Reduced-Basis Output Bound Methods for Parametrized Partial Differential Equations

by

Dimitrios V. Rovas

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2003

Author .....
Department of Mechanical Engineering
October 11, 2002

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

Accepted by .....
Ain Sonin
Chairman, Department Committee on Graduate Students

Reduced-Basis Output Bound Methods for Parametrized Partial Differential Equations

by

Dimitrios V. Rovas

Submitted to the Department of Mechanical Engineering
on October 11, 2002, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

An efficient and reliable method for the prediction of outputs of interest of partial differential equations with affine parameter dependence is presented. To achieve efficiency we employ the reduced-basis method: a weighted residual Galerkin-type method, where the solution is projected onto low-dimensional spaces with certain problem-specific approximation properties. Reliability is obtained by a posteriori error estimation methods — relaxations of the standard error-residual equation that provide inexpensive but sharp and rigorous bounds for the error in outputs of interest. Special affine parameter dependence of the differential operator is exploited to develop a two-stage off-line/on-line blackbox computational procedure. In the on-line stage, for every new parameter value, we calculate the output of interest and an associated error bound. The computational complexity of the on-line stage of the procedure scales only with the dimension of the reduced-basis space and the parametric complexity of the partial differential operator; the method is thus ideally suited for the repeated and rapid evaluations required in the context of parameter estimation, design, optimization, and real-time control.

The theory and corroborating numerical results are presented for: symmetric coercive problems (e.g. problems in conduction heat transfer), parabolic problems (e.g. unsteady heat transfer), noncoercive problems (e.g. the reduced-wave, or Helmholtz, equation), the Stokes problem (e.g flow of highly viscous fluids), and certain nonlinear equations (e.g. eigenvalue problems).

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

Acknowledgments

This thesis would not have been possible without the help and contributions of many people. First, I would like to thank my thesis advisor Prof. Anthony T. Patera. He taught me many things, including but by no means limited to numerical analysis and the ways of research. He provided me with guidance and support, and opened up for me new areas of thinking. His insights as well as his humor were very much appreciated. To him I owe more, both intellectually and humanly than I can ever repay.

My sincere thanks to Professors Robert A. Brown, Bora Mikic and Jaime Peraire for serving on my thesis committee, and for their careful criticisms and comments regarding this thesis. I would also like to thank Prof. Yvon Maday, for hosting me for six months in the University of Paris VI, and for providing a stimulating working environment. Many thanks should go also to Dr. Luc Machiels and Prof. Einar Rønquist, for the significant help they provided when I was starting with my research, and for their friendship.

I am also grateful to the other "GAP" members: Alex, Christophe, Ivan, Karen, Thomas and Yuri. I will never forget the long discussions about science, "life, the universe and everything" — thank you for the time we spent together. Also Mrs. Debra Blanchard, whose supportive and encouraging attitude made our work a more pleasant experience.

Surviving in this "brave new world" would not have been easy without a few good friends. I would like to thank all the people here in Boston, that supported me through the bad times and shared the good times. Also my many friends back in Greece, that I so much missed during the last four years.

Last but not least, I would like to express my love and gratitude to my parents, Vasilis and Kassiani, and my brother Panagiotis for believing in me and accepting the long separation. Without their support, love and encouragement, I would not have been able to pursue my dreams. This work is dedicated to them...

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

In engineering and science, numerical simulation has an increasingly important role. The systems or components in consideration are often modeled using a set of partial differential equations and related boundary conditions; then, a discrete form of the mathematical problem is derived and a solution is obtained by numerical solution methods. As the physical problems become more complicated and the mathematical models more involved, current computational resources prove inadequate.

Especially in the field of optimization or design, where the evaluation of many different possible configurations is required — corresponding to different choices of the design parameters, — even for modest-complexity problems, the computational cost is unacceptably high. Especially for design problems we resort to more traditional approaches: the design goals and constraints are prescribed, and then empirical or semi-empirical approaches are employed to solve the design problem. Numerical simulation is used at the final stages only, as a validation tool. In this case, the results are oftentimes less than satisfactory, relying on crude assumptions, intuition or even luck. To more efficiently utilize the existing computational resources, reliable methods that reduce the complexity of the problem while at the same time preserve all relevant information, are becoming very important.

1.1.1 Input-Output relationship

Central to every design, optimization, or control problem is the evaluation of an "input-output" relationship. The set of input parameters μ, which we will collectively denote as "inputs," identify a particular configuration of the system or component. These inputs may represent design or decision variables, such as geometry or physical properties — for example, in optimization studies; control variables, such as actuator power — for example in real-time applications; or characterization variables, such as physical properties — for example in inverse problems. The output parameters $s(\mu)$, which we'll collectively denote as "outputs," are performance indices for the particular input μ — for example maximum temperatures, stresses, flow rates. These outputs are typically expressed as functionals of field variables associated with a set of parametrized partial differential equations which describe the physical behavior of the system or component. Then we are interested in calculating the outputs $s(\mu) = \mathcal{F}(\mu)$, for many different inputs/configurations μ chosen from a parameter space $\mathcal{D} \subset \mathbb{R}^P$ (P is the number of input parameters). Here, $\mathcal{F}$ encompasses the mathematical description of the physical problem.

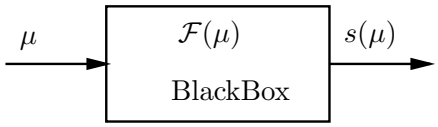

Figure 1-1: Input-Output relationship.

For the evaluation of $\mathcal{F}$ the underlying equations have to be solved. Usually, an analytical solution is not easy to obtain, rather a discretization procedure like the finite-element method, is used; then $\mathcal{F}$ is replaced by $\mathcal{F}_h$, a discrete form amenable to numerical solution. The basic premise, is that as the discretization "length" $h \to 0$, then $\mathcal{F}_h \to \mathcal{F}$, and consequently $s_h(\mu) \to s(\mu),\ \forall \mu \in \mathcal{D}$ *but* as $h \to 0$ the cost of evaluating $\mathcal{F}_h$ becomes prohibitive. Especially in the context of design, control, or parameter identification where "real-time" response or many "input-output" evaluations are required, a balance between computational cost and accuracy/certainty is essential.

1.1.2 Computational Method

Identifying the problem in the high dimensionality of the discrete problems, model-order reduction techniques have been developed. The critical observation is that instead of using projection spaces with general approximation properties — like in finite-element or wavelet methods — we choose problem-specific approximation spaces and use these for the discretization of the original problem. Using such spaces, we can hope to construct a model that represents with sufficient accuracy the physical problem of interest using a significantly smaller number of degrees of freedom. Depending on the choice of the global approximation spaces many possible reductions are available.

The computational methods developed in this work permit, for a restricted but important class of problems, *rapid* and *reliable* evaluation of this partial-differential-equation-induced input-output relationship *in the limit of many queries* — that is, in the design, optimization, control, and characterization contexts. In designing new methods, certain qualities must be considered:

- *Efficiency* is crucial for the problems in consideration. To achieve efficiency, we shall pursue the reduced-basis method; a weighted-residual Galerkin-type method, where the solution is projected onto low-dimensional spaces with certain problem-specific approximation properties.

- *Relevance.* Usually in a design or optimization procedure we are not interested in the field solution or norms of it, but rather in certain design measures such as the drag coefficient in the case of flow past a bluff body, or the average temperature on a surface in the case of heat conduction. The methods developed as part of this work give accurate approximations to these outputs of interest, defined as functional outputs of the field solution.

- *Reliability.* To quantify the error introduced by the reduced-basis method, *a posteriori* error analysis techniques must be invoked. A crucial part of this work is the development of procedures for obtaining rigorous and sharp upper and lower bounds directly for the outputs of interest.

1.2 Model Problem

To motivate and illustrate the various aspects of our method we consider the problem of steady-state heat conduction in a thermal fin. In this section, using the model problem, we present the main ingredients of the method, compare with more traditional approaches, and present some indicative results.

1.2.1 Problem Description

Consider the thermal fin, shown in Figure 1-2, designed to effectively remove heat from a surface. The two-dimensional fin consists of a vertical central "post" and four horizontal "subfins"; the fin conducts heat from a prescribed uniform flux "source" at the root, Γ_{root}, through the large-surface-area subfins to surrounding flowing air.

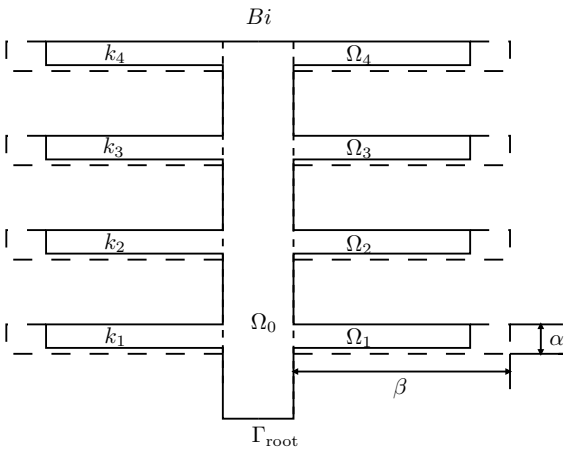

Figure 1-2: Two-dimensional thermal fin

The fin is characterized by a seven component parameter vector or "input", $\mu = (\mu^1, \mu^2, \ldots, \mu^7)$, where $\mu^i = k_i,\ i = 1, \ldots, 4$, $\mu^5 = \text{Bi}$, $\mu^6 = \alpha$, and $\mu^7 = \beta$; μ may take on any value in a specified design space $\mathcal{D} \subset \mathbb{R}^7$. Here k_i is the thermal conductivity of the i^{th} subfin (normalized relative to the post conductivity $k_i \equiv 1$); Bi is the Biot number, a non-dimensional heat transfer coefficient reflecting convective transport to the air at the fin surfaces; and α and β are the thickness and length of the subfins (normalized relative to the post width). The total height of the fin is fixed $H = 4$ (relative to the post width). For our parameter space we choose $\mathcal{D} = [0.1, 10.0]^4 \times [0.01, 1.0] \times [0.1, 0.5] \times [2.0 \times 3.0]$, that is, $0.1 \leq k_i \leq 10.0,\ i = 1, \ldots, 4$ for the conductivities, $0.01 \leq \text{Bi} \leq 1.0$ for the Biot number, and $0.1 \leq \alpha \leq 0.5$, $2.0 \leq \beta \leq 3.0$ for the geometric parameters.

We consider two quantities of interest or "outputs". The first output is T_{root}, the average temperature at the root of the fin normalized by the prescribed heat flux into the fin root. The

particular output relates directly to the cooling efficiency of the fin — lower values of T_{root} imply better performance. The second output is the volume of the fin, V, which represents weight and material cost — lower values are preferred. In general, better performance — lower temperature — requires larger fin volume (e.g., larger α) or materials with higher conductivity; in both cases the production cost of the fin would increase accordingly. Hence there are design trade-offs that must be investigated.

1.2.2 Governing Equations

The temperature distribution $u(\mu)$, is obtained by solution of the following elliptic partial differential equation:

$$-k_i \, \nabla^2 u_i(\mu) = 0 \text{ in } \Omega_i, \; i = 0, \ldots, 4, \tag{1.1}$$

where ∇^2 is the Laplacian operator, and $u_i(\mu) \equiv u(\mu)|_{\Omega_i}$ refers to the restriction of $u(\mu)$ to Ω_i. Here Ω_i is the region of the fin with conductivity k_i, $i = 0, \ldots, 4$: Ω_0 is thus the central post, and Ω_i, $i = 1, \ldots, 4$, corresponds to the four subfins. We must also ensure continuity of temperature and heat flux at the conductivity-discontinuity interfaces $\Gamma_i \equiv \partial\Omega_0 \cap \partial\Omega_i$, $i = 1, \ldots, 4$, where $\partial\Omega_i$ denotes the boundary of Ω_i:

$$\left.\begin{array}{rcl} u_0(\mu) & = & u_i(\mu) \\ -(\nabla u_0(\mu) \cdot \hat{n}_i) & = & -k_i(\nabla u_i(\mu) \cdot \hat{n}_i) \end{array}\right\} \text{ on } \Gamma_i, \; i = 1, \ldots, 4;$$

here $\hat{n}_i$ is the outward normal on $\partial\Omega_i$. Finally, we introduce a Neumann boundary condition on the fin root:

$$-(\nabla u_0(\mu) \cdot \hat{n}_o) = -1 \text{ on } \Gamma_{\text{root}},$$

which models the heat source; and a Robin boundary condition:

$$-k_i(\nabla u_i(\mu) \cdot \hat{n}_i) = \text{Bi } u_i(\mu) \text{ on } \Gamma_{\text{ext}\,i}, \; i = 0, \ldots, 4,$$

which models the convective heat losses. Here $\Gamma_{\text{ext}\,i}$ is that part of the boundary of Ω_i exposed to the fluid.

For every choice of the design parameter-vector μ — which determines the k_i, Bi, and also the fin geometry through α and β — solution of the equations above yields the temperature distribution $u(\mu)$. The average temperature at the root, T_{root}, can then be obtained from $s(\mu) \equiv T_{\text{root}} = \ell^O(u(\mu))$, where

$$\ell^O(v) = \int_{\Gamma_{\text{root}}} v, \tag{1.2}$$

(recall Γ_{root} is of length unity). The volume, V, can be calculated using a simple algebraic relationship $V(\mu) = 4 + 8\alpha\beta$.

The thermal fin problem exercises many aspects of our methods as there is a relatively large number of input parameters that appear in the problem equations and boundary conditions. The variations in geometry are treated in an indirect way by mapping the parameter-dependent solution domain Ω to a fixed reference domain $\hat{\Omega}$. The geometry variations enter then in the problem as parameter-dependent effective orthotropic conductivities. The output or the inhomogeneities in the equations above are not parameter-dependent, by the choice of our non-dimensional variables — this will simplify the presentation and the notation, without loss of generality.

1.2.3 Discretization — Finite Element Method

Finite Element Mesh

Obtaining a solution to the continuous problem (1.1) using analytical techniques, is not easy. Instead, the finite-element method — among many other possible choices — is used to obtain numerically an accurate approximation the exact solution. The point of departure for the finite-element method is an integral re-statement of the equations, called the weak form. The weak form has several advantages: it allows for more general solution spaces, the boundary and continuity conditions are integrated in the problem formulation; see [107] for more details. The problem can then be written as: find $u(\mu) \in Y$ the solution of

$$\mathcal{A}(\mu)u(\mu) = F; \tag{1.3}$$

with $\mathcal{A}$ a linear (distributional) operator, and F a linear functional. The precise definition of Y, $\mathcal{A}$ and F are alluded to the following chapters.

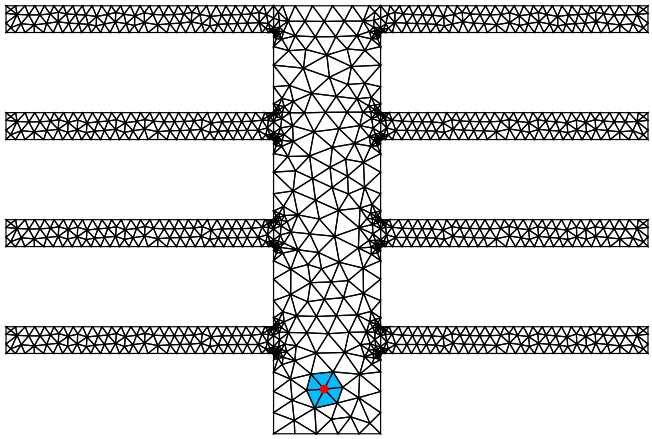

Figure 1-3: Finite element mesh.

For the solution of (1.3), a triangulation $\mathcal{T}_h$ of the computational domain is introduced, as in Figure 1-3. We assume that the triangles, also referred to as elements, cover the computational domain $\hat{\Omega}$, $\bar{\hat{\Omega}} = \cup_{T_h \in \mathcal{T}_h} \bar{T}_h$ ($\bar{T}_h$ is the closure T_h) and that each of the elements do not overlap, $T_h^i \cap T_h^j = 0,\ \forall T_h^i,\ T_h^j \in \mathcal{T}_h$. The subscript h denotes the diameter of the triangulation defined as:

$$h = \sup_{T_h \in \mathcal{T}_h} \sup_{x,y \in T_h} |x - y|; \tag{1.4}$$

here $|\cdot|$ is the Euclidean norm.

Discrete Problem

Using then the triangulation $\mathcal{T}_h$, we define the space Y_h as the space of continuous functions which are piecewise linear over each of the elements $T_h \in \mathcal{T}_h$:

$$Y_h = \{v \in C^0(\hat{\Omega}) | v|_{T_h} \in \mathbb{P}^1(T_h),\ \forall T_h \in \mathcal{T}_h\}. \tag{1.5}$$

If $\mathcal{N}$ is the number of nodes in the triangulation, we introduce the functions $\varphi_i \in Y_h$, such that $\varphi_i(x_j) = \delta_{i\,j},\ i = 1, \ldots, \mathcal{N}$, where x_j are the coordinates of node j, and $\delta_{i\,i} = 1$ if $i = j$, or $\delta_{i\,j} = 0$ if $i \neq j$. Each function φ_i has compact support over the region defined by the elements surrounding node i (shaded area on Figure 1-3). Then, it is not hard to see, that

these functions form a complete basis for the finite element space Y_h. And Y_h can also be defined in terms of this basis:

$$Y_h = \text{span}\{\varphi_i,\ i = 1, \ldots, \mathcal{N}\}. \tag{1.6}$$

Since φ_i is a basis for Y_h, any function $v_h \in Y_h$ can then be written as $v_h = \sum_{i=1}^{\mathcal{N}} v_{h\,i}\varphi_i$, where $v_{h\,i} = v_h(x_i)$ the value of v_h at the node i. From this last expression, we see that we need $\mathcal{N}$ values at the nodes of the triangulation to define each function in V_h. Therefore V_h is a finite-dimensional space with $\dim V_h = \mathcal{N}$. Different choices for the finite-element spaces are possible, for example we can choose to approximate the function using higher order polynomials over each of the elements; these and other choices are discussed in [20].

Using a Galerkin projection in the space spanned by the φ_i, we compute an approximation $u_h \in Y_h$ to the solution $u \in Y$, from:

$$\underline{\mathcal{A}}_h(\mu)\underline{u}_h(\mu) = \underline{F}_h; \tag{1.7}$$

here $\underline{\mathcal{A}}_h$ is an $\mathcal{N} \times \mathcal{N}$ matrix, and $\underline{u}_h(\mu)$ a vector for which $u_{h\,i}(\mu) = u_h(x_i; \mu)$, with x_i the coordinates of the node i. Solving the linear system above, we obtain the nodal values $\underline{u}_h(\mu)$, and therefore $u_h(\mu) = \sum_{i=1}^{\mathcal{N}} u_{h\,i}(\mu)\varphi_i$. The output approximation $s_h(\mu)$ can then be easily computed from:

$$s_h(\mu) = \ell^O(u_h(\mu)). \tag{1.8}$$

Computational Complexity

We see that the original problem has been replaced by a finite-dimensional one. The *a priori* convergence theory for this type of finite-elements and assuming sufficient regularity of the solution $u(\mu)$, suggests that the error in the output $|s(\mu) - s_h(\mu)|$ will converge as h^2, where h is defined in (1.4). Moreover as $h \to 0$, we get $u_h(\mu) \to u(\mu)$ and $s_h(\mu) \to s(\mu)$. The above *a priori* result suggests also, that to decrease the error in the output by a factor $C > 0$, we need to increase the number of elements and therefore $\mathcal{N}$ roughly by the same factor. We see that as the requirements for accuracy increase or the geometric complexity increases, we

need higher $\mathcal{N}$ to obtain accurate and reliable results (to ascertain the accuracy we need *a posteriori* error estimators). Moreover, in the presence of singularities or boundary layers, local refinement is essential, further increasing the required degrees of freedom.

The discussion above suggests that even for relatively simple problems, $\mathcal{N}$ can be large. For the thermal fin problem, $\mathcal{N} \sim \mathcal{O}(10^3)$, but it is not uncommon for $\mathcal{N}$ to be $\mathcal{O}(10^6)$ or higher. We also see the difficulty, as $\mathcal{N}$ increases, so does the size of the linear system (1.7), that has to be inverted. By virtue of the compact support of φ_i, the matrix $\underline{\mathcal{A}}_h$ is sparse and therefore iterative solvers can be used to obtain a solution. The computational complexity scales as $\mathcal{O}(\mathcal{N}^a)$, where a depends on the condition number of the problem (which increases quadratically with $1/h$). Especially in contexts where repeated solution of (1.7) is required, the computational requirements soon become unacceptably large.

1.2.4 Reduced-Basis Method

Low-dimensional approximation

Identifying the problem in the high dimensionality of the finite-element spaces, we look for ways to further reduce the computational complexity. The large number of degrees of freedom required in the case of finite-element methods, is attributed to the particular choice of basis functions, which have general approximation properties for functions in Y. To further reduce the computational complexity we look for spaces with approximation properties specific to the problem of interest.

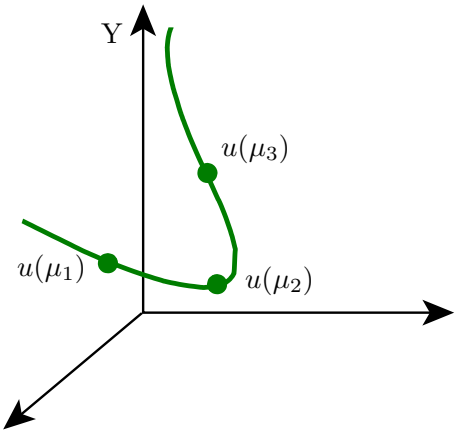

Figure 1-4: Low-dimensional manifold

Our method of choice is the reduced-basis method, first introduced in [74]. The critical observation is that the solution and the output evolve in a low-dimensional manifold induced by the parametric dependence of the problem. Central to reduced-basis methods, is constructing an approximation to this manifold. In our approach, slightly different from

earlier approaches, we construct linear reduced-basis spaces comprising of solutions to (1.3) at different parameter points. We then use these spaces to find an approximation $u_N(\mu)$ to the exact solution.

Earlier approaches viewed the reduced-basis method as a combined projection and continuation method. A different view, suitable for our purposes, is that of multi-dimensional parameter-space interpolation. The required interpolation weights are obtained by solving suitably defined low-dimensional problems chosen to minimize the approximation error measured in problem-specific energy norms.

Reduced-basis space

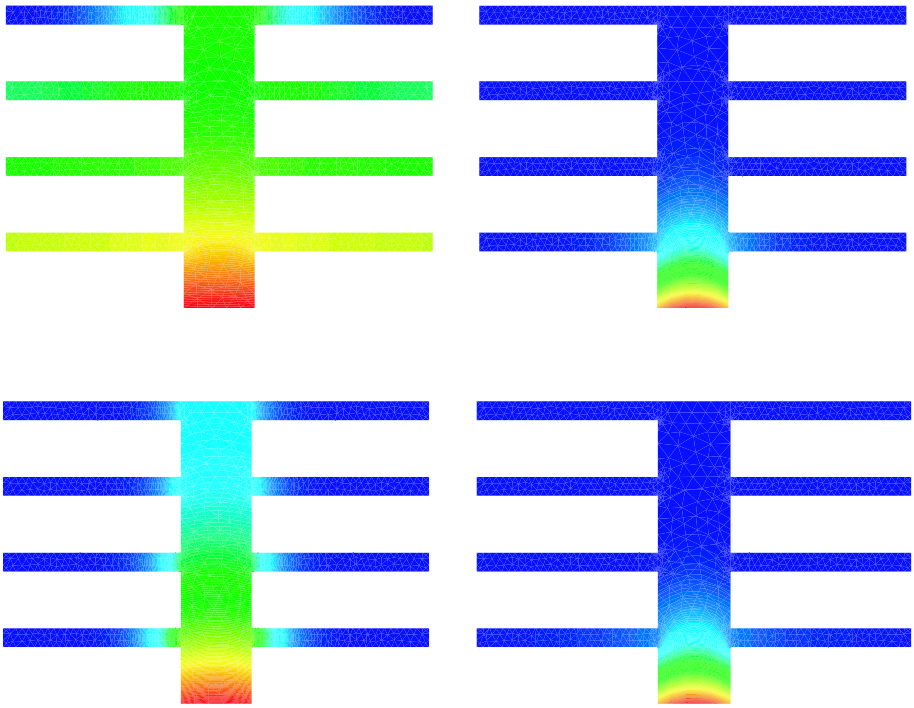

Figure 1-5: Basis functions for W_N

To construct the reduced-basis space we choose N points — N is small, typically $\mathcal{O}(10)$ — $\mu_i \in \mathcal{D},\ i = 1, \ldots, N$. We then compute the solution of (1.3) for each of these points and construct the reduced basis space W_N:

$$W_N = \text{span}\,\{u(\mu_i),\ i = 1, \ldots, N\} \equiv \text{span}\,\{\zeta_i,\ i = 1, \ldots, N\}\,; \tag{1.9}$$

the ζ_i form a basis for the space W_N. By the construction above, and assuming linear independence of the basis functions, the dimension of W_N will be $\dim W_N = N$. Then using a Galerkin projection we compute $\underline{u}_N(\mu)$, the solution of:

$$\underline{\mathcal{A}}_N(\mu)\underline{u}_N(\mu) = \underline{F}_N; \tag{1.10}$$

note that, $\underline{u}_N(\mu)$ can be understood as the interpolation weights mentioned above. The reduced-basis approximation to solution can then be computed from $u_N(\mu) = \sum_{i=1}^{N} u_{N\,i}\zeta_i$, and for the output $s_N(\mu) = \ell^O(u_N(\mu))$.

The *a priori* convergence theory, and extensive numerical tests, suggest that the convergence of the reduced-basis approximation to the exact will be very fast. In fact, exponential convergence is observed in all the numerical tests. This suggests that even with a very modest N, we can hope to achieve good accuracy. The linear system above can be formed and solved very efficiently in the case where the operator depends affinely on the parameters. In this case we can separate the computational steps into two stages:

- The *off-line* stage, in which the reduced-basis space is constructed and some preprocessing is performed. This is an expensive step, that needs to be performed only once, requiring solutions of finite-element problems.

- The *on-line* stage, in which for each new parameter value, the reduced-basis approximation for the output of interest is calculated.

The on-line stage is "blackbox" in the sense that there is no longer any reference to the original problem formulation: the computational complexity of this stage scales only with the dimension of the reduced-basis space and the parametric complexity of the partial differential operator. The "blackbox" nature of the on-line component of the procedure has other advantages. In particular, the on-line code is simple, non-proprietary, and completely decoupled from the (often complicated) off-line "truth" code. This is particularly important in multidisciplinary design optimization, in which various models and approximations must be integrated.

1.2.5 Output Bounds

The computational relaxation introduced in the previous section, allows us to compute very efficiently accurate approximations to the solution and the output of interest. Thanks to the expected rapid convergence N could, in theory, be chosen quite small. However, in practice we do not know how small N should be: this will depend of the desired accuracy, the choice of μ_i in the construction of the reduced-basis spaces, the output of interest and the particular problem in question; in some cases $N = 5$ may suffice, while in other cases $N = 100$ may still be insufficient. In the face of this uncertainty, either too many or too few basis functions will be retained: the former results in computational inefficiency; the later in unacceptable uncertainty. For the successful application of reduced-basis methods it is therefore critical that we can ascertain the accuracy of our predictions; we develop in the next chapters, rigorous error-estimation approaches, directly for outputs of interest, to *a posteriori* validate the accuracy of our predictions.

We prove that these estimators $s_N^+(\mu)$ and $s_N^-(\mu)$ are upper and lower bounds, respectively, to the "true" output $s_h(\mu)$ that would be obtained by solution of the expensive finite-element problem:

$$s_N^+(\mu) \leq s_h(\mu) \leq s_N^-(\mu). \tag{1.11}$$

Unlike the exact value, these error estimators can be computed inexpensively — with a complexity that scales only with the dimension of the reduced-basis space.

In reality the error in the output has two components:

$$|s(\mu) - s_N(\mu)| \leq |s(\mu) - s_h(\mu)| + |s_h(\mu) - s_N(\mu)|;$$

the first related to the discretization error (see in Section 2.1.2); and the second to the reduced-basis error. In practice, both of these errors have to be estimated for reliability in our predictions. Estimation of the discretization error has been treated extensively in the literature; see [87] for a review. For our purposes, we assume that h is chosen very conservatively such that $s_h(\mu) \approx s(\mu)$ and the dominant error is due to the reduced-basis approximation.

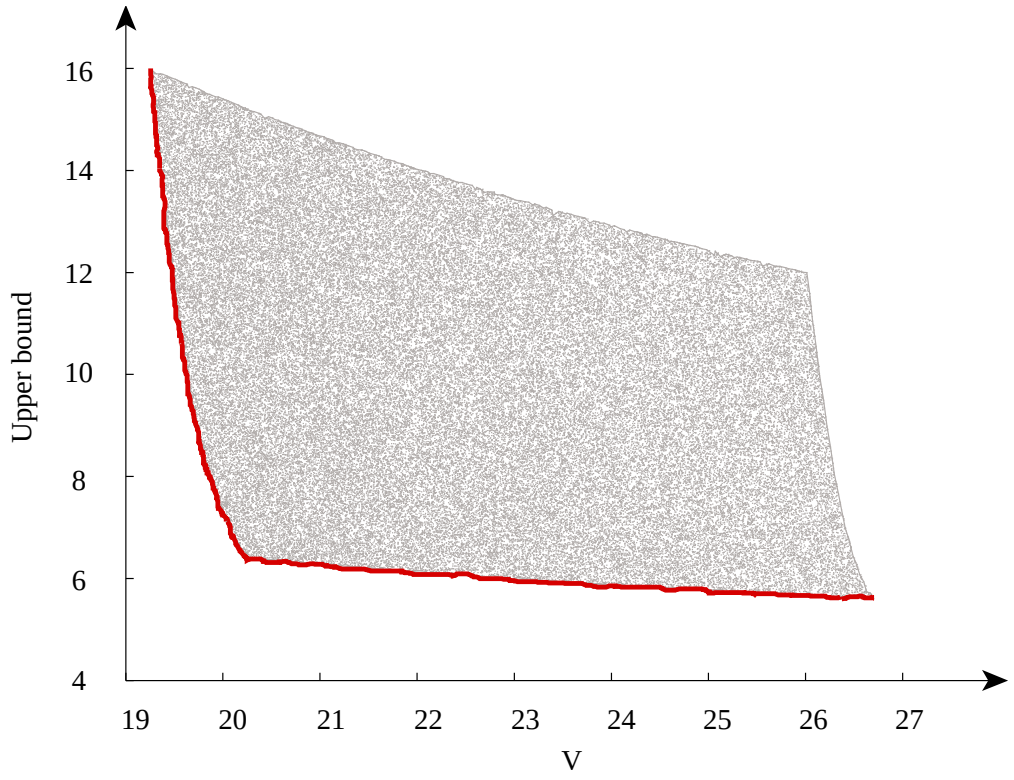

Figure 1-6: Pareto efficient frontier.

1.2.6 Design Exercise — Pareto curve

We close this section with a more applied example. We fix all parameters except α and β so that $\mathcal{D} = [2.0, 3.0] \times [0.1, 0.5]$. As a "design exercise" we construct the achievable set — all those $(V(\mu), s(\mu))$ pairs associated with some (α, β) in $\mathcal{D}$; the result, based on many evaluations of $(V(\mu), s_N^+(\mu))$ for different values of $(\alpha, \beta) \in \mathcal{D}$, is shown in Figure 1-6. We present the results in terms of $s_N^+(\mu)$ rather than $s_N(\mu)$ to ensure that the actual temperature $s_h(\mu)$ will always be lower than our predictions (that is, conservative); and we choose N such that $s_N^+(\mu)$ is always within 0.1% of $s_h(\mu)$ to ensure that the design process is not misled by inaccurate predictions. Given the obvious preferences of lower volume and lower temperature, the designer will be most interested in the lower left boundary of the achievable set — the Pareto efficient frontier; although this boundary can of course be found without constructing the entire achievable set, many evaluations of the outputs will still be required. As regards computational cost, the calculation of $s_N^+(\mu)$ is roughly 24 times faster

than direct calculation of $s_h(\mu)$. The computational savings are quite modest; for more complex problems, savings of the $\mathcal{O}(100)$ or even $\mathcal{O}(1000)$ should be expected.

1.3 Outline

The discussion above suggests that for design, optimization and control problems reduced-basis output bound methods are attractive alternatives to more traditional approaches. In the following chapters we develop rigorously and with more details these methods. More specifically: first, we investigate what the definition of the reduced-basis spaces and the projection operator should be, and how these choices affect the accuracy and stability of our approximations; second, we develop error estimation procedures, directly for the outputs of interest; and finally, corroborating numerical results are presented. In all cases, we give computational complexity estimates and implementation details.

The issues above are investigated in conjunction with the mathematical properties of the underlying partial differential operator. In our presentation, we consider the following classes of problems: coercive — for example, heat conduction problems; elliptic non-coercive — for example problems in acoustics; parabolic problems — for example unsteady heat conduction; eigenvalue problems; and Stokes problems — for example, highly viscous fluid flow.

In the next chapter, we review some of the earlier work related to model-order reduction and in particular to reduced-basis methods; also, we give a few mathematical preliminaries required in the following. Then we develop the reduced-basis method for the different classes of problems: in Chapter 3 for coercive problems; in Chapter 4 for parabolic problems; in Chapter 5 for non-coercive problems, like the reduced-wave (Helmholtz) equation; in Chapter 6 for the Stokes problem; and in Chapter 7 for eigenvalue problems. We conclude in Chapter 8, with some suggestions for future work.

