and upper bounds to $\sigma(\mu)$; we have $\sigma_{\text{LB}}(\mu) \leq \sigma(\mu) \leq \sigma_{\text{UB}}(\mu), \forall \mu \in \mathcal{D}$.

Relative to our earlier band-stop filter problem, the increased parameter dimensionality, $P = 4$, is an issue that merits some discussion. For $P = 4$ parameters, we cannot apply our SCM greedy algorithm blindly. From Section 7.5, we know that the computational cost of solving the LPs represents the largest portion of the computational overhead for the greedy algorithm. Since the SCM greedy algorithm consists of repeatedly solving LPs over an exhaustive set Ξ_{train} to identify the constraint set C_K, we need to be careful about the size of Ξ_{train}. The increased parameter dimensionality already dictates the use of a much larger Ξ_{train} relative to our $P = 2$ parameters. Therefore, the *final* size of C_K will have serious implications on the feasibility of the greedy algorithm if used as-is. If the final size of C_K is very large, then a large Ξ_{train} will drastically affect the *offline* computational performance of the algorithm because of the large number of LPs that have to be solved in every iteration of the greedy algorithm.

We first discuss modifications to the SCM greedy algorithm, then compute the inf-sup lower bounds using the modified algorithm. Given $\mu \in \mathcal{D}$, we shall denote the inf-sup lower bound approximation for the inlet (respectively, outlet) system as $\beta_{\mathrm{LB,in}}(\mu)$ (respectively, $\beta_{\mathrm{LB,out}}(\mu)$). We also identify for all $\mu \in \mathcal{D}$, $\sigma_{\mathrm{UB,in}}(\mu)$ (respectively, $\sigma_{\mathrm{LB,out}}(\mu)$) and $\sigma_{\mathrm{LB,in}}(\mu)$ (respectively, $\sigma_{\mathrm{LB,out}}(\mu)$) as the upper and lower bounds of $\sigma(\mu)$ for the inlet (respectively, outlet) system.

Offline Convergence

To better understand the complexity of the problem and the effect of $P = 4$ parameters, we choose to do simpler inf-sup lower bound approximations for the relatively more difficult inlet system. We present in Figure 9-15 (a)–(d) the obtained inf-sup lower bounds for problems where we vary only one parameter at a time. We note that a simple Cartesian product of the required one-dimensional constraint sets would give us $K = 4320$ constraints for the $P = 4$ parameter case; it is quite possible that more might be needed.

Figure 9-15 clearly shows that the bulk of the constraints will be required to handle the variation in the inf-sup due to changes in k^2 and h - relatively few are required for variations in Z_R and Z_I. As Figure 9-15(b) shows, we require $K = 16$ constraints to ensure satisfaction of the offline termination criteria for the 1-D slice in h compared to the $K = 3$ constraints required for the 1-D slices in Z_R and Z_I — this is slightly counter-intuitive since the variation of $\sigma(\mu)$ with h seems to be much simpler relative to its variation with Z_R or Z_I. However, we should note that the range of h is larger, and the actual magnitude of $\sigma(\mu)$ also varies over a larger ranger for variations in h (2.2E-03 to 2.8E-03) when compared to variations in Z_R (2.19E-03 to 2.28E-03) or Z_I (2.19E-03 to 2.22E-03). Furthermore, 1-D slices in h for other values of k^2, Z_R and Z_I also require similar number of SCM constraints. We confirm our hypthesis that the number of constraints for fixed Z_R and Z_I will be large by obtaining inf-sup lower bound approximations for problems where we vary only k^2 and h. In Figure 9-16, we present the inf-sup lower bound approximations generated for $\mu_3 \equiv Z_R = 5$ and $\mu_4 \equiv Z_I = 1.8$. We confirm (approximately) the correctness of the Cartesian product estimate for K (expected ≈ 480 constraints); although it underestimates the number of required constraints a little — we require $K = 518$. Table 9.20 summarizes the results for some other two-dimensional slices in (k^2, h) space for specific choices of Z_R and Z_I respectively. We require a minimum of $K = 487$ constraints for $Z_R = 4.8$ and $Z_I = 1.8$; and a maximum of $K = 530$

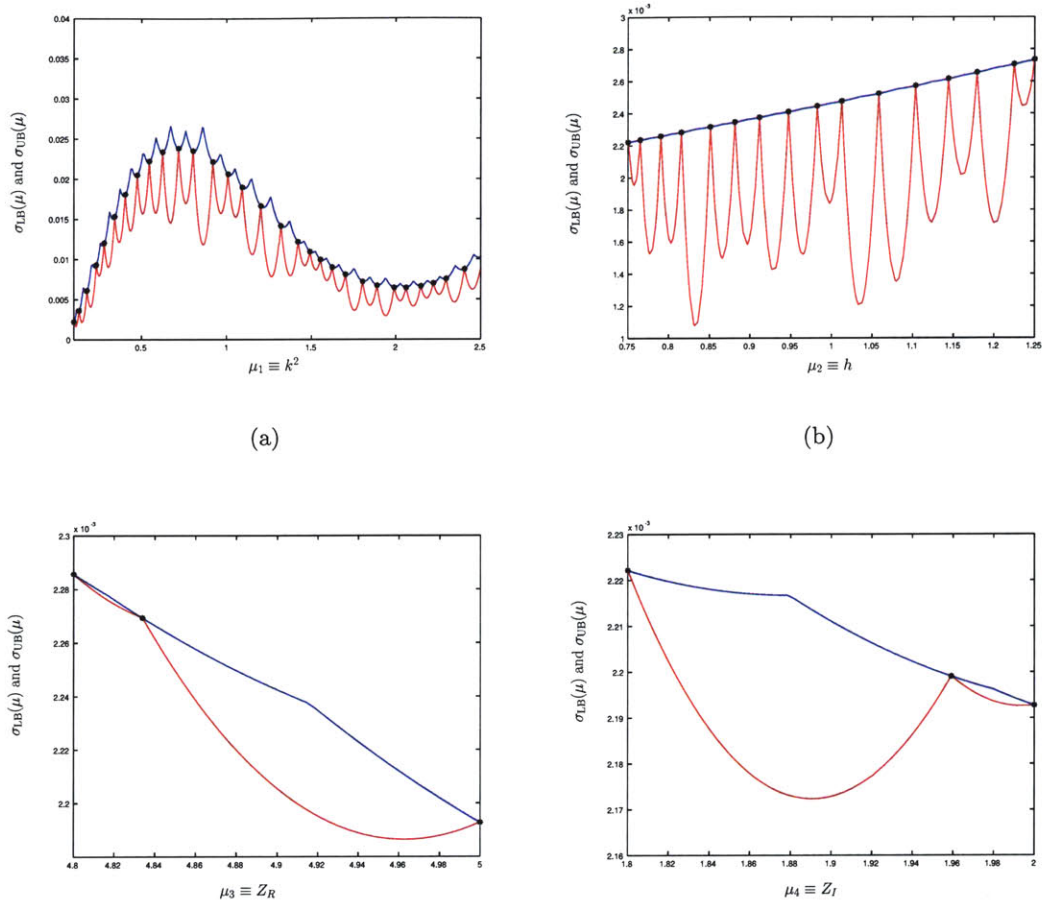


Figure 9-15: Inf-Sup lower bounds using the successive constraint method for the low-pass filter inlet system (a) $k^2 \in [0.1, 2.5]$, fixed $h = 0.75$, $Z_R = 4.8$ and $Z_I = 1.8$; (b) $h \in [0.75, 1.25]$, fixed $k^2 = 0.1$, $Z_R = 5.0$, and $Z_I = 2.0$; (c) $Z_R \in [4.8, 5.0]$, fixed $k^2 = 2.0$, $h = 1.25$, and $Z_I = 2.0$; and (d) $Z_I \in [1.8, 2.0]$, fixed $k^2 = 2.0$, $h = 1.25$, and $Z_R = 5.0$. The size of the training sample set $\Xi_{\mathrm{train}} = 500$ and termination criteria for the greedy algorithm $\epsilon_{\mathrm{tol,min}} = 0.75$. The lower bounds $\sigma_{\mathrm{LB}}(\mu)$ are shown in red, the upper bounds $\sigma_{\mathrm{UB}}(\mu)$ in blue and the black "$*$"s represent the constraints $\omega_k \in \mathcal{C}_K, k = 1, \ldots, K$. The number of constraints K (the size of $\mathcal{C}_K$) is $K = 30$ for Case (a), $K = 16$ for Case (b), and $K = 3$ for Cases (c) and (d). We use $\mathcal{M}_\alpha = 10$ and $\mathcal{M}_+ = 5$ in all cases.

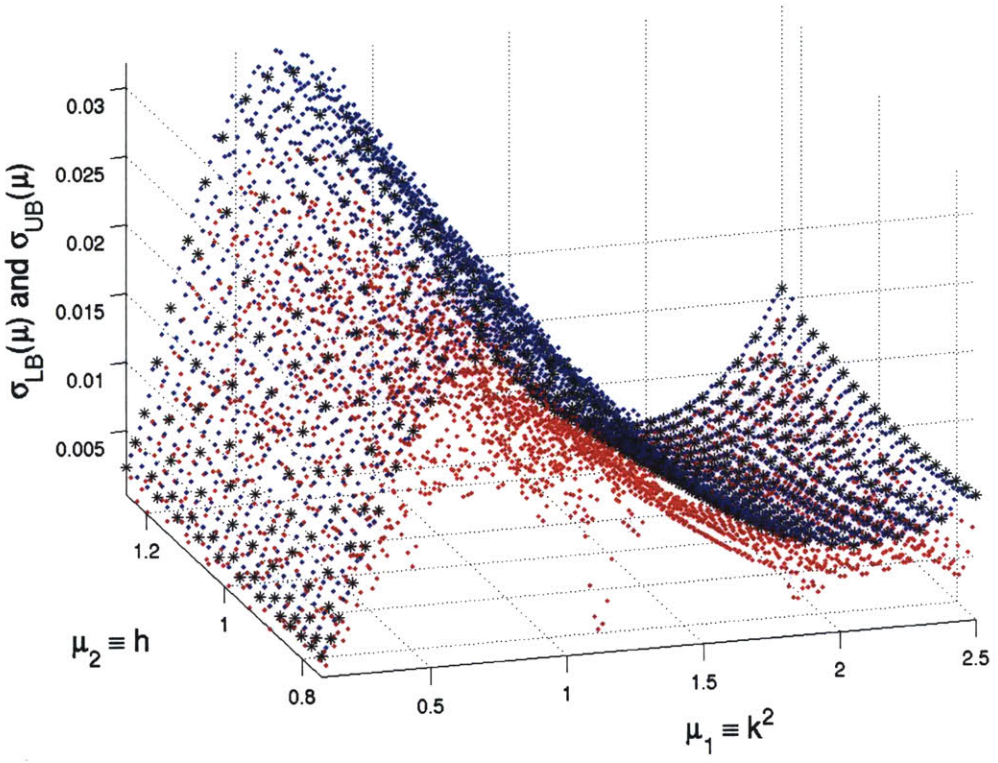


Figure 9-16: Inf-Sup lower bounds using the successive constraint method for the low-pass filter inlet system: we vary $k^2 \in [0.1, 2.5]$ and $h \in [0.75, 1.25]$, and fix $Z_R = 5.0$ and $Z_I = 1.8$. The size of the training sample set $\Xi_{\text{train}} = 3000$ and termination criteria for the greedy algorithm $\epsilon_{\text{tol,min}} = 0.9$. The lower bounds $\sigma_{\text{LB}}(\mu)$ are shown in red, the upper bounds $\sigma_{\text{UB}}(\mu)$ in blue and the black "$*$"s represent the constraints $\omega_k \in \mathcal{C}_K, k = 1, \ldots, K$. We obtain $K = 518$ for our chosen sample set. We specify $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$ for the offline algorithm.

constraints for $Z_R = 5.0$ and $Z_I = 2.0$.