Chapter 2

Background

Before we proceed with the development of reduced-basis output bound methods, we give in this chapter some relevant background information. The issue of reducing complexity while preserving all relevant information, has been a very active research area in many disciplines. A characteristic of systems whose behavior is governed by partial differential equations is that the resulting state models, obtained by a discretization procedure, are of very high-dimension. Therefore some of the existing methods developed, for example in control systems theory, are not directly applicable. We summarize in Section 2.1, recent developments and relevant approaches. The references provided in the following and additional references at the end of this thesis, although by no means exhaustive, should cover most of the recent work. As model-order reduction methods are by definition pre-asymptotic, validation of the obtained results has been recognized to be a critical ingredient. Even though residual-based error measures have been suggested, no rigorous *a posteriori* error estimation procedures have been developed. In other contexts, like estimation of the discretization error in finite-element analysis, a plethora of *a posteriori* error estimation methods exist. Some of these methods are relevant for our problems; we discuss in section 2.1.2 the connection and differences between them. Finally, we review in section 2.2 some mathematical concepts that will be used extensively in the following.

2.1 Earlier Work

2.1.1 Model-Order Reduction

Proper Orthogonal Decomposition

We start our discussion with the proper orthogonal decomposition method, probably the most popular model-order reduction technique. Underlying this method is the solution of the following approximation problem [44]: given a (possibly large) set of vectors, identify the best approximating N-dimensional plane (subspace) such that the root-mean square L_2-projection error is minimized. A solution to this problem can be obtained using the singular value (or Karhunen-Loève) decomposition [48, 59]. The proper orthogonal decomposition has been applied and (re-)discovered in many different areas: system dynamics, stochastic processes, image processing, to name a few.

For reduction of physical systems, it has been extensively applied to time-dependent problems. In this case, time is considered as the varying parameter, and "snapshots" of the field variable (e.g. temperature, displacement) at different times — parameter points — are obtained using numerical or experimental procedures. The optimal N-dimensional approximation space (for N small) is constructed by applying the singular-value decomposition to these vectors, and keeping only the N singular vectors corresponding to the largest singular values. As the singular values are related to the total "energy" of the approximation, these modes can be identified as the ones preserving most of the energy. The reduced model is then obtained by using a Galerkin projection to the space spanned by these vectors.

The optimality property and generality of these ideas, has led to the successful application of the method in many areas: turbulent flows [60], fluid structure-interaction [22], non-linear structural mechanics [51], turbo-machinery flows [115]. Extension of these methods to general multi-parameter problems has been quite limited. The problem is that the singular values are not system invariants as they depend on the choice of "snapshots" and the particular configuration in consideration. It has been observed that reduced-order models obtained for one configuration were not optimal for other configurations; using such models often lead to inaccurate or, even worse, incorrect results. It has been suggested in [19] to give more

weight or preselect some of the vectors in the starting basis, leading to "weighted-POD" or "predefined-POD" methods, but the selection of the required weights is not automatic limiting the generality of such approaches.

An analysis of the model-truncation error suggests that the error can be attributed to two sources: first in the inability of the low-order model to reproduce the exact loading; and second, for the approximated loading, in the inability of the low-order model to recover the exact solution [49]; see also [90] for similar ideas. Using terms from control-systems theory, the first error is related to the controllability (primal) and the second to the observability (dual) of the low-order model. In a similar manner, for our methods, we use a combined primal-dual approach to estimate both of these errors. The notion that a truncation of the model should balance both of these errors, led to balanced-truncation methods [72]. For high-dimensional systems, computation of the required observability and controllability grammians is very expensive. A number of methodologies combining the proper orthogonal decomposition and the balanced-truncation method have been constructed [55, 115].

Reduced-Basis Methods

We turn now to reduced-basis approaches, upon which our method is also based. The reduced-basis method has been proposed in [6, 74] for the non-linear analysis of structures. In these approaches, only single-parameter problems were considered and the method was viewed as a continuation procedure. The method has been further investigated and extended by Noor [75, 76, 77, 78, 79, 80, 81, 82], where it was realized that the method could be applied for general multi-parameter problems. Much of the earlier work focused: first, on the selection and efficient computation of basis functions; and second, on validation of the efficiency and accuracy of reduced-basis approaches in a number of test problems.

As was mentioned in the introduction the reduced-basis method recognizes 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. In these earlier approaches, the approximation spaces for the low-dimensional manifold were typically defined

"locally" — relative to a particular parameter point. Fink and Rheinboldt [33] placed the method in this geometric setting and carried out an error analysis for a general class of single-parameter problems. Porsching [91] considered Lagrangian, Taylor and discrete least squares approximation spaces, and extended some of the *a priori* analysis. In [34] a general local error estimation theory for single-parameter problems was developed containing the earlier estimates as special cases. The extension of the error analysis to multi-parameter problems was presented in [101]. Finally, evaluation of the constants that appear on the error bounds was considered in [14]. The *a priori* theory as developed in the works above concludes that, close to the parameter point selected for the construction of the reduced-basis spaces, the error converges to zero exponentially fast with the number of basis functions used.

Reduced-basis approaches have been subsequently developed in many other areas. Peterson [89] applied it to fluid flow problems and the Navier-Stokes equations, and in [41, 40] it was used for control of fluid problems. Also an analysis was carried out for ordinary differential equations [92], and differential algebraic equations [56]. The reduced-basis approach as earlier articulated was local in parameter space in both practice and theory. As a result, the computational improvements — relative to conventional (say) finite-element approximation — were often quite modest [91]. Balmes [12], was the first to consider general multi-parameter problems. In his approach, similar to the one developed below, he suggests choosing the basis functions by sampling globally in parameter space. Finally, in [70] a combined reduced-basis domain-decomposition approach is proposed for the treatment of geometric parameters. Even though the importance of error estimation is emphasized in the literature, no rigorous *validation* methods have been developed.

The work here differs from these earlier efforts in several important ways: first, we develop (in some cases, provably [69]) *global* approximation spaces; second, we introduce rigorous *a posteriori error estimators*; and third, we exploit *off-line/on-line* computational decompositions (see [12] for an earlier application of this strategy within the reduced-basis context). These three ingredients allow us — for the restricted but important class of "parameter-affine" problems — to reliably decouple the generation and projection stages of reduced-basis approximation, thereby effecting computational economies of several orders of magnitude [94].

Other Methods

Krylov-subspace techniques like the Arnoldi or the Lanczos methods and their variants, have traditionally been used for the calculation of a small set of the extremal eigenvalues and eigenvectors for large-scale eigenproblems. But these are precisely the eigenvalues and eigenvectors of interest for model reduction. Many reduction approaches based on Krylov-subspace techniques have been developed; for an overview see [7, 45] and the references contained therein. The iterative nature of the algorithms, makes it difficult to develop error bounds; moreover, the stability of the reduced-order problem is not always guaranteed.

Finally, for the sake of completeness we mention that a number of other approaches — not based on model-order reduction — for the efficient and reliable evaluation of "input-output" relationships are available: from "fast loads" (e.g., [18, 30]) to matrix pertubation theories (e.g., [4, 116]) to continuation methods (e.g., [5, 100].)

2.1.2 *A posteriori* error estimation

The issue of *a posteriori* error estimation and, more generally, validation of the numerical predictions has received considerable attention in the finite-element literature. The problem of interest there is related to the choice of mesh to be used for the definition of the finite-element spaces. Following the discussion on Section 1.2.3, it is understood that there are certain trade-offs associated with the choice of the finite-element mesh: on one hand, a conservative choice, ensures high accuracy but also the computational costs become formidable; on the other hand, the choice of a relatively coarse mesh ensures efficiency but the accuracy is dubious. More to that, for a specific choice of mesh, the obtained accuracy is not easy to calculate as it depends on the topology of the mesh, the particular problem in consideration, the choice of finite-element spaces, or even the way we choose to measure the error. We can also relate a number of other problems like, for example, the choice of elements to be refined in adaptive refinement or, more generally the choice of "optimal" meshes (i.e. meshes which for a given accuracy minimize computational cost). For all these problems, the ability to estimate and therefore control, the discretization error is critical.

The extensive *a priori* theory can not be used as the provided error bounds depend

on norms of the exact solution which, in general, is not known. Rather, the *a posteriori* error estimators give bounds which depend on computable quantities, like residuals. The study of these types of error estimators started in the 70s with the first paper by Babuska and Rheinboldt [10], and since then the literature has grown appreciably; a review can be found in [3]. Most error estimators developed give bounds for abstract norms of the error. Relevant to this thesis are *a posteriori* error estimators directly for outputs of interest; see for example [87, 88] for relevant work.

The parallel with the discussion in Section 1.2.5 for the reduced-basis method should be clear: instead of the finite-element mesh and the discretization error, we have the parameter space "discretization" (in the sense, of the choice of μ_i in (1.9)), and the reduced-basis approximation error; refinement of the mesh, corresponds to adding more basis-functions in the definition of W_N (1.9). But there are also differences, the most important being the parameter-dependence of the operator, consideration of which is not required in the finite-element case. Even though the methodologies are distinctively different; some of the general ideas [88] for *a posteriori* error estimation are common.

2.2 Mathematical Preliminaries

In this section, we introduce some notation and review some basic definitions that will be used extensively in the following. To start, let $\Omega \subset \mathbb{R}^d$, $d = 1, \ldots, 3$ be an open domain with Lipschitz-continuous boundary. The following function spaces can be defined:

Spaces of Continuous Functions.

Definition 1. *Choose k a non-negative integer, and define $C^k(\bar{\Omega})$ as:*

$$C^k(\bar{\Omega}) = \{v |\ D^\alpha v \text{ is bounded and uniformly continuous on } \Omega,\ 0 \leq |\alpha| \leq k\}\,; \tag{2.1}$$

where α is a multi-index and

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

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

$$\|v\|_{C^k(\bar{\Omega})} = \max_{0\leq|\alpha|\leq k} \sup_{x\in\Omega} |D^\alpha v(x)|.$$

Also, recall that $C_0^\infty(\Omega)$ is the space of continuous, infinitely differentiable functions with compact support, i.e. vanishing outside a bounded open set $\Omega' \subset \Omega$. In general, we will use the subscript 0 to indicate spaces with functions of compact support.

Lebesgue Spaces

Definition 2. *We choose $1 \leq p \leq \infty$, and define $L^p(\Omega)$ as:*

$$L^p(\Omega) = \begin{cases} \left\{ v | \int_\Omega |v|^p \, dx < \infty \right\}, & 1 \leq p < \infty \\ \left\{ v | \operatorname*{ess\,sup}_{x\in\Omega} |v(x)| \leq \infty \right\}, & p = \infty \end{cases}; \tag{2.2}$$

these spaces are also Banach spaces, with an associated norm:

$$\|v\|_{L^p(\Omega)} = \left(\int_\Omega |v|^p \, dx \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty$$
$$\|v\|_{L^\infty(\Omega)} = \operatorname*{ess\,sup}_{x\in\Omega} |v(x)|, \quad p = \infty$$

We assume here (and in the following) that $\int_\Omega$ is the Lebesgue integral. Also, in theory, v is not a function but rather an (equivalence) class of functions that differ over a set of measure zero. The essential supremum in the definitions above is the greatest lower bound C' of the set of all constants C, such that $|v(x)| \leq C$ almost everywhere on Ω.

Sobolev Spaces

Definition 3. *Choose k a non-negative integer, and $1 \leq p \leq \infty$, the Sobolev spaces $W^{k,p}(\Omega)$ are then defined:*

$$W^{k,p}(\Omega) = \begin{cases} \{ v | \ D^\alpha v \in L^p(\Omega), \ \forall \alpha : |\alpha| \leq k \}, & 1 \leq p < \infty \\ \{ v | \ D^\alpha v \in L^\infty(\Omega), \ \forall \alpha : |\alpha| \leq k \}, & p = \infty, \end{cases} \tag{2.3}$$

these spaces are Banach spaces with an associated norm:

$$\|v\|_{W^{k,p}(\Omega)} = \left(\sum_{|\alpha| \leq k} \int_{\Omega} |D^{\alpha} v|^{p} \, dx \right)^{\frac{1}{p}}, \; 1 \leq p < \infty$$
$$\|v\|_{W^{k,\infty}(\Omega)} = \max_{|\alpha| \leq k} \operatorname*{ess\,sup}_{x \in \Omega} |D^{\alpha} v(x)|, \; p = \infty.$$

The Sobolev spaces are the natural setting for the variational formulation of partial differential equations. The derivatives here should be interpreted in the proper distributional sense [39]. Choosing $k = 0$ we see that $W^{0,p}(\Omega) \equiv L^{p}(\Omega)$, and the Lebesgue spaces, are included in the Sobolev Spaces. Of particular interest in the following, is also the choice $p = 2$ which is a family of Hilbert Spaces.

Hilbert Spaces

Definition 4. *Choose k a non-negative integer, then the Hilbert Spaces $H^{k}(\Omega)$ are defined:*

$$H^{k}(\Omega) = \left\{ v | \; D^{\alpha} v \in L^{2}(\Omega), \; \forall \alpha : |\alpha| \leq k \right\}; \tag{2.4}$$

these spaces are Hilbert spaces with a norm:

$$\|v\|_{H^{k}(\Omega)} = \left(\sum_{|\alpha| \leq k} \int_{\Omega} |D^{\alpha} v|^{2} \, dx \right)^{\frac{1}{2}},$$

which is induced by the following inner product:

$$(w, v)_{H^{k}(\Omega)} = \sum_{|\alpha| \leq k} \int_{\Omega} D^{\alpha} w \cdot D^{\alpha} v \, dx.$$

The Hilbert spaces will be used extensively in the following, note that from the Lebesgue spaces only $L^{2}(\Omega)$ is a Hilbert space. Hilbert spaces is the natural generalization of Euclidean spaces in the functional setting. The fact that the norm is induced by an inner-product,

implies that the Cauchy-Schwarz inequality holds:

$$|(w,v)_{H^k(\Omega)}| \leq \|w\|_{H^k(\Omega)}\|v\|_{H^k(\Omega)}.$$

Dual Hilbert Spaces

For a general Hilbert space Z, we denote the associated inner product and induced norm by $(\cdot,\cdot)_Z$ and $\|\cdot\|_Z$ respectively; we identify the corresponding dual space Z', with norm $||\cdot||_{Z'}$ given by:

$$\|f\|_{Z'} = \sup_{v\in Z} \frac{f(v)}{\|v\|_z}.$$

The dual space Z' comprises of all the functionals $f : Z \to \mathbb{R}$ for which the norm $\|f\|_{Z'}$ is bounded. This space is also a Hilbert space and if $Z = H^k(\Omega)$ we will denote the dual (and for good reasons) $Z' = H^{-k}(\Omega)$. In general:

$$H^k(\Omega) \subset \cdots \subset H^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega) \subset \cdots \subset H^{-k}(\Omega).$$

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

$$(\rho_f^Z, v)_Z = f(v),\ \forall v \in Z.$$

It is then readily deduced that

$$\rho_f^Z = \arg\sup_{v\in Z} \frac{f(v)}{\|v\|_Z},$$

and

$$\|f\|_{Z'} = \|\rho_f^Z\|_Z,$$

which we will use repeatedly in what follows.

The duality pairing between members of Z' and Z will be denoted by ${}_{Z'}\langle\cdot,\cdot\rangle_Z$, and unless no confusion arises we will write $\langle\cdot,\cdot\rangle$.

Time-Dependent Spaces

Definition 5. *Let $T > 0$ we then define, for $1 \leq q < \infty$*

$$L^q(0,T;W^{k,p}(\Omega)) = \left\{ v : (0,T) \to W^{k,p}(\Omega) | v \text{ is measurable and } \int_0^T \|v(t)\|_{W^{k,p}(\Omega)}\, dt < \infty \right\} \tag{2.5}$$

with the norm:

$$\|v\|_{L^q(0,T;W^{k,p}(\Omega))} = \left(\int_0^T \|v(t)\|^q_{W^{k,p}(\Omega)}\, dt \right)^{\frac{1}{q}}.$$

In a similar fashion we can define $C^0([0,T];W^{k,p}(\Omega))$ and more generally the Sobolev spaces $W^{k,p}(0,T;W^{s,q}(\Omega))$; see [57] for more details.

Chapter 3

Coercive Problems

We start our presentation with the case of coercive elliptic problems. In Section 3.1, we introduce an abstract problem formulation and an illustrative instantiation for the model problem of Section 1.2. In Section 3.2 we describe, for coercive symmetric problems and "compliant" outputs, the reduced-basis approximation; and in Section 3.3 we present the associated *a posteriori* error estimation procedure. In Section 3.4 we consider the extension of our approach to noncompliant outputs and nonsymmetric operators, and finally in Section 3.5 we give some numerical results.

3.1 Problem Statement

3.1.1 Abstract Formulation

We consider a suitably regular domain $\Omega \subset \mathbb{R}^d$, $d = 1$, 2, or 3, and associated function space $H_0^1(\Omega) \subset Y \subset H^1(\Omega)$. The inner product and norm associated with Y are given by $(\cdot,\cdot)_Y$ and $\|\cdot\|_Y = (\cdot,\cdot)^{1/2}$, respectively. We also define a parameter set $\mathcal{D} \in \mathbb{R}^P$, a particular point in which will be denoted μ. Note that Ω does *not* depend on the parameter.

We then introduce a bilinear form $a\colon\ Y \times Y \times \mathcal{D} \to \mathbb{R}$, and linear forms $f\colon\ Y \to \mathbb{R}$, $\ell\colon\ Y \to \mathbb{R}$. We shall assume that a is continuous,

$$a(w, v; \mu) \le \gamma(\mu)\|w\|_Y\|v\|_Y \le \gamma_0\|w\|_Y\|v\|_Y,\ \forall \mu \in \mathcal{D}; \tag{3.1}$$

furthermore, we asume that a is coercive: there exists $\alpha(\mu) > 0$ such that

$$0 < \alpha_0 \leq \alpha(\mu) = \inf_{w \in Y} \frac{a(w, w; \mu)}{\|w\|_Y^2}, \ \forall \mu \in \mathcal{D}, \tag{3.2}$$

and symmetric, $a(w, v; \mu) = a(v, w; \mu)$, $\forall w, v \in Y^2$, $\forall \mu \in \mathcal{D}$. We also require that the linear forms f and ℓ be bounded; in Sections 3.2 and 3.3 we additionally assume a "compliant" output, $f(v) = \ell^O(v)$, $\forall v \in Y$.

We shall also make certain assumptions on the parametric dependence of a, f, and ℓ^O. In particular, we shall suppose that, for some finite (preferably small) integer Q, a may be expressed as

$$a(w, v; \mu) = \sum_{q=1}^{Q} \sigma^q(\mu) \ a^q(w, v), \ \forall w, v \in Y^2, \ \forall \mu \in \mathcal{D}, \tag{3.3}$$

for some $\sigma^q : \ \mathcal{D} \to \mathbb{R}$ and $a^q : \ Y \times Y \to \mathbb{R}$, $q = 1, \ldots, Q$. This "separability," or "affine," assumption on the parameter dependence is crucial to computational efficiency; however, certain relaxations are possible — see in [106]. For simplicity of exposition, we assume that f and ℓ^O do not depend on μ; in actual practice, affine dependence is readily admitted.

Our abstract problem statement is then: for any $\mu \in \mathcal{D}$, find $s(\mu) \in \mathbb{R}$ given by

$$s(\mu) = \ell^O(u(\mu)), \tag{3.4}$$

where $u(\mu) \in Y$ is the solution of

$$a(u(\mu), v; \mu) = f(v), \ \forall v \in Y. \tag{3.5}$$

In the language of the introduction, a is our partial differential equation (in weak form), μ is our parameter, $u(\mu)$ is our field variable, and $s(\mu)$ is our output. For simplicity of exposition, we may on occasion suppress the explicit dependence on μ.

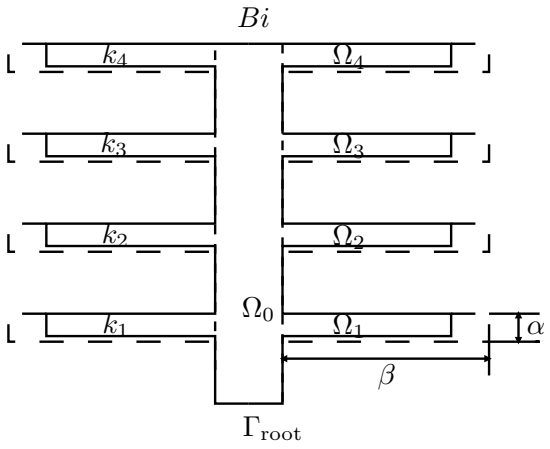

Figure 3-1: Two-Dimensional Thermal Fin.

3.1.2 Particular Instantiation

Thermal Fin

In this example we consider the two-dimensional thermal fin problem, discussed extensively in Section 1.2 — see also, Figure 3-1. The physical model is simple conduction, and the strong form of the governing equations was given in 1.2.2. The starting point for variational solution methods is the weak form: the (non-dimensional) temperature field in the fin, u, satisfies

$$\sum_{i=0}^{4} \int_{\tilde{\Omega}_i} k_i \tilde{\nabla}\tilde{u} \cdot \tilde{\nabla}\tilde{v} \; + \int_{\partial\tilde{\Omega}\setminus\Gamma_{\text{root}}} \text{Bi}\; \tilde{u}\tilde{v} \; = \int_{\Gamma_{\text{root}}} \tilde{v}, \; \forall \; \tilde{v} \in H^1(\tilde{\Omega}), \tag{3.6}$$

where $\tilde{\Omega}_i$ is that part of the domain with conductivity $\tilde{k}^i$, and $\partial\tilde{\Omega}$ denotes the boundary of $\tilde{\Omega}$. We now apply a continuous piecewise-affine transformation from $\tilde{\Omega}$ to a fixed (μ-independent) reference domain Ω (dashed and solid lines on Figure 3-1, respectively). The abstract problem statement (3.5) is then recovered. Recall that here $\mu = \{k_1,\ k_2,\ k_3,\ k_4,\ \text{Bi},\ \alpha,\ \beta\}$, and $\mu \in \mathcal{D} \subset \mathbb{R}^{P=7}$; with $k_1, \ldots, k_4$ the thermal conductivities of the "subfins" relative to the thermal conductivity of the fin base; Bi the non-dimensional form of the heat transfer coefficient; and, $\alpha,\ \beta$ the length and thickness of each of the "subfins" relative to the length of the fin root. It is readily verified that the bilinear form a is continuous, coercive, and symmetric; and that the "affine" assumption (3.3) obtains for $Q = 16$ (two-dimensional case). Note that the geometric variations are reflected, via the mapping, in the $\sigma^q(\mu)$. For our output of

interest, $s(\mu)$, we consider the non-dimensional average temperature at the root of the fin. This output may be expressed as $s(\mu) = \ell^O(u(\mu))$, where $\ell^O(v) = \int_{\Gamma_{\text{root}}} v$. It is readily shown that this output functional is bounded and also "compliant": $\ell^O(v) = f(v),\ \forall v \in Y$.

3.2 Reduced-Basis Approach

We recall that in this section, as well as in Section 3.3, we assume that a is continuous, coercive, symmetric, and affine in μ — see (3.3); and that $\ell^O(v) = f(v)$, which we denote "compliance."

3.2.1 Reduced-Basis Approximation

We first introduce a sample in parameter space, $S_N = \{\mu_1, \ldots, \mu_N\}$, where $\mu_i \in \mathcal{D}$, $i = 1, \ldots, N$; see Section 3.2.2 for a brief discussion of point distribution. We then define our Lagrangian [91] reduced-basis approximation space as $W_N = \text{span}\,\{\zeta_n \equiv u(\mu_n),\ n = 1, \ldots, N\}$, where $u(\mu_n) \in Y$ is the solution to (3.5) for $\mu = \mu_n$. In actual practice, $u(\mu_n)$ is replaced by an appropriate finite-element approximation on a suitably fine truth mesh; we shall discuss the associated computational implications in Section 3.2.3. Our reduced-basis approximation is then: for any $\mu \in \mathcal{D}$, find $s_N(\mu) = \ell(u_N(\mu))$, where $u_N(\mu) \in W_N$ is the solution of

$$a(u_N(\mu), v; \mu) \;=\; \ell(v), \quad \forall\ v \in W_N. \tag{3.7}$$

Non-Galerkin projections are also possible, they will be discussed in Chapter 5.

3.2.2 *A Priori* Convergence Theory

Optimality

We consider here the convergence rate of $u_N(\mu) \to u(\mu)$ and $s_N(\mu) \to s(\mu)$ as $N \to \infty$. To begin, it is standard to demonstrate optimality of $u_N(\mu)$ in the sense that

$$\|u(\mu) - u_N(\mu)\|_Y \le \sqrt{\frac{\gamma(\mu)}{\alpha(\mu)}} \inf_{w_N \in W_N} \|u(\mu) - w_N\|_Y. \tag{3.8}$$

(We note that, in the coercive case, stability of our ("conforming") discrete approximation is not an issue; the noncoercive case is decidedly more delicate (see Chapter 5).) Furthermore, for our compliance output,

$$\begin{aligned} s(\mu) &= s_N(\mu) + \ell(u - u_N) = s_N(\mu) + a(u, u - u_N; \mu) \\ &= s_N(\mu) + a(u - u_N, u - u_N; \mu) \end{aligned} \tag{3.9}$$

from symmetry and Galerkin orthogonality. It follows from (3.8) that

$$s(\mu) - s_N(\mu) \leq c \inf_{w_N \in W_N} \|u(\mu) - w_N\|_Y^2$$

and the error in the output converges as the square of the error in the best approximation. Also from coercivity, we notice that $s_N(\mu)$ is a lower bound for $s(\mu)$, $s_N(\mu) \leq s(\mu)$.

Best Approximation

Regarding the dependence of the error in the best approximation as a function of N, the analysis presented in [69] applies. The theory is restricted to the case in which $P = 1$, $\mathcal{D} = [0, \mu_{\max}]$ and suggests (under weak assumptions), that for $N > N_{\text{crit}}(\ln \mu_{\max})$,

$$\inf_{w_N \in W_N} \|u(\mu) - w_N(\mu)\|_X \leq c_1 \exp\left\{\frac{-(N-1)}{c_2}\right\}, \ \forall \mu \in \mathcal{D}; \tag{3.10}$$

for the precise definitions of N_{crit} and c_1, c_2, see [69]. The important thing to notice is that exponential convergence is proved, uniformly (globally) for all μ in $\mathcal{D}$, with only very weak (logarithmic) dependence on the range of the parameter ($\mu_{\max}$).

The proof exploits a parameter-space (non-polynomial) interpolant as a surrogate for the Galerkin approximation. As a result, the bound is not always "sharp": in practice, we observe many cases in which the Galerkin projection is considerably better than the associated interpolant; optimality (3.8) chooses to "illuminate" only certain points μ_n, automatically selecting a best "sub-approximation" amongst all possibilities — we thus see why reduced-basis *state-space* approximation of $s(\mu)$ via $u(\mu)$ is preferred to simple *parameter-space* interpolation of $s(\mu)$ via $(\mu_n, s(\mu_n))$ pairs. We note, however, that the logarithmic

N	$\frac{\lvert s(\mu) - s_N(\mu)\rvert}{s(\mu)}$	$\frac{\Delta_N(\mu)}{s(\mu)}$	$\eta_N(\mu)$
10	1.29×10^{-2}	8.60×10^{-2}	2.85
20	1.29×10^{-3}	9.36×10^{-3}	2.76
30	5.37×10^{-4}	4.25×10^{-3}	2.68
40	8.00×10^{-5}	5.30×10^{-4}	2.86
50	3.97×10^{-5}	2.97×10^{-4}	2.72
60	1.34×10^{-5}	1.27×10^{-4}	2.54
70	8.10×10^{-6}	7.72×10^{-5}	2.53
80	2.56×10^{-6}	2.24×10^{-5}	2.59

Table 3.1: Error, error bound, and effectivity as a function of N, at a particular representative point $\mu \in \mathcal{D}$, for the two-dimensional thermal fin problem (compliant output).

point distribution implicated by the interpolant-based arguments is *not* simply an artifact of the proof: in numerous numerical tests, the logarithmic distribution performs considerably (and in many cases, provably) better than other more obvious candidates, in particular for large ranges of the parameter.

Similar exponential behavior is observed for more general problems. Consider for example the thermal fin problem. We present in Table 3.1 the error $|s(\mu) - s_N(\mu)|/s(\mu)$ as a function of N, at a particular representative point μ in $\mathcal{D}$. The μ_n for the construction of the reduced-basis space are chosen "log-randomly" over $\mathcal{D}$: we sample from a multivariate uniform probability density on $\log(\mu)$. We observe that, the error is remarkably small even for very small N; and that, in both cases, very rapid convergence obtains as $N \to \infty$.

3.2.3 Computational Procedure

The theoretical and empirical results of Sections 3.2.1 and 3.2.2 suggest that N may, indeed, be chosen very small. We now develop off-line/on-line computational procedures that exploit this dimension reduction.

We first express $u_N(\mu)$ as

$$u_N(\mu) = \sum_{j=1}^{N} u_{N\,j}(\mu)\,\zeta_j = (\underline{u}_N(\mu))^T\,\underline{\zeta}, \tag{3.11}$$

where $\underline{u}_N(\mu) \in \mathbb{R}^N$; we then choose for test functions $v = \zeta_i$, $i = 1, \ldots, N$. Inserting these representations into (3.7) yields the desired algebraic equations for $\underline{u}_N(\mu) \in \mathbb{R}^N$,

$$\underline{A}_N(\mu)\underline{u}_N(\mu) = \underline{F}_N, \tag{3.12}$$

in terms of which the output can then be evaluated as $s_N(\mu) = \underline{F}_N^T \underline{u}_N(\mu)$. Here $\underline{A}_N(\mu) \in \mathbb{R}^{N\times N}$ is the SPD matrix with entries $A_{N\,i,j}(\mu) \equiv a(\zeta_j, \zeta_i; \mu)$, $1 \le i, j \le N$, and $\underline{F}_N \in \mathbb{R}^N$ is the "load" (and "output") vector with entries $F_{N\,i} \equiv f(\zeta_i)$, $i = 1, \ldots, N$.

We now invoke (3.3) to write

$$A_{N\,i,j}(\mu) = a(\zeta_j, \zeta_i; \mu) = \sum_{q=1}^{Q} \sigma^q(\mu) a^q(\zeta_j, \zeta_i), \tag{3.13}$$

or

$$\underline{A}_N(\mu) = \sum_{q=1}^{Q} \sigma^q(\mu) \underline{A}_N^q,$$

where the $\underline{A}_N \in \mathbb{R}^{N\times N}$ are given by $A^q_{N\,i,j} = a^q(\zeta_j, \zeta_i)$, $i \le i, j \le N$, $1 \le q \le Q$. The off-line/on-line decomposition should be clear. In the *off-line* stage, we compute the $u(\mu_n)$ and form the $\underline{A}_N^q$ and $\underline{F}_N$: this requires N (expensive) "a" finite-element solutions and $\mathcal{O}(QN^2)$ finite-element-vector inner products. In the *on-line* stage, for any given new μ, we first form $\underline{A}_N$ from (3.13), then solve (3.12) for $\underline{u}_N(\mu)$, and finally evaluate $s_N(\mu) = \underline{F}_N^T \underline{u}_N(\mu)$: this requires $\mathcal{O}(QN^2) + \mathcal{O}(\frac{2}{3}N^3)$ operations and $\mathcal{O}(QN^2)$ storage.

Thus, as required, the incremental, or marginal, cost to evaluate $s_N(\mu)$ for any given new μ — as proposed in a design, optimization, or inverse-problem context — is very small: first, because N is very small, typically $\mathcal{O}(10)$ — thanks to the good convergence properties of W_N; and second, because (3.12) can be very rapidly assembled and inverted — thanks to the off-line/on-line decomposition (see [12] for an earlier application of this strategy within the reduced-basis context). For the problems discussed in this thesis, the resulting computational savings relative to standard (well-designed) finite-element approaches are significant — at least $\mathcal{O}(10)$, typically $\mathcal{O}(100)$, and often $\mathcal{O}(1000)$ or more.

3.3 A Posteriori Error Estimation: Output Bounds

From Section 3.2 we know that, in theory, we can obtain $s_N(\mu)$ very inexpensively: the on-line computational effort scales as $\mathcal{O}(\frac{2}{3}N^3) + \mathcal{O}(QN^2)$; and N can, *in theory*, be chosen quite small. However, *in practice*, we do not know *how* small N can be chosen. Surprisingly, *a posteriori* error estimation has received relatively little attention within the reduced-basis framework [79], even though reduced-basis methods are particularly in need of accuracy assessment: the spaces are *ad hoc* and pre-asymptotic, thus admitting relatively little intuition, "rules of thumb," or standard approximation notions. Recall that, in this section, we continue to assume that a is coercive and symmetric, and that ℓ is "compliant."

The approach described in this section is a particular instance of a general "variational" framework for *a posteriori* error estimation of outputs of interest. However, the reduced-basis instantiation described here differs significantly from earlier applications to finite-element discretization error [67, 65] and iterative solution error [85] both in the choice of (energy) relaxation and in the associated computational artifice.

3.3.1 Formulation

We assume that we are given a positive function $g(\mu) : \mathcal{D} \rightarrow \mathbb{R}_+$, and a continuous, coercive, symmetric (μ-independent) bilinear form $\hat{a} : Y \times Y \rightarrow \mathbb{R}$, such that

$$c\|v\|_Y^2 \leq g(\mu)\hat{a}(v,v) \leq a(v,v;\mu), \ \forall v \in Y, \ \forall \mu \in \mathcal{D} \tag{3.14}$$

for some positive real constant c. We then find $\hat{e}(\mu) \in Y$ such that

$$g(\mu)\hat{a}(\hat{e}(\mu), v) = R(v; u_N(\mu); \mu), \ \forall v \in Y, \tag{3.15}$$

where for a given $w \in Y$, $R(v;w;\mu) = \ell(v) - a(w,v;\mu)$ is the weak form of the residual. Our lower and upper output estimators are then evaluated as

$$s_N^-(\mu) \equiv s_N(\mu), \ \text{and} \ s_N^+(\mu) \equiv s_N(\mu) + \Delta_N(\mu), \tag{3.16}$$

respectively, where

$$\Delta_N(\mu) \equiv g(\mu)\hat{a}(\hat{e}(\mu), \hat{e}(\mu)) \tag{3.17}$$

is the estimator gap.

3.3.2 Properties

We shall prove in this section that $s_N^-(\mu) \leq s(\mu) \leq s_N^+(\mu)$, and hence that $|s(\mu) - s_N(\mu)| = s(\mu) - s_N(\mu) \leq \Delta_N(\mu)$. Our lower and upper output estimators are thus lower and upper output *bounds*; and our output estimator gap is thus an output *bound* gap — a rigorous bound for the error in the output of interest. It is also critical that $\Delta_N(\mu)$ be a relatively *sharp* bound for the true error: a poor (overly large) bound will encourage us to refine an approximation which is, in fact, already adequate — with a corresponding (unnecessary) increase in off-line and on-line computational effort. We shall prove in this section that $\Delta_N(\mu) \leq \frac{\gamma_0}{c}(s(\mu) - s_N(\mu))$, where γ_0 and c are defined in (3.1) and (3.14), respectively. Our two results of this section can thus be summarized as

$$1 \leq \eta_N(\mu) \leq C, \ \forall N, \ \forall \mu \in \mathcal{D} \tag{3.18}$$

where

$$\eta_N(\mu) = \frac{\Delta_N(\mu)}{s(\mu) - s_N(\mu)} \tag{3.19}$$

is the effectivity, and C is a constant independent of N or $\mu \in \mathcal{D}$. We shall denote the left (bounding property) and right (sharpness property) inequalities of (3.18) as the lower effectivity and upper effectivity inequalities, respectively.

We first prove the lower effectivity inequality (bounding property):

Lemma 3.3.1. *For $s_N^-(\mu)$ and $s_N^+(\mu)$ defined in (3.16),*

$$s_N^-(\mu) \leq s(\mu) \leq s_N^+(\mu), \ \forall \mu \in \mathcal{D},$$

Proof. The lower bound property follows directly from the discussion in Section 3.2.2. To

prove the upper bound property, we first observe that

$$R(v; u_N; \mu) = a(u(\mu) - u_N(\mu), v; \mu) = a(e(\mu), v; \mu),$$

where $e(\mu) \equiv u(\mu) - u_N(\mu)$; we may thus rewrite (3.15) as

$$g(\mu)\hat{a}(\hat{e}(\mu), v) = a(e(\mu), v; \mu), \ \forall v \in Y.$$

We thus obtain

$$\begin{aligned} g(\mu)\hat{a}(\hat{e}, \hat{e}) &= g(\mu)\hat{a}(\hat{e} - e, \hat{e} - e) + 2g(\mu)\hat{a}(\hat{e}, e) - g(\mu)\hat{a}(e, e) \\ &= g(\mu)\hat{a}(\hat{e} - e, \hat{e} - e) + (a(e, e; \mu) - g(\mu)\hat{a}(e, e)) + a(e, e; \mu) \\ &\geq a(e, e; \mu) \end{aligned} \tag{3.20}$$

where $g(\mu)\hat{a}(\hat{e}(\mu) - e(\mu), \hat{e}(\mu) - e(\mu)) \geq 0$, and $a(e(\mu), e(\mu); \mu) - g(\mu)\hat{a}(e(\mu), e(\mu)) \geq 0$ from (3.14). Invoking (3.9) and (3.20), we then obtain

$$s(\mu) - s_N(\mu) = a(e(\mu), e(\mu); \mu) \leq g(\mu)\hat{a}(\hat{e}(\mu), \hat{e}(\mu));$$

and thus $s(\mu) \leq s_N(\mu) + g(\mu)\hat{a}(\hat{e}(\mu), \hat{e}(\mu)) \equiv s_N^+(\mu)$, as desired. □

We next prove the upper effectivity inequality (sharpness property):

Lemma 3.3.2. *For the effectivity $\eta_N(\mu)$, defined in (3.19),*

$$\eta_N(\mu) = \frac{\Delta_N(\mu)}{s(\mu) - s_N(\mu)} \leq \frac{\gamma_0}{c}, \ \forall N, \ \forall \mu \in \mathcal{D}.$$

Proof. To begin, we appeal to a-continuity and $g(\mu)\hat{a}$-coercivity to obtain

$$a(\hat{e}(\mu), \hat{e}(\mu); \mu) \leq \frac{\gamma_0 g(\mu)}{c}\hat{a}(\hat{e}(\mu), \hat{e}(\mu)). \tag{3.21}$$

But from the modified error equation (3.15) we know that

$$g(\mu)\hat{a}(\hat{e}(\mu),\hat{e}(\mu)) = R(\hat{e}(\mu);\mu) = a(e(\mu),\hat{e}(\mu);\mu).$$

Invoking the Cauchy-Schwartz inequality, we obtain

$$\begin{aligned} g(\mu)\hat{a}(\hat{e},\hat{e}) = a(e,\hat{e};\mu) &\leq (a(\hat{e},\hat{e};\mu))^{1/2}(a(e,e;\mu))^{1/2} \\ &\leq \left(\frac{\gamma_0}{c}\right)^{1/2}(g(\mu)\hat{a}(\hat{e},\hat{e}))^{1/2}(a(e,e;\mu))^{1/2}; \end{aligned}$$

the desired result then directly follows from (3.9) and (3.17). □

We now provide empirical evidence for (3.18). In particular, we present in Table 3.1 the bound gap and effectivities for the thermal fin example. Clearly $\eta_N(\mu)$ is always greater than unity for any N, and bounded — indeed, quite close to unity — as $N \to \infty$.

3.3.3 Computational Procedure

Finally, we turn to the computational artifice by which we can efficiently compute $\Delta_N(\mu)$ in the on-line stage of our procedure. We again exploit the affine parameter dependence, but now in a less transparent fashion. To begin, we rewrite the "modified" error equation, (3.15), as

$$\hat{a}(\hat{e}(\mu),v) = \frac{1}{g(\mu)}\left(\ell(v) - \sum_{q=1}^{Q}\sum_{j=1}^{N}\sigma^q(\mu)u_{N\,j}(\mu)a^q(\zeta_j,v)\right),\ \forall v \in X,$$

where we have appealed to our reduced-basis approximation (3.11) and the affine decomposition (3.3). It is immediately clear from linear superposition that we can express $\hat{e}(\mu) \in Y$ as

$$\hat{e}(\mu) = \frac{1}{g(\mu)}\left(\hat{z}_0 + \sum_{q=1}^{Q}\sum_{j=1}^{N}\sigma^q(\mu)u_{N\,j}(\mu)\hat{z}_j^q\right), \tag{3.22}$$

where $\hat{z}_0 \in Y$ satisfies $\hat{a}(\hat{z}_0,v) = \ell(v),\ \forall\ v \in Y$, and $\hat{z}_j^q \in Y, j = 1,\ldots,N,\ q = 1,\ldots,Q$, satisfies $\hat{a}(\hat{z}_j^q,v) = -a^q(\zeta_j,v),\ \forall\ v \in Y$. Inserting (3.22) into our expression for the upper

bound, $s_N^+(\mu) = s_N(\mu) + g(\mu)\hat{a}(\hat{e}(\mu), \hat{e}(\mu))$, we obtain

$$s_N^+(\mu) = s_N(\mu) + \frac{1}{g(\mu)}\left(c_0 + 2\sum_{q=1}^{Q}\sum_{j=1}^{N}\sigma^q(\mu)u_{N\,j}(\mu)\Lambda_j^q \;+\; \sum_{q=1}^{Q}\sum_{q'=1}^{Q}\sum_{j=1}^{N}\sum_{j'=1}^{N}\sigma^q(\mu)\sigma^{q'}(\mu)u_{N\,j}(\mu)u_{N\,j'}(\mu)\Gamma_{jj'}^{qq'}\right) \tag{3.23}$$

where $c_0 = \hat{a}(\hat{z}_0, \hat{z}_0)$, $\Lambda_j^q = \hat{a}(\hat{z}_0, \hat{z}_j^q)$, and $\Gamma_{jj'}^{qq'} = \hat{a}(\hat{z}_j^q, \hat{z}_{j'}^{q'})$.

The off-line/on-line decomposition should now be clear. In the *off-line* stage we compute $\hat{z}_0$ and $\hat{z}_j^q$, $j = 1, \ldots, N$, $q = 1, \ldots, Q$, and then form c_0, Λ_j^q, and $\Gamma_{jj'}^{qq'}$: this requires $QN + 1$ (expensive) "$\hat{a}$" finite element solutions, and $\mathcal{O}(Q^2N^2)$ finite-element-vector inner products. In the *on-line* stage, for any given new μ, we evaluate s_N^+ as expressed in (3.23): this requires $\mathcal{O}(Q^2N^2)$ operations and $\mathcal{O}(Q^2N^2)$ storage (for c_0, Λ_j^q, and $\Gamma_{jj'}^{qq'}$). As for the computation of $s_N(\mu)$, the marginal cost for the computation of $s_N^{\pm}(\mu)$ for any given new μ is quite small — in particular, it is *independent* of the dimension of the truth finite element approximation space Y.

There are a variety of ways in which the off-line/on-line decomposition and output error bounds can be exploited. A particularly attractive mode incorporates the error bounds into an on-line adaptive process, in which we successively approximate $s_N(\mu)$ on a sequence of approximation spaces $W_{N_j'} \subset W_N, N_j' = N_0 2^j$ — for example, $W_{N_j'}$ may contain the N_j' sample points of S_N closest to the new μ of interest — until $\Delta_{N_j'}$ is less than a specified error tolerance. This procedure both minimizes the on-line computational effort and reduces conditioning problems — while simultaneously ensuring accuracy and certainty.

The essential advantage of the approach described in this section is the guarantee of rigorous bounds. There are, however, certain disadvantages related to the choice of $g(\mu)$ and $\hat{a}$. In many cases, simple inspection suffices: for example, in our thermal fin problem of Section 3.1.2, $g(\mu) = \min_{q=1,\ldots,Q} \sigma^q(\mu)$ and $\hat{a}(w, v) = \sum_{q=1}^{Q} a^q(w, v)$ yields the very good effectivities summarized in Table 3.1. In other cases, however, there is no self-evident (or readily computed [68]) good choice. For example when $g(\mu)$ is very small, then the effectivities will be unacceptably large. The remedy in these cases, is to replace condition

(3.14) with a more general spectral condition. The development of these spectral conditions, and "bound conditioners" satisfying such conditions, is given in [113].

3.4 Noncompliant Outputs and Nonsymmetric Operators

In Sections 3.2 and 3.3 we formulated the reduced-basis method and associated error estimation procedure for the case of compliant outputs, $\ell(v) = f(v),\ \forall v \in Y$. We describe here the formulation and theory for more general linear bounded output functionals; moreover, the assumption of symmetry (but not yet coercivity) is relaxed, permitting treatment of a wider class of problems — a representative example is the convection-diffusion equation, in which the presence of the convective term renders the operator nonsymmetric. We first present the reduced-basis approximation, now involving a dual or adjoint problem; we then formulate the associated *a posteriori* error estimators; and we conclude with a few illustrative results.

As a preliminary, we first generalize the abstract formulation of Section 3.1.1. As before, we define the "primal" problem as in (3.5), however we of course no longer require symmetry. But we also introduce an associated adjoint or "dual" problem: for any $\mu \in X$, find $\psi(\mu) \in X$ such that

$$a(v, \psi(\mu); \mu) = -\ell^O(v),\ \forall v \in X; \tag{3.24}$$

recall that $\ell^O(v)$ is our output functional.

3.4.1 Reduced-Basis Approximation

To develop the reduced-basis space, we first choose — randomly or log-randomly as described in Section 3.2.2 — a sample set in parameter space, $S_{N/2} = \{\mu_1, \ldots, \mu_{N/2}\}$, where $\mu_i \in \mathcal{D},\ i = 1, \ldots, N/2$ (N even); we next define an "integrated" Lagrangian reduced-basis approximation space, $W_N = \text{span}\{(u(\mu_n), \psi(\mu_n)),\ n = 1, \ldots, N/2\}$.

For any $\mu \in \mathcal{D}$, our reduced basis approximation is then obtained by standard Galerkin projection onto W_N (though for highly nonsymmetric operators minimum residual and

Petrov-Galerkin projections are attractive — stabler — alternatives). To wit, for the primal problem, we find $u_N(\mu) \in W_N$ such that

$$a(u_N(\mu), v; \mu) = f(v), \ \forall v \in W_N;$$

and for the adjoint problem, we define $\psi_N(\mu) \in W_N$ such that

$$a(v, \psi_N(\mu); \mu) = -\ell^O(v), \ \forall v \in W_N.$$

The reduced-basis output approximation is then calculated from $s_N(\mu) = \ell^O(u_N(\mu))$.

Turning now to the *a priori* theory, it follows from standard arguments that $u_N(\mu)$ and $\psi_N(\mu)$ are "optimal" in the sense that

$$\begin{aligned} \|u(\mu) - u_N(\mu)\|_Y &\leq \left(1 + \frac{\gamma(\mu)}{\alpha(\mu)}\right) \inf_{w_N \in W_N} \|u(\mu) - w_N\|_Y, \\ \|\psi(\mu) - \psi_N(\mu)\|_Y &\leq \left(1 + \frac{\gamma(\mu)}{\alpha(\mu)}\right) \inf_{w_N \in W_N} \|\psi(\mu) - w_N\|_Y. \end{aligned}$$

The best approximation analysis is then similar to that presented in Section 3.2.2. As regards our output, we now have

$$\begin{aligned} |s(\mu) - s_N(\mu)| &= |\ell^O(u(\mu)) - \ell^O(u_N(\mu))| = |a(u - u_N, \psi; \mu)| \\ &= |a(u - u_N, \psi - \psi_N; \mu)| \\ &\leq \gamma_0 \|u - u_N\|_X \|\psi - \psi_N\|_Y \end{aligned} \tag{3.25}$$

from Galerkin orthogonality, the definition of the primal and the adjoint problems, and the Cauchy-Schwartz inequality. We now understand why we include the $\psi(\mu_n)$ in W_N: to ensure that $\|\psi(\mu) - \psi_N(\mu)\|_Y$ is small. We thus recover the "square" effect in the convergence rate of the output, albeit (and unlike the symmetric case) at the expense of some additional computational effort — the inclusion of the $\psi(\mu_n)$ in W_N; typically, even for the very rapidly convergent reduced-basis approximation, the "fixed error-minimum cost" criterion favors the adjoint enrichment.

For simplicity of exposition (and to a certain extent, implementation), we present here

the "integrated" primal-dual approximation space. However, there are significant computational and conditioning advantages associated with a "non-integrated" approach, in which we introduce *separate* primal ($u(\mu_n)$) and dual ($\psi(\mu_n)$) approximation spaces for $u(\mu)$ and $\psi(\mu)$, respectively. Note in the "non-integrated" case we are obliged to compute $\psi_N(\mu)$, since to preserve the output error "square effect" we must modify our predictor with a residual correction, $f(\psi_N(\mu)) - a(u_N(\mu), \psi_N(\mu); \mu)$ — see the next chapters for more details. Both the "integrated" and "non-integrated" approaches admit an off-line/on-line decomposition similar to that described in Section 3.2.3 for the compliant, symmetric problem; as before, the on-line complexity and storage are independent of the dimension of the very fine ("truth") finite element approximation.

3.4.2 Method I *A Posteriori* Error Estimators

We extend here the method developed in Section 3.3.2 to the more general case of noncompliant and nonsymmetric problems. We begin with the formulation.

We first find $\hat{e}^{\rm pr}(\mu) \in Y$ such that

$$g(\mu)\hat{a}(\hat{e}^{\rm pr}(\mu), v) = R^{\rm pr}(v; u_N(\mu); \mu), \ \forall\ v \in Y,$$

where $R^{\rm pr}(v; w; \mu) \equiv f(v) - a(w, v; \mu), \ \forall v \in X$; and $\hat{e}^{\rm du}(\mu) \in Y$ such that

$$g(\mu)\hat{a}(\hat{e}^{\rm du}(\mu), v) = R^{\rm du}(v; \psi_N(\mu); \mu), \ \forall\ v \in Y,$$

where $R^{\rm du}(v; w; \mu) \equiv -\ell(v) - a(v, w; \mu), \ \forall v \in Y$. We then define

$$\bar{s}_N(\mu) = s_N(\mu) - \frac{g(\mu)}{2}\hat{a}(\hat{e}^{\rm pr}(\mu), \hat{e}^{\rm du}(\mu)), \ \text{and} \tag{3.26}$$

$$\Delta_N(\mu) = \frac{g(\mu)}{2}\left[\hat{a}(\hat{e}^{\rm pr}(\mu), \hat{e}^{\rm pr}(\mu))\right]^{\frac{1}{2}}\left[\hat{a}(\hat{e}^{\rm du}(\mu), \hat{e}^{\rm du}(\mu))\right]^{\frac{1}{2}}. \tag{3.27}$$

Finally, we evaluate our lower and upper estimators as

$$s_N^{\pm}(\mu) = \bar{s}_N(\mu) \pm \Delta_N(\mu). \tag{3.28}$$

Note that, as before, $g(\mu)$ and $\hat{a}$ still satisfy (3.14); and that, furthermore, (3.14) will only involve the *symmetric* part of a. We define the effectivity as

$$\eta_N(\mu) \;=\; \frac{\Delta_N(\mu)}{|s(\mu) - s_N(\mu)|}; \tag{3.29}$$

note that $s(\mu) - s_N(\mu)$ now has no definite sign.

We now prove that our error estimators are bounds (the lower effectivity inequality):

Proposition 1. *For $s_N^-(\mu)$ and $s_N^+(\mu)$ defined in (3.28) then*

$$s_N^-(\mu) \le s(\mu) \le s_N^+(\mu), \;\; \forall N, \;\; \forall \mu \in \mathcal{D}.$$

Proof. To begin, we define $\hat{e}^{\pm}(\mu) = \hat{e}^{\rm pr}(\mu) \mp \frac{1}{\kappa}\hat{e}^{\rm du}(\mu)$, and note that, from the coercivity of $\hat{a}$,

$$\kappa g(\mu)\hat{a}(e^{\rm pr} - \frac{1}{2}\hat{e}^{\pm}, e^{\rm pr} - \frac{1}{2}\hat{e}^{\pm}) = \kappa g(\mu)\hat{a}(e^{\rm pr}, e^{\rm pr}) + \frac{\kappa g(\mu)}{4}\hat{a}(\hat{e}^{\pm}, \hat{e}^{\pm}) - \kappa g(\mu)\hat{a}(\hat{e}^{\pm}, e^{\rm pr}) \ge 0, \tag{3.30}$$

where $e^{\rm pr}(\mu) = u(\mu) - u_N(\mu)$, $e^{\rm du}(\mu) = \psi(\mu) - \psi_N(\mu)$, and κ is a positive real number. From the definition of $\hat{e}^{\pm}(\mu)$ and $\hat{e}^{\rm pr}(\mu)$, $\hat{e}^{\rm du}(\mu)$, we can express the "cross-term" as

$$\begin{aligned} g(\mu)\hat{a}(\hat{e}^{\pm}, e^{\rm pr}) &= R^{\rm pr}(e^{\rm pr}; u_N; \mu) \mp \frac{1}{\kappa} R^{\rm du}(e^{\rm pr}; \psi_N; \mu) \\ &= a(e^{\rm pr}, e^{\rm pr}; \mu) \mp \frac{1}{\kappa} a(e^{\rm pr}, e^{\rm du}; \mu) \\ &= a(e^{\rm pr}, e^{\rm pr}; \mu) \pm \frac{1}{\kappa}(s(\mu) - s_N(\mu)), \end{aligned} \tag{3.31}$$

since

$$\begin{aligned} R^{\rm pr}(e^{\rm pr}; u_N; \mu) &= a(u, e^{\rm pr}; \mu) - a(u_N, e^{\rm pr}; \mu) = a(e^{\rm pr}, e^{\rm pr}; \mu), \\ R^{\rm du}(e^{\rm pr}; \psi_N; \mu) &= a(e^{\rm pr}, \psi; \mu) - a(e^{\rm pr}, \psi_N; \mu) = a(e^{\rm pr}, e^{\rm du}; \mu), \end{aligned}$$

and

$$\begin{aligned}\ell^O(\mu) - \ell^O(u_N) &= -a(u - u_N, \psi; \mu) \\ &= -a(u - u_N, \psi - \psi_N; \mu) \quad \text{(using Galerkin orthogonality)} \\ &= -a(e^{\text{pr}}, e^{\text{du}}; \mu).\end{aligned}$$

We then substitute (3.31) into (3.30) to obtain

$$\begin{aligned}\pm(s(\mu) - s_N(\mu)) &\le -\kappa \left(a(e^{\text{pr}}, e^{\text{pr}}; \mu) - g(\mu)\hat{a}(e^{\text{pr}}, e^{\text{pr}})\right) + \frac{\kappa g(\mu)}{4}\hat{a}(\hat{e}^{\pm}, \hat{e}^{\pm}) \\ &\le \frac{\kappa g(\mu)}{4}\hat{a}(\hat{e}^{\pm}, \hat{e}^{\pm}),\end{aligned}$$

since $\kappa > 0$ and $a(e^{\text{pr}}(\mu), e^{\text{pr}}(\mu); \mu) - g(\mu)\hat{a}(e^{\text{pr}}(\mu), e^{\text{pr}}(\mu)) \ge 0$ from (3.14).
Expanding $\hat{e}^{\pm}(\mu) = \hat{e}^{\text{pr}}(\mu) \mp \frac{1}{\kappa}\hat{e}^{\text{du}}(\mu)$ then gives

$$\pm(s(\mu) - s_N(\mu)) \le \frac{g(\mu)}{4}\left[\kappa\hat{a}(\hat{e}^{\text{pr}}, \hat{e}^{\text{pr}}) + \frac{1}{\kappa}\hat{a}(\hat{e}^{\text{du}}, \hat{e}^{\text{du}}) \mp 2\hat{a}(\hat{e}^{\text{pr}}, \hat{e}^{\text{du}})\right],$$

or

$$\pm\left(s(\mu) - (s_N(\mu) - \frac{g(\mu)}{2}\hat{a}(\hat{e}^{\text{pr}}, \hat{e}^{\text{du}}))\right) \le \frac{\kappa g(\mu)}{4}\hat{a}(\hat{e}^{\text{pr}}, \hat{e}^{\text{pr}}) + \frac{g(\mu)}{4\kappa}\hat{a}(\hat{e}^{\text{du}}, \hat{e}^{\text{du}}). \tag{3.32}$$

We now choose $\kappa(\mu)$ as

$$\kappa(\mu) = \left(\frac{\hat{a}(\hat{e}^{\text{du}}(\mu), \hat{e}^{\text{du}}(\mu))}{\hat{a}(\hat{e}^{\text{pr}}(\mu), \hat{e}^{\text{pr}}(\mu))}\right)^{\frac{1}{2}}$$

so as to minimize the right-hand side (3.32); we then obtain

$$|s(\mu) - \bar{s}_N(\mu)| \le \Delta_N(\mu), \tag{3.33}$$

and hence $s_N^-(\mu) \le s(\mu) \le s_N^+(\mu)$. □

We now turn to the upper effectivity inequality (sharpness property). If the primal and dual errors are a-orthogonal, or become increasingly orthogonal as N increases, then the effectivity will not, in fact, be bounded as $N \to \infty$. However, if we make the (plausible)

hypothesis that $|s(\mu)-s_N(\mu)| \geq \underline{C}\|e^{\rm pr}(\mu)\|_Y\|e^{\rm du}(\mu)\|_Y$, then it is simple to demonstrate that

$$\eta_N(\mu) \leq \frac{\gamma_0^2}{2\underline{C}c}. \tag{3.34}$$

In particular, it is an easy matter to demonstrate that

$$g^{1/2}(\mu)\left(\hat{a}\left(\hat{e}^{\rm pr}(\mu),\hat{e}^{\rm pr}(\mu)\right)\right)^{1/2} \leq \frac{\gamma_0}{c^{1/2}}\|e^{\rm pr}(\mu)\|_Y$$

(note we lose a factor of $\gamma_0^{1/2}$ relative to the symmetric case); similarly,

$$g^{1/2}(\mu)\left(\hat{a}\left(\hat{e}^{\rm du}(\mu),\hat{e}^{\rm du}(\mu)\right)\right)^{1/2} \leq \frac{\gamma_0}{c^{1/2}}\|e^{\rm du}(\mu)\|_Y.$$

The desired result then directly follows from the definition of $\Delta_N(\mu)$ and our hypothesis on $|s(\mu) - s_N(\mu)|$.

3.4.3 Blackbox Method

Finally, turning to computational issues, we note that the off-line/on-line decomposition described in Sections 3.2.3 and 3.3.3 for compliant symmetric problems directly extends to the noncompliant, nonsymmetric case — except that we must compute the norm of both the primal and dual "modified errors," with a concomitant doubling of computational effort. The details of the blackbox technique follow. For convenience we define $\mathcal{N}$ as the set $\{1,\ldots,N\}$, and $\mathcal{Q}$ as the set $\{1,\ldots,Q\}$.

Off-line Stage

1. Calculate $u(\mu_i)$ and $\psi(\mu_i), i = 1,\ldots,N/2$, to form W_N.
2. Compute $\underline{A}^q \in \mathbb{R}^{N\times N}$ as $A^q_{ij} = a^q(\zeta_j,\zeta_i), \forall i,j \in \mathcal{N}^2$ and $\forall q \in \mathcal{Q}$.
3. Solve for $\hat{z}^{0\,\rm pr} \in Y$ and $\hat{z}^{0\,\rm du} \in Y$ from $\hat{a}(\hat{z}^{0\,\rm pr},v) = f(v),\ \forall v \in Y$, and $\hat{a}(\hat{z}^{0\,\rm du},v) = -\ell^O(v),\ \forall v \in Y$, respectively. Also, compute $\hat{z}^q_j \in Y$ from $\hat{a}(\hat{z}^q_j,v) = -a^q(\zeta_j,v),\ \forall v \in Y$, $\forall j \in \mathcal{N}$ and $\forall q \in \mathcal{Q}$.
4. Calculate and store $c_0^{\rm pr} = \hat{a}(\hat{z}^{0\,\rm pr},\hat{z}^{0\,\rm pr})$; $c_0^{\rm du} = \hat{a}(\hat{z}^{0\,\rm du},\hat{z}^{0\,\rm du})$; $c_0^{\rm prdu} = \hat{a}(\hat{z}^{0\,\rm pr},\hat{z}^{0\,\rm du})$; $F^{\rm pr}_{N,j} =$

$f(\zeta_j)$ and $F^{\rm du}_{N,j} = \ell^O(\zeta_j)$, $\forall j \in \mathcal{N}$; $\Lambda^{q\,\rm pr}_j = \hat{a}(\hat{z}^{0\,\rm pr}, \hat{z}^q_j)$ and $\Lambda^{q\,\rm du}_j = \hat{a}(\hat{z}^{0\,\rm du}, \hat{z}^q_j)$, $\forall j \in \mathcal{N}$ and $\forall q \in \mathcal{Q}$; $\Gamma^{pq}_{ij} = \hat{a}(\hat{z}^p_i, \hat{z}^q_j)$, $\forall i, j \in \mathcal{N}^2$ and $\forall p, q \in \mathcal{Q}^2$.

This stage requires $(NQ+N+2)$ Y-linear system solves; $(N^2Q^2+2NQ+3)$ $\hat{a}$-inner products; and $2N$ evaluations of linear functionals.

On-line Stage

For each new desired design point $\mu \in \mathcal{D}$ we then compute the reduced-basis prediction and error bound based on the quantities computed in the off-line stage.

1. Form $\underline{A}_N = \sum_{q=1}^Q \sigma^q(\mu)\underline{A}^q$ and solve for $\underline{u}_N \equiv \underline{u}_N(\mu) \in \mathbb{R}^N$ and $\underline{\psi}_N \equiv \underline{\psi}_N(\mu) \in \mathbb{R}^N$ from $\underline{A}_N\ \underline{u}_N = \underline{F}^{\rm pr}_N$ and $\underline{A}^T_N\ \underline{\psi}_N = -\underline{F}^{\rm du}_N$, respectively.

2. Evaluate the bound average and bound gap as

$$\bar{s}_N = (\underline{F}^{\rm du}_N)^T \underline{u}_N - \frac{1}{2g(\mu)}\Big(\sum_{i=1}^N\sum_{j=1}^N\sum_{p=1}^Q\sum_{q=1}^Q u_{N\,i}\psi_{N\,j}\sigma^p(\mu)\sigma^q(\mu)\Gamma^{pq}_{ij} + \sum_{j=1}^N\sum_{q=1}^Q \psi_{N\,j}\sigma^q(\mu)\Lambda^{q\,\rm pr}_j + \sum_{j=1}^N\sum_{q=1}^Q u_{N\,j}\sigma^q(\mu)\Lambda^{q\,\rm du}_j + c^{\rm prdu}_0\Big),$$

and

$$\Delta_N(\mu) = \frac{1}{2\ g(\mu)} \times \Big(\sum_{i=1}^N\sum_{j=1}^N\sum_{p=1}^Q\sum_{q=1}^Q u_{N\,i}u_{N\,j}\sigma^p(\mu)\sigma^q(\mu)\Gamma^{pq}_{ij} + 2\sum_{j=1}^N\sum_{q=1}^Q u_{N\,j}\sigma^q(\mu)\Lambda^{q\,\rm pr}_j + c^{\rm pr}_0\Big)^{\frac{1}{2}} \times \Big(\sum_{i=1}^N\sum_{j=1}^N\sum_{p=1}^Q\sum_{q=1}^Q \psi_{N\,i}\psi_{N\,j}\sigma^p(\mu)\sigma^q(\mu)\Gamma^{pq}_{ij} + 2\sum_{j=1}^N\sum_{q=1}^Q \psi_{N\,j}\sigma^q(\mu)\Lambda^{q\,\rm du}_j + c^{\rm du}_0\Big)^{\frac{1}{2}}.$$

respectively.

For each μ, $\mathcal{O}(N^2Q^2 + N^3)$ operations are required to obtain the reduced-basis solution and the bounds. Since $\dim(W_N) \ll \dim(Y)$, the cost to compute $s_N(\mu)$, $\bar{s}_N(\mu)$, and $\Delta_N(\mu)$ in the on-line stage will typically be much less than the cost to directly evaluate $u(\mu)$ and $s(\mu) = \ell^O(u(\mu))$.

3.5 Numerical Results

We presented in Table 3.1 for the thermal fin example, the behavior of the relative error, the bound gap, and the effectivities as a function of N. We see that even for small N, the accuracy is very good; furthermore, convergence with N is quite rapid. This is particularly noteworthy given the high-dimensional parameter space; even with $N = 50$ points we have less than two points (effectively) in each parameter coordinate. We also note that the effectivity remains roughly constant with increasing N: the estimators are not only bounds, but relatively sharp bounds — good predictors when N is "large enough." The behavior we observe at this particular value of μ is representative of most points in (a random sample over) $\mathcal{D}$, however there can certainly be points where the effectivity is larger.

3.5.1 Thermal fin — Shape optimization

We conclude this Section with a more practical application: suppose we wish to find the configuration which yields a base (e.g., chip) temperature of s_* (say 1.8) to within $\epsilon = .01$ by varying only the height α of the radiators. To start, we choose a relatively large number of basis functions in the design space $\mathcal{D}$ defined above, and perform the off-line stage of the blackbox method. For efficiency in the on-line stage, we then enlist only a subset of these basis functions — those which are closer in the design space to the desired evaluation point — and refine when higher accuracy is required. A binary chop algorithm, summarized 3-2, is implemented to effect the coupled approximation-optimization; we assume monotonicity for simplicity of exposition.

In the particular test case shown in Table 3.2, we begin with $N = 10$ points and set $N^+ = 10$ as well; we initialize $\alpha_l = 0.1$ and $\alpha_r = 0.5$. During the optimization process, refinement is effected twice, such that a total of $N = 30$ basis functions are invoked (considerably less than the 50 available). The savings are significant, yet we are still ensured, thanks to the bounds, that our design requirement is met to the desired tolerance of $\epsilon = .01$. One can also apply a dynamic adaptation strategy in which only a minimal number of basis functions are generated (initially) in the off-line stage: if these prove inadequate, we return to the off-line stage for additional basis functions and also revision of the necessary matrices and inner

for i = 1:max_iterations **do**
  Choose $\overline{\alpha} := (\alpha_l + \alpha_r)/2$
  Blackbox for $\overline{\alpha} \Rightarrow s_N^+,\ s_N^-$
  $d_1 := \max\left(|s_* - s_N^+|, |s_* - s_N^-|\right)$
  $d_2 := \min\left(|s_* - s_N^+|, |s_* - s_N^-|\right)$
  if $d_2 > \epsilon$ **then**
    if $s_N^+ > s_*$ and $s_N^- > s_*$ **then**
      $\alpha_l := \overline{\alpha}$
    else if $s_N^+ < s_*$ and $s_N^- < s_*$ **then**
      $\alpha_r := \overline{\alpha}$
    else
      $N := N + N^+$
    end if
  end if
  if $d_1 < \epsilon$ **then**
    Stop.
  else
    $N := N + N^+$
  end if
end for

Figure 3-2: Optimization Algorithm

products.

i	$\bar{\alpha}$	s_N^+	s_N^-	α_l	α_r
1	0.3	1.683	1.753	0.1	0.5
2	0.2	1.716	2.056	0.1	0.3
3	0.2	1.766	1.807	0.1	0.3
4	0.2	1.771	1.778	0.1	0.3
5	0.15	1.817	1.840	0.1	0.2
6	0.175	1.792	1.806	.15	0.2

Table 3.2: Shape Optimization

If we choose a tighter tolerance ϵ, or if we wish to investigate many different set points s_*, or if we perform the optimization permitting all 7 design parameters to vary, we would of course greatly increase the number of output predictions required — and hence greatly increase the efficiency of the reduced-basis blackbox technique relative to conventional approaches.