We also summarize the results obtained for the same two-dimensional slices in (k^2, h) space for the outlet system in Table 9.21. Note that the inf-sup is much better behaved for the expanded outlet system - we specify $\mathcal{M}_\alpha = 20$, $\mathcal{M}_+ = 10$ and termination tolerance $\epsilon_{\text{tol,min}} = 0.9$.

Note that as we increase the number of parameters P, we also need to increase $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ correspondingly. The increased parameter dimensionality requires higher number of constraints: for example, if we rquire $\mathcal{M}_+ = 2$ positivity constraints in 1-D, we will need $\mathcal{M}_+ = 4$ in 2-D, and $\mathcal{M}_+ = 8$ in 3-D. The effect of the increased *offline* cost will be magnified by the increase in the size of the offline sample: to maintain the same spacing of parameter points in higher parameter

Z_R	Z_I	K	$\max_{\mu\in\Xi_{\text{train}}} \frac{\sigma_{\text{UB}}(\mu)-\sigma_{\text{LB}}(\mu)}{\sigma_{\text{UB}}(\mu)}$
5.0	1.8	518	8.36E-01
4.8	1.8	487	8.89E-01
4.8	2.0	510	8.64E-01
5.0	2.0	530	8.98E-01

Table 9.20: Inf-Sup lower bound approximation for the $P = 2$ parameter low-pass filter inlet system. We fix Z_R and Z_I and vary $k^2 \in (0.1, 2.5)$ and $h \in (0.75, 1.25)$. $\dim(\Xi_{\text{train}}) = 3000$ and $\epsilon_{\text{tol,min}} = 0.9$ for $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$.

Z_R	Z_I	K	$\max_{\mu\in\Xi_{\text{train}}} \frac{\sigma_{\text{UB}}(\mu)-\sigma_{\text{LB}}(\mu)}{\sigma_{\text{UB}}(\mu)}$
5.0	1.8	205	8.999E-01
4.8	1.8	204	8.991E-01
4.8	2.0	210	8.886E-01
5.0	2.0	212	8.981E-01

Table 9.21: Inf-Sup lower bound approximation for the $P = 2$ parameter low-pass filter outlet system. We fix Z_R and Z_I and vary $k^2 \in (0.1, 2.5)$ and $h \in (0.75, 1.25)$. $\dim(\Xi_{\text{train}}) = 2500$ and $\epsilon_{\text{tol,min}} = 0.9$ for $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$.

dimensions, the number of samples has to grow exponentially (like 2^P). We avoid the 2^P effect by choosing lesser number of random samples (compared to a uniform grid in parameter space) but compensate by increasing $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ accordingly. We choose $\mathcal{M}_\alpha = 50$ and $\mathcal{M}_+ = 20$ for $P = 4$; this choice is indeed very problem-specific and reflects our belief in the relative importance of the two parameters — increasing $\mathcal{M}_\alpha$ is more beneficial than increasing $\mathcal{M}_+$.

Construction of $\mathcal{C}_K$: Modified Greedy Algorithm

We note that on average each LP with $\mathcal{M}_\alpha = 50$ and $\mathcal{M}_+ = 20$ requires roughly $15-20$ milliseconds. Using a sample-set Ξ_{train} of size $n_{\text{train}} = 10000$ samples in the *offline* stage would require ≈ 3.5 minutes for each *offline* iteration. Our greedy SCM algorithm applied as-is would still require ≈ 11 days to run the LPS to generate a constraint set $\mathcal{C}_K$ of size $K = 4500$. This calculation does not even include the cost of the inf-sup calculation at the constraint points, the evaluation of the constraint infimizer coefficients, and the evaluation of upper bounds. Factoring in the additional essential costs and assuming roughly equal computational times for the LP and all other activites, we can conclude that the greedy algorithm will be prohibitively expensive for the low-pass filter problem with $P = 4$ parameters. Furthermore, this simple calculation was based on using only 10000 samples which in $P = 4$-dimensional parameter space represents a fairly coarse mesh at best.

A simpler way to mitigate the effect of n_{train} is to first run the greedy algorithm on relatively small training sample sets before attempting to augment the space with samples from larger sample sets. We now describe the modified SCM algorithm in more detail.

We specify the number of training sample sets $\mathcal{I}$, the smallest training set size $n_{\text{tr,small}}$ and termination error tolerance $\epsilon_{\text{tol,min}}$. Generate $\mathcal{I}$ training sets $\Xi_{\text{tr},1}, \ldots, \Xi_{\text{tr},\mathcal{I}}$ of size $n_{\text{tr},1}, \ldots, n_{\text{tr},\mathcal{I}}$. Choose $\mu_1^* \in \Xi_{\text{tr},1}$ at random as the first μ-sample to be added into the set $\mathcal{C}_K$. Denote

$$C_1^* = \{\mu_1^*\} \tag{9.83}$$

and associated relative error bound

$$\epsilon_k^*(\mu; \Xi_{\text{tr},1}) = \frac{\sigma_{\text{UB}}(\mu) - \sigma_{\text{LB}}(\mu)}{\sigma_{\text{UB}}(\mu)}, \ \forall k, \forall \mu \in \Xi_{\text{tr},1}. \tag{9.84}$$

The modified algorithm then proceeds as follows

```
for j = 1 : I
    μ*_k = arg max_{μ∈Ξ_tr,j} ε*_{k−1}(μ; Ξ_tr,j);
    ε*_k = ε*_{k−1}(μ*_k; Ξ_tr,j);
    if ε*_k > ε_tol,min
        C*_k = C*_{k−1} ∪ μ*_k;
        S*_UB,k = { S*_UB,k−1 ; z(μ*_k) = arg min_{z∈ℝ^Q̂} J(μ*_k; z) };
    else;
        j = j + 1;
        start loop over next training set
    end;
end.
```

We specify $\mathcal{I} = 10$, $n_{\text{tr,small}} = 2000$ and $\epsilon_{\text{tol,min}} = 0.9$. The first 9 sample sets represent a 3×3 uniform grid in (Z_R, Z_I) space ($Z_R = (4.8, 4.9.5.0)$ and $Z_I = (1.8, 1.9, 2.0)$) — for fixed Z_R and Z_I

we have $n_{\text{tr,small}} = 2000$ points in (k^2, h) space. We start by constructing $\mathcal{C}_K$ for $Z_R = 4.8$ and $Z_I = 1.8$; we then augment $\mathcal{C}_K$ with sample points from $\Xi_2, \ldots, \Xi_9$ respectively. We end by testing our $\mathcal{C}_K$ set for completeness over the tenth sample-set $\Xi_{\text{tr},10}$ of size $n_{\text{tr},10} = 20000$ consisting of random samples.

Since we expect the bulk of the SCM constraints to deal with k^2 and h dependence, the first nine sample sets represent a uniform grid over (Z_R, Z_I) space. We used the observation that for fixed k^2 and h, 1-D slices in Z_R and Z_I require roughly three constraints (i.e, sample points) per parameter direction. The Cartesian product in (Z_R, Z_I) space would give us approximately 9 constraints for fixed k^2 and h. The nine 2-D slices in (k^2, h) space thus partition the $P = 4$ parameter space effectively. The last sampleset is large and randomly generated - this is the stage where we find out how good the SCM constraint sets are, our objective is to minimize the number of new constraints we have to generate to satisfy the termination criterion over this large set.

We readily admit the possibility that modified SCM algorithm could yield a larger constraint set, compared to the SCM constraint set obtained from a single large sample set — given our earlier discussion on the computational cost per iteration over a large sample set, we believe this approach is justified. We also believe that other approaches to optimally select sequences of samples to reduce both computational overhead as well as constraint set size could be pursued — however, these approaches would have to be tailored to the particular problem and would require extensive exploration of the parametric dependence of $\sigma(\mu)$. Another approach that would definitely work would be to construct a large constraint set upfront by choosing a certain set of samples where the inf-sup and associated infimizers are calculated initially before any augmentation procedure is initiated; this approach also would likely lead to larger constraint sets.

The smaller sample set size, $n_i = n_{\text{tr,small}} = 2000, i = 1, \ldots, 9$ and the nature of the samples (effectively samples in (μ_1, μ_2) space) imply we can use $\mathcal{M}_\alpha = 20$ and $\mathcal{M}_+ = 10$ for the first 9 sample sets. For the last and largest sample set of size 20000 samples, we specify $\mathcal{M}_\alpha = 50$ and $\mathcal{M}_+ = 20$; we expect that most of the samples will already satisfy our required tolerance criteria and the remaining *offline* iterations will proceed over a greatly reduced sample-set. For the $P = 4$ low-pass filter inlet system, we obtain $\dim(\mathcal{C}_K^{\text{in}}) = 4781$; the outlet system requires $\dim(\mathcal{C}_K^{\text{out}}) = 2078$ constraints.

For the inlet system, the greedy algorithm required on *average* $\approx$ 1 minute per iteration for

the first 9 sample-sets of size 2000 and ended with $k = 4417$ constraints; subseqeuntly the verification/augmentation process over the tenth and largest sample required on *average* ≈ 1.5 minutes ($\approx$ 17000 samples satisfied the tolerance criteria using $k = 4417$ constraints, and were excluded from the augmentation process) and added an additional 364 constraints — the total running time of the greedy algorithm was ≈ 4 days. For the outlet system, the greedy algorithm required on *average* ≈ 1 minute per iteration for the first 9 sample-sets ending with $k = 1861$ constraints; subsequently the verification/augmentation process over the tenth and largest sample required on *average* ≈ 1.3 minutes and added an additional 217 constraints — the total running time of the greedy algorithm was ≈ 1.5 days. We thus conclude that the modified algorithm allowed us to calculate the SCM constraint sets in significantly lesser time compared to our earlier (conservative) estimate of ≈ 11 days for the original SCM algorithm.

Inspite of our convergence over the last and largest sample set, it is important to remember that there is *always* the possibility that some samples might still fail our *offline* termination criterion *online*. In such cases, the samples that fail our offline termination criterion will to be collected, the SCM algorithm re-run over these failed samples, and C_K appropriately augmented.

Post-Processing of C_K

In the construction of C_K using the modified SCM algorithm, we had used $\mathcal{M}_\alpha = 50$ and $\mathcal{M}_+ = 20$ to augment our SCM constraint set from samples in the largest sample set. Here, we experiment with different combinations of $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ to check if we can reduce the online time for inf-sup lower bound approximations by reducing $\mathcal{M}_\alpha$ and $\mathcal{M}_+$. We present the results for the inlet system in Table 9.22; in Table 9.23 we present the results for the outlet system. We confirm that our offline termination criteria is satisfied by the post-processing sample set Ξ_{pp} for the restrictive $\mathcal{M}_\alpha$ and $\mathcal{M}_+$ chosen for the greedy algorithm. For both inlet and outlet systems, we cannot make any improvement and choose $\mathcal{M}_\alpha = 50$ and $\mathcal{M}_+ = 20$ for *online* computations.

The inf-sup lower bound for the inlet system, $\beta_{\text{LB,in}}(\mu)$ ($\equiv \sqrt{\sigma_{\text{LB,in}}(\mu)}$), is ≈ 2.7 times smaller than the exact inf-sup in the worst case; the associated computational cost is ≈ 230 times lesser on average than the exact inf-sup calculation.

The computation for the lower bound of the outlet system, $\beta^{\text{out}}_{\text{LB}}(\mu)$ ($\equiv \sqrt{\sigma_{\text{LB,out}}(\mu)}$), is ≈ 194 times faster on average than the computational cost for the inf-sup and ≈ 2.2 times smaller than

$\mathcal{M}_\alpha$	$\mathcal{M}_+$	T_{B}^{av}	$T_{\text{B}}^{\text{av}}/T_{\text{LB}}^{\text{av}}$	$\min_{\mu\in\Xi_{\text{pp}}} \frac{\sigma(\mu)}{\sigma_{\text{LB,in}}(\mu)}$	$\max_{\mu\in\Xi_{\text{pp}}} \frac{\sigma(\mu)}{\sigma_{\text{LB,in}}(\mu)}$	$\max_{\mu\in\Xi_{\text{pp}}} \epsilon_K^{\text{in}}(\mu)$
40	15	1.46E-02	238.68	-261.55	213.49	1.4923
40	20	1.49E-02	235.26	-261.55	213.49	1.4865
45	15	1.48E-02	236.62	-23.05	9.98	1.2763
45	20	1.51E-02	231.31	-6.82	34.18	1.2759
50	15	1.51E-02	231.67	-2.57	6.77	1.27
50	20	1.52E-02	230.07	1.08	6.76	8.97E-01

Table 9.22: Post-processing of $\mathcal{C}_K^{\text{in}}$ ($K = 4781$): approximation of the inf-sup lower bound for the low-pass filter inlet system for different $(\mathcal{M}_\alpha, \mathcal{M}_+)$. We also present *averaged* timings (obtained using TOMLAB's "linprog" function) over the set Ξ_{pp} ($\dim(\Xi_{\text{pp}}) = 5000$). Here $\sigma_{\text{UB,in}}(\mu)$ and $\sigma_{\text{LB,in}}(\mu)$ are the SCM upper and lower bounds for $\sigma(\mu)$ of the expanded inlet system. $T_{\text{LB}}^{\text{av}}$ is the average time to compute the lower bound in seconds; $T_{\text{B}}^{\text{av}}/T_{\text{LB}}^{\text{av}}$ is the ratio of the average times to compute the inf-sup and the lower bound respectively. Here $\epsilon_K^{\text{in}}(\mu) = \frac{\sigma_{\text{UB,in}}(\mu)-\sigma_{\text{LB,in}}(\mu)}{\sigma_{\text{UB,in}}(\mu)}$.

the exact inf-sup in the worst case.

$\mathcal{M}_\alpha$	$\mathcal{M}_+$	T_{B}^{av}	$T_{\text{B}}^{\text{av}}/T_{\text{LB}}^{\text{av}}$	$\min_{\mu\in\Xi_{\text{pp}}} \frac{\sigma(\mu)}{\sigma_{\text{LB,out}}(\mu)}$	$\max_{\mu\in\Xi_{\text{pp}}} \frac{\sigma(\mu)}{\sigma_{\text{LB,out}}(\mu)}$	$\max_{\mu\in\Xi_{\text{pp}}} \epsilon_K^{\text{out}}(\mu)$
40	15	1.44E-02	202.83	-11.61	58.84	1.16
40	20	1.43E-01	199.66	-11.42	58.84	1.09
45	15	1.47E-02	198.88	-11.61	41.53	1.08
45	20	1.50E-02	195.96	-11.42	41.53	1.07
50	15	1.50E-02	195.02	-11.24	19.11	1.06
50	20	1.52E-02	193.59	1.02	4.53	8.45E-01

Table 9.23: Post-processing of $\mathcal{C}_K^{\text{out}}$ ($K = 2078$): approximation of the inf-sup lower bound for the low-pass filter outlet system with different $\mathcal{M}_\alpha$ and $\mathcal{M}_+$. We also present *averaged* timings (obtained using TOMLAB's "linprog" function) over Ξ_{pp} ($\dim(\Xi_{\text{pp}}) = 5000$). Here $\sigma_{\text{UB,out}}(\mu)$ and $\sigma_{\text{LB,out}}(\mu)$ are the SCM upper and lower bounds for $\sigma(\mu)$ of the expanded outlet system. $T_{\text{LB}}^{\text{av}}$ is the average time to compute the lower bound in seconds; $T_{\text{B}}^{\text{av}}/T_{\text{LB}}^{\text{av}}$ is the ratio of the average times to compute the inf-sup and the lower bound respectively. Here $\epsilon_K^{\text{out}}(\mu) = \frac{\sigma_{\text{UB,out}}(\mu)-\sigma_{\text{LB,out}}(\mu)}{\sigma_{\text{UB,out}}(\mu)}$.

9.3.8 A Posteriori Error Estimation

In Section 9.2.8, we discussed the rapid convergence of the reduced basis outputs, and the properties of the obtained output bound effectivities for the band-stop filter in great detail. Since our earlier observations about the band-stop filter problem results remain valid for the results obtained here, we just summarize the convergence results for $\mathcal{S}_{\text{in},N}(\mu)$, $\mathcal{S}_{\text{out},N}(\mu)$, and the derived output $s_N(\mu)$.

Output Bounds: Online Convergence

We present in Tables 9.24 and 9.25 the error bounds and effectivities for the inlet and the outlet system respectively. As before, we present the results over the large verification set Ξ_{verif} of 5000 samples used for the reduced basis approximation in Tables 9.18 and 9.19. Note that relative to the results for the earlier band-stop filter example, here we choose to report *average* effectivities $\overline{\eta}$ instead of the ratio of maxima ρ_N^{err}. We note that the discussion on effectivities for outputs of the band-stop filter in Section 9.2.9 remains relevant here; similar observations can be made about the effectivities associated with the outputs of the low-pass filter problem as well.

Given $\mu \in \mathcal{D}$, $\eta_N^{\mathcal{S}_{\text{in}}}(\mu)$ (respectively, $\eta_N^{\mathcal{S}_{\text{in}}}(\mu)$), given by (9.59) (respectively, (9.60)), measures the *sharpness* of the error estimate $\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)$ (respectively, $\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)$) associated with the reduced basis error $|\mathcal{S}_{\text{in}}(\mu) - \mathcal{S}_{\text{in},N}(\mu)|$ (respectively, $|\mathcal{S}_{\text{out}}(\mu) - \mathcal{S}_{\text{out},N}(\mu)|$) of the inlet (respectively, outlet) system.

We report the maximum of the relative error in the output $E_{\text{in},N}(\mu)$ given by (9.29), the maximum over the relative error bound $\mathcal{E}_{\text{in},N}(\mu)$ given by (9.30), and the average effectivity $\overline{\eta}_N^{\mathcal{S}_{\text{in}}}$ as a function of N for the inlet system in Table 9.24. We also present in Table 9.24 the computational time to compute the error bound $\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)$ given by (9.27) in terms of the dual norm of the residual for the inlet system, $\epsilon_N^{\mathcal{S}_{\text{in}}}(\mu)$, and the associated inf-sup lower bound, $\beta_{\text{LB,in}}(\mu)$. We obtain relative accuracies of $O(10^{-5})$ for $N = 35$, the computational cost to evaluate $\mathcal{S}_{\text{in},N}(\mu)$ and $\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)$ is ≈ 21 times faster than the evaluation of the true output $s(\mu)$.

N	$\max_{\mu\in\Xi_{\text{verif}}} E_{\text{in},N}(\mu)$	$\max_{\mu\in\Xi_{\text{verif}}} \mathcal{E}_{\text{in},N}(\mu)$	$\overline{\eta}_N^{\mathcal{S}_{\text{in}}}$	Online Time		
				$\mathcal{S}_{\text{in},N}$	$\Delta_N^{\mathcal{S}_{\text{in}}}$	
					$\epsilon_N^{\mathcal{S}_{\text{in}}}$	$\beta_{\text{LB,in}}$
20	5.45E-02	3.53E+00	64.84	1.30E-03	8.7E-03	
25	8.04E-05	9.1E-03	112.93	2.30E-03	9.7E-03	
30	8.45E-07	8.58E-05	101.72	2.56E-03	1.18E-02	3.33E-02
35	5.13E-07	2.33E-05	45.53	2.70E-03	1.26E-02	
40	9.16E-08	1.18E-05	129.15	4.03E-03	1.31E-02	

Table 9.24: Low-Pass Filter: reduced basis convergence result and the online time to evaluate $\mathcal{S}_{\text{in},N}$, $\Delta_N^{\mathcal{S}_{\text{in}}}$ as a function of N; the timing results are normalized with respect to the time to calculate the "truth" output $s(\mu)$. $\Delta_N^{\mathcal{S}_{\text{in}}}(\mu) \equiv \frac{(\epsilon_N^{\mathcal{S}_{\text{in}}}(\mu))^2}{\beta_{\text{LB,in}}(\mu)}$, where $\epsilon_N^{\mathcal{S}_{\text{in}}}(\mu)$ and $\beta_{\text{LB,in}}(\mu)$ are the dual norm of the residual and the inf-sup lower bound for the inlet system.