Chapter 4

Parabolic Problems

4.1 Introduction

In this Chapter, we consider the extension of reduced-basis output bound methods, to problems described by *parabolic* partial differential equations. The essential new ingredient in the parabolic case is the presence of time in the formulation and solution of the problem. For the parametrization of the problem, time is considered as an additional parameter, albeit a special one as we will see in the development to follow. For the numerical solution of the problem the finite-element method is employed for the spatial discretization. For the temporal discretization the discontinuous Galerkin method [42, 109] is used; although not the only choice, the variational origin of the discontinuous Galerkin is desirable for the development and proof of the bounding properties. A procedure to efficiently calculate upper and lower estimators to the outputs of interest is developed. We prove that these estimators are bounds to the exact value for the output. These bounds can be calculated efficiently by assuming an (often-satisfied) form for the partial differential operator [66].

4.2 Problem Statement

To start, consider a bounded open domain $\Omega \subset \mathbb{R}^d,\ d = 1, 2, 3$ with Lipschitz-continuous boundary; if $T > 0$ is the final time and $I = (0, T)$ ($\bar{I} = [0, T]$) the time interval of interest,

we define the "space-time" domain $Q_T = I \times \Omega$. Furthermore, let V be a closed linear subspace of $H^1(\Omega)$, such that $H_0^1(\Omega) \subset V \subset H^1(\Omega)$. The space $L^2(I;V)$ can be defined as in Section 2.2. Similarly, we define $C^0(\bar{I};L^2(\Omega))$ the set of functions which are continuous (and therefore bounded) in time, and $L^2(\Omega)$ in space for $t \in \bar{I}$; also, we will use in the following $L^2(Q_T) \equiv L^2(I;L^2(\Omega))$, and $\mathcal{H} \equiv L^2(I;V) \cap C^0(\bar{I};L^2(\Omega))$ [57, 97]. For the parametric dependence, let P be the number of input parameters and $\mathcal{D} \subset \mathbb{R}^P$ the set of allowed configurations; a particular configuration will be denoted by $\mu \in \mathcal{D}$.

Let $f(\cdot;\mu) \in L^2(Q_T)$ and $u_0(\mu) \in L^2(\Omega)$ be known functions which depend on the parameter μ. The problem we are interested in solving is: given a $\mu \in \mathcal{D}$, find the solution $u(\cdot;\mu) \in \mathcal{H}$ to the equation:

$$\begin{aligned} (\partial_t u(t;\mu), v) + a(u(t;\mu), v;\mu) &= (f(t;\mu), v), \ \forall v \in V, \\ u(0;\mu) &= u_0(\mu); \end{aligned} \tag{4.1}$$

here $(\cdot,\cdot)$ denotes the $L^2(\Omega)$-inner product and $a(\cdot,\cdot;\mu) : V \times V \to \mathbb{R}$ is a continuous and coercive-in-V bilinear form, uniformly in $\mu \in \mathcal{D}$. Equation (4.1) has to be understood in the proper distributional sense for $t \in I$. Under the assumptions above the problem is parabolic and a unique solution $u(\cdot;\mu) \in \mathcal{H}$ exists for all $\mu \in \mathcal{D}$ [97]. We should also mention that a solution to (4.1) exists under weaker assumptions than the ones presented above (e.g. $f(\cdot;\mu) \in L^2(I;V')$, with V' the dual of V) — this generality is not required for our presentation. Also to keep the notation minimal, we assume that the $L^2(\Omega)$-inner product and the bilinear form $a(\cdot,\cdot;\mu)$ do not depend on time.

As was mentioned in Section 1.2, in practical applications the solution field $u(\cdot;\mu)$ is less important than relevant outputs of interest. We consider here the output of interest which is obtained from $s(\mu) \equiv \mathcal{S}(u(\cdot;\mu))$, with $\mathcal{S} : \mathcal{H} \to \mathbb{R}$ a linear functional

$$\mathcal{S}(v) = \int_I \left(\ell^O(t), v(t)\right) \, dt + (g^O, v(T^-));$$

with $v(t^\pm) = \lim_{s \to 0^+} v(t \pm s)$. Here $\ell^O(\cdot) \in L^2(Q_T)$ (or more generally, $\ell^O \in L^2(I;V')$) and $g^O \in L^2(\Omega)$ do not depend on μ — a parametric dependence of the output can be readily

treated.

It will be useful in the following to replace (4.1), with a space-time weak formulation: given $\mu \in \mathcal{D}$, find $u(\cdot;\mu) \in \mathcal{H}$ such that

$$\int_I (\partial_t u(t;\mu), v(t))\ dt + \int_I a(u(t;\mu), v(t);\mu)\, dt + (u(0^+;\mu), v(0^+)) = \int_I (f(t;\mu), v(t))\ dt + (u_0(\mu), v(0^+)), \quad (4.2)$$

$\forall v \in \mathcal{H}$. It is obvious that if $u(\cdot;\mu)$ is the solution of (4.1) then it is also a solution of (4.2). We can readily prove the following:

Lemma 4.2.1. *The problem in (4.2) is stable, and therefore $u(\cdot;\mu) \in \mathcal{H}$ is the unique weak solution to (4.2).*

Proof. Stability and therefore uniqueness, follows from the coercivity of the bilinear form $a(\cdot,\cdot;\mu)$,

$$\exists c > 0 \text{ such that } c\|v\|_{H^1(\Omega)} \le a(v,v;\mu), \forall v \in V, \forall \mu \in \mathcal{D};$$

which implies that

$$\begin{aligned} &\int_I (\partial_t v(t), v(t))\ dt + \int_I a(v(t), v(t);\mu)\, dt + (v(0^+), v(0^+)) \quad = \\ &\frac{1}{2}(v(T^-), v(T^-)) + \frac{1}{2}(v(0^+), v(0^+)) + \int_I a(v(t), v(t);\mu)\, dt \quad \ge \quad \underline{c}\|v\|^2_{L^2(I;H^1)},\ \forall v \in \mathcal{H},\ v \neq 0. \end{aligned}$$

□

We will also require in the following $\psi(\cdot;\mu) \in \mathcal{H}$ which is the solution of the following dual problem:

$$-\int_I (\partial_t \psi(t;\mu), v(t))\ dt + \int_I a(v(t), \psi(t;\mu);\mu)\, dt + (\psi(T^-;\mu), v(T^-)) = -\int_I \left(\ell^O(t), v(t)\right)\ dt - (g^O, v(T^-)),\ \forall v \in \mathcal{H}; \quad (4.3)$$

the importance of the dual problem will become clear in the analysis that follows. Notice that if we define $\tau = T - t$, (4.3) becomes parabolic — the dual problem evolves *backward*

in time. Therefore, under the requirements above for the primal problem a unique weak solution $\psi(\cdot;\mu)$ to (4.3) will exist.

In practice, for the solution of (4.2) and (4.3), we replace V by a finite but *high-dimensional* finite-element space V_h, so that $V_h \approx V$ ($\dim V_h = \mathcal{N}$). Given an input configuration μ, solution of the resulting system of ordinary differential equations (and relatedly, calculation of the output of interest), can be very expensive. We develop in the next section, a reduced-basis approach to significantly reduce the complexity of this problem.

4.3 Reduced-basis Approximation

We define $\tilde{\mu} = (t,\mu) \in \tilde{\mathcal{D}} \equiv I \times \mathcal{D}$, and introduce the following sample sets $S_N^{\text{pr}} = \{\tilde{\mu}_1^{\text{pr}},\ldots,\tilde{\mu}_N^{\text{pr}}\}$ and $S_M^{\text{du}} = \{\tilde{\mu}_1^{\text{du}},\ldots,\tilde{\mu}_M^{\text{du}}\}$. In general, $N \neq M$ and $\tilde{\mu}_i^{\text{pr}} \neq \tilde{\mu}_j^{\text{du}}$, $i = 1,\ldots,N$, $j = 1,\ldots,M$. We then compute the solution of (4.2) for all $\{\mu \in I \mid \exists t : (t,\mu) \in S_N^{\text{pr}}\}$, and of (4.3) for all $\left\{\mu \in I \;\middle|\; \exists t : (t,\mu) \in S_M^{\text{du}}\right\}$. Using these solutions we define the Lagrangian reduced-basis approximation spaces, as follows:

$$W_N^{\text{pr}} = \text{span}\{\zeta_i \equiv u(\tilde{\mu}_i^{\text{pr}}),\ i = 1,\ldots,N\},\ W_M^{\text{du}} = \text{span}\{\xi_i \equiv \psi(\tilde{\mu}_i^{\text{du}}),\ i = 1,\ldots,M\},$$

where $\dim W_N^{\text{pr}} = N$, and $\dim W_M^{\text{du}} = M$; by construction $W_N^{\text{pr}},\ W_M^{\text{du}} \subset V$. We can then define the following spaces,

$$\mathcal{H}_N^{\text{pr}} \equiv L^2(I;W_N^{\text{pr}}) \cap C^0(\bar{I};L^2(\Omega)),\ \text{and}\ \mathcal{H}_M^{\text{du}} \equiv L^2(I;W_M^{\text{du}}) \cap C^0(\bar{I};L^2(\Omega)).$$

In the construction of the reduced-basis spaces the choice of $\mu_i \in \mathcal{D}$ (and consequently $\tilde{\mu}_i$) for the sample sets S_N^{pr} and S_M^{du} is critical. Both the *a priori* theory [69] (in the context of elliptic problems) and extensive numerical tests [94] suggest that the points should be chosen "log-randomly" over $\mathcal{D}$: we sample from a multivariate uniform probability density on $\log(\mathcal{D})$. Especially for large ranges of the input parameters, this logarithmic distribution performs considerably better than other obvious candidates.

The reduced-basis approximation $u_N(\cdot;\mu)$ to $u(\cdot;\mu)$ is obtained by a standard Galerkin

projection: given a $\mu \in \mathcal{D}$, find $u_N(\cdot;\mu) \in \mathcal{H}_N^{\rm pr}$, such that

$$\int_I (\partial_t u_N(t;\mu), v(t))\ dt + \int_I a(u_N(t;\mu), v(t);\mu)\, dt + (u_N(0^+;\mu), v(0^+)) = \int_I (f(t;\mu), v(t))\ dt + (u_0(\mu), v(0^+)), \forall v \in \mathcal{H}_N^{\rm pr}. \quad (4.4)$$

The error to the approximation of $u(\cdot;\mu)$ by $u_N(\cdot;\mu)$ is $e^{\rm pr}(t;\mu) \equiv u(t;\mu) - u_N(t;\mu)$, and relatedly $\mathcal{R}^{\rm pr}(v;\mu)$ is the residual for the primal problem:

$$\begin{aligned}\mathcal{R}^{\rm pr}(v;\mu) &= \int_I (f(t;\mu), v(t))\ dt - \int_I (\partial_t u_N(t;\mu), v(t))\ dt \\ &\quad - \int_I a(u_N(t;\mu), v(t);\mu)\, dt - \big(u_N(0^+;\mu) - u_0(\mu), v(0^+)\big) \\ &= \int_I (\partial_t e^{\rm pr}(t;\mu), v(t))\ dt + \int_I a(e^{\rm pr}(t;\mu), v(t);\mu)\, dt + \big(e^{\rm pr}(0^+;\mu), v(0^+)\big)\,;\end{aligned} \quad (4.5)$$

the last line above follows from (4.2). Similarly, for the dual variable, we obtain an approximation $\psi_M(\cdot;\mu) \in \mathcal{H}_M^{\rm du}$ to $\psi(\cdot;\mu) \in \mathcal{H}$ from:

$$-\int_I (\partial_t \psi_M(t;\mu), v(t))\ dt + \int_I a(v(t), \psi_M(t;\mu);\mu)\, dt + (\psi_M(T^-;\mu), v(T^-)) = -\int_I \big(\ell^O(t), v(t)\big)\ dt - (g^O, v(T^-)), \forall v \in \mathcal{H}_M^{\rm du}. \quad (4.6)$$

The residual for the dual problem $\mathcal{R}^{\rm du}(v;\mu)$ is then:

$$\begin{aligned}\mathcal{R}^{\rm du}(v;\mu) &= -\int_I \big(\ell^O(t), v(t)\big)\ dt + \int_I (\partial_t \psi_M(t;\mu), v(t))\ dt \\ &\quad - \int_I a(v(t), \psi_M(t;\mu);\mu)\, dt - \big(\psi_M(T^-;\mu) + g^O, v(T^-)\big) \\ &= -\int_I \big(\partial_t e^{\rm du}(t;\mu), v(t)\big)\ dt + \int_I a(v(t), e^{\rm du}(t;\mu);\mu)\, dt + \big(e^{\rm du}(T^-;\mu), v(T^-)\big)\,;\end{aligned} \quad (4.7)$$

from (4.3) and defining $e^{\rm du}(t;\mu) = \psi(t;\mu) - \psi_M(t;\mu)$.

Using now the reduced-basis solutions to the primal and dual problems, we can obtain

an approximation to the output of interest $s_N(\mu)$ from:

$$\begin{aligned} s_N(\mu) &\equiv \mathcal{S}(u_N(\cdot;\mu)) - \mathcal{R}^{\text{pr}}(\psi_M(\cdot;\mu);\mu) \\ &= \int_I \left(\ell^O(t), u_N(t;\mu)\right)\, dt + (g^O, u_N(T^-;\mu)) - \mathcal{R}^{\text{pr}}(\psi_M(\cdot;\mu);\mu). \end{aligned} \tag{4.8}$$

Regarding the convergence of the output approximation (4.8), we have the following:

Lemma 4.3.1. *Let*

$$\begin{aligned} \varepsilon_M^{\text{du}} \quad = \quad \inf_{\chi_M \in \mathcal{H}_M^{\text{du}}} \Big\{ &\left[\|e^{\text{pr}}\|_{L^\infty(I;L^2)} + \|e^{\text{pr}}\|_{L^2(I;H^1)}\right] \times \left[\|\psi - \chi_M\|_{L^\infty(I;L^2)} + \|\psi - \chi_M\|_{L^2(I;H^1)}\right] \\ &+ \|e^{\text{pr}}\|_{L^2(I;L^2)} \|\psi - \chi_M\|_{H^1(I;L^2)} \Big\}, \end{aligned}$$

then

$$|s(\mu) - s_N(\mu)| \le C \left[\|e^{\text{pr}}\|_{L^\infty(I;L^2)} + \|e^{\text{pr}}\|_{L^2(I;H^1)}\right] \times \left[\|e^{\text{du}}\|_{L^\infty(I;L^2)} + \|e^{\text{du}}\|_{L^2(I;H^1)}\right] + C\varepsilon_M^{\text{du}}. \tag{4.9}$$

Proof. We start with an auxiliary result that will also be required below,

$$\begin{aligned} s(\mu) - s_N(\mu) &= \int_I \left(\ell^O, u\right)\, dt + (g^O, u(T^-)) \\ &\quad - \int_I \left(\ell^O, u_N\right)\, dt - (g^O, u_N(T^-)) + \mathcal{R}^{\text{pr}}(\psi_M;\mu) \\ &= \int_I \left(\ell^O, e^{\text{pr}}\right)\, dt + (g^O, e^{\text{pr}}(T^-)) + \mathcal{R}^{\text{pr}}(\psi_M;\mu) \\ &= \int_I (\partial_t \psi, e^{\text{pr}})\, dt - \int_I a(e^{\text{pr}}, \psi;\mu)\, dt - (\psi(T^-), e^{\text{pr}}(T^-)) + \mathcal{R}^{\text{pr}}(\psi_M;\mu) \\ &= -\int_I (\partial_t e^{\text{pr}}, \psi)\, dt - \int_I a(e^{\text{pr}}, \psi;\mu)\, dt - (e^{\text{pr}}(0^+), \psi(0^+)) + \mathcal{R}^{\text{pr}}(\psi_M;\mu) \\ &= -\mathcal{R}^{\text{pr}}(e^{\text{du}};\mu); \end{aligned} \tag{4.10}$$

using (4.3), integration by parts, (4.5) and linearity of the primal residual. From (4.10)

$$|s(\mu) - s_N(\mu)| = \left| \int_I \left(\partial_t e^{\text{pr}}, e^{\text{du}}\right)\, dt + \int_I a(e^{\text{pr}}, e^{\text{du}};\mu)\, dt + (e^{\text{pr}}(0^+), e^{\text{du}}(0^+)) \right|; \tag{4.11}$$

we look at each of the terms on the right-hand side separately. Let $\chi_M(\cdot) \in \mathcal{H}_M^{\mathrm{du}}$, with $\chi_M(0^+) = \psi_M(0^+)$. Then

$$\begin{aligned}\left|\int_I \left(\partial_t e^{\mathrm{pr}}, e^{\mathrm{du}}\right)\, dt\right| &= \left|\int_I (\partial_t e^{\mathrm{pr}}, \psi - \chi_M + \chi_M - \psi_M)\, dt\right| \\ &\leq \left|\int_I (\partial_t e^{\mathrm{pr}}, \psi - \chi_M)\, dt\right| + \left|\int_I (\partial_t e^{\mathrm{pr}}, \psi_M - \chi_M)\, dt\right|.\end{aligned}$$

For the first term above we use integration by parts to get:

$$\left|\int_I (\partial_t e^{\mathrm{pr}}, \psi - \chi_M)\, dt\right| \leq C\left[\|e^{\mathrm{pr}}\|_{L^\infty(I;L^2)}\|\psi - \chi_M\|_{L^\infty(I;L^2)} + \|e^{\mathrm{pr}}\|_{L^2(I;L^2)}\|\psi - \chi_M\|_{H^1(I;L^2)}\right];$$

and for the second term from (4.5) and using the Galerkin orthogonality property (since $\psi_M(t) - \chi_M(t) \in W_M^{\mathrm{du}}$), we get:

$$\begin{aligned}\left|\int_I (\partial_t e^{\mathrm{pr}}, \psi_M - \chi_M)\, dt\right| &= \left|\int_I a(e^{\mathrm{pr}}, \psi_M - \chi_M; \mu)\, dt + (e^{\mathrm{pr}}(0^+), \underbrace{\psi_M(0^+) - \chi_M(0^+)}_{=0})\right| \\ &\leq \gamma\|e^{\mathrm{pr}}\|_{L^2(I;H^1)}\|\psi_M - \psi + \psi - \chi_M\|_{L^2(I;H^1)} \\ &\leq \gamma\|e^{\mathrm{pr}}\|_{L^2(I;H^1)}\left(\|\psi_M - \psi\|_{L^2(I;H^1)} + \|\psi - \chi_M\|_{L^2(I;H^1)}\right),\end{aligned}$$

with γ the continuity constant of $a(\cdot,\cdot;\mu)$. Combining the expressions above:

$$\left|\int_I \left(\partial_t e^{\mathrm{pr}}, e^{\mathrm{du}}\right)\, dt\right| \leq \gamma\|e^{\mathrm{pr}}\|_{L^2(I;H^1)}\|e^{\mathrm{du}}\|_{L^2(I;H^1)} + C\varepsilon_M^{\mathrm{du}}. \tag{4.12}$$

The second and third terms in (4.11) can be bounded using the continuity of the bilinear form a and the Cauchy-Schwartz inequality, giving

$$|s(\mu) - s_N(\mu)| \leq \left|\int_I \left(\partial_t e^{\mathrm{pr}}, e^{\mathrm{du}}\right)\, dt\right| + \gamma\|e^{\mathrm{pr}}\|_{L^2(I;H^1)}\|e^{\mathrm{du}}\|_{L^2(I;H^1)} + \|e^{\mathrm{pr}}\|_{L^\infty(I;L^2)}\|e^{\mathrm{du}}\|_{L^\infty(I;L^2)}. \tag{4.13}$$

The desired result follows directly from (4.12) and (4.13). □

The previous lemma gives an *a priori* bound on the convergence of the output approximation, defined in (4.8), to its exact value; as we see from (4.9), a term appears involving $\varepsilon_M^{\mathrm{du}}$

— a measure of how well members of the reduced-basis space $\mathcal{H}^{\rm du}_M$ approximate the solution to the adjoint problem — as well as norms of the error to the dual problem $e^{\rm du}$. Had we used $\mathcal{S}(u_N(\cdot;\mu))$ instead of (4.8) to calculate the output approximation, the corresponding bound would depend on norms of the primal error $e^{\rm pr}$ only. As M increases, the term involving the dual errors will become smaller and, given the approximation properties of $W^{\rm du}_M$, will converge to zero; this suggests faster convergence of the adjoint-corrected output and use of (4.8) is justified.

In the calculation above we have, in effect, replaced V (or V_h) with $W^{\rm pr}_N$ for the primal and $W^{\rm du}_M$ for the dual problem. These reduced-basis spaces have approximation properties specific to the problem of interest, so only a small number of basis functions need to be retained to accurately represent the solution. Significant computational savings are affected, since the computational complexity scales as $N(=\dim W^{\rm pr}_N)$ and $M(=\dim W^{\rm du}_M)$ instead of $\mathcal{N}(=\dim V_h)$, and $N,\ M$ will be small — typically $O(10)$ — and independent of $\mathcal{N}$. As $N,\ M\to\infty$, and given the specific choice of the approximation spaces, $u_N(\cdot;\mu)\to u(\cdot;\mu)$, $\psi_M(\cdot;\mu)\to\psi(\cdot;\mu)$, and $s_N(\mu)\to s(\mu)$ will converge to the exact values very fast.

4.4 A posteriori error estimation

The computational relaxation introduced in the previous section, allows us to compute very efficiently accurate approximations to the solution and the output of interest. Thanks to the expected rapid convergence N and M could, in theory, be chosen quite small. However, in practice we do not know how small N and M can be: this will depend of the desired accuracy, the choice of $\tilde{\mu}_i$ in the construction of the reduced-basis spaces, the output of interest and the particular problem in question; in some cases $N,\ M=5$ may suffice, while in other cases $N,\ M=100$ may still be insufficient. In the face of this uncertainty, either too many or too few basis functions will be retained: the former results in computational inefficiency; the later in unacceptable uncertainty. It is therefore critical that we can ascertain the accuracy of our predictions; we develop next, a rigorous error-estimation approach, directly for outputs of interest, to *a posteriori* validate the accuracy of our predictions.

To begin assume that we may find a function $g(\mu):\mathcal{D}\to\mathbb{R}_+$, and a symmetric, continuous

and coercive bilinear form $\hat{a} : V \times V \to \mathbb{R}$ such that

$$c\|v\|_1 \leq g(\mu)\hat{a}(v,v) \leq a(v,v;\mu),\ \forall v \in V,\ \forall \mu \in \mathcal{D}; \tag{4.14}$$

we understand $g(\mu)$ as a lower bound to the $\hat{a}$-coercivity constant.

We then compute the "reconstructed" errors $\hat{e}^{\rm pr}(\cdot;\mu) \in \mathcal{H}$ and $\hat{e}^{\rm du}(\cdot;\mu) \in \mathcal{H}$ such that

$$\begin{aligned} g(\mu)\int_I \hat{a}(\hat{e}^{\rm pr}(t;\mu), v(t))\,dt &= \mathcal{R}^{\rm pr}(v;\mu),\ \forall v \in \mathcal{H},\ \text{and} \\ g(\mu)\int_I \hat{a}(\hat{e}^{\rm du}(t;\mu), v(t))\,dt &= \mathcal{R}^{\rm du}(v;\mu),\ \forall v \in \mathcal{H}. \end{aligned} \tag{4.15}$$

Note that a unique solution exists for problems (4.15), by an application of the Riesz-Frechet representation theorem since $\mathcal{R}^{\rm pr}$ and $\mathcal{R}^{\rm du}$ are continuous linear functionals on the Hilbert space $L^2(I;V')$ and $\int_I \hat{a}(\cdot,\cdot)\,dt$ is a scalar product in $L^2(I;V)$. An estimate for the output is then computed, $s_B(\mu)$:

$$s_B(\mu) = s_N(\mu) - \frac{g(\mu)}{2}\int_I \hat{a}(\hat{e}^{\rm pr}(t;\mu), \hat{e}^{\rm du}(t;\mu))\,dt; \tag{4.16}$$

and a bound gap $\Delta(\mu)$:

$$\Delta(\mu) = \frac{g(\mu)}{2}\left[\int_I \hat{a}(\hat{e}^{\rm pr}(t;\mu), \hat{e}^{\rm pr}(t;\mu))\,dt\right]^{\frac{1}{2}}\left[\int_I \hat{a}(\hat{e}^{\rm du}(t;\mu), \hat{e}^{\rm du}(t;\mu))\,dt\right]^{\frac{1}{2}}. \tag{4.17}$$

Finally, upper and lower output estimators can be calculated from $s^{\pm}(\mu) = s_B(\mu) \pm \Delta(\mu)$. We now prove that these estimators $s^{\pm}(\mu)$ are always rigorous *bounds* to the true output $s(\mu)$. In the proof that follows, unless it is essential, we will not explicitly indicate dependence on the variables t and μ.

Proposition 2. *Let $s_B(\mu)$ be the output approximation, defined in (4.16), and $\Delta(\mu)$ the bound gap, defined in (4.17). If we then define $s^{\pm}(\mu) = s_B(\mu) \pm \Delta(\mu)$ then*

$$s^-(\mu) \leq s(\mu) \leq s^+(\mu),\ \forall \mu \in \mathcal{D};$$

that is, $s^+(\mu)$ and $s^-(\mu)$ are rigorous *upper and lower bounds to the true output $s(\mu)$.*

Proof. To start, notice that $\mathcal{R}^{\rm pr}(e^{\rm du};\mu) = \mathcal{R}^{\rm du}(e^{\rm pr};\mu)$ since:

$$\begin{aligned}
\mathcal{R}^{\rm pr}(e^{\rm du};\mu) &= \int_I \left(\partial_t e^{\rm pr}, e^{\rm du}\right)\, dt + \int_I a(e^{\rm pr}, e^{\rm du};\mu)\, dt + \left(e^{\rm pr}(0^+), e^{\rm du}(0^+)\right) \\
&= -\int_I \left(\partial_t e^{\rm du}, e^{\rm pr}\right)\, dt + \int_I a(e^{\rm pr}, e^{\rm du};\mu)\, dt + \left(e^{\rm pr}(T^-), e^{\rm du}(T^-)\right) \\
&= \mathcal{R}^{\rm du}(e^{\rm pr};\mu);
\end{aligned}$$

using integration by parts, and the definition of the primal (4.5) and dual (4.7) residuals. Therefore from (4.10),

$$-\mathcal{R}^{\rm du}(e^{\rm pr};\mu) = s(\mu) - s_N(\mu). \tag{4.18}$$

We can now start the proof of the bounding property, and define $\hat{e}^{\pm} = \hat{e}^{\rm pr} \mp \frac{1}{\kappa}\hat{e}^{\rm du}$, with $\kappa > 0$. From the coercivity of $\hat{a}$, we have:

$$\begin{aligned}
\kappa g(\mu) \int_I \hat{a}(e^{\rm pr} - \frac{1}{2}\hat{e}^{\pm}, e^{\rm pr} - \frac{1}{2}\hat{e}^{\pm}) = \\
\kappa g(\mu) \int_I \hat{a}(e^{\rm pr}, e^{\rm pr}) + \kappa\frac{g(\mu)}{4} \int_I \hat{a}(\hat{e}^{\pm}, \hat{e}^{\pm}) - \kappa g(\mu) \int_I \hat{a}(\hat{e}^{\pm}, e^{\rm pr}) \geq 0.
\end{aligned} \tag{4.19}$$

Since $\hat{e}^{\pm} = \hat{e}^{\rm pr} \mp \frac{1}{\kappa}\hat{e}^{\rm du}$, and using (4.15) we get:

$$g(\mu) \int_I \hat{a}(\hat{e}^{\pm}, e^{\rm pr})\, dt = \mathcal{R}^{\rm pr}(e^{\rm pr};\mu) \mp \frac{1}{\kappa}\mathcal{R}^{\rm du}(e^{\rm pr};\mu). \tag{4.20}$$

But:

$$\begin{aligned}
\mathcal{R}^{\rm pr}(e^{\rm pr};\mu) &= \int_I (\partial_t e^{\rm pr}, e^{\rm pr})\, dt + \int_I a(e^{\rm pr}, e^{\rm pr};\mu)\, dt + \left(e^{\rm pr}(0^+), e^{\rm pr}(0^+)\right) \\
&\geq \frac{1}{2}\underbrace{(e^{\rm pr}(T^-), e^{\rm pr}(T^-))}_{>0} + \frac{1}{2}\underbrace{(e^{\rm pr}(0^+), e^{\rm pr}(0^+))}_{>0} + \int_I a(e^{\rm pr}, e^{\rm pr};\mu)\, dt \\
&\geq g(\mu) \int_I \hat{a}(e^{\rm pr}, e^{\rm pr})\, dt,
\end{aligned}$$

since from (4.14) we have $\int_I a(e^{\rm pr}, e^{\rm pr};\mu)\, dt \geq g(\mu)\int_I \hat{a}(e^{\rm pr}, e^{\rm pr})\, dt$. Replacing in (4.20) for

$\mathcal{R}^{\rm pr}(e^{\rm pr};\mu)$ the expression we just obtained, and (4.18) for $\mathcal{R}^{\rm du}(e^{\rm pr};\mu)$, we have:

$$-\kappa g(\mu)\int_I \hat{a}(\hat{e}^{\pm},e^{\rm pr})\,dt \le -\kappa g(\mu)\int_I \hat{a}(e^{\rm pr},e^{\rm pr})\,dt \mp (s(\mu)-s_N(\mu)). \tag{4.21}$$

Combining now (4.19) and (4.21), we get

$$\pm(s(\mu)-s_N(\mu)) \le \frac{\kappa g(\mu)}{4}\int_I \hat{a}(\hat{e}^{\pm},\hat{e}^{\pm})\,dt.$$

Expanding $\hat{e}^{\pm} = \hat{e}^{\rm pr} \mp \frac{1}{\kappa}\hat{e}^{\rm du}$ we have

$$\pm(s(\mu)-s_N(\mu)) \le \frac{g(\mu)}{4}\left[\kappa\int_I \hat{a}(\hat{e}^{\rm pr},\hat{e}^{\rm pr})\,dt + \frac{1}{\kappa}\int_I \hat{a}(\hat{e}^{\rm du},\hat{e}^{\rm du})\,dt \mp 2\int_I \hat{a}(\hat{e}^{\rm pr},\hat{e}^{\rm du})\,dt\right]$$

and from the definition of $s_B(\mu) = s_N(\mu) - \frac{g(\mu)}{2}\int_I \hat{a}(\hat{e}^{\rm pr},\hat{e}^{\rm du})\,dt$,

$$\pm(s(\mu)-s_B(\mu)) \le \frac{\kappa g(\mu)}{4}\int_I \hat{a}(\hat{e}^{\rm pr},\hat{e}^{\rm pr})\,dt + \frac{g(\mu)}{4\kappa}\int_I \hat{a}(\hat{e}^{\rm du},\hat{e}^{\rm du})\,dt. \tag{4.22}$$

Since κ is an arbitrary positive constant, we choose it as:

$$\kappa = \left(\frac{\int_I \hat{a}(\hat{e}^{\rm du},\hat{e}^{\rm du})\,dt}{\int_I \hat{a}(\hat{e}^{\rm pr},\hat{e}^{\rm pr})\,dt}\right)^{\frac{1}{2}},$$

so that the right-hand side in (4.22) is minimized. Then

$$\pm(s(\mu)-s_B(\mu)) \le \frac{g(\mu)}{2}\left[\int_I \hat{a}(\hat{e}^{\rm pr},\hat{e}^{\rm pr})\,dt\right]^{\frac{1}{2}}\left[\int_I \hat{a}(\hat{e}^{\rm du},\hat{e}^{\rm du})\,dt\right]^{\frac{1}{2}};$$

which from the definition of $\Delta(\mu)$ becomes $\pm(s(\mu)-s_B(\mu)) \le \Delta(\mu)$, or

$$s^-(\mu) \equiv s_B(\mu) - \Delta(\mu) \le s(\mu) \le s_B(\mu) + \Delta(\mu) \equiv s^+(\mu).$$

□

So following the previous proposition, instead of using the exact value for the output $s(\mu)$, we can use the output prediction $s_B(\mu)$ and the bound gap $\Delta(\mu)$. The basic premise

is that these two quantities can be computed more efficiently than the exact output. This is indeed the case when a certain decomposition exists for all the parameter-dependent linear and bilinear forms [61]. More specifically, assume that for $t \in I,\ \mu \in \mathcal{D}$ and for $Q_{a,\ f,\ u} \in \mathbb{N}$ the following "affine" decomposition exists:

$$a(w, v; \mu) = \sum_{q=1}^{Q_a} \sigma_a^q(\mu) a^q(w, v),\ \forall w, v \in V^2,\ f(t; \mu) = \sum_{q=1}^{Q_f} \sigma_f^q(t; \mu) f^q,\ u_0(\mu) = \sum_{q=1}^{Q_u} \sigma_u^q(\mu) u_0^q; \tag{4.23}$$

with $\sigma^q_{a,\ f,\ u}$ functions which depend on μ and t, whereas the $a^q,\ f^q$, and u_0^q *do not*. For a large class of problems such a decomposition exists; certain relaxations are possible for locally non-affine problems [94].