We report the maximum of the relative error in the output $E_{\text{out},N}(\mu)$ given by (9.32), the

maximum over the relative error bound $\mathcal{E}_{\text{out},N}(\mu)$ given by (9.33), and the average effectivity $\overline{\eta}_N^{\mathcal{S}_{\text{out}}}$ as a function of N for the outlet system in Table 9.25. We also present in Table 9.14 the computational time to compute the error bound $\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)$ given by (9.28) in terms of the dual norm of the residual for the inlet system, $\epsilon_N^{\mathcal{S}_{\text{out}}}(\mu)$, and the associated inf-sup lower bound, $\beta_{\text{LB,out}}(\mu)$. We obtain relative accuracies of $O(10^{-5})$ for $N = 35$, the computational cost to evaluate $\mathcal{S}_{\text{out},N}(\mu)$ and $\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)$ is ≈ 21 times faster than the evaluation of the true output $s(\mu)$.

N	$\max_{\mu\in\Xi_{\text{verif}}} E_{\text{out},N}(\mu)$	$\max_{\mu\in\Xi_{\text{verif}}} \mathcal{E}_{\text{out},N}(\mu)$	$\overline{\eta}_N^{\mathcal{S}_{\text{out}}}$	Online Time		
				$\mathcal{S}_{\text{out},N}$	$\Delta_N^{\mathcal{S}_{\text{out}}}$	
					$\epsilon_N^{\mathcal{S}_{\text{out}}}$	$\beta_{\text{LB,out}}$
20	2.00E-03	7.31E-02	35.81	1.3E-03	9.2E-03	
25	3.72E-04	5.1EE-03	13.59	2.3E-03	9.8E-03	
30	1.32E-05	1.92E-04	14.44	2.6E-03	1.08E-02	3.33E-02
35	6.84E-07	9.57E-06	13.98	2.75E-03	1.23E-02	
40	1.03E-07	1.54E-06	14.91	4.00E-03	1.29E-02	

Table 9.25: Low-Pass Filter: reduced basis convergence result and the online time to evaluate $\mathcal{S}_{\text{out},N}$, $\Delta_N^{\mathcal{S}_{\text{out}}}$ as a function of N; the timing results are normalized with respect to the time to calculate the "truth" output $s(\mu)$. $\Delta_N^{\mathcal{S}_{\text{out}}}(\mu) \equiv \frac{(\epsilon_N^{\mathcal{S}_{\text{out}}}(\mu))^2}{\beta_{\text{LB,out}}(\mu)}$, where $\epsilon_N^{\mathcal{S}_{\text{out}}}(\mu)$ and $\beta_{\text{LB,out}}(\mu)$ are the dual norm of the residual and the inf-sup lower bound for the outlet system.

Finally, in Table 9.26, we present the results for the actual output of interest, the transmission coefficient. The error bound reported is the maximum of the relative error bound, $\Delta_N^s(\mu)/|s(\mu)|$ over the same verification sample set Ξ_{verif} used for the inlet and outlet systems. We note that we cannot obtain $s_N(\mu)$ and $\Delta_N^s(\mu)$ for any choice of N — since these are derived outputs and output bounds, we need to guarantee that the actual error estimators $\Delta_N^{\mathcal{S}_{\text{in}}}(\mu)$ and $\Delta_N^{\mathcal{S}_{\text{out}}}(\mu)$ are small enough: we require $\Delta_N^{\mathcal{S}_{\text{in}}}(\mu) \leq \mathcal{S}_{\text{in},N}(\mu)$ and $\Delta_N^{\mathcal{S}_{\text{out}}}(\mu) \leq \mathcal{S}_{\text{out},N}(\mu)$. We denote $\overline{\eta}_N^s$ the average of the effectivity, $\eta_N^s(\mu)$ over Ξ_{verif}. Note that we reject from our sample Ξ_{verif} those parameter points for which the error bound $\Delta_N^s(\mu)$ smaller than machine precision, as for these parameter points the calculation is contaminated by round-off. We observe relatively good effectivity, our $\overline{\eta}_N^s$ is usually $O(20)$. We obtain relative accuracies of $O(10^{-5})$ for $N = 35$, the computational cost to evaluate $s_N(\mu)$ and $\Delta_N^s(\mu)$ is ≈ 10 times faster than the evaluation of the true output $s(\mu)$.

N	$\max_{\mu\in\Xi_{\text{verif}}} \frac{\lvert s(\mu)-s_N(\mu)\rvert}{s(\mu)}$	$\max_{\mu\in\Xi_{\text{verif}}} \frac{\Delta_N^s(\mu)}{s(\mu)}$	$\overline{\eta}_N^s$
25	1.71E-04	7.8E-03	45.41
30	8.02E-06	1.53E-04	19.09
35	2.08E-06	9.53E-05	45.64
40	1.02E-06	2.95E-05	28.88

Table 9.26: Low-Pass Filter: reduced basis convergence of $s_N(\mu)$ and $\Delta_N^s(\mu)$ as a function of N.

Chapter 10

Conclusions

In this final chapter we conclude the thesis by providing final observations regarding the reduced basis methods we described and give an outline for the directions of the future work.

10.1 Summary

The central themes of this thesis have been the development of reduced-basis methods for strongly non-coercive problems governed by parametrized elliptic partial differential equations and their applications to acoustic contexts requiring the *rapid* calculation of quadratic outputs of interest.

We first introduced the abstract formulation for the acoustics waveguide problems studied in detail in this thesis. We built upon existing ideas [21] and developed a radiation boundary condition which allows us to treat propagating (and evanescent) modes correctly at outflow and truncate the semi-infinite domain without loss of accuracy. The radiation/outflow boundary condition by enabling efficient domain truncation permits us to consider a larger range of acoustic waveguide problems. The radiation boundary condition requires the solution of an eigenproblem to calculate eigenmodes at outflow; however, thanks to the offline-online decomposition in reduced basis methods our computational times are not affected.

Next we outlined the expanded formulation which enables the treatment of quadratic outputs (functionals) as *linear* and *compliant* outputs (functionals) of a slightly different problem with twice the number of unknowns relative to the original problem. The treatment is very general and can be applied readily to the evaluation of *useful* quadratic outputs in other areas of research — for

example, the detection of cracks and in particular analysis of subsequent fatigue growth of cracks and potential brittle failure by allowing accurate estimation of the Stress Intensity Factor [52]. Our "output of interest" (Transmission Coefficient) is in some sense a derived output; its evaluation requires the calculation of two quadratic outputs – the average pressure intensities at the inlet and the outlet. Thus, for the sake of treating both quadratic outputs as linear compliant outputs, we introduced two separate formulations — the expanded inlet and outlet systems.

We then introduced basic but very important concepts of the reduced-basis approach, laying out a solid foundation for several subsequent chapters. The essential components of the approach are (*i*) rapidly uniformly convergent reduced-basis approximations — Galerkin projection onto the reduced-basis space W_N spanned by solutions of the governing partial differential equation at N (optimally) selected points in parameter space; (*ii*) *a posteriori* error estimation — relaxations of the residual equation that provide inexpensive yet sharp and rigorous bounds for the error in the outputs; and (*iii*) offline/online computational procedures — stratagems that exploit affine parameter dependence to decouple the generation and projection stages of the approximation process; and (*iv*) the output-bound based greedy sampling algorithm that constructs the optimal reduced basis *offline*. The operation count for the *online* stage — in which, given a new parameter value, we calculate the output and associated error bound — depends only on N (typically small) and the parametric complexity of the problem. Finally, we also brought in some additional ingredients: the changes to the newly orthogonalized basis and the offline sampling algorithm to deal with the "expanded" system for quadratic outputs.

We next presented the successive constraint method for the construction of rigorous and efficient (online-inexpensive) lower bound for the critical inf-sup stability constant — a generalized minimum singular value — that appears in the denominator of our *a posteriori* error bounds. The lower bound construction is applicable to linear coercive and non-coercive problems. The method, based on an Offline-Online strategy relevant in the reduced basis many-query and real-time context, reduces the Online calculation of the inf-sup to a small Linear Program: the objective is a parametric expansion of the underlying Rayleigh quotient; the constraints reflect stability information at optimally selected parameter points. We discussed the different ingredients in the construction of the inf-sup lower bound and introduced a greedy algorithm for the construction of the constraint set essential to the development of the bound. We *rigorously* proved the validity of the bound and

presented analytical and numerical examples to confirm the theoretical results and demonstrate the efficacy of the lower bound construction for strongly non-coercive problems.

We then combined the reduced-basis approximation and the inf-sup lower bound to obtain our *rigorous* a posteriori error estimators for non-coercive elliptic problems with quadratic outputs. The developed error estimators can easily be reduced to simpler forms to tackle coercive problems or even non-coercive problems with linear outputs. We noted that these error estimators for quadratic outputs treated as linear compliant outputs, generally obtain considerably *sharper* error bounds relative to earlier primal-dual based error estimators [52, 53].

Finally, we integrated the different methodological pieces: the radiation boundary condition at outflow, the expanded formulation for quadratic outputs, the successive constraint-based inf-sup lower bound, and the *a posteriori* error bounds to calculate *realistic* outputs of interest — the transmission coefficient, important for the characterization of acoustic filters.

We considered two representative filter examples — a band-stop filter and a low-pass filter. For the band-stop filter, we noted the difficulty in obtaining lower bounds when the inf-sup is near singular; we thus obtained much larger values of K for the expanded inlet system relative to the outlet system. We explored the twin tradeoffs of offline and online cost of the SCM algorithm relative to the band-stop filter example. For the low-pass filter problem, we outlined a modified greedy algorithm for the SCM to deal with $P = 4$ parameters. We then demonstrated the rapid convergence of the reduced basis approximation, the quality of our inf-sup lower bounds and the *sharpness* of our *a posteriori* error estimators for both problems.

For both examples, we obtained $O(20)$ computational savings for the quadratic outputs (i.e, the inlet and outlet average pressure intensities), and $O(10)$ computational savings for the derived output, the transmission coefficient. We note that the biggest chunk of the computational cost — both offline and online — is still related to the inf-sup lower bound approximation. However, our methods will obtain significantly larger computational economies when $\mathcal{N}$ is larger because our *online* approximation is $\mathcal{N}$-independent. For the problems treated here, $\mathcal{N}$ is small relative to practical three-dimensional problems with far more complex geometries.

The examples presented in these chapters, and the numerical tests performed throughout this thesis, demonstrate that these methods can be very useful for strongly non-coercive problems with near-singular inf-sup stability constants with quadratic outputs. There are a large number

of engineering problems that fall in the ambit of the proposed methods — we only described a few of them in the acoustics context. The significant (expected for larger problems) savings in computation cost, relative to conventional finite element methods, will allow us to consider a large class of problems in both the *real-time* and *many-query* contexts.

10.2 Future Work

We now conclude with some suggestions for future work that could be carried out based on the contributions presented in this thesis. Some of our suggestions relate to the problems posed in the acoustics context, others are more general. We first indicate the suggestions related to the acoustics context.

The first suggestion is related to the choice of the parametrization μ and its effect on the form of the radiation boundary condition. While the radiation boundary condition described in Chapter 4 is very general, there has been a critical assumption that has considerably simplified the formulation of the acoustic waveguide problems. We present once again our μ-dependent eigen problem on the boundary to highlight our assumption: find $(\Xi_i(\mu), \gamma_i(\mu))$ such that

$$\int_{\Gamma_{\text{out}}} \sigma_y(y;\mu) \frac{\partial \Xi_i(\mu)}{\partial y} \frac{\partial \overline{v}}{\partial y} - \phi(y;\mu) \Xi_i(\mu)\, \overline{v} = \gamma_i(\mu) \int_{\Gamma_{\text{out}}} \sigma(x;\mu) \Xi_i(\mu)\, \overline{v},\ \forall v \in X^e|_{\Gamma_{\text{out}}} \tag{10.1}$$

and

$$\int_{\Gamma_{\text{out}}} \sigma(x;\mu) \Xi_i(\mu)\, \overline{\Xi}_j(\mu) = \delta_{ij}; \tag{10.2}$$

where $\sigma_x(y;\mu), \sigma_y(y;\mu) \in L^\infty(]0,1[) > 0$, $\phi(y;\mu) \in L^\infty(]0,1[$, and $\mu \in \mathcal{D} \in \mathbb{R}^P$. In the problems we have discussed, $\sigma_x(\mu)$, $\sigma_y(\mu)$ and $\phi(\mu)$ were chosen such that the eigen vectors Ξ_i were μ-independent even though the γ_i were μ-dependent. This allowed us to calculate the bilinear terms on the radiation outflow condition (3.57) directly. To tackle more general acoustics formulations, we need to be able to handle μ-dependent Ξ_i and its concomitant effect on the bilinear form. While there are ways to handle the μ-dependence — a collateral reduced basis for the boundary eigen problem [64] — there are serious implications on the calculation of the inf-sup lower bound and the development of *a posteriori* error estimators.

The second suggestion is related to the efficiency of the Linear Program that is central to the utility of the successive constraint method. For the numerical examples described in the thesis, the

Linear Program is the most computationally expensive part of the online output bound calculation. As shown in Chapter 9, for the $P = 4$ parameter low-pass filter, we resorted to decompose the parameter space into a series of 2-D slices to limit the offline cost of solving large number of Linear Programs. Irrespective of parameter dimensionality, the computational cost of the LP drastically slows down the offline construction of $\mathcal{C}_K$ and the online approximation of the inf-sup lower bound. Our computational savings will be much greater if we can solve the Linear Programs more efficiently. Hence, we need to identify or develop more efficient LP solvers for our Linear Programs — this is admittedly rather difficult because the LPs are already solved extremely efficiently. Instead, we need to explore dual techniques for *faster* calculation of the inf-sup lower bound.

The third suggestion is related to the development of *rigorous* error estimators for problems with *non-affine* or *non-linear* parameter dependence. While the empirical interpolation method [14, 43] provides an affine expansion for non-affine or non-linear parameter-dependent bilinear and linear forms; the nature of the affine expansion does have an adverse effect on the development of inf sup lower bounds. There is a need to extend the successive constraint method so that it can handle *non-affine* and *non-linear* parameter dependence. To apply the method to acoustic waveguide problems that involve non-affine geometric variation (e.g., for shape optimization, or more refined buried object characterization), the SCM approach will need to be extended.

The development of these methods allows for the solution for general non-coercive elliptic partial differential equations. However, as the problems become more complex, there is a growing need to *automate* the processes — the *offline* effort to enable inf-sup lower bound approximation and reduced basis-space construction — that make efficient *online* response feasible. While complete automation is probably infeasible, there is a need to identify *smarter* approaches to sampling the parameter space that help reduce offline computational and storage costs; construct *better* inf-sup lower bounds *faster* and identify efficient eigenvalue techniques.

Furthermore, there is a need to quantify the level of parametric complexity that can be adequately handled by the reduced-basis methods. It is important to identify the number of parameters P that we can feasibly consider — how large can P be if our techniques remain viable? It is undeniably the case that ultimately we should anticipate exponential scaling (of both N and $\mathcal{C}_K$ related to our inf-sup lower bounds) as P increases, with a concomitant increase in offline cost. However, for modest P we hope that the growth of N will only be polynomial and not exponential: this has

been empirically verified for the case of linear coercive problems with compliant outputs [95]. But if we are to explore the $N - P$ dependence more fully, practical considerations like good sampling algorithms and the efficient construction and verification procedures for our inf-sup lower bound construction will have to be addressed.

Our final suggestion is related to the application of this work to engineering design, optimization, and analysis. Although we applied our method to specific real-world problems, the work cannot be directly applied in *any* setting or to *any* non-coercive problem without sophisticated knowledge of reduced-basis methods. There is a need to *automate* our methods to such a degree that engineers can *deploy* it as *black-box* methods to tackle complex problems, such as (1) non-destructive evaluation of materials and structures relevant to the structural health monitoring of aeronautical and mechanical systems (e.g., aging aircraft, oil pipelines, and nuclear power plant), and in (2) detection of unexploded ordnance detection (e.g., of mines), underwater surveillance (e.g., of submarines), and tomographic scans(e.g., of biological tissues). These practical large-scale applications bring many new opportunities and exciting challenges. On one hand, we expect the computational savings will be even greater for problems with more complex geometry and physical modeling. On the other hand, these problems often require very high dimension of the "truth" approximation space associated with the underlying PDE and large number of parameters. The treatment of these challenging problems will certainly require both theoretical and algorithmic progress on our methods as described above.

Appendix A

Derivation of $\mathcal{P}_{\text{in}} = \mathcal{P}_{\text{out}}$ for a non-dissipative acoustic waveguide

A.1 Dimensional Strong Form

We consider a suitably regular (smooth) domain $\tilde{\Omega} \subset \mathbb{R}^2$ with Lipschitz-continuous boundary $\tilde{\Gamma}$.

Consider a semi-infinite acoustic waveguide coupled with other acoustic elements — for example, a pressure release tube, Helmholtz resonator, or an expansion chamber, as shown in Figure A-1. The governing equations are

$$\nabla^2 \tilde{p} + \frac{\tilde{\omega}^2}{\tilde{c}^2} \tilde{p} = 0 \tag{A.1}$$

on the domain $\tilde{\Omega}$.

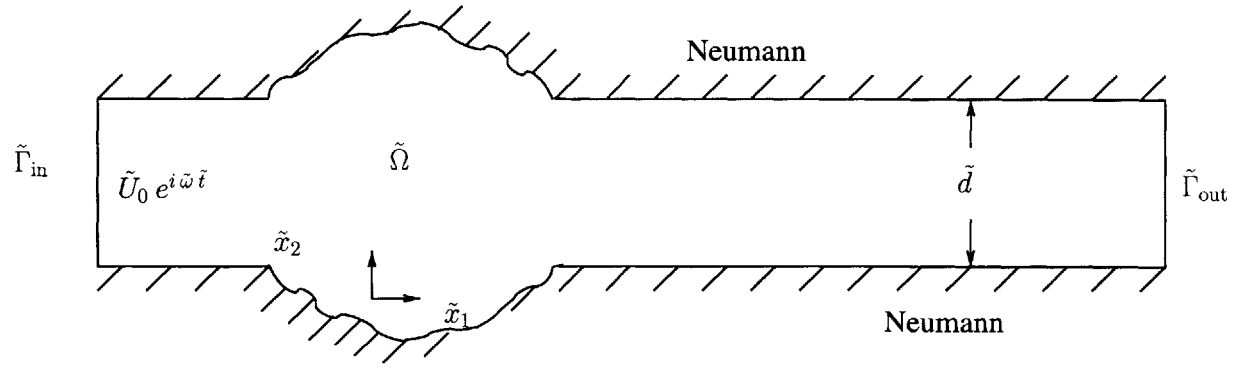


Figure A-1: Acoustic waveguide coupled with other acoustic elements shown as a wavy expanded section in the middle.

The boundary $\tilde{\Gamma}$ ($= \tilde{\Gamma}_{\rm in} \cup \tilde{\Gamma_{\rm out}} \cup \tilde{\Gamma}_N$) has an imposed velocity condition specified at the inlet,

$$\frac{\partial \tilde{p}}{\partial \tilde{n}}|_{\tilde{\Gamma}_{\rm in}} = {\rm i}\,\tilde{\rho}\,\tilde{\omega}\,\tilde{U}_0, \tag{A.2}$$

homogeneous Neumann boundary conditions on Γ_N

$$\frac{\partial \tilde{p}}{\partial \tilde{n}}|_{\tilde{\Gamma}_N} = 0, \tag{A.3}$$

and the radiation boundary condition at the outlet.