Using (4.23) and following the same steps as in [66], a two-stage computational procedure can be developed: *Off-line* the reduced-basis space is formed and a database with certain auxiliary quantities is created; this is a relatively expensive preprocessing step which needs to be performed only once. *On-line*, for each new μ, using the database: the reduced-basis problem is formed and solved; the reduced-basis solution is used to compute the output approximation; and finally, the output bounds are calculated. The incremental cost for each on-line step is minimal and scales only with the dimension $N,\ M$ of the reduced-basis spaces and the parametric complexity $Q_{a,\ f,\ u}$ of the linear and bilinear forms.

The definition of the reduced-basis spaces comprising of snapshots to the solution at different parameter points is not the only possibility. An alternative approach is to construct the reduced-basis spaces by using the entire time-dependent solution at certain parameter values. A space-time Galerkin projection can then be used to obtain the reduced-basis problems. Moreover the *a posteriori* error estimator, defined above, could be easily adapted to this case. There are certain advantages in this alternative approach; for example, instead of solving the low-dimensional parabolic problems (4.4) and (4.6), one has to solve linear systems of small dimension. Also, there is some simplification in the computation of the error estimator. On the other hand, during the preprocessing/off-line stage the computational cost and required memory storage become much higher, making overall this second approach less attractive.

4.5 Time Discretization — Discontinuous Galerkin Method

In the previous section we presented the general theory without any reference to the time-discretization procedure. Here we consider one possible time-discretization method, the discontinuous Galerkin method. The discontinuous Galerkin method was first introduced in the context of time-dependent problems by Jamet [42], and was further analyzed [71, 103]. The variational origin of the discontinuous Galerkin method, will allow us to extend the *a posteriori* error estimation method developed in the previous section for the discrete-in-time approximation.

Consider a set of $L+1$ points in $\bar{I} = [0,T]$ such that $t_0 \equiv 0 < t_1 < t_2 < \ldots < t_L \equiv T$ is a partition $\mathcal{I}$ of I in intervals $I_l = (t_{l-1}, t_l),\ l \in \mathcal{L} \equiv \{1,\ldots,L\}$. The diameter for each I_l will be $\Delta\tau^l = t_l - t_{l-1},\ l \in \mathcal{L}$. We then define the spaces $\mathbb{P}^q(I_l;V) = \{v : I_l \to V |\ v(t) = \sum_{s=0}^q v_s t^s,\ v_s \in V\} \subset L^2(I_l;V) \cap C^0(\bar{I}_l; L^2(\Omega)), \forall l \in \mathcal{L}$, and $V^q(\mathcal{I};V) = \{v \in L^2(I;V) \mid v|_{I_l} \in \mathbb{P}^q(I_l;V),\ \forall I_l \in \mathcal{I}\}$. Obviously, if $v \in V^q(\mathcal{I};V)$ then the function can be discontinuous at the points $t_l,\ l \in \mathcal{L}$. We further define the jump at these points as $[v]_l = v(t_l^+) - v(t_l^-),\ l \in \{0,\ldots,L\}$, with $v(t_l^\pm) = \lim_{s\to 0^+} v(t_l \pm s)$. The problem is then to compute using the discontinuous Galerkin method a solution $u^q(\cdot;\mu) \in V^q(\mathcal{I};V)$ — which is a *discontinuous* approximation to $u(\cdot;\mu)$ of (4.2) — from:

$$\int_I (\partial_t u^q(t;\mu), v(t))\ dt + \int_I a(u^q(t;\mu), v(t);\mu)\ dt + \sum_{l\in\mathcal{L}} ([u^q(\cdot;\mu)]_{l-1}, v(t_{l-1}^+)) = \int_I (f(t;\mu), v(t))\ dt, \quad (4.24)$$

$\forall v \in V^q(\mathcal{I};V)$; with $[u^q(\cdot;\mu)]_0 = u^q(0^+;\mu) - u_0(\mu)$ (or $u^q(0^-;\mu) = u_0(\mu)$). In (4.24) we can solve separately for each I_l; continuity is imposed only weakly due to the presence of the additional jump terms. For the dual problem, we can compute a solution $\psi^q(\cdot;\mu) \in V^q(\mathcal{I};V)$

from:

$$-\int_I (\partial_t \psi^q(t;\mu), v(t))\ dt + \int_I a(v(t), \psi^q(t;\mu);\mu)\, dt - \sum_{l\in\mathcal{L}} ([\psi^q(\cdot;\mu)]_l, v(t_l^-)) = \\ -\int_I \left(\ell^O(t), v(t)\right)\ dt, \quad (4.25)$$

$\forall v \in V^q(\mathcal{I}; V)$; with $[\psi^q(\cdot;\mu)]_L = -g^O - \psi^q(T^-;\mu)$ (or $\psi^q(T^+;\mu) = -g^O$). The output of interest $s^q(\mu)$ can then be calculated using $u^q(\cdot;\mu)$, from:

$$s^q(\mu) = \int_I \left(\ell^O(t), u^q(t;\mu)\right)\ dt + (g^O, u^q(T^-;\mu)) = \sum_{l\in\mathcal{L}} \int_{I_l} \left(\ell^O(t), u^q(t;\mu)\right)\ dt + (g^O, u^q(T^-;\mu)). \quad (4.26)$$

The reduced-basis spaces are formed similarly to the continuous case, by obtaining "snapshots"of the solution to the primal and dual problems for all points in the sets $S_N^{\rm pr}$ and $S_M^{\rm du}$ respectively:

$$W_N^{\rm pr} = \text{span}\{\zeta_i \equiv u^q(\tilde{\mu}_i^{\rm pr}), i = 1, \ldots, N,\ \tilde{\mu}_i^{\rm pr} \in S_N^{\rm pr}\},$$
$$W_M^{\rm du} = \text{span}\{\xi_i \equiv \psi^q(\tilde{\mu}_i^{\rm du}), i = 1, \ldots, M,\ \tilde{\mu}_i^{\rm du} \in S_M^{\rm du}\}.$$

The reduced-basis approximation to $u^q(t;\mu)$ can be obtained by a standard Galerkin projection: for a given $\mu \in \mathcal{D}$, find $u_N^q(\cdot;\mu) \in V^q(\mathcal{I}; W_N^{\rm pr})$, such that

$$\int_I (\partial_t u_N^q(t;\mu), v(t))\ dt + \int_I a(u_N^q(t;\mu), v(t);\mu) dt + ([u_N^q(\cdot;\mu)]_{l-1}, v(t_{l-1}^+)) = \int_I (f(t;\mu), v(t))\ dt,$$

$\forall v \in V^q(\mathcal{I}; W_N^{\rm pr})$ with $[u_N^q(\cdot;\mu)]_0 = u_N^q(0^+;\mu) - u_0(\mu)$; similarly, we define the dual problem and obtain $\psi_M^q(\cdot;\mu) \in V^q(\mathcal{I}; W_M^{\rm du})$. The primal and dual residuals are defined as: $\mathcal{R}^{{\rm pr}\,q}(v;\mu) = \sum_{l\in\mathcal{L}} \mathcal{R}_l^{{\rm pr}\,q}(v;\mu)$ with $\mathcal{R}_l^{{\rm pr}\,q}(v;\mu)$ the residual for the primal problem in the

time interval I_l:

$$\begin{aligned}
\mathcal{R}_l^{\mathrm{pr}\,q}(v;\mu) &= \int_{I_l} (f(t;\mu), v(t))\,dt - \int_{I_l} (\partial_t u_N^q(t;\mu), v(t))\,dt \\
&\quad - \int_{I_l} a(u_N^q(t;\mu), v(t);\mu)dt - ([u_N^q(\cdot;\mu)]_{l-1}, v(t_{l-1}^+)) \\
&= \int_{I_l} (\partial_t e^{\mathrm{pr}\,q}(t;\mu), v(t))\ dt + \int_{I_l} a(e^{\mathrm{pr}\,q}(t;\mu), v(t);\mu)\,dt + ([e^{\mathrm{pr}\,q}(\cdot;\mu)]_{l-1}, v(t_{l-1}^+));
\end{aligned}$$

where $e^{\mathrm{pr}\,q}(t;\mu) \equiv u^q(t;\mu) - u_N^q(t;\mu)$, the error in the primal variable. The residual for the dual problem $\mathcal{R}_l^{\mathrm{du}\,q}(v;\mu)$ is defined as:

$$\begin{aligned}
\mathcal{R}_l^{\mathrm{du}\,q}(v;\mu) &= -\int_{I_l} \left(\ell^O(t), v(t)\right) dt + \int_{I_l} (\partial_t \psi_M^q(t;\mu), v(t))\,dt \\
&\quad - \int_{il} a(v(t), \psi_M^q(t;\mu);\mu)dt + ([\psi_M^q(\cdot;\mu)]_l, v(t_l^-)) \\
&= -\int_{I_l} \left(\partial_t e^{\mathrm{du}\,q}(t;\mu), v(t)\right)\ dt + \int_{I_l} a(v(t), e^{\mathrm{du}\,q}(t;\mu);\mu)\,dt - ([e^{\mathrm{du}\,q}(\cdot;\mu)]_l, v(t_l^-));
\end{aligned}$$

from (4.25) and defining $e^{\mathrm{du}\,q}(t;\mu) = \psi^q(t;\mu) - \psi_M^q(t;\mu)$, the error in the dual variable. An approximation to the output of interest $s_N^q(\mu)$ can then be obtained from:

$$s_N^q(\mu) = \sum_{l\in\mathcal{L}} \left[\int_{I_l} \left(\ell^O(t), u_N^q(t;\mu)\right)\ dt - \mathcal{R}_l^{\mathrm{pr}\,q}(\psi_M^q(\cdot;\mu);\mu)\right] + (g^O, u_N^q(T^-;\mu)). \tag{4.27}$$

Turning now to the *a posteriori* error estimator, we compute "representations" of the error $\hat{e}^{\mathrm{pr}\,q}(\cdot;\mu) \in V^q(\mathcal{I};V)$ with $\hat{e}_l^{\mathrm{pr}\,q}(\cdot;\mu) \equiv \hat{e}^{\mathrm{pr}\,q}(\cdot;\mu)|_{I_l}$, and $\hat{e}^{\mathrm{du}\,q}(\cdot;\mu) \in V^q(\mathcal{I};V)$ with $\hat{e}_l^{\mathrm{du}\,q}(\cdot;\mu) \equiv \hat{e}^{\mathrm{du}\,q}(\cdot;\mu)|_{I_l}$ such that:

$$\begin{aligned}
g(\mu)\int_{I_l} \hat{a}(\hat{e}_l^{\mathrm{pr}\,q}(t;\mu), v(t))\,dt &= \mathcal{R}_l^{\mathrm{pr}\,q}(v;\mu),\ \forall v \in \mathbb{P}^q(I_l;V),\ \forall I_l \in \mathcal{I} \text{ and} \\
g(\mu)\int_{I_l} \hat{a}(\hat{e}_l^{\mathrm{du}\,q}(t;\mu), v(t))\,dt &= \mathcal{R}_l^{\mathrm{du}\,q}(v;\mu),\ \forall v \in \mathbb{P}^q(I_l;V),\ \forall I_l \in \mathcal{I}.
\end{aligned} \tag{4.28}$$

For the error estimator we first calculate the output approximation, $s_B^q(\mu)$:

$$s_B^q(\mu) = s_N^q(\mu) - \frac{g(\mu)}{2}\sum_{l\in\mathcal{L}} \int_{I_l} \hat{a}(\hat{e}_l^{\mathrm{pr}\,q}(t;\mu), \hat{e}_l^{\mathrm{du}\,q}(t;\mu))\,dt; \tag{4.29}$$

and the bound gap $\Delta^q(\mu)$ is defined as:

$$\Delta^q(\mu) = \frac{g(\mu)}{2}\left[\sum_{l\in\mathcal{L}}\int_{I_l}\hat{a}(\hat{e}_l^{\mathrm{pr}\,q}(t;\mu),\hat{e}_l^{\mathrm{pr}\,q}(t;\mu))\,dt\right]^{\frac{1}{2}}\left[\sum_{l\in\mathcal{L}}\int_{I_l}\hat{a}(\hat{e}_l^{\mathrm{du}\,q}(t;\mu),\hat{e}_l^{\mathrm{du}\,q}(t;\mu))\,dt\right]^{\frac{1}{2}}. \tag{4.30}$$

Finally, as before, symmetric upper and lower output estimators can be calculated from $s^{\pm\,q}(\mu) = s_B^q(\mu) \pm \Delta^q(\mu)$. We can then prove the following:

Proposition 3. *Let $s^q(\mu)$ be the exact value of the output for the semi-discrete problem, defined in (4.26). If we define $s_B^q(\mu)$ and $\Delta^q(\mu)$ as in (4.29) and (4.30), respectively, then $s^{\pm\,q}(\mu) = s_B^q(\mu) \pm \Delta^q(\mu)$ are upper and lower bounds to the true output:*

$$s^{-\,q}(\mu) \le s^q(\mu) \le s^{+\,q}(\mu), \forall \mu \in \mathcal{D}.$$

Proof. We first obtain some results for $\sum_{l\in\mathcal{L}}\mathcal{R}_l^{\mathrm{pr}\,q}(e^{\mathrm{pr}\,q};\mu)$ and $\sum_{l\in\mathcal{L}}\mathcal{R}_l^{\mathrm{du}\,q}(e^{\mathrm{pr}\,q};\mu)$ that will be required in the following. First we look in the error for the output, which from the definition of $s^q(\mu)$ (4.26) and $s_N^q(\mu)$ (4.27) becomes:

$$s^q(\mu) - s_N^q(\mu) = \sum_{l\in\mathcal{L}}\left[-\int_{I_l}(\partial_t e^{\mathrm{pr}\,q},\psi^q)\;dt - \int_{I_l}a(e^{\mathrm{pr}\,q},\psi^q;\mu)\,dt + \mathcal{R}_l^{\mathrm{pr}\,q}(\psi_M^q(\cdot;\mu);\mu)\right] + I_1;$$

using (4.25) and integration by parts. The additional terms I_1 can be simplified, as follows:

$$\begin{aligned}
I_1 = &\sum_{l\in\mathcal{L}}\left[([\psi^q]_l, e^{\mathrm{pr}\,q}(t_l^-)) + (\psi^q(t_l^-), e^{\mathrm{pr}\,q}(t_l^-)) - (\psi^q(t_{l-1}^+), e^{\mathrm{pr}\,q}(t_{l-1}^+))\right] + (g^O(\mu), e^{\mathrm{pr}\,q}(T^-))\\
= &\sum_{l\in\mathcal{L}}\left[(\psi^q(t_l^+), e^{\mathrm{pr}\,q}(t_l^-)) - (\psi^q(t_{l-1}^+), e^{\mathrm{pr}\,q}(t_{l-1}^-)) - (\psi^q(t_{l-1}^+), [e^{\mathrm{pr}\,q}]_{l-1})\right]\\
&+(g^O(\mu), e^{\mathrm{pr}\,q}(T^-))\\
= &\,(g^O(\mu), e^{\mathrm{pr}\,q}(T^-)) + (\underbrace{\psi^q(T^+)}_{-g^O(\mu)}, e^{\mathrm{pr}\,q}(T^-)) - (\psi^q(0^+), \underbrace{e^{\mathrm{pr}\,q}(0^-)}_{0}) - \sum_{l\in\mathcal{L}}([e^{\mathrm{pr}\,q}]_{l-1}, \psi^q(t_{l-1}^+));
\end{aligned}$$

using in the second line the definition of the jump operator $e^{\mathrm{pr}\,q}(t_{l-1}^+) = [e^{\mathrm{pr}\,q}]_{l-1} + e^{\mathrm{pr}\,q}(t_{l-1}^-)$;

and in the last line, $\psi^q(T^+;\mu) = -g^O$ and $e^{\mathrm{pr}\,q}(0^-;\mu) = 0$. Therefore,

$$\begin{aligned} s^q(\mu) - s_N^q(\mu) = \sum_{l\in\mathcal{L}} \Bigg[& -\int_{I_l} (\partial_t e^{\mathrm{pr}\,q}, \psi^q)\ dt \\ & - \int_{I_l} a(e^{\mathrm{pr}\,q}, \psi^q;\mu)\, dt - ([e^{\mathrm{pr}\,q}]_{l-1}, \psi^q(t_{l-1}^+)) + \mathcal{R}_l^{\mathrm{pr}\,q}(\psi_M^q(\cdot;\mu);\mu); \Bigg] \qquad (4.31) \\ = & -\sum_{l\in\mathcal{L}} \mathcal{R}_l^{\mathrm{pr}\,q}(e^{\mathrm{du}\,q};\mu). \end{aligned}$$

But $\sum_{l\in\mathcal{L}} \mathcal{R}_l^{\mathrm{pr}\,q}(e^{\mathrm{du}\,q};\mu) = \sum_{l\in\mathcal{L}} \mathcal{R}_l^{\mathrm{du}\,q}(e^{\mathrm{pr}\,q};\mu)$, since

$$\begin{aligned} \sum_{l\in\mathcal{L}} \mathcal{R}_l^{\mathrm{pr}\,q}(e^{\mathrm{du}\,q};\mu) = & \sum_{l\in\mathcal{L}} \Bigg[\int_{I_l} \left(\partial_t e^{\mathrm{pr}\,q}, e^{\mathrm{du}\,q}\right)\ dt + \int_{I_l} a(e^{\mathrm{pr}\,q}, e^{\mathrm{du}\,q};\mu)\, dt \qquad (4.32) \\ & + ([e^{\mathrm{pr}\,q}]_{l-1}, e^{\mathrm{du}\,q}(t_{l-1}^+)) \Bigg] \\ = & \sum_{l\in\mathcal{L}} \left[-\int_{I_l} \left(\partial_t e^{\mathrm{du}\,q}, e^{\mathrm{pr}\,q}\right)\ dt + \int_{I_l} a(e^{\mathrm{pr}\,q}, e^{\mathrm{du}\,q};\mu)\, dt \right] + I_2 \\ = & \sum_{l\in\mathcal{L}} \left[-\int_{I_l} \left(\partial_t e^{\mathrm{du}\,q}, e^{\mathrm{pr}\,q}\right)\ dt + \int_{I_l} a(e^{\mathrm{pr}\,q}, e^{\mathrm{du}\,q};\mu)\, dt - ([e^{\mathrm{du}\,q}]_l, e^{\mathrm{pr}\,q}(t_l^-)) \right] \\ = & \sum_{l\in\mathcal{L}} \mathcal{R}_l^{\mathrm{du}\,q}(e^{\mathrm{pr}\,q};\mu); \qquad (4.33) \end{aligned}$$

from integration by parts and the definitions of the primal and dual residuals. The additional terms I_2 are calculated below:

$$\begin{aligned} I_2 = & \sum_{l\in\mathcal{L}} \left[(e^{\mathrm{pr}\,q}(t_l^-), e^{\mathrm{du}\,q}(t_l^-)) - (e^{\mathrm{pr}\,q}(t_{l-1}^+), e^{\mathrm{du}\,q}(t_{l-1}^+)) + (e^{\mathrm{pr}\,q}(t_{l-1}^+) - e^{\mathrm{pr}\,q}(t_{l-1}^-), e^{\mathrm{du}\,q}(t_{l-1}^+)) \right] \\ = & \sum_{l\in\mathcal{L}} \left[-(e^{\mathrm{pr}\,q}(t_l^-), [e^{\mathrm{du}\,q}]_l) + (e^{\mathrm{pr}\,q}(t_l^-), e^{\mathrm{du}\,q}(t_l^+)) - (e^{\mathrm{pr}\,q}(t_{l-1}^-), e^{\mathrm{du}\,q}(t_{l-1}^+)) \right] \\ = & -\sum_{l\in\mathcal{L}} ([e^{\mathrm{du}\,q}]_l, e^{\mathrm{pr}\,q}(t_l^-)); \end{aligned}$$

again the definition of the jump operator has been used, and $e^{\mathrm{pr}\,q}(0^-) = e^{\mathrm{du}\,q}(T^+) = 0$. Combining (4.31) and (4.32) we obtain:

$$-\sum_{l\in\mathcal{L}} \mathcal{R}_l^{\mathrm{du}\,q}(e^{\mathrm{pr}\,q};\mu) = \left(s^q(\mu) - s_N^q(\mu)\right). \qquad (4.34)$$

Turning now to $\sum_{l\in\mathcal{L}} \mathcal{R}_l^{\mathrm{pr}\,q}(e^{\mathrm{pr}\,q};\mu)$, we first compute I_3

$$\begin{aligned} I_3 &= \sum_{l\in\mathcal{L}} \Big[\frac{1}{2}(e^{\mathrm{pr}\,q}(t_l^-), e^{\mathrm{pr}\,q}(t_l^-)) - \frac{1}{2}(e^{\mathrm{pr}\,q}(t_{l-1}^+), e^{\mathrm{pr}\,q}(t_{l-1}^+)) \\ &\quad +(e^{\mathrm{pr}\,q}(t_{l-1}^+) - e^{\mathrm{pr}\,q}(t_{l-1}^-), e^{\mathrm{pr}\,q}(t_{l-1}^+))\Big] \\ &= \sum_{l\in\mathcal{L}} \frac{1}{2}\| \left[e^{\mathrm{pr}\,q}(t_{l-1})\right] \|^2_{L^2(\Omega)} + \frac{1}{2}(e^{\mathrm{pr}\,q}(T^-), e^{\mathrm{pr}\,q}(T^-)) - \frac{1}{2}(\underbrace{e^{\mathrm{pr}\,q}(0^-)}_{0}, e^{\mathrm{pr}\,q}(0^-)); \end{aligned}$$

as before we used here the definition of the jump operator and simple algebraic manipulations. But then,

$$\begin{aligned} \sum_{l\in\mathcal{L}} \mathcal{R}_l^{\mathrm{pr}\,q}(e^{\mathrm{pr}\,q};\mu) &= \sum_{l\in\mathcal{L}} \left[\int_{I_l} (\partial_t e^{\mathrm{pr}\,q}, e^{\mathrm{pr}\,q})\ dt + \int_{I_l} a(e^{\mathrm{pr}\,q}, e^{\mathrm{pr}\,q};\mu)\, dt + ([e^{\mathrm{pr}\,q}]_{l-1}, e^{\mathrm{pr}\,q}(t_{l-1}^+))\right] \\ &= I_3 + \sum_{l\in\mathcal{L}} \int_{I_l} a(e^{\mathrm{pr}\,q}, e^{\mathrm{pr}\,q};\mu)\, dt \geq g(\mu) \int_I \hat{a}(e^{\mathrm{pr}\,q}, e^{\mathrm{pr}\,q})\, dt \end{aligned} \tag{4.35}$$

since I_3 is the sum of non-negative terms, and also using (4.14). We turn now to the proof of the bounding properties, and as before, for $\kappa > 0$ we define $\hat{e}^{\pm\,q} = \hat{e}^{\mathrm{pr}\,q} \mp \frac{1}{\kappa}\hat{e}^{\mathrm{du}\,q}$. From the coercivity of $\hat{a}$,

$$\begin{gathered} \kappa g(\mu) \int_I \hat{a}(e^{\mathrm{pr}\,q} - \frac{1}{2}\hat{e}^{\pm\,q}, e^{\mathrm{pr}\,q} - \frac{1}{2}\hat{e}^{\pm\,q})\, dt \geq 0 \\ \kappa g(\mu) \int_I \hat{a}(e^{\mathrm{pr}\,q}, e^{\mathrm{pr}\,q})\, dt + \frac{\kappa g(\mu)}{4} \int_0^T \hat{a}(\hat{e}^{\pm\,q}, \hat{e}^{\pm\,q})\, dt - \kappa g(\mu) \int_I \hat{a}(\hat{e}^{\pm\,q}, e^{\mathrm{pr}\,q})\, dt \geq 0. \end{gathered} \tag{4.36}$$

From the definition of $\hat{e}^{\pm\,q}$, $\hat{e}^{\mathrm{pr}\,q}$, and $\hat{e}^{\mathrm{du}\,q}$ we have:

$$g(\mu) \int_I \hat{a}(\hat{e}^{\pm\,q}, e^{\mathrm{pr}\,q})\, dt = \sum_{l\in\mathcal{L}} \mathcal{R}_l^{\mathrm{pr}\,q}(e^{\mathrm{pr}\,q};\mu) \mp \frac{1}{\kappa} \sum_{l\in\mathcal{L}} \mathcal{R}_l^{\mathrm{du}\,q}(e^{\mathrm{pr}\,q};\mu).$$

Using (4.35) to replace $\sum_{l\in\mathcal{L}} \mathcal{R}_l^{\mathrm{pr}\,q}(e^{\mathrm{pr}\,q};\mu)$, and (4.34) to replace $\sum_{l\in\mathcal{L}} \mathcal{R}_l^{\mathrm{du}\,q}(e^{\mathrm{pr}\,q};\mu)$, we get

$$-\kappa g(\mu) \int_I \hat{a}(\hat{e}^{\pm\,q}, e^{\mathrm{pr}\,q})\, dt \leq -\kappa g(\mu) \int_I \hat{a}(e^{\mathrm{pr}\,q}, e^{\mathrm{pr}\,q})\, dt \mp (s^q(\mu) - s_N^q(\mu))\,. \tag{4.37}$$

Replacing now (4.37) in (4.36), we get

$$\pm\left(s^q(\mu) - s_N^q(\mu)\right) \leq \frac{\kappa g(\mu)}{4} \int_I \hat{a}(\hat{e}^{\pm\, q}, \hat{e}^{\pm\, q})\, dt.$$

The rest of the proof follows similarly to that of the continuous case. □

To measure the quality of the computed bounds, we define the *a posteriori* effectivity index $\eta(\mu)$, as the ratio of the computed error over the true error in the output prediction

$$\eta(\mu) = \frac{\Delta^q(\mu)}{|s^q(\mu) - s_B^q(\mu)|}.$$

The previous Proposition suggests that the prediction of the error in the output will overestimate the true error and therefore the effectivity will always be larger than one, $\eta(\mu) \geq 1,\ \forall \mu \in \mathcal{D}$. Large effectivities indicate that the computed error bound largely overestimates the true error and therefore the bounds obtained are not sharp. This implies that for a given accuracy, the error estimator suggests the use of a higher number of basis functions than are actually required and the computational cost will be unnecessarily high. For efficiency, it is therefore desired that the effectivities will be as close to one as possible. The choice of $\hat{a}$ and $g(\mu)$ is critical to obtain good effectivities; for a discussion see [94]. For better effectivities, it is possible to choose different $\hat{a}$ and $g(\mu)$ which satisfy (4.14) only in subregions of the parameter domain. Also the more general bound conditioners, developed in [113] for elliptic coercive problems, can also be extended to the parabolic case.

4.6 Results

We consider the problem of designing the thermal fin of Figure 4-1 to cool (say) an electronic component at the fin base, Γ_{root}; the description of the problem is given in the introduction.

Initially, the non-dimensional temperature is $u_0(\mu) = 0$. A uniform heat flux is applied at the root of the fin at $t = 0$ and remains on until the final time $t = T \equiv 3$. The temperature increases until it reaches the final value $u(T^-; \mu)$. On the original domain the bilinear form

is given by, $\hat{a}(w,v;\mu) = \int_{\hat{\Omega}_0} \nabla w \cdot \nabla v + \sum_{i=1}^{4} k_i \int_{\hat{\Omega}_i} \nabla w \cdot \nabla v + \mathrm{Bi} \int_{\partial\hat{\Omega}\setminus\Gamma_{\mathrm{root}}} wv$; with $\hat{\Omega}_0$ the fin central-post domain, and $\hat{\Omega}_i$ the ith radiator domain. We then map the original domain $\hat{\Omega}$ to a reference geometry Ω, shown by solid lines in Figure 4-1. The original bilinear form $\hat{a}(w,v;\mu)$ is replaced by $a(w,v;\mu)$ defined in the fixed domain Ω — the variable geometry appears as domain-dependent effective orthotropic conductivities and Bi numbers. Similarly, the L^2-inner product $(w,v)_{L^2(\hat{\Omega})}$ is replaced by $(w,v)_{L^2(\Omega)} \equiv b(w,v;\mu)$, defined on the fixed domain — the variable geometry also makes the L^2-inner product parameter-dependent. We consider two outputs: the first, is the mean temperature of the base Γ_{root} averaged over the time interval $(0,T)$:

$$s^1(\mu) \equiv s^1(u(\cdot;\mu)) = \frac{1}{T}\int_I \int_{\Gamma_{\mathrm{root}}} u(t;\mu)\, dS\, dt;$$

the second, is the mean temperature in the shaded region Ω_{out} (with area $A_{\Omega_{\mathrm{out}}}$) at the final time $t = T$:

$$s^2(\mu) \equiv s^2(u(\cdot;\mu)) = \frac{1}{A_{\Omega_{\mathrm{out}}}} \int_{\Omega_{\mathrm{out}}} u(T^-;\mu)\, dS.$$

Both outputs are, to a certain extent, indicators of the cooling performance of the fin.

Taking advantage of the natural domain decomposition afforded by our mapping, it is not difficult to cast the problem such that the affine decomposition assumption is verified:

$$a(w,v;\mu) = \sum_{q=1}^{Q_a} \sigma_a^q(\mu) a^q(w,v), \ \forall w,v \in V^2, \text{ and}$$

$$b(w,v;\mu) = \sum_{q=1}^{Q_b} \sigma_b^q(\mu) b^q(w,v), \ \forall w,v \in L^2(\Omega);$$

with $Q_a = 16$ and $Q_b = 3$. Choosing $\hat{a}(w,v) = \sum_{q=1}^{Q_a} a^q(w,v) = \int_\Omega \nabla w \cdot \nabla v + \int_{\partial\Omega\setminus\Gamma_{\mathrm{root}}} wv$, and $g(\mu) = \min_{q\in\{1,\ldots,Q_a\}} \sigma_a^q(\mu)$ (the $\sigma_a^q(\mu)$ are all bounded below by a positive constant), we are able to verify (4.14).

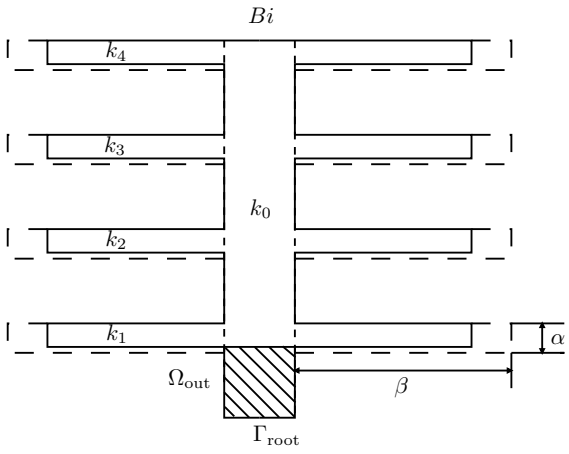

Figure 4-1: Two-dimensional thermal fin

Thus all the requirements are honored, and the bound
method can be applied.

We choose the total (non-dimensional) height of the thermal fin $\hat{H} = 4$, and the length and height of the radiators $\hat{\alpha} = 2.5$ and $\hat{\beta} = 0.25$ respectively; the reference geometry $\hat{\Omega}$ is thus completely defined. To obtain the "exact" solution: first, for the spatial discretization, we introduce a very fine triangulation $\mathcal{T}_h$ and define the finite-element space $V \approx V_h = \{v \in H^1(\hat{\Omega}) | v|_{T_h} \in \mathbb{P}^1, \forall T_h \in \mathcal{T}_h\}$ with piecewise linear polynomials over each of the elements T_h; and second, for the temporal discretization, the discontinuous Galerkin method is used with $q = 0$ and the time interval $\bar{I} = [0, 3]$ is partitioned into $L = 30$ intervals of uniform length $\Delta\tau^l \equiv \Delta\tau = 0.1, \ \forall l \in \mathcal{L}$. (The same parameters are used for the reduced-basis problems.)

Next in the definition of our problem, is the specification of the ranges for each of the input parameters. We choose a parameter space as follows: $\mathcal{D} = [0.01, 100.0]^4 \times [0.001, 10.0] \times [0.2, 0.6] \times [2.3, 2.8]$, that is $0.01 \leq k_{1,2,3,4} \leq 100.0$, $0.001 \leq \text{Bi} \leq 10.0$, $0.2 \leq \alpha \leq 0.6$ and $2.3 \leq \beta \leq 2.8$. Points in this parameter space — for example, for the construction of the sample sets S_N^{pr} and S_M^{du} — are obtained by sampling "log-randomly" (see Section 4.3). A point $\mu \in \mathcal{D}$, describes a particular configuration. For example, $\mu_t = \{0.4, 0.6, 0.8, 1.2, 0.1, 0.3, 2.8\}$ represents a thermal fin with $k_1 = 0.4$, $k_2 = 0.6$, $k_3 = 0.8$, $k_4 = 1.2$, $\text{Bi} = 0.1$, $\alpha = 0.3$, and $\beta = 2.8$; this particular configuration will be used as a test point μ_t in the following numerical experiments.

For the construction of the primal reduced-basis space we sample $\mathcal{D}$ and obtain a number of points μ_i^{pr}. For each of these points the primal problem is solved and the reduced-basis vectors are obtained by taking "snapshots" of the solution at different times. The sampling times or the number of snapshots can vary arbitrarily from one configuration to the next; in the following, for each configuration, four "snapshots" were obtained at $t = 1\Delta\tau, \ 10\Delta\tau, \ 20\Delta\tau, \ \text{and} \ 30\Delta\tau$. For example if $N = 20$, five different configurations were considered, each giving four basis vectors for the construction of the reduced-basis space. For the dual reduced-basis space the same procedure is followed, solving the dual problem for a different set of parameter points and taking "snapshots" of the solution at $t = 29\Delta\tau, \ 20\Delta\tau, \ 10\Delta\tau, \ \text{and} \ 0\Delta\tau$.

As a first test, we study the convergence of the reduced-basis solution to the exact one.