The input power $\tilde{\mathcal{P}}_{\rm in}$ is given by

$$\tilde{\mathcal{P}}_{\rm in} = \frac{\tilde{\omega}}{2\,\pi} \int_0^{2\,\pi} \int_{\tilde{\Gamma}_{\rm in}} Re(\tilde{p})\, Re(\tilde{u}); \tag{A.4}$$

where $\tilde{u}$ is the dimensional velocity. We can similarly write the output power $\tilde{\mathcal{P}}_{\rm in}$ as

$$\tilde{\mathcal{P}}_{\rm out} = \frac{\tilde{\omega}}{2\,\pi} \int_0^{2\,\pi} \int_{\tilde{\Gamma}_{\rm out}} Re(\tilde{p})\, Re(\tilde{u}); \tag{A.5}$$

We can re-write the expression for $\tilde{\mathcal{P}}_{\rm in}$ as

$$\begin{aligned}
\tilde{\mathcal{P}}_{\rm in} &= \frac{\tilde{\omega}}{2\,\pi} \int_0^{2\,\pi} \int_{\tilde{\Gamma}_{\rm in}} Re(\tilde{p})\, Re(\tilde{u}) && \text{(A.6)}\\
&= \frac{\tilde{\omega}}{2\,\pi} \int_0^{2\,\pi} \int_{\tilde{\Gamma}_{\rm in}} (\frac{\tilde{p}\, e^{{\rm i}\,\tilde{\omega}\,\tilde{t}} + \overline{\tilde{p}}\, e^{-{\rm i}\,\tilde{\omega}\,\tilde{t}}}{2})(\frac{\tilde{U}_0\, e^{{\rm i}\,\tilde{\omega}\,\tilde{t}} + \overline{\tilde{U}}_0\, e^{-{\rm i}\,\tilde{\omega}\,\tilde{t}}}{2}) && \text{(A.7)}\\
&= \frac{1}{2} \int_{\tilde{\Gamma}_{\rm in}} (\frac{\tilde{p} + \overline{\tilde{p}}}{2})\, \tilde{U}_0; && \text{(A.8)}
\end{aligned}$$

here $\overline{\tilde{p}}$ (respectively $\overline{\tilde{U}}_0$) indicates the complex conjugate of $\tilde{p}$ (respectively $\tilde{U}_0$).

We can re-write the expression for $\tilde{\mathcal{P}}_{\rm out}$ as

$$\begin{aligned}
\tilde{\mathcal{P}}_{\rm out} &= \frac{\tilde{\omega}}{2\,\pi} \int_0^{2\,\pi} \int_{\tilde{\Gamma}_{\rm out}} Re(\tilde{p})\, Re(\tilde{u}) && \text{(A.9)}\\
&= \frac{\tilde{\omega}}{2\,\pi} \int_0^{2\,\pi} \int_{\tilde{\Gamma}_{\rm out}} Re(\tilde{p})\, Re(-\frac{1}{{\rm i}\,\tilde{\omega}\,\tilde{\rho}} \frac{\partial \tilde{p}}{\partial \tilde{n}}) && \text{(A.10)}\\
&= -\frac{\tilde{\omega}}{2\,\pi} \int_0^{2\,\pi} \int_{\tilde{\Gamma}_{\rm out}} Re(\tilde{p})\, Re(\frac{1}{\rm i} \frac{\partial \tilde{p}}{\partial \tilde{n}})\, \frac{1}{\tilde{\rho}\tilde{\omega}} && \text{(A.11)}
\end{aligned}$$

A.2 Non-dimensional Strong Form

We non-dimensionalize the problem with respect to the depth $\tilde{d}$ of the waveguide element: denote $k = \frac{\tilde{\omega}\,\tilde{d}}{\tilde{c}}$ as the non-dimensional wave-number, and $p = \frac{\tilde{p}}{\tilde{\rho}\,\tilde{c}\,\tilde{U}_0}$ as the non-dimensional pressure per unit depth. The governing Helmholtz acoustic equation is given by

$$\nabla^2 p + k^2 p = 0; \tag{A.12}$$

with homogeneous Neumann conditions

$$\frac{\partial p}{\partial n} = 0 \tag{A.13}$$

on Γ_N; and an imposed velocity condition at the inlet

$$\frac{\partial p}{\partial n}|_{\Gamma_{\text{in}}} = \mathrm{i}\,k. \tag{A.14}$$

We truncate the semi-infinite domain and specify the radiation boundary condition at the outlet Γ_{out} as

$$\frac{\partial p}{\partial n}|_{\Gamma_{\text{out}}} = -\sum_{n=1}^{N_{\text{prop}}} \mathrm{i}\,\alpha_n \,(\int_{\Gamma_{\text{out}}} p\,\overline{\Xi}_n)\,\Xi_n. \tag{A.15}$$

N_{prop} and the length of the truncated domain are functions of the range of k^2 for a particular problem and need to be chosen correctly. These issues are discussed in more detail in Chapter 5.

We denote $\mathcal{P}_{\text{in}}$ as the non-dimensional input power per unit depth. Using (A.8) we obtain

$$\mathcal{P}_{\text{in}} = \frac{\tilde{\mathcal{P}}_{\text{in}}}{\rho\,\tilde{c}\,\tilde{U}_0^2\,\tilde{d}} = \frac{1}{2}\int_{\Gamma_{\text{in}}} \mathrm{Re}(p). \tag{A.16}$$

We denote $\mathcal{P}_{\text{out}}$ as the non-dimensional output power per unit depth

$$\mathcal{P}_{\text{out}} = \frac{\tilde{\mathcal{P}}_{\text{out}}}{\tilde{c}\,\tilde{U}_0^2\,\tilde{d}}. \tag{A.17}$$

Substituting (A.11) in (A.17) we obtain

$$\begin{aligned}
\mathcal{P}_{\text{out}} &= \frac{\tilde{\mathcal{P}}_{\text{out}}}{\tilde{c}\,\tilde{U}_0^2\,\tilde{d}} \\
&= -\frac{\tilde{\omega}}{2\,\pi}\int_0^{2\,\pi}\int_{\Gamma_{\text{out}}} \text{Re}(\frac{\tilde{p}}{\tilde{\rho}\,\tilde{c}\,\tilde{U}_0})\,\text{Re}(\frac{1}{\text{i}\,\tilde{\rho}\,\tilde{c}\,\tilde{U}_0}\frac{\partial\tilde{p}}{\partial n})\,\frac{1}{\tilde{\rho}\tilde{\omega}}\,\frac{\tilde{\rho}\,\tilde{c}}{\tilde{d}} \\
&= -\frac{\tilde{\omega}}{2\,\pi}\int_0^{2\,\pi}\frac{1}{k}\int_{\Gamma_{\text{out}}} \text{Re}(p)\,\text{Re}(\frac{1}{\text{i}}\,\frac{\partial p}{\partial n}) \\
&= -\frac{\tilde{\omega}}{2\,\pi}\int_0^{2\,\pi}\frac{1}{k}\int_{\Gamma_{\text{out}}} (\frac{\sum_n p_n\,\Xi_n\,e^{\text{i}\,\tilde{\omega}\,\tilde{t}} + \overline{p}_n\,\overline{\Xi}_n\,e^{-\text{i}\,\tilde{\omega}\,\tilde{t}}}{2}) \\
&\qquad\qquad (\frac{\sum_n \frac{1}{\text{i}}\frac{\partial p_n}{\partial n}\,\Xi_n\,e^{\text{i}\,\tilde{\omega}\,\tilde{t}} - \frac{1}{\text{i}}\frac{\partial \overline{p}_n}{\partial n}\,\overline{\Xi}_n\,e^{-\text{i}\,\tilde{\omega}\,\tilde{t}}}{2}) \\
&= -\frac{1}{4\,k}\sum_n (\frac{1}{\text{i}}\,\overline{p}_n\,\frac{\partial p_n}{\partial n} - \frac{1}{\text{i}}\,p_n\,\frac{\partial \overline{p}_n}{\partial n}) \\
&= -\frac{1}{2\,k}\sum_n \text{Re}(\frac{1}{\text{i}}\,\overline{p}_n\,\frac{\partial p_n}{\partial n}); \qquad \text{(A.18)}
\end{aligned}$$

here p is given by the infinite modal sum

$$p = \sum_{n=1}^{\infty} c_n e^{-\text{i}\,\alpha_n\,x}\,\Xi_n(y) = \sum_n p_n\,\Xi_n, \text{ where } \int_{\Gamma_{\text{out}}} \Xi_m\,\overline{\Xi}_n = \delta_{mn}. \qquad \text{(A.19)}$$

We then note that for $n > N_{\text{prop}}$, the evanescent modes are real,

$$\alpha_n = -\text{i}\,|\alpha_n|,\ p_n \equiv e^{-|\alpha_n|\,x}, \qquad \text{(A.20)}$$

thus, $(\frac{1}{\text{i}}\,\overline{p}_n\,\frac{\partial p_n}{\partial n})$ is pure imaginary. We can then drop the evanescent modes, and write $\mathcal{P}_{\text{out}}$ (A.18) as

$$\mathcal{P}_{\text{out}} = \frac{1}{2\,k}\sum_{n=1}^{N_{\text{prop}}} \alpha_n\,(|\int_{\Gamma_{\text{out}}} p\,\overline{\Xi}_n|\,)^2 \qquad \text{(A.21)}$$

We now prove that $\mathcal{P}_{\text{in}} = \mathcal{P}_{\text{out}}$. We multiply (A.12) by $\overline{p}$ and integrate over Ω to obtain

$$\int_\Omega \nabla\,\overline{p}\cdot\nabla\,p\; - k^2\int_\Omega \overline{p}\,p - \int_{\Gamma_{\text{in}}} \overline{p}\,\frac{\partial p}{\partial n} - \int_{\Gamma_{\text{out}}} \overline{p}\,\frac{\partial p}{\partial n} = 0. \qquad \text{(A.22)}$$

We impose our boundary conditions (A.14) and (A.15) to obtain

$$\int_{\Omega} \nabla \overline{p} \cdot \nabla p \; - k^2 \int_{\Omega} \overline{p}\, p \; - \mathrm{i}\, k \int_{\Gamma_{\mathrm{in}}} \overline{p} + \mathrm{i} \sum_{n=1}^{N_{\mathrm{prop}}} \alpha_n \Big(\int_{\Gamma_{\mathrm{out}}} p \; \overline{\Xi}_n \Big) \Big(\int_{\Gamma_{\mathrm{out}}} \overline{p} \; \Xi_n \Big) = 0. \tag{A.23}$$

Hence, looking at just the imaginary part,

$$\begin{aligned} k \int_{\Gamma_{\mathrm{in}}} \mathrm{Re}(\overline{p}) &= \sum_{n=1}^{N_{\mathrm{prop}}} \alpha_n \Big(\int_{\Gamma_{\mathrm{out}}} p \; \overline{\Xi}_n \Big) \Big(\int_{\Gamma_{\mathrm{out}}} \overline{p} \; \Xi_n \Big) \\ \frac{1}{2} \int_{\Gamma_{\mathrm{in}}} \mathrm{Re}(\overline{p}) &= \frac{1}{2\,k} \sum_{n=1}^{N_{\mathrm{prop}}} \alpha_n \Big(\Big| \int_{\Gamma_{\mathrm{out}}} p\, \overline{\Xi}_n \Big| \Big)^2 \\ \mathcal{P}_{\mathrm{in}} &= \mathcal{P}_{\mathrm{out}} \end{aligned} \tag{A.24}$$

Bibliography

[1] S. Abarbanel and D. Gottlieb. A mathematical analysis of the pml method. *J. Comput. Phys.*, 134:357–363, 1995.

[2] Erik Abenius, Fredrik Edelvik, and Christer Johansson. Waveguide truncation using upml in the finite-element time-domain method. *Uppsala Universitet Technical Report 2005-026*, September 2005.

[3] R. A. Adams. *Sobolev Spaces.* Academic Press, 1975.

[4] O.M. Alifanov. *Inverse Heat Transfer Problems.* Springer, New York, 1994.

[5] B. O. Almroth, P. Stern, and F. A. Brogan. Automatic choice of global shape functions in structural analysis. *AIAA Journal*, 16:525–528, May 1978.

[6] X. Antoine, H. Barucq, and A. Bendali. Bayliss-turkel-like radiation conditions on surfaces of arbitrary shape. *Journal of Mathematical Analysis and Applications*, 229:184–221, 1999.

[7] I. Babuska and J. Osborn. Eigenvalue problems. In *Handbook of numerical analysis*, volume II, pages 641–787. Elsevier, 1991.

[8] I. Babuška, F. Ihlenburg, E. T. Paik, and S. A. Sauter. A generalized finite element method for solving the helmholtz equation in two dimensions with minimum pollution. *Comput. Methods Appl. Mech. Engrg.*, 128:325–359, 1995.

[9] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst. *Templates for the Solution of Algebraic Eigenvalue Problems.* SIAM, 2000.

[10] E. Balmes. Parametric families of reduced finite element models: Theory and applications. *Mechanical Systems and Signal Processing*, 10(4):381–394, 1996.

[11] H.T. Banks. Parameter identification techniques for physiological control systems. In F. Hoppensteadt, editor, *Mathematical Aspects of Physiology*, volume 19 of *Lectures in Applied Mathematics*, pages 361–383. AMS, Providence, RI, 1981.

[12] H.T. Banks and J.M. Crowley. Parameter identification in continuum models. *Journal of Astronautical Science*, 33:85–94, 1985.

[13] H.T. Banks and K. Kunisch. *Estimation Techniques for Distributed Parameter Systems.* Systems & Control: Foundations & Applications. Birkhäuser, 1989.

[14] M. Barrault, N. C. Nguyen, Y. Maday, and A. T. Patera. An "empirical interpolation" method: Application to efficient reduced-basis discretization of partial differential equations. *C. R. Acad. Sci. Paris, Série I.*, 339:667–672, 2004.

[15] A. Barrett and G. Reddien. On the reduced basis method. *Z. Angew. Math. Mech.*, 75(7):543–549, 1995.

[16] K.J. Bathe. *Finite Element Procedure.* Prentice-Hall, Inc., 1996.

[17] A. Bayliss and E. Turkel. Radiation boundary conditions for wave like equations. *Comm. Pure Appl. Math.*, 42:430–451, 1982.

[18] J.V. Beck and K.J. Arnold. *Parameter Estimation.* Wiley, 1977.

[19] J.V. Beck, B. Blackwell, and C.R. St. Clair Jr. *Inverse Heat Conduction.* Wiley, New York, 1985.

[20] V. Bellucci, P. Flohr, C. O. Paschereit, and F. Magni. On the use of helmholtz resonators for damping acoustic pulsations in industrial gas turbines. *J. of Engineering for Gas Turbines and Power*, 126:271–275, 2004.

[21] A. Bendali and Ph. Guillaume. Non-reflecting boundary conditions for waveguides. *Mathematics of Computation*, 68(225):123–144, 1999.

[22] D. T. Blackstock. *Fundamentals of Physical Acoustics.* J. Wiley and Sons, Inc., 2000.

[23] B. Chalmond. *Modeling and Inverse Problems in Image Analysis.* Springer, 2003.

[24] V. Chandrasekharan, M. Sheplak, and L. Cattafesta. Experimental study of acoustic impedance of mems-based micro-perforated liners (aiaa-2006-2401). In *12th AIAA/CEAS Aeroacoustics Conference (27th AIAA Aeroacoustics Conference), Cambridge, Massachusetts*, May 2006.

[25] D. G. Childers and C. Ahn. Modeling the glottal volume-velocity for three voice types. *J. Acoust. Soc. Am.*, 97:505–519, 1995.

[26] D. G. Childers and H. T. Hu. Speech synthesis by glottal excited prediction. *J. Acoust. Soc. Am.*, 96:2026–2036, 1994.

[27] S.T. Clegg and R.B. Roemer. Reconstruction of experimental hyperthermia temperature distributions: Application of state and parameter estimation. *ASME Journal of Biomechanical Engineering*, 115:380–388, 1993.

[28] D. Colton, J. Coyle, and P. Monk. Recent developments in inverse acoustic scattering theory. *SIAM Review*, 42:369–414, 2000.

[29] D. Colton, K. Giebermann, and P. Monk. A regularized sampling method for solving three dimensional inverse scattering problems. *SIAM J. Sci. Comput.*, 21:2316–2330, 2000.

[30] D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering Theory.* Springer, 1998.

[31] A. Craggs. The application of the transfer matrix and matrix condensation methods with finite elements to duct acoustics. *J. Sound and Vibration*, 132:241–254, 1989.

[32] Tie Jun Cui, Yao Qin, Gong-Li Wang, and Weng Cho Chew. Low-frequency detection of two-dimensional buried objects using high-order extended born approximations. *Inverse Problems*, 20:S41–S62, 2004.

[33] W. Desch, F. Kappel, and K. Kunisch, editors. *Control and Estimation of Distributed Parameter Systems*, volume 126 of *International Series of Numerical Mathematics.* Birkhäuser, 1998.

[34] L. Desmons, J. Hardy, and Y. Auregan. Determination of the acoustical source characteristics of an internal combustion engineby using several calibrated loads. *Journal of Sound and Vibration*, 179(5):869–878, 1995.

[35] C. Drioli. A flow waveform-matched glottal model. *J. Acoust. Soc. Am.*, 117(5):3184–3195, 2005.

[36] G. Fant, J. Liljencrants, and Q. Lin. A four-parameter model of glottal flow. *Speech Transmiss. Lab. Q. Prog. Stat. Rep.*, pages 1–13, 1985.

[37] J. P. Fink and W. C. Rheinboldt. On the error behavior of the reduced basis technique for nonlinear finite element approximations. *Z. Angew. Math. Mech.*, 63:21–28, 1983.

[38] M. Frohlich, D. Michaelis, and H. W. Strube. Sim-simulataneous inverse filtering and matching of a glottal flow model for acoustic speech. *J. Acoust. Soc. Am.*, 110:479–488, 2001.

[39] M. A. Goberna and M. A. Lopez. *Linear Semi-Infinite Optimization.* J. Wiley, New York, 1998.

[40] Ch. I. Goldstein. A finite element method for solving helmholtz type equations in waveguides and other unbounded domains. *Math. of Comp.*, 43(190):309–324, 1982.

[41] G. H. Golub and C. F. Van Loan. *Matrix Computations.* The John Hopkins University Press, 1989.

[42] M. Grepl. *Reduced-Basis Approximations and* A Posteriori *Error Estimation for Parabolic Partial Differential Equations.* PhD thesis, Massachusetts Institute of Technology, May 2005.

[43] M. A. Grepl, Y. Maday, N. C. Nguyen, and A. T. Patera. Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. *M2AN (Math. Model. Numer. Anal.)*, 2006.

[44] M. A. Grepl, N. C. Nguyen, K. Veroy, A. T. Patera, and G. R. Liu. Certified rapid solution of partial differential equations for real-time parameter estimation and optimization. In *Proceedings of the 2nd Sandia Workshop of PDE-Constrained Optimization: Towards Real-Time and On-Line PDE-Constrained Optimization*, SIAM Computational Science and Engineering Book Series, 2005. To appear.

[45] Steve Griffin, Steven A. Lane, and Steve Huybrechts. Coupled helmholtz resonators for acoustic attenuation. *Journal of Vibration and Acoustics*, 123(1):11–17, 2001.

[46] M. D. Gunzburger. *Finite Element Methods for Viscous Incompressible Flows: A Guide to Theory, Practice, and Algorithms.* Academic Press, Boston, 1989.

[47] I. Harari and T. J. R. Hughes. Finite element method for the helmholtz equation in an exterior domain: Model problems. *Comp. Meth. Appl. Mech.*, 87:59–96, 1991.

[48] G.F. Hawkins, E.C. Johnson, and J.P. Nokes. Detecting manufacturing flaws in composite retrofits. *SPIE*, 3587:97–104, 1999.

[49] A. S. Hersh, B. E. Walker, and J. W. Celano. Helmholtz resonator impedance model, partl: Nonlinear behavior. *AIAA*, 41(5):795–808, 2003.

[50] S. B. Horowitz, T. Nishida, L. N. Cattafesta, and M. Sheplak. Characterization of a compliant-backplate helmholtz resonator for an electromechanical acoustic liner. *International Journal of Aeroacoustics*, 1(2):185–203 (23), 2002.

[51] I. Huang. A theoretical study of duct noise control by flexible panels. *J. Acoust. Soc. Am.*, 106:1801–1809, 1999.

[52] D. B. P. Huynh and A. T. Patera. Reduced basis approximation and a posteriori error estimation for stress intensity factors. *Int. J. Numer. Meth. Engng (Submitted)*, 2006.

[53] D. B. P. Huynh, J. Peraire, A. T. Patera, and G. R. Liu. Real-time reliable prediction of linear-elastic mode-i stress intensity factors for failure analysis. *Singapore MIT Alliance Conference*, 2006.

[54] F. Ihlenburg and I. Babuška. Finite element solution of the helmholtz equation with high wave number part i: The h-version of the fem. *Computers Math. Applic.*, 30:9–37, 1995.

[55] F. Ihlenburg and I. Babuška. Finite element solution of the helmholtz equation with high wave number part ii: The h-p version of the fem. *SIAM J. Numer. Anal.*, 34:315–358, 1997.

[56] K. Ishikaza and J. L. Flanagan. Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell Syst. Tech. J.*, 51:1233–1268, 1972.

[57] K. Ito and S. S. Ravindran. A reduced-order method for simulation and control of fluid flows. *Journal of Computational Physics*, 143(2):403–425, July 1998.

[58] M. Jones, M. Tracy, W. Watson, and T. Parrott. Effects of liner geometry on acoustic impedance. In *AIAA-2002-2446 8th AIAA/CEAS Aeroacoustics Conference and Exhibit, Breckenridge, Colorado*, June 17-19 2002.

[59] D. Klatt and L. Klatt. Analysis, synthesis and perception of voice quality variations among female and male talkers. *J. Acoust. Soc. Am.*, 87:820–857, 1994.