$N = M$	$\frac{\lvert s_N^1(\mu_t) - s^1(\mu_t)\rvert}{s^1(\mu_t)}$	$\frac{\lvert s_N^2(\mu_t) - s^2(\mu_t)\rvert}{s^2(\mu_t)}$
8	$1.22e-01$	$2.03e-01$
20	$3.15e-03$	$6.41e-03$
40	$1.18e-04$	$4.61e-04$
60	$3.91e-05$	$3.02e-05$
80	$1.16e-06$	$2.56e-06$
100	$7.42e-07$	$1.22e-06$
120	$1.74e-08$	$2.46e-07$

Table 4.1: Relative error by the reduced-basis prediction of the outputs of interest for different values of $N = M$.

For this, we sample log-randomly the parameter space $\mathcal{D}$ and construct reduced-basis spaces of increasing dimension $N = M$. Using these spaces we compute, for the test point μ_t, the reduced-basis solution and the two outputs of interest. In Table 4.1, the error in the prediction of the adjoint-corrected output relative to the exact value, is shown for increasing values of N. We can see that, for both outputs, the output prediction converges very fast to the exact value, albeit at a different rate for each output. If, for example, a 1% accuracy is required — which is sufficient in many engineering applications, — then *only* $N = 20$ basis functions would be sufficient. This implies that the *incremental cost* for each new output evaluation is very small; depending on the dimension of the space V_h, the computational savings can be of several orders of magnitude. For sufficiently large values of N, M the vectors that comprise the reduced basis spaces are closely related and this leads to ill-conditioning problems. Indeed in our case increasing N, M above 120, ill-conditioning leads first to deterioration of the convergence rate and eventually to incorrect results. The issue of ill-conditioning in the reduced-basis context is very important, but an analysis will not be further pursued; first, because we are usually interested in the pre-asymptotic region (small values of N); and second, because even for the conservative triangulation used here, the discretization error is of the same order of magnitude as the reduced-basis error when $N = M = 80$ — using higher values for N is not relevant except, maybe, for testing the convergence rate.

The choice of the sample set S_N^{pr}, is critical for the approximation properties of W_N^{pr}. For

N	20	40	60	80	100
M	100	80	60	40	20
$\Delta^1(\mu_t)$	$1.10e-03$	$7.62e-04$	$8.39e-04$	$1.22e-03$	$9.98e-04$
$\Delta^2(\mu_t)$	$2.78e-03$	$2.10e-03$	$2.33e-03$	$3.25e-03$	$3.47e-03$
$\eta^1(\mu_t)$	34.2	53.0	13.7	41.7	106.1
$\eta^2(\mu_t)$	184.8	22.3	16.5	33.1	88.2

Table 4.2: Bound gap and effectivities for the two outputs of interest, for different choices of $N = \dim W_N^{\text{pr}}$ and $M = \dim W_M^{\text{du}}$ (N+M=120).

the same N, different choices for S_N^{pr} can give different reduced-basis spaces and consequently different output approximations; relatedly the approximation error can vary significantly for different test points $\mu \in \mathcal{D}$. Moreover, even for the same sample set S_N^{pr}, the error for different outputs can be quite different. For example, in Table 4.1, for $N = 40$ and the particular point μ_t, the error in the prediction of the second output is four times larger compared to the error in the first output. Ascertaining the accuracy of our predictions without, of course, computing the exact solution, is therefore critical for the successful application of the reduced-basis method; the importance of efficient and reliable methods to *a posteriori* estimate the error in our predictions should be clear.

We turn now to the *a posteriori* error estimator procedure and investigate its behavior in the context of the model problem. To calculate the bounds, we need to solve using the reduced-basis method both the primal problem of dimension N, and the dual problem of dimension M. These dimensions determine the accuracy of the approximation to each of the problems and can, in principle, be chosen independently. To understand how this choice affects the accuracy of the predictions, we fix the total dimension $N + M = 120$ and choose different combinations for N and M. In Table 4.2, for the particular point μ_t and the two outputs of interest, the bound gap $\Delta(\mu_t)$ and the effectivity $\eta(\mu_t)$ are presented for different choices of N and M.

To understand the behavior the bound gap, recall that it is defined (4.30) as the product of norms of representations to the primal $\hat{e}^{\text{pr}}$ and dual errors $\hat{e}^{\text{du}}$ — which are directly related to the true errors. As N increases the error in the primal solution becomes smaller, while at the same time, M is decreasing and the error in the dual solution becomes larger.

Therefore, as we can verify from Table 4.2, the bound gap does not change appreciably for the different N and M. The small variations can also be attributed to the different selection of basis functions in the formation of the reduced-basis spaces. On the other hand, the dual correction term in the output approximation (4.27), ensures that the output will be more accurate when either N or M are large. In these cases the error in the output is small and given that the bound gap does not change significantly, justifies the higher effectivities. The discussion above suggests that, for a given accuracy — as dictated by the bound gap $\Delta(\mu)$, — we can choose N or M arbitrarily, such that the total number of basis functions is constant. On one side, we have the case $M = 0$ ($N = 0$), which corresponds to a pure primal (dual) problem; on the other, we can have a mixed approach with $N = M$. The computational cost for the second case is roughly two times smaller in the off-line stage and four times smaller in the on-line. The use of both the primal and the adjoint problems is thus dictated by computational efficiency considerations.

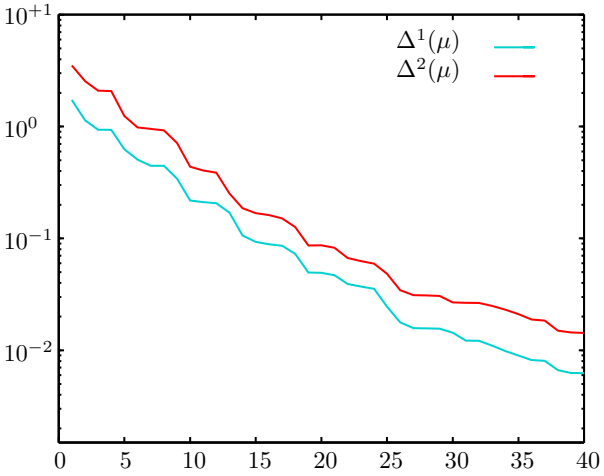

Figure 4-2: Convergence of the bound gap as a function of N(=M), for the point μ_t.

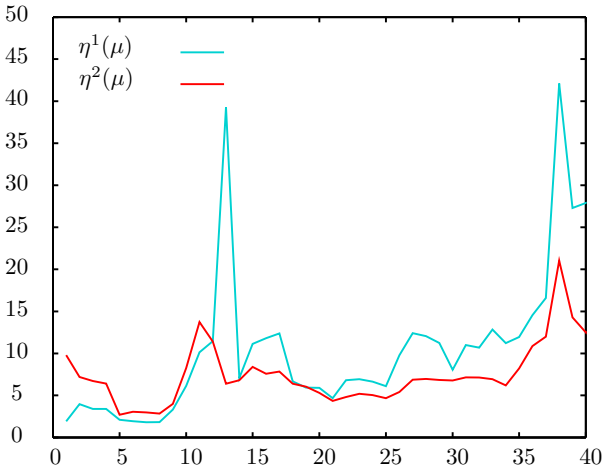

Figure 4-3: Effectivity as a function of N(=M) for the point μ_t.

As a final test, we choose $N = M$ and for the test point μ_t, we vary the dimension of the reduced-basis spaces. The behavior of the bound gap as a function of $N = M$ is shown in Figure 4-2, and of the effectivity in Figure 4-3. Despite the relatively high dimension of the parameter space, we observe the good accuracy and rapid convergence of the bound gap. Also, given that the effectivity remains bounded for all values of N, we conclude that the bound gap converges at the same rate as the true error in the output. This suggests

that instead of using the high-dimensional model to evaluate outputs for different parameter points, we can replace it with a reduced-basis model. Due to the rapid convergence only a few basis functions are required and therefore we can obtain high efficiency. In addition, we recover certainty as the error bounds validate the accuracy of the reduced-basis predictions. In terms of computational effort, the off-line stage requires, typically, a few hundred solutions of the continuous problem — depending on the number of basis functions and the parametric complexity of the bilinear forms. But then the on-line cost, for each new configuration $\mu \in \mathcal{D}$ is typically more than a hundred or a thousand times smaller — depending on the dimension of V_h. The computational advantages in the limit of many evaluations, should be obvious. More realistic applications, as well as integration of these components in an optimization or design framework will be addressed in a future paper.

Chapter 5

Noncoercive Problems

5.1 Problem description

5.1.1 Problem statement

Given a Hilbert space Y of dimension $\mathcal{N}$ (possibly infinite), a linear functional $\ell \in Y'$, and a parameter μ in a set $\mathcal{D} \subset \mathbb{R}^P$, we look for $u(\mu) \in Y$ such that

$$a(u(\mu), v; \mu) = \ell(v), \ \forall v \in Y, \tag{5.1}$$

where $a(\cdot, \cdot; \mu)$ is a bilinear form the assumptions on which are detailed below. We further prescribe an *output functional* $\ell^O \in Y'$, in terms of which we can evaluate our output of interest $s(\mu)$ as

$$s(\mu) = \ell^O(u(\mu)). \tag{5.2}$$

We also introduce a dual, or adjoint, problem associated with ℓ^O: find $\psi(\mu) \in Y$ such that

$$a(v, \psi(\mu); \mu) = -\ell^O(v), \ \forall v \in Y.$$

The relevance of this dual problem will become clear in what follows. It is relatively simple to permit μ-dependence in ℓ and ℓ^O as well, however for clarity of exposition we do not consider this here.

We shall assume (though this is essential for only some of our arguments) that a is symmetric,

$$a(w, v; \mu) = a(v, w; \mu), \ \forall w, v \in Y^2, \ \forall \mu \in \mathcal{D}.$$

We further assume that $a(w, v; \mu)$ is uniformly continuous,

$$|a(w, v; \mu)| \leq \gamma \|w\|_Y \|v\|_Y, \ \forall w, v \in Y^2, \ \forall \mu \in \mathcal{D},$$

and that $a(w, v; \mu)$ satisfies a uniform inf-sup condition,

$$0 < \beta_0 \leq \beta(\mu) = \inf_{w \in Y} \sup_{v \in Y} \frac{a(w, v; \mu)}{\|w\|_Y \|v\|_Y} = \inf_{w \in Y} \frac{\|a(w, \cdot; \mu)\|_{Y'}}{\|w\|_Y}, \ \forall \mu \in \mathcal{D}; \tag{5.3}$$

it is classical that these latter two conditions are required for the well-posedness of our primal and dual problems. Finally, we shall make the assumption that our bilinear form a is affine in the parameter μ in the sense that, for some finite integer Q,

$$a(w, v; \mu) = \sum_{q=1}^{Q} \sigma^q(\mu) a^q(w, v), \ \forall w, v \in Y^2, \tag{5.4}$$

where the a^q are bilinear forms. The assumption (5.4) permits our blackbox approach; non-blackbox variants of the methods described here — in which (5.4) is relaxed — can also be developed.

5.1.2 Inf-sup supremizers and infimizers

We can rephrase (5.3) as: for every $w \in Y$, there exists an element $T_\mu w$ in Y, such that

$$\beta(\mu) \|w\|_Y \|T_\mu w\|_Y \leq a(w, T_\mu w; \mu), \tag{5.5}$$

where $T_\mu w$ is the *supremizer* associated with $\|a(w, \cdot; \mu)\|_{Y'}$. It follows from Section 2.2 that $T_\mu w = \rho^Y_{a(w, \cdot; \mu)}$, that is,

$$(T_\mu w, v)_Y = a(w, v; \mu), \ \forall v \in Y.$$

It is thus clear that $T_\mu : Y \to Y$ is a linear operator; we can also readily show that T_μ is symmetric (since a is symmetric); furthermore, T_μ is bounded, since

$$\|T_\mu w\|_Y^2 = a(w, T_\mu w; \mu) \leq \gamma \|w\|_Y \|T_\mu w\|_Y,$$

and hence

$$\frac{\|T_\mu w\|_Y}{\|w\|_Y} \leq \gamma. \tag{5.6}$$

Finally, we can now express our inf-sup parameter in terms of T_μ as:

$$\beta(\mu) = \inf_{w \in Y} \frac{\|T_\mu w\|_Y}{\|w\|_Y} = \frac{\|T_\mu \chi(\mu)\|_Y}{\|\chi(\mu)\|_Y},$$

where

$$\chi(\mu) = \arg \inf_{w \in Y} \frac{\|T_\mu w\|_Y}{\|w\|_Y}$$

is our *infimizer*; we thus also have

$$\beta(\mu) = \frac{a(\chi(\mu), T_\mu \chi(\mu))}{\|\chi(\mu)\|_Y \|T_\mu \chi(\mu)\|_Y}.$$

It will be useful in the subsequent analysis to recognize that $\beta(\mu)$ and $\chi(\mu)$ are related to the minimum eigenvalue and associated eigenfunction of a symmetric positive-definite eigenproblem. In particular we look for $(\theta(\mu), \lambda(\mu)) \in (Y \times \mathbb{R})$ solution of

$$\mathcal{A}(\theta(\mu), v; \mu) = \lambda(\mu)(\theta(\mu), v)_Y, \ \forall v \in Y, \text{ and } \|\theta(\mu)\|_Y = 1, \tag{5.7}$$

where

$$\mathcal{A}(w, v; \mu) = (T_\mu w, T_\mu v)_Y, \ \forall w, v \in Y^2; \tag{5.8}$$

we denote the resulting eigenpairs as $(\theta_i(\mu), \lambda_i(\mu)), \ i = 1, \ldots$, with $\lambda_{\min} \equiv \lambda_1 \leq \lambda_2 \leq \cdots$. It follows immediately from Rayleigh quotient arguments that

$$\lambda_{\min}(\mu) = \min_{w \in Y} \frac{\mathcal{A}(w, w; \mu)}{(w, w)_Y} = \min_{w \in Y} \frac{\|T_\mu w\|_Y^2}{\|w\|_Y^2} = \beta^2(\mu),$$

and thus $\beta(\mu) = \sqrt{\lambda_{\min}(\mu)}$ and $\theta_{\min}(\mu) \equiv \theta_1(\mu) = \chi(\mu)$.

To further understand the relationship between the infimizers and supremizers, we consider a second symmetric positive-definite eigenproblem: find $(\Upsilon \times \omega) \in (Y \times \mathbb{R})$ such that

$$2\gamma(\Upsilon(\mu), v)_Y = \omega(\mu)\mathcal{B}(\Upsilon(\mu), v; \mu), \ \forall v \in Y \tag{5.9}$$

where

$$\mathcal{B}(w, v; \mu) = 2\gamma(w, v)_Y - a(w, v; \mu);$$

note that $\mathcal{B}$ is symmetric and coercive. By the usual arguments (and appropriate normalization), $2\gamma(\Upsilon_i, \Upsilon_j)_Y = \omega_i \mathcal{B}(\Upsilon_i, \Upsilon_j) = \omega_i \delta_{ij}$, with $0 < \omega_1 \leq \omega_2 \leq \cdots$; here δ_{ij} is the Kronecker delta symbol, and (Υ_i, ω_i) refers to the i^{th} eigenpair associated with (5.9). We can then write

$$(T_\mu w, v)_Y = 2\gamma(w, v)_Y - \mathcal{B}(w, v; \mu),$$

expand

$$w = \sum_{i=1}^{\mathcal{N}} c_i \Upsilon_i,$$

and exploit orthogonality to deduce that

$$T_\mu w = \sum_{i=1}^{\mathcal{N}} d_i \Upsilon_i$$

for $d_i = 2\gamma c_i(\omega_i - 1)/\omega_i$. Thus

$$\frac{\|T_\mu w\|_Y^2}{\|w\|_Y^2} = \frac{4\gamma^2 \sum_{i=1}^{\mathcal{N}} \left| \frac{\omega_i(\mu) - 1}{\omega_i(\mu)} \right|^2 \omega_i(\mu) c_i^2}{\sum_{i=1}^{\mathcal{N}} \omega_i(\mu) c_i^2},$$

again by orthogonality. We conclude that

$$\beta(\mu) = 2\gamma \left| \frac{\omega_{i^*(\mu)}(\mu) - 1}{\omega_{i^*(\mu)}(\mu)} \right|,$$

and $\chi(\mu) = c_{i^*(\mu)}\Upsilon_{i^*(\mu)}$, where

$$i^*(\mu) = \arg\min_{i\in\{1,\ldots,\mathcal{N}\}} \left| \frac{\omega_i(\mu) - 1}{\omega_i(\mu)} \right|. \tag{5.10}$$

We thus observe that

$$T_\mu \chi = 2\gamma \left(\frac{\omega_{i^*(\mu)}(\mu) - 1}{\omega_{i^*(\mu)}(\mu)} \right) \chi,$$

which states that $T_\mu\chi$ and χ are *collinear*; in general, $T_\mu w$ and w will be linearly dependent only if w is proportional to a *single* eigenfunction Υ_i.

5.2 Reduced-basis output bound formulation

5.2.1 Approximation spaces

Infimizing spaces W_N

We select M_1 points in our parameter set $\mathcal{D}$, $\mu_m \in \mathcal{D},\ m = 1, \ldots, M_1$, the collection of which we denote

$$\mathcal{S}_{M_1} = \{\mu_1, \ldots, \mu_{M_1}\}.$$

We then introduce the "Lagrangian" space [91],

$$W^u_{M_1} = \text{span}\{u(\mu_m),\ \forall \mu_m \in \mathcal{S}_{M_1}\}. \tag{5.11}$$

Similarly, we choose M_2 points in $\mathcal{D}$, possibly different from the ones above, and define $\mathcal{S}_{M_2}$ and

$$W^\psi_{M_2} = \text{span}\{\psi(\mu_m),\ \forall \mu_m \in \mathcal{S}_{M_2}\}; \tag{5.12}$$

and also M_3 points in $\mathcal{D}$ to define $\mathcal{S}_{M_3}$ and

$$W^\chi_{M_3} = \text{span}\{\chi(\mu_m),\ \forall \mu_m \in \mathcal{S}_{M_3}\}. \tag{5.13}$$

These three spaces are associated with our primal solutions, dual solutions, and infimizers, respectively.

We shall consider two approximation spaces W_N. In the first, we set $N = M_1 + M_2$ and choose $W_N = W_N^0$, where

$$\begin{aligned} W_N^0 &= W_{M_1}^u + W_{M_2}^\psi \\ &= \text{span}\{u(\mu_i), \psi(\mu_j),\ \forall \mu_i \in \mathcal{S}_{M_1},\ \forall \mu_j \in \mathcal{S}_{M_2}\} \\ &\equiv\ \text{span}\{\zeta_1, \ldots, \zeta_N\}. \end{aligned} \tag{5.14}$$

In the second case we set $N = M_1 + M_2 + M_3$ and choose $W_N = W_N^1$, where

$$\begin{aligned} W_N^1 &= W_{M_1}^u + W_{M_2}^\psi + W_{M_3}^\chi \\ &=\ \text{span}\{u(\mu_i), \psi(\mu_j), \chi(\mu_k),\ \forall \mu_i \in \mathcal{S}_{M_1},\ \forall \mu_j \in \mathcal{S}_{M_2},\ \forall \mu_k \in \mathcal{S}_{M_3}\} \\ &\equiv\ \text{span}\{\zeta_1, \ldots, \zeta_N\}. \end{aligned} \tag{5.15}$$

(Obviously the ζ_N — the reduced-basis functions — take different meanings in the two cases.) The role of each of the components of the W_N shall become clear later in our development.

Remark 5.2.1. Compliance. *In the case in which $\ell^O = \ell$, then $\psi(\mu) = -u(\mu)$; if $\mathcal{S}_{M_1} \cap \mathcal{S}_{M_2} \neq \emptyset$ we need to redefine W_N^0 and W_1^N to remove any linearly-dependent vectors. This will of course result in computational savings. Note that if ℓ^O is close to ℓ and $\mathcal{S}_{M_1} \cap \mathcal{S}_{M_2} \neq \emptyset$ then W_N^0 of (5.14) and W_N^1 of (5.15) can lead to ill-conditioned systems.*

We shall see shortly that the W_N will play the role of the infimizing space.

Supremizing spaces V_N

We will also need supremizing spaces. To that end, we introduce $V_N \subseteq Y$, with $(\cdot,\cdot)_{V_N} = (\cdot,\cdot)_Y$ and hence $\|\cdot\|_{V_N} = \|\cdot\|_Y$. To define the supremizing space, we compute $z^{n,q} \equiv \rho^Y_{a^q(\zeta_n,\cdot)}$ for $1 \le n \le N$, and $1 \le q \le Q$ (where Q and a^q are defined in (5.4)); more specifically we compute

$$(z^{n,q}, v)_Y = a^q(\zeta_n, v),\ \forall v \in Y,\ 1 \le q \le Q,\ 1 \le n \le N. \tag{5.16}$$

Our first choice for the supremizing space is then $V_N = Z_N(\mu)$, with

$$Z_N(\mu) \equiv \text{span}\{\sum_{q=1}^{Q} \sigma^q(\mu) z^{n,q},\ n = 1, \ldots, N\}. \tag{5.17}$$

We make a few observations: first, notice that the supremizing space is related to infimizing space (through the choice of ζ_i); second, unlike earlier definitions of reduced-basis spaces, the supremizing space is now parameter dependent — this will require modifications is the computational procedure; and third, we notice that even though we need NQ functions, the $z^{n,q}$, the supremizing space has dimension N. The definition above might not seem very motivated at this point, it should become clear in the following sections.

We shall also consider two other possibilities, in particular: $V_N = W_N$ ($= W_N^0$ or W_N^1), and hence of dimension N; and $V_N = Y$, and hence of dimension $\mathcal{N}$.

5.2.2 Output Prediction

We next define, for all $w_N \in W_N$ and $\varphi_N \in W_N$, the primal and dual residuals, $R^{\text{pr}}(\cdot; w_N; \mu) \in Y'$ and $R^{\text{du}}(\cdot; \varphi_N; \mu) \in Y'$, respectively:

$$\begin{aligned} R^{\text{pr}}(v; w_N; \mu) &\equiv \ell(v) - a(w_N, v; \mu),\ \forall v \in Y, \\ R^{\text{du}}(v; \varphi_N; \mu) &\equiv -\ell^O(v) - a(v, \varphi_N; \mu),\ \forall v \in Y. \end{aligned}$$

It follows from our primal and dual problem statements that

$$\begin{aligned} R^{\text{pr}}(v; w_N; \mu) &= a(u - w_N, v; \mu), \\ R^{\text{du}}(v; \varphi_N; \mu) &= a(v, \psi - \varphi_N; \mu), \end{aligned} \tag{5.18}$$

which is the standard residual-error relation evoked in most *a posteriori* frameworks.

We then look for $u_N(\mu) \in W_N,\ \psi_N(\mu) \in W_N$, such that

$$u_N(\mu) = \arg\inf_{w_N \in W_N} \|R^{\text{pr}}(\cdot; w_N; \mu)\|_{(V_N)'} = \arg\inf_{w_N \in W_N} \sup_{v \in V_N} \frac{R^{\text{pr}}(v; w_N; \mu)}{\|v\|_Y}, \tag{5.19}$$

and

$$\psi_N(\mu) = \arg\inf_{\varphi_N \in W_N} \|R^{\rm du}(\cdot;\varphi_N;\mu)\|_{(V_N)'} = \arg\inf_{\varphi_N \in W_N} \sup_{v \in V_N} \frac{R^{\rm du}(v;\varphi_N;\mu)}{\|v\|_Y}, \tag{5.20}$$

which is a minimum-residual (or least-squares) projection; see also [14] for discussion of various projections within the reduced-basis context.

Our output approximation is then given by:

$$s_N(\mu) = \ell^O(u_N(\mu)) - R^{\rm pr}(\psi_N(\mu); u_N(\mu); \mu); \tag{5.21}$$

the additional adjoint terms will improve the accuracy [64, 85, 90].

It will be convenient to express our minimum-residual approximation in terms of affine supremizing operators $P_\mu^N : W_N \to V_N, D_\mu^N : W_N \to V_N$, defined by

$$P_\mu^N w_N = \rho^{V_N}_{R^{\rm pr}(\cdot;w_N;\mu)},$$
$$D_\mu^N w_N = \rho^{V_N}_{R^{\rm du}(\cdot;w_N;\mu)},$$

that is

$$(P_\mu^N w_N, v)_Y = R^{\rm pr}(v; w_N; \mu), \ \forall v \in V_N, \tag{5.22}$$

$$(D_\mu^N w_N, v)_Y = R^{\rm du}(v; w_N; \mu), \ \forall v \in V_N, \tag{5.23}$$

for any $w_N \in W_N$. In particular, it follows from Section 2.2 that we can now write

$$u_N(\mu) = \arg\inf_{w_N \in W_N} \|P_\mu^N w_N\|_Y, \quad \psi_N(\mu) = \arg\inf_{\varphi_N \in W_N} \|D_\mu^N \varphi_N\|_Y; \tag{5.24}$$

the μ-dependence is through P_μ^N and D_μ^N.

5.2.3 Error bound prediction

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

$$\beta_N(\mu) = \inf_{w_N \in W_N} \sup_{v \in V_N} \frac{a(w_N, v; \mu)}{\|w_N\|_Y \|v\|_Y} = \inf_{w_N \in W_N} \frac{\|a(w_N; \cdot; \mu)\|_{(V_N)'}}{\|w_N\|_Y}. \tag{5.25}$$

We can rephrase (5.25) as: for any $w_N \in W_N$, there exists an element $T_\mu^N w_N$ in V_N, such that

$$\beta_N(\mu)\|w_N\|_Y\|T_\mu^N w_N\|_Y \leq a(w_N, T_\mu^N w_N; \mu),\ \forall w_N \in W_N, \tag{5.26}$$

where $T_\mu^N w_N$ is the supremizer associated with $\|a(w_N, \cdot; \mu)\|_{(V_N)'}$. It follows from Section 2.2 that $T_\mu^N : W_N \to V_N$ is given by $\rho^{V_N}_{a(w_N,\cdot;\mu)}$, or more explicitly,

$$(T_\mu^N w_N, v)_Y = a(w_N, v; \mu),\ \forall v \in V_N,$$

for any $w_N \in W_N$. We can now express $\beta_N(\mu)$ as

$$\beta_N(\mu) = \inf_{w_N \in W_N} \frac{\|T_\mu^N w_N\|_Y}{\|w_N\|_Y} = \frac{\|T_\mu^N \chi_N(\mu)\|_Y}{\|\chi_N(\mu)\|_Y},$$

where

$$\chi_N(\mu) = \arg\inf_{w_N \in W_N} \frac{\|T_\mu^N w_N\|_Y}{\|w_N\|_Y}$$

is our infimizer over W_N; we thus also have

$$\beta_N(\mu) = \frac{a(\chi_N(\mu), T_\mu^N \chi_N(\mu))}{\|\chi_N(\mu)\|_Y \|T_\mu^N \chi_N(\mu)\|_Y}.$$

Then, given $u_N(\mu), \psi_N(\mu)$, and a real constant σ, $0 < \sigma < 1$, we compute

$$\Delta_N(\mu) = \frac{1}{\sigma\beta_N(\mu)} \|R^{\rm pr}(\cdot; u_N(\mu); \mu)\|_{Y'} \|R^{\rm du}(\cdot; \psi_N(\mu); \mu)\|_{Y'}, \tag{5.27}$$

which will serve as our *a posteriori* error bound for $|(s - s_N)(\mu)|$.

Remark 5.2.2. Output Bounds. *We can of course translate our error bound $\Delta_N(\mu)$ into (symmetric) upper and lower bounds for $s(\mu)$, $s_N^+(\mu) = s_N(\mu) + \Delta_N(\mu)$, $s_N^-(\mu) = s_N(\mu) - \Delta_N(\mu)$. For coercive problems the output bounds are in fact nonsymmetric — due to a shift which also effectively reduces the bound gap by a factor of two relative to the noncoercive case.*

5.3 Error analysis

In Section 5.3.1 we analyze the accuracy of our reduced-basis output prediction $s_N(\mu)$, and in Section 5.3.2 we consider the properties of our error estimator $\Delta_N(\mu)$. In both Section 5.3.1 and 5.3.2 we make certain hypotheses about $\beta_N(\mu)$ that we then discuss in Section 5.3.3. Note that, for convenience of exposition, we shall not always explicitly indicate the μ dependence of all quantities.

5.3.1 A priori theory

We first prove that our discrete approximation is well-defined, as summarized in

Lemma 5.3.1. *If $\beta_N(\mu) \geq \tilde{\beta}_0 > 0,\ \forall \mu \in \mathcal{D}$, then the discrete problems (5.19) and (5.20) are well-posed.*

Proof. We consider the primal problem (5.19); analysis of the dual problem (5.20) is similar. To begin, we recall that $\rho_\ell^{V_N} \in V_N$ satisfies

$$(\rho_\ell^{V_N}, v)_Y = \ell(v),\ \forall v \in V_N.$$

It thus follows that, for any $w_N \in W_N$,

$$P_\mu^N w_N = \rho_\ell^{V_N} - T_\mu^N w_N;$$

from our minimum-residual statement (5.24) we then know that $u_N \in W_N$ satisfies

$$(T_\mu^N u_N, T_\mu^N v)_Y = \ell(T_\mu^N v),\ \forall v \in W_N. \tag{5.28}$$

We now choose $v = u_N$ in (5.28) and note that, since $T_\mu^N u_N$ is the supremizer over V_N associated with u_N,

$$(T_\mu^N u_N, T_\mu^N u_N)_Y = a(u_N, T_\mu^N u_N; \mu) \geq \beta_N \|u_N\|_Y \|T_\mu^N u_N\|_Y,$$

and thus

$$\|u_N\|_Y \leq \frac{1}{\beta_N}\|\ell\|_{Y'} \leq \frac{1}{\tilde{\beta}_0}\|\ell\|_{Y'}.$$

We have thus proven stability; uniqueness follows in the usual way by considering two candidate solutions. □

The following lemma, connects minimum-residual with more standard Galerkin or Petrov-Galerkin methods:

Lemma 5.3.2. *If $\beta_N(\mu) \geq \tilde{\beta}_0 > 0$ and $V_N = Z_N(\mu)$, the minimum-residual statement is equivalent to a standard Petrov-Galerkin approximation: $u_N(\mu) = u_N^{\mathrm{PG}}(\mu)$, where $u_N^{\mathrm{PG}}(\mu) \in W_N$ satisfies*

$$R^{\mathrm{pr}}(v; u_N^{\mathrm{PG}}(\mu); \mu) = 0, \ \forall v \in Z_N(\mu); \tag{5.29}$$

an analogous result applies for the dual.

Proof. For $V_N = Z_N(\mu)$ and from standard arguments we know that, if $\beta_N(\mu) \geq \tilde{\beta}_0 > 0$, the Petrov-Galerkin approximation (5.29) admits a unique solution $u_N^{\mathrm{PG}}(\mu)$. But since $\|P_\mu^N u_N^{\mathrm{PG}}\|_Y = 0$, u_N^{PG} must be the (unique) residual minimizer, and hence $u_N = u_N^{\mathrm{PG}}$. □

Remark 5.3.3. *Using the same argument we can prove that for $V_N = W_N$, then $u_N(\mu) = u_N^{\mathrm{Gal}}(\mu)$, where $u_N^{\mathrm{Gal}}(\mu) \in W_N$ satisfies*

$$R^{\mathrm{pr}}(v; u_N^{\mathrm{Gal}}(\mu); \mu) = 0, \ \forall v \in W_N. \tag{5.30}$$

So, for specific choices of V_N, we recover from the minimum-residual statement the Galerkin and Petrov-Galerkin as special cases. We can then prove that $u_N(\mu), \psi_N(\mu)$ are optimal. Indeed, we have

Lemma 5.3.4. *If $\beta_N(\mu) \geq \tilde{\beta}_0 > 0, \ \forall \mu \in \mathcal{D}$, then*

$$\|u(\mu) - u_N(\mu)\|_Y \leq \left(1 + \frac{2\gamma}{\tilde{\beta}_0}\right) \inf_{w_N \in W_N} \|u(\mu) - w_N\|_Y,$$

with an analogous result for the dual.