[60] S. Lall, J. E. Marsden, and S. Glavaski. A subspace approach to balanced truncation for model reduction of nonlinear control systems. *Int. J. Robust Nonlinear Control*, 12:519–535, 2002.

[61] J. B. Lawrie and I. M. M. Guled. On tuning a reactive silencer by varying the position of an internal membrane. *J. Acoust. Soc. Am.*, 120:780–790, 2006.

[62] J. F. Lindsay and S. Katz. *Dynamics of Physical Circuits and Systems.* Champaign, Illinois: Matrix Publishers, Inc., 1978.

[63] G.P. Akilov L.V. Kantorovich. *Functional Analysis in Normed Spaces.* The Macmillan Company, 1964.

[64] L. Machiels, Y. Maday, I. B. Oliveira, A. T. Patera, and D. V. Rovas. Output bounds for reduced-basis approximations of symmetric positive definite eigenvalue problems. *C. R. Acad. Sci. Paris, Série I*, 331(2):153–158, July 2000.

[65] Y. Maday, A. T. Patera, and D. V. Rovas. A blackbox reduced-basis output bound method for noncoercive linear problems. In D. Cioranescu and J.-L. Lions, editors, *Nonlinear Partial Differential Equations and Their Applications, Collége de France Seminar Volume XIV*, pages 533–569. Elsevier Science B.V., 2002.

[66] Y. Maday, A. T. Patera, and D.V. Rovas. Petrov-Galerkin reduced-basis approximations to noncoercive linear partial differential equations. In progress.

[67] Y. Maday, A. T. Patera, and G. Turinici. Global *a priori* convergence theory for reduced-basis approximation of single-parameter symmetric coercive elliptic partial differential equations. *C. R. Acad. Sci. Paris, Série I*, 335(3):289–294, 2002.

[68] T. H. Melling. The acoustic impedance of perforates at medium and high sound pressure levels. *J. Sound. Vib.*, 29(1):9–12, 1973.

[69] B.C. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, 26(1):17–32, 1981.

[70] M. L. Munjal. *Acoustics of Ducts and Mufflers.* J. Wiley, New York, 1987.

[71] A.W. Naylor and G.R. Sell. *Linear Operator Theory in Engineering and Science*, volume 40 of *Applied Mathematical Sciences.* Springer-Verlag, New York, 1982.

[72] N. C. Nguyen. *Reduced-Basis Approximation and* A Posteriori *Error Bounds for Nonaffine and Nonlinear Partial Differential Equations: Application to Inverse Analysis.* PhD thesis, Singapore-MIT Alliance, National University of Singapore, July 2005.

[73] N. C. Nguyen, K. Veroy, and A. T. Patera. Certified real-time solution of parametrized partial differential equations. In S. Yip, editor, *Handbook of Materials Modeling*, pages 1523–1558. Springer, 2005.

[74] A. K. Noor and J. M. Peters. Reduced basis technique for nonlinear analysis of structures. *AIAA Journal*, 18(4):455–462, April 1980.

[75] D. Noreland. Impedance boundary conditions for acoustic waves in a duct with a step discontinuity. *Uppsala Universitet Technical Report 2003-032*, May 2003.

[76] J. Tinsley Oden and L. F. Demkowicz. *Applied Functional Analysis.* CRC Press, 1996.

[77] M.N. Özişik and H.R.B. Orlande. *Inverse Heat Transfer.* Taylor & Francis, New York, 2000.

[78] W. Pascher and R. Pregla. Analysis of rectangular waveguide junctions by the method of lines. *IEEE. Trans. on Microwaves Theory and Techniques*, 43:2649–2653, 1995.

[79] K. S. Peat and K. L. Rathi. A finite element analysis of the convected acoustic wave motion in dissipative silencers. *J. Sound. Vib.*, 184:529–545, 1995.

[80] J. S. Peterson. The reduced basis method for incompressible viscous flow calculations. *SIAM J. Sci. Stat. Comput.*, 10(4):777–786, July 1989.

[81] T. A. Porsching. Estimation of the error in the reduced basis method solution of nonlinear equations. *Mathematics of Computation*, 45(172):487–496, October 1985.

[82] J.K. Potocki and H.S. Tharp. Reduced-order modeling for hyperthermia control. *IEEE Transactions on Biomedical Engineering*, 39:1265–1273, 1992.

[83] C. Prud'homme and A. T. Patera. Reduced-basis output bounds for approximately parametrized elliptic coercive partial differential equations. *Computing and Visualization in Science*, 6(2–3):147–162, March 2004.

[84] C. Prud'homme, D. Rovas, K. Veroy, Y. Maday, A. T. Patera, and G. Turinici. Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods. *Journal of Fluids Engineering*, 124(1):70–80, March 2002.

[85] S. Ramamoorthy, K. Grosh, and T. G. Nawar. Structural acoustic silencers-design and experiment. *Acoustical Society of America Journal*, 114:2812–2824, November 2003.

[86] W. C. Rheinboldt. On the theory and error estimation of the reduced basis method for multi-parameter problems. *Nonlinear Analysis, Theory, Methods and Applications*, 21(11):849–858, 1993.

[87] M. Rossi. *Acoustics and Electroacoustics.* Artech House: Norwood, MA, 1988.

[88] D. V. Rovas, L. Machiels, and Y. Maday. Reduced-basis output bound methods for parabolic problems. *IMA Journal of Applied Mathematics*, 2004. Submitted.

[89] D.V. Rovas. *Reduced-Basis Output Bound Methods for Parametrized Partial Differential Equations.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, October 2002.

[90] J. M. Sabatier and N. Xiang. Systematic investigation on acoustic-to-seismic responses of landmines buried in soil. *J. Acoust. Soc. Am.*, 107:2896, 1999.

[91] Z. S. Sacks, D. M. Kingsland, R. Lee, and J. F. Lee. A perfectly matched anisotropic absorber for use as an absorbing boundary condition. *IEEE Trans. Antennas Propagat.*, 43(12):1460–1463, 1995.

[92] A. Selamet, M. B. Xu, I. J. Lee, and N. T. Huff. Analytical approach for sound attenuation in perforated dissipative silencers. *J. Acoust. Soc. Am.*, 115(5):2091–2099, May 2004.

[93] A. Selamet, M. B. Xu, I. J. Lee, and N. T. Huff. Helmholtz resonator lined with absorbing material. *Acoustical Society of America Journal*, 117:725–733, February 2005.

[94] Ahmet Selamet and Iljae Lee. Helmholtz resonator with extended neck. *The Journal of the Acoustical Society of America*, 113(4):1975–1985, 2003.

[95] S. Sen. Reduced basis approximations and *a posteriori* error estimation for many-parameter heat conduction problems. *Numerical Heat Transfer, Part B Fundamentals*, 2007. In preparation.

[96] S. Sen, K. Veroy, D.B.P. Huynh, S. Deparis, N.C. Nguyen, and A.T. Patera. Natural norm a posteriori error estimators for reduced basis approximations. *JCP (J. Comp. Phys.)*, 2006. In press.

[97] Sang-Hyun Seo and Yang-Hann Kim. Silencer design by using array resonators for low-frequency band noise reduction. *The Journal of the Acoustical Society of America*, 118(4):2332–2338, 2005.

[98] L. Sirovich. Turbulence and the dynamics of coherent structures, part 1: Coherent structures. *Quarterly of Applied Mathematics*, 45(3):561–571, October 1987.

[99] L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America A*, 4(3):519–524, March 1987.

[100] Y. Solodukhov. *Reduced-Basis Methods Applied to Locally Non-Affine and Non-Linear Partial Differential Equations.* PhD thesis, Massachusetts Institute of Technology, 2005.

[101] M. M. Sondhi and J. Schroeter. A hybrid time-frequency domain articulatory speech synthesizer. *IEEE Trans. Acoust., Speech, Signal Process. ASSP-35*, pages 955–967, 1987.

[102] H. Strik. Automatic parametrization of differentiated glottal flow: Comparing methods by means of synthetic flow pulses. *J. Acoust. Soc. Am.*, 103:2659–2669, 1998.

[103] N. Sun, N.-Z. Sun, M. Elimelech, and J.N. Ryan. Sensitivity analysis and parameter identifiability for colloid transport in geochemically heterogeneous porous media. *Water Resources Research*, 37(2):209–222, 2001.

[104] A. Tarantola. *Inverse problem theory and methods for model parameter estimation.* Siam, 2005.

[105] R. Taylor, F. Liu, S. Horowitz, K. Ngo, and T. Nishida. Technology development for electromechanical acoustic liners. In *ACTIVE 04, Williamsburg, Virginia*, September 2004.

[106] L. N. Trefethen and D. Bau III. *Numerical Linear Algebra.* SIAM, 1997.

[107] S. V. Tsynkov. Numerical solution of problems on unbounded domains: A review. *Appl. Numer. Math.*, 27:465–532, 1998.

[108] K. Veroy. *Reduced-Basis Methods Applied to Problems in Elasticity: Analysis and Applications.* PhD thesis, Massachusetts Institute of Technology, 2003.

[109] K. Veroy and A. T. Patera. Certified real-time solution of the parametrized steady incompressible Navier-Stokes equations; Rigorous reduced-basis *a posteriori* error bounds. *International Journal for Numerical Methods in Fluids*, 47:773–788, 2005.

[110] K. Veroy, C. Prud'homme, D. V. Rovas, and A. T. Patera. *A posteriori* error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations (AIAA Paper 2003-3847). In *Proceedings of the 16th AIAA Computational Fluid Dynamics Conference*, June 2003.

[111] K. Veroy, D. Rovas, and A. T. Patera. *A Posteriori* error estimation for reduced-basis approximation of parametrized elliptic coercive partial differential equations: "Convex inverse" bound conditioners. *Control, Optimisation and Calculus of Variations*, 8:1007–1028, June 2002. Special Volume: A tribute to J.-L. Lions.

[112] K. Willcox and J. Peraire. Balanced model reduction via the proper orthogonal decomposition. In *15th AIAA Computational Fluid Dynamics Conference.* AIAA, June 2001.

[113] Z. Wu and J. Fang. Numerical implementation and performance of perfectly matched layer boundary conditions for waveguide structure. *IEEE. Trans. on Microwaves Theory and Techniques*, 43:2676–2683, 1995.

[114] W.W.-G. Yeh. Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resource Research*, 22(2):95–108, 1986.

[115] Y. Q. Zeng and Q. H. Liu. Acoustic detection of buried objects in 3-d fluid saturated porous media: Numerical modeling. *IEEE Transactions on Geoscience and Remote Sensing*, 39(6):216–218, 1998.