Proof. Since for any $w_N \in W_N$, $w_N - u_N$ is an element of W_N, we have from (5.26) that

$$\begin{aligned}
&\beta_N \|w_N - u_N\|_Y \|T_\mu^N(w_N - u_N)\|_Y \\
&\qquad \le a(w_N - u_N, T_\mu^N(w_N - u_N); \mu) \\
&\qquad = a(w_N - u + u - u_N, T_\mu^N(w_N - u_N); \mu) \\
&\qquad \le |a(w_N - u, T_\mu^N(w_N - u_N); \mu)| + |a(u - u_N, T_\mu^N(w_N - u_N); \mu| \\
&\qquad = |R^{\mathrm{pr}}(T_\mu^N(w_N - u_N); w_N; \mu)| + |R^{\mathrm{pr}}(T_\mu^N(w_N - u_N); u_N; \mu)| \\
&\qquad \le (\|P_\mu^N w_N\|_Y + \|P_\mu^N u_N\|_Y)\ \|T_\mu^N(w_N - u_N)\|_Y \\
&\qquad \le 2\|P_\mu^N w_N\|_Y \|T_\mu^N(w_N - u_N)\|_Y,
\end{aligned} \tag{5.31}$$

where the last three steps follow from (5.18), (5.22), and (5.24), respectively. We now take $v = P_\mu^N w_N \in V_N$ in (5.22) and apply (5.18) and continuity to obtain

$$\|P_\mu^N w_N\|_Y \le \gamma \|u - w_N\|_Y, \tag{5.32}$$

which then yields, with (5.31),

$$\|w_N - u_N\|_Y \le \frac{2\gamma}{\beta_N} \|u - w_N\|_Y, \ \forall w_N \in W_N. \tag{5.33}$$

The desired result then follows by expressing $\|u - u_N\|_Y$ as $\|u - w_N + w_N - u_N\|_Y$ and applying the triangle inequality, (5.33), and our hypothesis on $\beta_N(\mu)$. □

Remark 5.3.5. *In the case of (Petrov-)Galerkin we can show an improved result:*

$$\|u(\mu) - u_N(\mu)\|_Y \le \left(1 + \frac{\gamma}{\tilde{\beta}_0}\right) \inf_{w_N \in W_N} \|u(\mu) - w_N\|_Y.$$

Finally, we prove that our output prediction is optimal in the following Proposition:

Proposition 4. *If* $\beta_N(\mu) \ge \tilde{\beta}_0 > 0, \ \forall \mu \in \mathcal{D}$,

$$|(s - s_N)(\mu)| \le \gamma \|u - u_N\|_Y \|\psi - \psi_N\|_Y;$$

if furthermore $V_N \supseteq W_N$, which is satisfied for the choices $V_N = Y$ or $V_N = W_N$, we show that:

$$|(s - s_N)(\mu)| \leq \gamma \left(1 + \frac{4\gamma}{\tilde{\beta}_0}\right) \inf_{w_N \in W_N} \|u - w_N\|_Y \inf_{\varphi_N \in W_N} \|\psi - \varphi_N\|_Y.$$

Proof. We have that

$$\begin{aligned} |(s - s_N)(\mu)| &= |\ell^O(u) - \ell^O(u_N) + \ell(\psi_N) - a(u_N, \psi_N; \mu)| \\ &= |-a(u - u_N, \psi; \mu) + a(u - u_N, \psi_N; \mu)| \\ &= |a(u - u_N, \psi - \psi_N; \mu)| \\ &\leq \gamma \|u - u_N\|_Y \|\psi - \psi_N\|_Y, \end{aligned} \tag{5.34}$$

which proves the first result.

We also know that, for all $\varphi_N \in W_N, w_N \in W_N$,

$$\begin{aligned} &|a(u - u_N, \psi - \psi_N; \mu)| \\ &\quad = |a(u - u_N, \psi - \varphi_N + \varphi_N - \psi_N; \mu)| \\ &\quad \leq |a(u - u_N, \psi - \varphi_N; \mu)| + |a(u - u_N, \varphi_N - \psi_N; \mu)| \\ &\quad \leq \gamma \|u - u_N\|_Y \|\psi - \varphi_N\|_Y + |R^{\mathrm{pr}}(\varphi_N - \psi_N; u_N; \mu)| \\ &\quad \leq \gamma \left(1 + \frac{2\gamma}{\tilde{\beta}_0}\right) \|u - w_N\|_Y \|\psi - \varphi_N\|_Y + \|\varphi_N - \psi_N\|_Y \sup_{v \in V_N} \frac{R^{\mathrm{pr}}(v; u_N; \mu)}{\|v\|_Y}, \end{aligned} \tag{5.35}$$

where we have evoked continuity, (5.18), Lemma 5.3.4, and $W_N \subseteq V_N$. But from (5.22), (5.24), (5.32), and the dual version of (5.33)

$$\begin{aligned} \sup_{v \in V_N} \frac{R^{\mathrm{pr}}(v; u_N; \mu)}{\|v\|_Y} \|\varphi_N - \psi_N\|_Y &\leq \|P_\mu^N u_N\|_Y \frac{2\gamma}{\tilde{\beta}_0} \|\psi - \varphi_N\|_Y \\ &\leq \gamma \|u - w_N\|_Y \frac{2\gamma}{\tilde{\beta}_0} \|\psi - \varphi_N\|_Y, \end{aligned}$$

which with (5.34) and (5.35) proves the second result. □

Remark 5.3.6. *In the case of (Petrov-)Galerkin we can show the following improved result:*

$$|(s - s_N)(\mu)| \leq \gamma \left(1 + \frac{\gamma}{\tilde{\beta}_0}\right) \inf_{w_N \in W_N} \|u - w_N\|_Y \inf_{\varphi_N \in W_N} \|\psi - \varphi_N\|_Y.$$

Notice that in the previous estimates only the infimizing space W_N appears. As we will see in Section 5.3.3 the choice of the supremizing space V_N is related to stability, whereas as we see here the choice of infimizing space W_N is related to approximation.

Proposition 4 indicates in what sense reduced-basis methods yield optimal interpolations in parameter space. We could of course predict $s(\mu)$ at some new value of μ as some interpolant or fit to the $s(\mu_m), m = 1, \ldots, M$; however, it is not clear how to choose, or whether one has chosen, the best combination of the $s(\mu_m)$, in particular in higher dimensional parameter spaces. In contrast, Proposition 4 states that, by predicting $s(\mu)$ via a state space (W_N), and by ensuring stability ($\tilde{\beta}_0 > 0$ independent of N), we obtain in some sense a best approximation — the correct weights for each of the reduced-basis components. With sufficient smoothness, this best approximation will converge very rapidly with increasing N [33, 91] (see also Section 5.3.3). Note the importance of $W_{M_2}^{\psi}$ in W_N in ensuring that $\inf_{\varphi_N \in W_N} \|\psi - \varphi_N\|_Y$ is small — had we included only $W_{M_1}^{u}$ in W_N, this would not be the case, since reduced-basis spaces have no general approximation properties (that is, for arbitrary functions in Y).

Of course, Proposition 4 only tells us that we are doing as well as possible; it does not tell us *how* well we are doing — our *a posteriori* estimators are required for that purpose.

5.3.2 A posteriori theory

We can directly show that, under certain hypotheses, our error estimators are indeed *error bounds*.

Proposition 5. *If $\beta_N \to \beta$ as $N \to \infty$, then there exists an $N^*(\mu)$ such that, $\forall N \geq N^*(\mu)$,*

$$|(s - s_N)(\mu)| \leq \Delta_N(\mu),$$

for $\Delta_N(\mu)$ as given in (5.27).

Proof. We first note, evoking symmetry and our inf-sup condition (5.5), that

$$\begin{aligned}\beta(\mu)\|\psi-\psi_N\|_Y\|T_\mu(\psi-\psi_N)\|_Y &\le a(T_\mu(\psi-\psi_N),\psi-\psi_N;\mu)\\ &= R^{\mathrm{du}}(T_\mu(\psi-\psi_N);\psi_N;\mu)\\ &\le \|R^{\mathrm{du}}(\cdot;\psi_N;\mu)\|_{Y'}\|T_\mu(\psi-\psi_N)\|_Y,\end{aligned}$$

or

$$\|\psi-\psi_N\|_Y \le \frac{1}{\beta(\mu)}\|R^{\mathrm{du}}(\cdot;\psi_N;\mu)\|_{Y'}.$$

We then write, from (5.34) of Proposition 4,

$$\begin{aligned}|(s-s_N)(\mu)| &= |a(u-u_N,\psi-\psi_N;\mu)|\\ &= |R^{\mathrm{pr}}(\psi-\psi_N;u_N;\mu)|\\ &\le \|R^{\mathrm{pr}}(\cdot;u_N;\mu)\|_{Y'}\|\psi-\psi_N\|_Y\\ &\le \frac{1}{\beta(\mu)}\|R^{\mathrm{pr}}(\cdot;u_N;\mu)\|_{Y'}\|R^{\mathrm{du}}(\cdot;\psi_N;\mu)\|_{Y'}.\end{aligned}$$

The result then directly follows: for $\sigma<1$, we have from our hypothesis on β_N that $\sigma\beta_N(\mu)\le\beta(\mu)$ for N sufficiently large, say $N\ge N^*(\mu)$, and thus

$$\begin{aligned}|(s-s_N)(\mu)| &\le \frac{1}{\beta(\mu)}\|R^{\mathrm{pr}}(\cdot;u_N;\mu)\|_{Y'}\|R^{\mathrm{du}}(\cdot;\psi_N;\mu)\|_{Y'}\\ &\le \frac{1}{\sigma\beta_N(\mu)}\|R^{\mathrm{pr}}(\cdot;u_N;\mu)\|_{Y'}\|R^{\mathrm{du}}(\cdot;\psi_N;\mu)\|_{Y'}\\ &= \Delta_N(\mu),\end{aligned}$$

for $N\ge N^*(\mu)$. □

It is not only important to determine that $\Delta_N(\mu)$ is a bound for the error, but also that it is a *good* bound. As a measure of bound quality, we introduce the usual *a posteriori* effectivity,

$$\eta_N(\mu)=\frac{\Delta_N(\mu)}{|s(\mu)-s_N(\mu)|}. \tag{5.36}$$

Under the hypothesis of Proposition 5 we know that $\eta_N(\mu)\ge 1$ as $N\to\infty$, providing us

with the desired bounds; to ensure that the bound is tight, we would also like to verify that $\eta_N(\mu) \leq$ Const (independent of N) as $N \to \infty$. We have no proof for this result, but it is certainly plausible given the demonstration of Proposition 5, and is in fact confirmed by the numerical experiments of Section 5.5.

5.3.3 The discrete inf-sup parameter

It should be clear that good behavior of the discrete inf-sup parameter is the essential hypothesis of Propositions 4 and 5. If $\beta_N(\mu)$ vanishes, or becomes very small relative to $\beta(\mu)$, Propositions 4 and 5 indicate we risk that $|(s - s_N)(\mu)|$ and $\Delta_N(\mu)$ will both become very large: accuracy of our predictions thus requires $\beta_N(\mu) \geq \tilde{\beta}_0 > 0$. However, too much stability is not desirable, either. If $\beta_N(\mu)$ is large compared to $\beta(\mu)$ as $N \to \infty$, Proposition 5 indicates we risk that $\Delta_N(\mu)$ will not bound $|(s - s_N)(\mu)|$: certainty in our predictions thus requires $\beta_N(\mu)$ close to $\beta(\mu)$. It is clear that the best behavior is $\beta_N \to \beta$ from above as $N \to \infty$.

We now discuss several possible choices for V_N, W_N, and the extent to which each — either provably or intuitively — meets our desiderata.

5.3.4 The choice $V_N = Y, W_N = W_N^1$ — Method 1

It is simple in this case to prove stability:

Lemma 5.3.7. *For $V_N = Y$ (and any space $W_N \subset Y$),*

$$\beta_N(\mu) \geq \beta(\mu) \geq \beta_0 > 0,$$

for all $\mu \in \mathcal{D}$.

Proof. We have

$$\beta_N(\mu) = \inf_{w_N \in W_N} \frac{\|T_\mu^N w_N\|_Y}{\|w_N\|_Y} = \inf_{w_N \in W_N} \frac{\|T_\mu w_N\|_Y}{\|w_N\|_Y} \geq \inf_{w \in Y} \frac{\|T_\mu w\|_Y}{\|w\|_Y} = \beta(\mu) \geq \beta_0 > 0,$$

as desired. □

Thus, for $V_N = Y$, the hypothesis of Proposition 4 is satisfied with $\tilde{\beta}_0 = \beta_0$; we are guaranteed stability. To ensure accuracy of the inf-sup parameter — and hence asymptotic error *bounds* from Proposition 5 — we shall first need

Lemma 5.3.8. *If $\mathcal{S}_{M_3}$ is chosen such that for $\mu_m \in \mathcal{S}_{M_3}$*

$$\sup_{\mu\in\mathcal{D}} \inf_{m\in\{1,\ldots,M\}} \|\mu - \mu_m\| \to 0$$

as $M \to \infty$, and if $\chi(\mu)$ is sufficiently smooth in the sense that

$$\| \sup_{\mu\in\mathcal{D}} \|\nabla_\mu \chi\| \|_Y < \infty,$$

then

$$\inf_{w_N \in W_N^1} \|\chi(\mu) - w_N\|_Y \to 0, \ \forall \mu \in \mathcal{D}, \tag{5.37}$$

as M_3 (and hence N)$\to \infty$. Note $\|\cdot\|$ refers to the usual Euclidean norm.

Proof. Recalling that $\chi(\mu)$, the infimizer, is defined in (5.7), we next introduce $\tilde{\chi}_N(\mu) \in W_N^1$ as

$$\tilde{\chi}_N(\mu) = \chi(\mu_{m^*(\mu)}), \ m^*(\mu) = \arg\min_{m\in\{1,\ldots,M\}} |\mu - \mu_m|.$$

Thus

$$\begin{aligned}\|\chi(\mu) - \tilde{\chi}_N(\mu)\|_Y &\le (\inf_{m\in\{1,\ldots,M\}} \|\mu - \mu_m\|) \, \| \sup_{\mu\in\mathcal{D}} \|\nabla_\mu \chi\| \, \|_Y \\ &\le (\sup_{\mu\in\mathcal{D}} \inf_{m\in\{1,\ldots,M\}} \|\mu - \mu_m\|) \| \sup_{\mu\in\mathcal{D}} \|\nabla_\mu \chi\| \, \|_Y, \ \forall \mu \in \mathcal{D},\end{aligned}$$

and therefore for all $\mu \in \mathcal{D}$,

$$\begin{aligned}\inf_{w_N\in W_N} \|\chi(\mu) - w_N\|_Y &\le \|\chi(\mu) - \tilde{\chi}_N(\mu)\|_Y \\ &\le (\sup_{\mu\in\mathcal{D}} \inf_{m\in\{1,\ldots,M\}} \|\mu - \mu_m\|) \, \| \sup_{\mu\in\mathcal{D}} \|\nabla_\mu \chi\| \, \|_Y,\end{aligned}$$

which tends to zero as M_3 (and hence N) tends to infinity from our hypotheses on $\mathcal{S}_{M_3}$ and the smoothness of $\chi(\mu)$. □

Clearly, with sufficient smoothness, we can develop higher order interpolants [91], suggesting correspondingly higher rates of convergence. For our purposes here, (5.37) suffices; the method itself will choose a best approximation, typically much closer to $\chi(\mu)$ than our simple candidate above. The essential point is the inclusion of $W^{\chi}_{M_3}$ in W^1_N, which provides the necessary approximation properties within our reduced-basis space. We can now prove that, for $V_N = Y,\ W_N = W^1_N$, $\beta_N(\mu)$ is an accurate approximation to $\beta(\mu)$.

Proposition 6. *For $V_N = Y,\ W_N = W^1_N$,*

$$\beta_N(\mu_m) = \beta(\mu_m),\ \ m = 1, \ldots, M_3,\ \ \mu_m \in \mathcal{S}_{M_3}. \tag{5.38}$$

Furthermore, under the hypotheses of Lemma 5.3.8, there exists a C independent of N and an $N^{**}(\mu)$ such that

$$|\beta(\mu) - \beta_N(\mu)| \le C\frac{\gamma^2}{2\beta(\mu)} \inf_{w_N \in W_N} \|\chi(\mu) - w_N\|^2_Y,\ \ \forall N \ge N^{**}(\mu), \tag{5.39}$$

and thus from Lemma 5.3.8,

$$\beta_N(\mu) \to \beta(\mu)\ \ as\ N \to \infty,\ \ \forall \mu \in \mathcal{D}. \tag{5.40}$$

Proof. To prove (5.38), we note that, since $\chi(\mu_m) \in W_N$,

$$\beta_N(\mu_m) = \inf_{w_N \in W_N} \frac{\|T_{\mu_m} w_N\|_Y}{\|w_N\|_Y} \le \frac{\|T_{\mu_m}\chi(\mu_m)\|_Y}{\|\chi(\mu_m)\|_Y} = \beta(\mu_m);$$

but $\beta_N(\mu_m) \ge \beta(\mu_m)$ from Lemma 5.3.7, and thus $\beta_N(\mu_m) = \beta(\mu_m)$. To prove (5.39), we introduce the discrete eigenproblem analogous to (5.7): find $(\theta_N, \lambda_N) \in (W^1_N \times \mathbb{R})$ such that

$$\mathcal{A}(\theta_N(\mu), v; \mu) = \lambda_N(\mu)(\theta_N(\mu), v)_Y,\ \ \forall v \in W^1_N, \quad \|\theta_N(\mu)\|_Y = 1;$$

by arguments similar to those of Section 5.1.2 it is simple to show that $\beta_N(\mu) = \sqrt{\lambda_{N\ \min}(\mu)}$. We can now apply the standard theory for Galerkin approximation of symmetric positive-definite eigenproblems. To wit, from Theorem 9 of [9] and (5.37) of our Lemma 5.3.8, there

exists an $N^{**}(\mu)$ such that, $\forall N \geq N^{**}(\mu)$,

$$|\lambda_{\min}(\mu) - \lambda_{N\ \min}(\mu)| \leq C\mathcal{A}(\theta_{\min} - w_N, \theta_{\min} - w_N; \mu),\ \forall w_N \in W_N,$$

for C independent of N. (One can in fact show that C may be taken as $(1 + 2\beta^3)^2$.)

But from (5.6) and (5.8) of Section 5.1.2, we know that

$$\mathcal{A}(\theta_{\min} - w_N, \theta_{\min} - w_N; \mu) = \|T_\mu(\theta_{\min} - w_N)\|_Y^2 \leq \gamma^2 \|\theta_{\min} - w_N\|_Y^2.$$

The result (5.39) then follows by recalling that $\theta_{\min} = \chi$, $(\beta_N)^2 = \lambda_{N\ \min}$, $(\beta)^2 = \lambda_{\min}$, and noting that

$$|(\beta_N)^2 - \beta^2| = |(\beta_N - \beta)|\,|(\beta_N + \beta)| \geq |(\beta_N - \beta)|2\beta$$

since $\beta_N \geq \beta$ from Lemma 5.3.7. □

The hypothesis of Proposition 5 is thus verified for the case $V_N = Y$, $W_N = W_N^1$. The quadratic convergence of $\beta_N(\mu)$ is very important: it suggests an accurate prediction for $\beta(\mu)$ — and hence bounds — even if W_N is rather marginal.

5.3.5 The Choice $V_N = Y$, $W_N = W_N^0$ — Method 2

In this case the $\chi(\mu_m), m = 1, \ldots, M$, are no longer members of W_N. We see that Lemma 5.3.7 still obtains, and thus the method is stable — in fact, always *at least as stable* as $W_N = W_N^1$. Furthermore, since W_N^0 still contains $W_{M_1}^u$ and $W_{M_2}^\psi$, we expect $\|u - u_N\|_Y$ and $\|\psi - \psi_N\|_Y$ to be small, and hence from Proposition4 $|(s - s_N)(\mu)|$ should also be small. There is no difficulty at the level of *stability* or *accuracy of our output*.

However, Lemma 5.3.8 can no longer be proven. Thus not only is (5.38) of Proposition 6 obviously not applicable, but — and even more importantly — (5.40) no longer obtains: we can not expect $\beta_N(\mu)$ to tend to $\beta(\mu)$ as $N \to \infty$. In short, the scheme may be too stable, $\beta_N(\mu)$ may be too large, and hence for any fixed $\sigma < 1$ we may not obtain bounds even as $N \to \infty$. In short, in contrast to the choice $W_N = W_N^1$, the choice $W_N = W_N^0$ no longer ensures that $\beta_N(\mu)$ is sufficiently accurate.

In practice, however, $\beta_N(\mu)$ may be sufficiently close to $\beta(\mu)$ that $\sigma\beta_N(\mu) \leq \beta(\mu)$ for some suitably small σ. To understand why, we observe that, in terms of our eigenpairs (Υ_i, ω_i) of Section 5.1.2,

$$u(\mu_m) = \sum_{i=1}^{\mathcal{N}} \frac{\ell(\Upsilon_i)}{\omega_i(\mu_m) - 1} \Upsilon_i. \tag{5.41}$$

For "generic" ℓ, $u(\mu_m)$ will thus contain a significant component of $\Upsilon_{i^*(\mu_m)}$ and hence $\chi(\mu_m)$. It is possible to construct ℓ such that $\ell(\Upsilon_{i^*}) = 0$, and hence we cannot in general *count on* $\chi(\mu_m)$ being predominantly present in W_N^0; however, in practice, ℓ will typically be broadband, and thus $W_N = W_N^0$ may sometimes be sufficient. Obviously, for greater certainty that our error bound is, indeed, a bound, $W_N = W_N^1$ is unambiguously preferred over $W_N = W_N^0$.

5.3.6 The Choice $V_N = W_N^1$, $W_N = W_N^1$ — Method 3

We know from Remark 5.3.3 that this case corresponds to Galerkin approximation, but with $W_{M_3}^{\chi}$ present in our spaces. We first note that not only does Lemma 5.3.7 not apply, but unfortunately we can prove that for $\mu_m \in \mathcal{S}_{M_3}$, $\beta_N(\mu_m) \leq \beta(\mu_m)$, $m = 1, \ldots, M$:

$$\beta(\mu_m) = \sup_{v \in Y} \frac{a(\chi(\mu_m), v; \mu_m)}{\|\chi(\mu_m)\|_Y \|v\|_Y} \geq \sup_{v \in W_N^1} \frac{a(\chi(\mu_m), v; \mu_m)}{\|\chi(\mu_m)\|_Y \|v\|_Y} \geq \inf_{w \in W_N^1} \sup_{v \in W_N^1} \frac{a(w, v; \mu_m)}{\|w\|_Y \|v\|_Y} = \beta_N(\mu_m), \tag{5.42}$$

since $\chi(\mu_m) \in W_N^1 \subset Y$. Stability and accuracy of the output could thus be an issue, though not necessarily so if $\beta_N(\mu)$ is close to $\beta(\mu)$. As regards the accuracy of $\beta_N(\mu)$, Lemma 5.3.8 still applies, however (5.38), (5.39) (and hence (5.40)) of Proposition 6 can no longer be readily proven.

Nevertheless, in practice, $\beta_N(\mu)$ may be quite close to $\beta(\mu)$. To understand why, we recall from Section 5.1.2 that $\chi(\mu_m)$ is not only our infimizer, but also proportional to $T_{\mu_m}\chi(\mu_m)$. It follows that *if* $\chi(\mu_m)$ is the most dangerous mode in the sense that

$$\sup_{v \in W_N^1} \frac{a(\chi(\mu_m), v; \mu_m)}{\|\chi(\mu_m)\|_Y \|v\|_Y} \leq \sup_{v \in W_N^1} \frac{a(w, v; \mu_m)}{\|w\|_Y \|v\|_Y}, \ \forall w \in W_N^1, \tag{5.43}$$

then

$$\beta_N(\mu_m) = \sup_{v \in W_N^1} \frac{a(\chi(\mu_m), v; \mu_m)}{\|\chi(\mu_m)\|_Y \|v\|_Y} \geq \frac{a(\chi(\mu_m), T_{\mu_m}\chi(\mu_m); \mu_m)}{\|\chi(\mu_m)\|_Y \|T_{\mu_m}\chi(\mu_m)\|_Y} = \beta(\mu_m),$$

since both $\chi(\mu_m)$ and $T_{\mu_m}\chi(\mu_m)$ are in W_N^1; note that (5.43) is a conjecture, since the supremizing space here is W_N^1, *not* Y as in Section 5.1.2. Under our assumption (5.43) we thus conclude from (5.42) that

$$\beta_N(\mu_m) = \beta(\mu_m). \tag{5.44}$$

By similar arguments we might expect $\beta_N(\mu)$ to be quite accurate even for general $\mu \in \mathcal{D}$, as both $\chi(\mu)$ and $T_\mu\chi$ are well represented in the basis. (From this discussion we infer that a Petrov-Galerkin formulation is desirable — see Section 5.3.8.) The above arguments are clearly speculative. In order to more rigorously guide our choice of V_N, we can prove an illustrative relationship between the Galerkin $V_N = W_N^1$, $W_N = W_N^1$ (superscript "Gal") and minimum residual $V_N = Y$, $W_N = W_N^1$ (superscript "MR") approximations:

Proposition 7. *For all $\mu \in \mathcal{D}$,*

$$\Delta_N^{\mathrm{MR}}(\mu) \leq \Delta_N^{\mathrm{Gal}}(\mu), \tag{5.45}$$

where $\Delta_N^{\mathrm{MR}}(\mu)$ and $\Delta_N^{\mathrm{Gal}}(\mu)$ refer to (5.27) for the minimum-residual and Galerkin cases, respectively.

Proof. We first note that

$$\beta_N^{\mathrm{Gal}}(\mu) = \inf_{w \in W_N^1} \sup_{v \in W_N^1} \frac{a(w, v; \mu)}{\|w\|_Y \|v\|_Y} \leq \inf_{w \in W_N^1} \sup_{v \in Y} \frac{a(w, v; \mu)}{\|w\|_Y \|v\|_Y} = \beta_N^{\mathrm{MR}}(\mu), \tag{5.46}$$

for all $\mu \in \mathcal{D}$. We then note from (5.24) that

$$\begin{aligned}
\Delta_N^{\mathrm{MR}} &\equiv \frac{1}{\sigma \beta_N^{\mathrm{MR}}} \|R^{\mathrm{pr}}(\cdot; u_N^{\mathrm{MR}}; \mu)\|_{Y'} \|R^{\mathrm{du}}(\cdot; \psi_N^{\mathrm{MR}}; \mu)\|_{Y'} \\
&= \frac{1}{\sigma \beta_N^{\mathrm{MR}}} \|P_\mu^N u_N^{\mathrm{MR}}\|_Y \|D_\mu^N \psi_N^{\mathrm{MR}}\|_Y \\
&\leq \frac{1}{\sigma \beta_N^{\mathrm{MR}}} \|P_\mu^N w_N\|_Y \|D_\mu^N \varphi_N\|_Y, \ \forall w_N \in W_N^1, \ \forall \varphi_N \in W_N^1,
\end{aligned}$$

where $P_\mu^N : W_N^1 \to Y$ and $D_\mu^N : W_N^1 \to Y$ are here defined for $V_N = Y$. Thus

$$\begin{aligned}\Delta_N^{\mathrm{MR}}(\mu) &\le \frac{1}{\sigma\beta_N^{\mathrm{MR}}}\|P_\mu^N u_N^{\mathrm{Gal}}\|_Y \|D_\mu^N \psi_N^{\mathrm{Gal}}\|_Y \\ &= \frac{1}{\sigma\beta_N^{\mathrm{MR}}}\|R^{\mathrm{pr}}(\cdot; u_N^{\mathrm{Gal}}; \mu)\|_{Y'} \|R^{\mathrm{du}}(\cdot; \psi_N^{\mathrm{Gal}}; \mu)\|_{Y'} \\ &\le \frac{1}{\sigma\beta_N^{\mathrm{Gal}}}\|R^{\mathrm{pr}}(\cdot; u_N^{\mathrm{Gal}}; \mu)\|_{Y'} \|R^{\mathrm{du}}(\cdot; \psi_N^{\mathrm{Gal}}; \mu)\|_{Y'} = \Delta_N^{Gal}(\mu),\end{aligned}$$

where the last step follows from (5.46). □

We thus see that, in general, $V_N = Y$ will provide sharper error estimates: minimum residual is in fact equivalent to minimum error bound. Conversely, we might expect the Galerkin approximation to be more conservative, providing bounds even when the minimum-residual approach may not (e.g. for N very small).

The Galerkin method with $W_N = W_N^1$ thus has some redeeming features. However, there is the possibility of a loss of accuracy in both $s_N(\mu)$ and $\Delta_N(\mu)$, reflected in (5.45) and (5.46) of Proposition 7.

5.3.7 The Choice $V_N = W_N^0$, $W_N = W_N^0$ — Method 4

This case corresponds to Galerkin approximation, but now with $W_{M_3}^\chi$ absent. Here (5.42) no longer applies: $\beta_N(\mu)$ may be greater or less than $\beta(\mu)$. Furthermore, accuracy of $\beta_N(\mu)$ now relies on two fortuitous events — the "selective amplification" of (5.41) and the "most dangerous mode" of (5.43). Again, in practice, the method may perform well, but it is now more likely that either $\beta_N(\mu)$ will be too small and hence $(s - s_N)(\mu)$ and $\Delta_N(\mu)$ too large, or $\beta_N(\mu)$ will be too large and hence $\eta_N(\mu) < 1$ (no bounds).

We are able to prove a result analogous to Proposition 7, but now comparing $V_N = Y$, $W_N = W_N^0$ to $V_N = W_N^0$, $W_N = W_N^0$: the $\beta_N(\mu)$ (respectively $\Delta_N(\mu)$) for the former will be larger (respectively smaller) than the corresponding quantities for the latter. We thus expect that $V_N = W_N^0$, $W_N = W_N^0$ will yield rather poor accuracy.

5.3.8 The Choice $V_N = Z_N(\mu),\ W_N = W_N^1$ — Method 5

Following Lemma 5.3.2, we see that this choice of infimizing and supremizing spaces corresponds to a Petrov-Galerkin approximation, with the infimizers included. In the following we use the superscript "PG" to specify Petrov-Galerkin approximations. Regarding stability we have the following

Proposition 8. *For $V_N = Z_N(\mu)$ and $W_N = W_N^1$*

$$\beta_N^{\mathrm{MR}}(\mu) = \beta_N^{\mathrm{PG}}(\mu), \tag{5.47}$$

for all $\mu \in \mathcal{D}$.

Proof. To start recall that for $w_N \in W_N^1$, $T_\mu w_N \in Y$ is obtained from:

$$(T_\mu w_N, v)_Y = a(w_N, v; \mu),\ \forall v \in Y.$$

Since $w_N \in W_N^1$, we write $w_N = \sum_{i=1}^N c_i \zeta_i$, and using the affine decomposition assumption (5.4) we have

$$\begin{aligned}(T_\mu w_N, v)_Y &= \sum_{q=1}^{Q} \sigma^q(\mu) a^q(w_N, v) \\ &= \sum_{q=1}^{Q}\sum_{i=1}^{N} \sigma^q(\mu) c_i a^q(\zeta_i, v) \\ &= \sum_{i=1}^{N} c_i (\sum_{q=1}^{Q} \sigma^q(\mu) z^{n,q}, v)_Y;\end{aligned}$$

from the definition of $z^{n,q}$, (5.16). Therefore,

$$T_\mu w_N = \sum_{i=1}^{N} c_i (\sum_{q=1}^{Q} \sigma^q(\mu) z^{n,q}), \tag{5.48}$$

and from the definition of $V_N = Z_N(\mu)$ we see that $T_\mu w_N \in V_N$. Therefore for $T_\mu^N w_N$ defined:

$$(T_\mu^N w_N, v)_Y = a(w_N, v; \mu),\ \forall v \in Z_N(\mu),$$

we conclude that $T_\mu^N w_N = T_\mu w_N$. We then have for the inf-sup parameter:

$$\beta_N^{\mathrm{PG}}(\mu) = \inf_{w_N \in W_N^1} \frac{\|T_\mu^N w_N\|_Y}{\|w_N\|_Y} = \inf_{w_N \in W_N^1} \frac{\|T_\mu w_N\|_Y}{\|w_N\|_Y} = \beta_N^{\mathrm{MR}}(\mu).$$

□

Unlike the minimum-residual, the construction of (5.16) ensures stability *only* for members of W_N. The results of Lemma 5.3.8 and Proposition 6, as well as the comparisons for stability with the other methods, apply here — in terms of stability, minimum-residual and Petrov-Galerkin are identical. Also, for the choice $W_N = W_N^0$, the method presented here is similar to Method 2.

Remark 5.3.9. *The critical ingredient in the previous Proposition is to ensure that $T_\mu w_N \in V_N$. The discussion in Section 5.1.2 suggests that, the infimizers $\chi(\mu_m)$ are also parallel to the supremizers $T_{\mu_m}\chi(\mu_m)$, $\forall \mu_m \in \mathcal{S}_{M_3}$. Thus, instead of computing the $z^{n,q}$ for the members of $W_{M_3}^\chi$, we can directly include $W_{M_3}^\chi$ in $Z_N(\mu)$. It is then easy to see that Proposition 8 will still be true. For this choice, significant savings in storage and computational effort should be expected.*

Regarding the solution and the error estimator we have the following:

Proposition 9. *For all $\mu \in \mathcal{D}$,*

$$\begin{aligned} &u_N^{\mathrm{MR}}(\mu) = u_N^{\mathrm{PG}}(\mu), \quad \psi_N^{\mathrm{MR}}(\mu) = \psi_N^{\mathrm{PG}}(\mu), \text{ and} \\ &\Delta_N^{\mathrm{MR}}(\mu) = \Delta_N^{\mathrm{PG}}(\mu). \end{aligned} \tag{5.49}$$

Proof. We first prove that $u_N^{\mathrm{MR}}(\mu) = u_N^{\mathrm{PG}}(\mu)$; the proof for the dual solution is similar. From the minimum-residual statement, $u_N^{\mathrm{MR}}(\mu)$, can be obtained as the solution of the following problem:

$$(T_\mu u_N^{\mathrm{MR}}, T_\mu w_N)_Y = \ell(T_\mu w_N), \ \forall w_N \in W_N;$$

where T_μ is defined as

$$(T_\mu w_N, v)_Y = a(w_N, v; \mu), \ \forall v \in Y.$$

From the stability and continuity of a, it is simple to establish that the mapping between $w_N \rightarrow T_\mu w_N$ is a bijection and therefore since W_N is an N-dimensional space, then $T_\mu W_N$ will also be an N-dimensional space. Moreover, we showed in Proposition 8, that for all $w_N \in W_N$, then $T_\mu w_N \in Z_N(\mu)$. Since $Z_N(\mu)$ is an N-dimensional space, we thus conclude that $T_\mu W_N \equiv Z_N(\mu)$.

Therefore combining the equations above we have that $u_N^{\mathrm{MR}}(\mu)$ satisfies:

$$\begin{aligned} a(u_N^{\mathrm{MR}}, T_\mu w_N) &= \ell(T_\mu w_N),\ \forall w_N \in W_N \Rightarrow \\ a(u_N^{\mathrm{MR}}, v_N) &= \ell(v_N),\ \forall v_N \in Z_N(\mu) \Rightarrow \\ R^{\mathrm{pr}}(v_N; u_N^{\mathrm{MR}}; \mu) &= 0,\ \forall v_N \in Z_N(\mu); \end{aligned} \tag{5.50}$$

which is nothing more than the definition of the Petrov-Galerkin approximation. Since the solution is unique (it is simple to prove stability), we conclude that $u_N^{\mathrm{MR}}(\mu) \equiv u_N^{\mathrm{PG}}(\mu)$.

For the bound gap we have from the preceding proof and (5.47) that:

$$\begin{aligned} \Delta_N^{\mathrm{MR}}(\mu) &= \frac{1}{\sigma \beta_N^{\mathrm{MR}}(\mu)} = \|R^{\mathrm{pr}}(\cdot; u_N^{\mathrm{MR}}; \mu)\|_{Y'} \|R^{\mathrm{du}}(\cdot; \psi_N^{\mathrm{MR}}; \mu)\|_{Y'} \\ &= \frac{1}{\sigma \beta_N^{\mathrm{PG}}(\mu)} = \|R^{\mathrm{pr}}(\cdot; u_N^{\mathrm{PG}}; \mu)\|_{Y'} \|R^{\mathrm{du}}(\cdot; \psi_N^{\mathrm{PG}}; \mu)\|_{Y'} \\ &= \Delta_N^{\mathrm{PG}}(\mu). \end{aligned}$$

□

We therefore conclude that Method 1 and Method 5 are, in fact, different interpretations/descriptions of the same method. Even though the minimum-residual interpretation is more intuitive, there are certain important advantages to the Petrov-Galerkin approach. First, it is possible to develop more general bound-conditioner-based *a posteriori* error estimation procedures [113]. This will lead to uniform (for all N) rather than asymptotic bounds (for $N > N^*(\mu)$), and also improved bound gaps and effectivities. This development will be considered in a future paper. Second, this method will also be used in the next chapter for the Stokes problem — there the minimum-residual interpretation is not possible.

5.4 Computational procedure

For clarity we shall present the details for the most "rigorous" and general schemes, $V_N = Y$, $W_N = W_N^1$ (Method 1 of Section 5.3.3) and $V_N = Z_N(\mu)$, $W_N = W_N^1$ (Method 5 of Section 5.3.3). The computational procedure — as should be expected — is the same for both interpretations. As we proceed, we indicate simplifications for the other schemes, and at the conclusion we give a comparison of computational complexity.

5.4.1 An algebraic representation

Preliminaries

We assume here that Y is finite dimensional, with associated basis $\xi_i, i = 1, \ldots, \mathcal{N}$. We also recall that W_N can be expressed as $W_N = \text{span}\{\zeta_i, i = 1, \ldots, N\}$; we implicitly assume independence of the reduced-basis functions. A member $w \in Y$ is expressed as $\underline{w}^t \underline{\xi}$, $\underline{w} \in \mathbb{R}^{\mathcal{N}}$; a member $w \in W_N$ is expressed as $\underline{w}^t \underline{\zeta}$, $\underline{w} \in \mathbb{R}^N$. Here t denotes the transpose.

We next introduce the matrices $\underline{A}^{Y,Y} \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$, $\underline{A}^{W_N,W_N} \in \mathbb{R}^{N\times N}$, $\underline{A}^{Y,W_N} \in \mathbb{R}^{\mathcal{N}\times N}$, $\underline{B}^{Y,Y} \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$, $\underline{B}^{W_N,W_N} \in \mathbb{R}^{N\times N}$, given by

$$
\begin{aligned}
A_{i,j}^{Y,Y}(\mu) &= a(\xi_j, \xi_i; \mu), \ 1 \le i, j \le \mathcal{N}, \\
A_{i,j}^{W_N,W_N}(\mu) &= a(\zeta_j, \zeta_i; \mu), \ 1 \le i, j \le N, \\
A_{i,j}^{Y,W_N}(\mu) &= a(\zeta_j, \xi_i; \mu), \ 1 \le i \le \mathcal{N}, 1 \le j \le N, \\
B_{i,j}^{Y,Y}(\mu) &= (\xi_j, \xi_i)_Y, \ 1 \le i, j \le \mathcal{N}, \\
B_{i,j}^{W_N,W_N}(\mu) &= (\zeta_j, \zeta_i)_Y, 1 \le i, j \le N.
\end{aligned}
$$

From these matrices we can derive four further matrices, $\underline{Z}^{Y,Y} \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$, $\underline{Z}^{W_N,Y} \in \mathbb{R}^{N\times\mathcal{N}}$,

$\underline{S}^{Y,Y} \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$, $\underline{S}^{W_N,W_N} \in \mathbb{R}^{N\times N}$, as

$$\begin{aligned}
\underline{Z}^{Y,Y}(\mu) &= \underline{A}^{Y,Y}(\mu)(\underline{B}^{Y,Y})^{-1}, \\
\underline{Z}^{W_N,Y}(\mu) &= (\underline{A}^{Y,W_N}(\mu))^t(\underline{B}^{Y,Y})^{-1}, \\
\underline{S}^{Y,Y}(\mu) &= (\underline{A}^{Y,Y}(\mu))^t(\underline{B}^{Y,Y})^{-1}\underline{A}^{Y,Y}(\mu), \\
\underline{S}^{W_N,W_N}(\mu) &= (\underline{A}^{Y,W_N}(\mu))^t(\underline{B}^{Y,Y})^{-1}\underline{A}^{Y,W_N}(\mu),
\end{aligned}$$

where ^{t} denotes matrix transpose. These matrices are representations — in terms of our bases — of the operators introduced earlier. For example, if $w \in Y$ and $v \in Y$ are expressed as $\underline{w}^t\underline{\xi}$ and $\underline{v}^t\underline{\xi}$, then $\underline{v}^t\underline{B}^{Y,Y}\ \underline{w}$ is $(w,v)_Y$; $(\underline{Z}^{Y,Y}(\mu))^t$ is our representation of T_μ; and for $w \in Y$ expressed as $\underline{w}^t\underline{\xi}$, $\underline{w}^t\ \underline{S}^{Y,Y}(\mu)\underline{w}$ is $(T_\mu w, T_\mu w)_Y$. It follows that

$$(\beta(\mu))^2 = \min_{\underline{w}\in\mathbb{R}^{\mathcal{N}}} \frac{\underline{w}^t\underline{S}^{Y,Y}(\mu)\underline{w}}{\underline{w}^t\underline{B}^{Y,Y}\underline{w}}. \tag{5.51}$$

Similarly, for $w \in W_N$ expressed as $\underline{w}^t\underline{\zeta}$, $\underline{w}^t\ \underline{S}^{W_N,W_N}\underline{w}$ is $(T_\mu^N w, T_\mu^N w)$, and

$$(\beta_N(\mu))^2 = \min_{\underline{w}\in\mathbb{R}^{N}} \frac{\underline{w}^t\underline{S}^{W_N,W_N}(\mu)\underline{w}}{\underline{w}^t\underline{B}^{W_N,W_N}\underline{w}}. \tag{5.52}$$

Note $\underline{S}^{W_N,W_N}(\mu)$ represents the normal equations associated with the least-squares approach, and $\beta_N(\mu)$ is the "generalized" smallest singular value of $\underline{A}^{Y,W_N}$; see [14] for an earlier discussion of singular values and stability in the reduced-basis context. Finally, we shall need the vectors $\underline{L}^{\mathrm{pr},Y} \in \mathbb{R}^{\mathcal{N}}$, $\underline{L}^{\mathrm{du},Y} \in \mathbb{R}^{\mathcal{N}}$, $\underline{L}^{\mathrm{pr},W_N} \in \mathbb{R}^{N}$, $\underline{L}^{\mathrm{du},W_N} \in \mathbb{R}^{N}$, defined by

$$\begin{aligned}
L_i^{\mathrm{pr},Y} &= \ell(\xi_i), 1 \le i \le \mathcal{N}, & L_i^{\mathrm{du},Y} &= \ell^O(\xi_i), 1 \le i \le \mathcal{N}, \\
L_i^{\mathrm{pr},W_N} &= \ell(\zeta_i), 1 \le i \le N, & L_i^{\mathrm{du},W_N} &= \ell^O(\zeta_i), 1 \le i \le N,
\end{aligned}$$

which are the obvious representations of our primal and dual linear functionals.

Reduced basis

We first find, for $m = 1, \ldots, M$, $\underline{u}_m \in \mathbb{R}^{\mathcal{N}}, \underline{\psi}_m \in \mathbb{R}^{\mathcal{N}}$, solution of

$$\underline{A}^{Y,Y}(\mu_m)\underline{u}_m = \underline{L}^{\mathrm{pr},Y} \quad (\text{or } \underline{S}^{Y,Y}(\mu_m)\underline{u}_m = \underline{Z}^{Y,Y}\underline{L}^{\mathrm{pr},Y}), \tag{5.53}$$

$$\underline{A}^{Y,Y}(\mu_m)\underline{\psi}_m = -\underline{L}^{\mathrm{du},Y} \quad (\text{or } \underline{S}^{Y,Y}(\mu_m)\underline{\psi}_m = -\underline{Z}^{Y,Y}\underline{L}^{\mathrm{du},Y}). \tag{5.54}$$

We further obtain $(\underline{\chi}_m, \kappa_{m\,\min})$ as the first eigenpair $(\underline{\theta}, \kappa) \in (\mathbb{R}^{\mathcal{N}} \times \mathbb{R})$ of the symmetric positive-definite problem

$$\underline{S}^{Y,Y}(\mu_m)\underline{\theta} = \kappa \underline{B}^{Y,Y}\underline{\theta}, \quad \underline{\theta}^t \underline{B}^{Y,Y}\underline{\theta} = 1. \tag{5.55}$$

(Note that in some cases it may be preferable to find $\underline{\chi}_m$ by considering the $(\cdot,\cdot)_Y - \mathcal{B}(\cdot,\cdot;\mu)$ eigenproblem of Section 2.3; an inverse iteration with shift (of unity, initially) may work well.) It can then readily be shown that $u(\mu_m), \psi(\mu_m), \chi(\mu_m)$, $m = 1, \ldots, M$, of (5.11), (5.12), and (5.13) are given by

$$u(\mu_m) = \sum_{j=1}^{\mathcal{N}} u_{m\,j}\xi_j \equiv (\underline{u}_m)^t \underline{\xi},$$

$$\psi(\mu_m) = \sum_{j=1}^{\mathcal{N}} \psi_{m\,j}\xi_j \equiv (\underline{\psi}_m)^t \underline{\xi},$$

$$\chi(\mu_m) = \sum_{j=1}^{\mathcal{N}} \chi_{m\,j}\xi_j \equiv (\underline{\chi}_m)^t \underline{\xi},$$

where this last result can be readily motivated from (5.51); furthermore, $\kappa_{m\,min} = (\beta(\mu_m))^2$, though we shall not have direct need of this result in the construction of W_N. Note that for $W_N = W_N^0$ (Methods 2 and 4) we may omit (5.55); this constitutes significant "off-line" savings see Section 5.4.2 below.

Output prediction

We first find, for given $\mu \in \mathcal{D}$, $\underline{u}_N(\mu) \in \mathbb{R}^N, \underline{\psi}_N(\mu) \in \mathbb{R}^N$, solution of

$$\begin{aligned} \underline{S}^{W_N,W_N}(\mu)\underline{u}_N(\mu) &= \underline{Z}^{W_N,Y}\underline{L}^{\mathrm{pr},Y} \\ \underline{S}^{W_N,W_N}(\mu)\underline{\psi}_N(\mu) &= -\underline{Z}^{W_N,Y}\underline{L}^{\mathrm{du},Y}; \end{aligned} \tag{5.56}$$

it is readily shown that $u_N(\mu)$ and $\psi_N(\mu)$ of (5.19) and (5.20) of Section 5.2 are given by

$$\begin{aligned} u_N(\mu) &= \sum_{j=1}^{N} u_{N\,j}(\mu)\zeta_j \\ \psi_N(\mu) &= \sum_{j=1}^{N} \psi_{N\,j}(\mu)\zeta_j. \end{aligned}$$

We can then evaluate $s_N(\mu)$ of (5.21) as

$$s_N(\mu) = (\underline{u}_N(\mu))^t\underline{L}^{\mathrm{du},W_N} - (\underline{\psi}_N(\mu))^t(\underline{L}^{\mathrm{pr},W_N} - \underline{A}^{W_N,W_N}(\mu)\underline{u}_N(\mu)). \tag{5.57}$$

Note that for the Galerkin formulations, $V_N = W_N$, we may replace (5.56) with $\underline{A}^{W_N,W_N}\underline{u}_N = \underline{L}^{\mathrm{pr},W_N}$; but since we will need $\underline{S}^{W_N,W_N}$ in the error prediction step below, this is not a significant simplification.

Error bound prediction

We first calculate $\beta_N(\mu) = \sqrt{\kappa_{N\ \min}(\mu)}$, where $\kappa_{N\ \min}(\mu)$ is the eigenvalue associated with the first eigenpair $(\underline{\theta}_N(\mu), \kappa_N(\mu)) \in (\mathbb{R}^N \times \mathbb{R})$ of

$$\underline{S}^{W_N,W_N}(\mu)\underline{\theta}_N(\mu) = \kappa_N(\mu)\underline{B}^{W_N,W_N}\underline{\theta}_N(\mu), \quad (\underline{\theta}_N(\mu))^t\underline{B}^{W_N,W_N}\underline{\theta}_N(\mu) = 1, \tag{5.58}$$

as motivated by (5.52) above. Note in this "integrated formulation" that the *same* reduced-basis matrix, $\underline{S}^{W_N,W_N}$, serves to determine u_N, ψ_N, and β_N.

We next compute $\underline{\tau}^{\rm pr}(\mu) \in \mathbb{R}^{\mathcal{N}}, \underline{\tau}^{\rm du}(\mu) \in \mathbb{R}^{\mathcal{N}}$, solution of

$$\underline{B}^{Y,Y}\underline{\tau}^{\rm pr}(\mu) = \underline{L}^{{\rm pr},Y} - \underline{A}^{Y,W_N}(\mu)\underline{u}_N(\mu) \tag{5.59}$$

$$\underline{B}^{Y,Y}\underline{\tau}^{\rm du}(\mu) = -\underline{L}^{{\rm du},Y} - \underline{A}^{Y,W_N}(\mu)\underline{\psi}_N(\mu). \tag{5.60}$$

It can be readily shown that $P_\mu^N u_N(\mu), D_\mu^N \psi_N(\mu)$ as defined by (5.22) and (5.23) of Section 5.3 (for $V_N = Y$) are given by

$$P_\mu^N u_N(\mu) = \sum_{j=1}^{\mathcal{N}} \tau_j^{\rm pr}(\mu)\xi_j \equiv (\underline{\tau}^{\rm pr}(\mu))^t \underline{\xi}$$

$$D_\mu^N \psi_N(\mu) = \sum_{j=1}^{\mathcal{N}} \tau_j^{\rm du}(\mu)\xi_j \equiv (\underline{\tau}^{\rm du}(\mu))^t \underline{\xi},$$

respectively. Note that these calculations (5.59), (5.60) are required even for the Galerkin approach: we must compute the Y' norm of the residual to estimate the error.

Lastly, it then follows that $\Delta_N(\mu)$ of (5.27) can be evaluated as

$$\Delta_N(\mu) = \frac{1}{\sigma\beta_N(\mu)}((\underline{\tau}^{\rm pr}(\mu))^t \underline{B}^{Y,Y}\underline{\tau}^{\rm pr}(\mu))^{1/2}((\underline{\tau}^{\rm du}(\mu))^t \underline{B}^{Y,Y}\underline{\tau}^{\rm du}(\mu))^{1/2}, \tag{5.61}$$

which completes the procedure.

5.4.2 Blackbox approach

Preliminaries

To describe the blackbox procedure, and demonstrate the $\mathcal{N}$-independence of the on-line stage, we shall need a few additional definitions. First, we recognize that the $\zeta_i, i = 1, \ldots, N$, can be represented in terms of the ξ_j, which we express as

$$\zeta_i = \underline{z}_i^t \underline{\xi},$$

where $\underline{z}_i \in \mathbb{R}^{\mathcal{N}}, i = 1, \ldots, N$. Second, we need to introduce the matrices $\underline{A}_q^{Y,Y} \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}, q = 1, \ldots, Q$, given by

$$(A_q^{Y,Y})_{i,j} = a^q(\xi_j, \xi_i), 1 \le i, j \le \mathcal{N},$$

where the $a^q(\cdot,\cdot)$ are defined (implicitly) in (5.4).

We shall summarize the computational effort in the *off-line* stage in terms of $\underline{A}^{Y,Y}$-solves, $\underline{S}^{Y,Y}$-eigensolves, $\underline{B}^{Y,Y}$-solves, $\underline{A}^{Y,Y}$-actions (matrix vector products), $\underline{B}^{Y,Y}$-actions, and Y-inner products (inner products between two $\mathcal{N}$-vectors). Note for simplicity we assume that $\underline{A}_q^{Y,Y}$-actions are roughly equivalent to $\underline{A}^{Y,Y}$-actions. In many problems of interest, in particular in which there is underlying sparsity in $\underline{A}^{Y,Y}$, $\underline{S}$-eigensolves will be the most expensive, then $\underline{A}$-solves, then $\underline{B}^{Y,Y}$-solves (less costly because the equations are symmetric positive-definite), and then the "actions" (often only $O(\mathcal{N})$) and Y-inner products.

We shall summarize the *on-line* computational effort directly in terms of N and Q (albeit somewhat imprecisely, sometimes considering a multiplication and an addition as a single operation). Note that in the on-line stage we are not compelled to exploit all N basis functions computed in the off-line stage, and thus N in the on-line stage may be replaced with $N_{\text{used}}(\mu)$, with the error bound $\Delta_N(\mu)$ guiding the choice of minimal N_{used}; this can significantly reduce the cost of the on-line predictions.

As regards storage, we shall report, in the off-line stage, both Temporary Storage (required just during the off-line stage) and Permanent Storage (quantities passed by the off-line stage to the on-line stage). All quantities stored in the on-line stage originate in the off-line stage. The simplifications to the procedure in the case of "compliance" should be clear.

Off-line stage

1. Compute reduced-basis vectors: $\underline{u}_m \in \mathbb{R}^{\mathcal{N}}, \underline{\psi}_m \in \mathbb{R}^{\mathcal{N}}, \underline{\chi}_m \in \mathbb{R}^{\mathcal{N}}, m = 1, \ldots, M$, from (5.53), (5.54), and (5.55). Recall that the $\underline{\chi}_m$ are not needed for $W_N = W_N^0$.

Complexity: $2M$ $\underline{A}^{Y,Y}$-solves, M $\underline{S}^{Y,Y}$-eigensolves.

2. Compute $\underline{B}^{W_N,W_N} \in \mathbb{R}^{N\times N}$ as

$$B_{i,j}^{W_N,W_N} = (\zeta_j, \zeta_i)_Y,$$

or

$$B_{i,j}^{W_N,W_N} = \underline{z}_i^t \underline{B}^{Y,Y} \underline{z}_j.$$

Complexity: $N\underline{B}^{Y,Y}$-actions, N^2 Y-inner products.

Temporary Storage: $N\mathcal{N}$.

Permanent Storage: N^2.

3. Compute $\underline{\mathcal{U}}_{qi} \in \mathbb{R}^{\mathcal{N}}, 1 \le q \le Q, 1 \le i \le N$, where

$$(\mathcal{U}_{qi})_k = a^q(\zeta_i, \xi_k), 1 \le k \le \mathcal{N},$$

or

$$\underline{\mathcal{U}}_{qi} = \underline{A}_q^{Y,Y} \underline{z}_i.$$

Complexity: NQ $\underline{A}^{Y,Y}$-actions.

Temporary Storage: $NQ\mathcal{N}$.

4. Compute $\underline{\mathcal{V}}_{qi} \in \mathbb{R}^{\mathcal{N}}, 1 \le q \le Q, 1 \le i \le N$, and $\underline{\mathcal{V}}_0^{\rm pr} \in \mathbb{R}^{\mathcal{N}}, \underline{\mathcal{V}}_0^{\rm du} \in \mathbb{R}^{\mathcal{N}}$, solutions of

$$\underline{B}^{Y,Y}\underline{\mathcal{V}}_{qi} = \underline{\mathcal{U}}_{qi},\ \underline{B}^{Y,Y}\underline{\mathcal{V}}_0^{\rm pr} = \underline{L}^{{\rm pr},Y},\ \underline{B}^{Y,Y}\underline{\mathcal{V}}_0^{\rm du} = \underline{L}^{{\rm du},Y}.$$

Complexity: $(NQ+2)$ $\underline{B}^{Y,Y}$-solves.

Temporary Storage: $(NQ+2)\mathcal{N}$.

5. Compute $\Gamma_{qq'ii'} (= \Gamma_{q'qi'i}), 1 \le q, q' \le Q, 1 \le i, i' \le N$, as

$$\Gamma_{qq'ii'} = \underline{\mathcal{U}}_{qi}^t \underline{\mathcal{V}}_{q'i'}.$$

Complexity: N^2Q^2 Y-inner products.

Permanent Storage: N^2Q^2.

6. Compute $\Lambda_{qi}^{\mathrm{pr}}, \Lambda_{qi}^{\mathrm{du}}, 1 \leq q \leq Q, 1 \leq i \leq N$, as

$$\Lambda_{qi}^{\mathrm{pr}} = \underline{\mathcal{U}}_{qi}^t \underline{\mathcal{V}}_0^{\mathrm{pr}}, \Lambda_{qi}^{\mathrm{du}} = \underline{\mathcal{U}}_{qi}^t \underline{\mathcal{V}}_0^{\mathrm{du}}.$$

Complexity: $2NQ$ Y-inner products.

Permanent Storage: $2NQ$.

7. Compute $c_0^{\mathrm{pr}} \in \mathbb{R}$, $c_0^{\mathrm{du}} \in \mathbb{R}$, as

$$c_0^{\mathrm{pr}} = (\underline{L}^{\mathrm{pr},Y})^t \underline{\mathcal{V}}_0^{\mathrm{pr}},\ c_0^{\mathrm{du}} = (\underline{L}^{\mathrm{du},Y})^t \underline{\mathcal{V}}_0^{\mathrm{du}}.$$

Complexity: 2 Y-inner products.

Permanent Storage: 2.

8. Compute $\underline{L}^{\mathrm{pr},W_N} \in \mathbb{R}^N, \underline{L}^{\mathrm{du},W_N} \in \mathbb{R}^N$, as

$$L_i^{\mathrm{pr},W_N} = (\underline{L}^{\mathrm{pr},Y})^t \underline{z}_i,\ L_i^{\mathrm{du},W_N} = (\underline{L}^{\mathrm{du},Y})^t \underline{z}_i,\ i = 1, \ldots, N.$$

Complexity: $2N$ Y-inner products.

Permanent Storage: $2N$.

9. Compute $\Xi_{qii'}, 1 \leq q \leq Q, 1 \leq i \leq N$, as

$$\begin{aligned} \Xi_{qii'} &= a^q(\zeta_i, \zeta_{i'}) \\ &= \underline{z}_{i'}^t \underline{\mathcal{U}}_{qi}. \end{aligned}$$

Complexity: N^2Q Y-inner products.

Permanent Storage: N^2Q.

On-line stage

Given a new value of the parameter $\mu \in \mathcal{D}$:

1. Form $\underline{S}^{W_N,W_N}(\mu) \in \mathbb{R}^{N\times N}$ as

$$S_{i,i'}^{W_N,W_N} = \sum_{q=1}^{Q}\sum_{q'=1}^{Q} \sigma^q(\mu)\sigma^{q'}(\mu)\Gamma_{qq'ii'},\ 1 \le i,i' \le N.$$

Complexity: N^2Q^2.

2. Form the necessary "right-hand" sides:

$$\sum_{j=1}^{\mathcal{N}} Z_{i,j}^{W_N,Y} L_j^{\mathrm{pr},Y} = \sum_{q=1}^{Q} \sigma^q(\mu)\Lambda_{qi}^{\mathrm{pr}},\ 1 \le i \le N,$$
$$\sum_{j=1}^{\mathcal{N}} Z_{i,j}^{W_N,Y} L_j^{\mathrm{du},Y} = \sum_{q=1}^{Q} \sigma^q(\mu)\Lambda_{qi}^{\mathrm{du}},\ 1 \le i \le N.$$

Complexity: $2NQ$.

3. Find $\underline{u}_N(\mu) \in \mathbb{R}^N, \underline{\psi}_N(\mu) \in \mathbb{R}^N, \beta_N(\mu) \in \mathbb{R}$ solution of (5.56) and (5.58).

Complexity: $\mathcal{O}(N^3)$.

4. Compute $s_N(\mu)$ of (5.57) as

$$s_N(\mu) = (\underline{L}^{\mathrm{du},W_N})^t \underline{u}_N(\mu) - (\underline{L}^{\mathrm{pr},W_N})^t \underline{\psi}_N(\mu) + \sum_{q=1}^{Q}\sum_{i=1}^{N}\sum_{i'=1}^{N} \sigma^q(\mu) u_{N\,i}(\mu)\psi_{N\,i'}(\mu)\Xi_{qii'}.$$

Complexity: $2N + N^2Q$.

5. Compute $\Delta_N(\mu)$ of (5.61) as

$$\begin{aligned}\Delta_N(\mu) \;=\; & \frac{1}{\sigma\beta_N(\mu)}\Big[c_0^{\rm pr} - 2\sum_{q=1}^{Q}\sum_{i=1}^{N}\sigma^q(\mu)u_{N\,i}(\mu)\Lambda_{qi}^{\rm pr} \\ & \quad + \sum_{q=1}^{Q}\sum_{q'=1}^{Q}\sum_{i=1}^{N}\sum_{i'=1}^{N}\sigma^q(\mu)\sigma^{q'}(\mu)u_{N\,i}(\mu)u_{N\,i'}(\mu)\Gamma_{qq'ii'}\Big]^{1/2} \times \\ & \Big[c_0^{\rm du} + 2\sum_{q=1}^{Q}\sum_{i=1}^{N}\sigma^q(\mu)\psi_{N\,i}(\mu)\Lambda_{qi}^{\rm du} \\ & \quad + \sum_{q=1}^{Q}\sum_{q'=1}^{Q}\sum_{i=1}^{N}\sum_{i'=1}^{N}\sigma^q(\mu)\sigma^{q'}(\mu)\psi_{N\,i}(\mu)\psi_{N\,i'}(\mu)\Gamma_{qq'ii'}\Big]^{1/2}.\end{aligned}$$

Complexity: $2(N^2Q^2 + NQ + 1)$.

We now briefly discuss the computational complexity of the different schemes. The first comparison is between minimum-residual ($V_N = Y$) (or $V_N = Z_N(\mu)$) and Galerkin ($V_N = W_N$) approaches. The important point to note is that the quantity — $\Gamma_{qq'ii'}$ — required by Method 1 ($V_N = Y$, $W_N = W_N^1$) (or Method 5, $V_N = Z_N(\mu)$, $W_N = W_N^1$) to form the projection matrix $\underline{S}^{W_N,W_N}(\mu)$ is the same quantity required by *all* the methods to compute the error bound $\Delta_N(\mu)$; in both capacities, $\Gamma_{qq'ii'}$ represents the calculation of the necessary Y' norm. In this sense (see Proposition 7 and Lemma 9) the arguably better scheme $V_N = Y$, and somewhat riskier scheme $V_N = W_N$, have *similar complexity*, and we contend that $V_N = Y$ is thus preferred. The second comparison is between W_N^0 and W_N^1. For the on-line component, *the difference is not large* — $N = 3M$ *vs.* $N = 2M$; however, for the off-line component, the calculations of $\chi(\mu_m)$ can indeed be onerous, and its omission thus welcome. However, there is a corresponding rather significant loss of security, since the accuracy of β_N is *no longer controlled*.

5.5 The Helmholtz problem

5.5.1 1-d Example

Formulation

We take here $Y = H_0^1(\Omega)$, where Ω is a suitably smooth domain in $\mathbb{R}^d$, $d = 1, 2$, or 3, with inner product $(\cdot,\cdot)_Y$ and norm $\|\cdot\|_Y$. It is important to remark that we may substitute for $(\cdot,\cdot)_Y$ any inner product which induces a norm equivalent to the $H^1(\Omega)$-norm — for example, a good preconditioner. The latter will of course greatly reduce the off-line computational cost, as $\underline{B}^{Y,Y}$-solves will now be much less expensive.

For our bilinear form we take

$$a(w, v; \mu) = \int_\Omega \nabla w \cdot \nabla v - g(x; \mu) w v \ ,$$

where we assume that $g(x; \mu)$ satisfies

$$|g(x; \mu)| \leq g_{\max}, \quad \forall x \in \Omega, \ \forall \mu \in \mathcal{D},$$

and furthermore can be expressed as

$$g(x; \mu) = \sum_{q=1}^{\overline{Q}} \sigma^q(\mu) G^q(x), \tag{5.62}$$

where $G^q \in L^\infty(\Omega), q = 1, \ldots, \overline{Q}$. The difficult case, on which we focus here, is of course when $g(x; \mu)$ is positive, as in the reduced-wave (Helmholtz) equation.

The decomposition (5.62) is, in fact, reasonably general. We shall consider the situation in which $P = 2$ with parameter $\mu = (k_1, k_2)$, $\overline{Q} = 2$, $\sigma^1(\mu) = k_1^2$, $\sigma^2(\mu) = k_2^2$, and

$$G^1(x) = \begin{cases} 1 & x \in \Omega_1 \\ 0 & x \in \Omega_2 \end{cases} \ ,$$

$$G^2(x) = \begin{cases} 0 & x \in \Omega_1 \\ 1 & x \in \Omega_2 \end{cases} \ ,$$

where $\overline{\Omega} = \overline{\Omega}_1 \cup \overline{\Omega}_2$; this represents variable "frequencies" in two subdomains. It can be shown that the regularity of χ required in Lemma 5.3.8 follows from the smoothness of the $\sigma^q(\mu)$ and the interpretation of χ as $\theta_{\min}$ of (5.7).

It is simple to see that our requirement (5.4) is readily satisfied for $Q = 3$: with $\sigma^q(\mu),\ q = 1, 2$ as defined above, and $\sigma^3(\mu) = 1$, with $a^q(w, v) = -\int_\Omega G^q(x)wv,\ q = 1, 2$, and $a^3(w, v) = \int_\Omega \nabla w \cdot \nabla v$. It is similarly easy to show that a is symmetric, and also uniformly continuous with (say) $\gamma = 1 + g_{\max}$. The inf-sup condition will be satisfied so long as we exclude from $\mathcal{D}$ neighborhoods of points μ for which there exists a w such that $a(w, v; \mu) = 0,\ \forall v \in Y$. In general, if the inf-sup condition (5.3) is thus satisfied, and ℓ and ℓ^O are any bounded linear functionals, then our theoretical results of Section 5.3.3 will obtain.

We make two points of a more practical nature. First, in practice, we will of course not know where resonances occur, and thus we will typically posit a parameter domain which does indeed contain several points at which the inf-sup condition does not hold. However, unless driven to such a point by a design or optimization process, it is unlikely that a particular μ will coincide exactly with an eigenvalue, and thus for some sufficiently small β_0 our hypotheses will be "in practice" satisfied. (Obviously the physical model may also be made more elaborate, for example by including damping that will regularize the resonances.) Second, in practice, we choose not $Y = H_0^1(\Omega)$, but rather $Y = X_\mathcal{N}$, a suitably fine (say finite element) approximation of finite (albeit very large) dimension $\mathcal{N}$. As we are more and more conservative in defining this "truth" approximation, that is, as we increase $\mathcal{N}$, the off-line computational effort will of course increase; however, thanks to the blackbox formulation, the on-line computational effort is *independent* of the dimension $\mathcal{N}$.

Numerical Results

We take here $d = 1$ and $\Omega =]0, 1[$ (though obviously the computational savings provided by the reduced-basis approach will only be realized for more complicated multidimensional ($d > 1$) problems). Our truth space $X_\mathcal{N}$ is a linear finite element approximation with 200 elements. We consider the two-parameter Helmholtz equation defined in Section 5.5.1, with

$\Omega_1 =]0, 0.5[$ and $\Omega_2 =]0.5, 1.0[$. For simplicity, we present a "compliance" case in which

$$\ell(v) = \ell^O(v) = \int_{0.45}^{0.55} v,$$

corresponding to an imposed (oscillatory) distributed force for the input and an associated average displacement amplitude measurement for the output.

In the below we shall consider the four methods associated with the four choices of spaces of Sections 5.3.4, 5.3.5, 5.3.6, and 5.3.7. Note that, following the discussion on Section 5.3.8, we will make no distinction between the choices $V_N = Y$, $W_N = W_N^1$ and $V_N = Z_N(\mu)$, $W_N = W_N^1$ — both of these will be denoted as Method 1 in the following. Throughout this section we take $\sigma = (1.1)^{-1}$: it follows from Proposition 5 that a sufficient (though not necessary) condition for bounds is that β_N be within 10% of β. For most of the results of this section, we choose an effectively one–dimensional parameter space $\mathcal{D}$ which is the subset of $\mathcal{D}' =]11, 11[\times]1, 20[$ in which neighborhoods of the two resonance points $\mu \equiv (k_1, k_2) = (11, 7.5)$ and $\mu \equiv (k_1, k_2) = (11, 14.4)$ have been excised such that $\beta_0 = 0.02$. (Of course, in practice, we would not know the location of these resonance points, and we would thus consider $\mathcal{D} = \mathcal{D}'$ — which would only satisfy our inf-sup stability condition, "in practice," as discussed in the previous section.)

To begin, we fix $M = 3$, and hence $N = 2M = 6$ since we are in "compliance," with $\mu_1 = (11, 2)$, $\mu_2 = (11, 8)$ and thus $\mu_3 = (11, 14)$, and thus $\mathcal{S}_M = \{\mu_1, \mu_2, \mu_3\}$; we shall denote this the "$M = 3$" case. We first investigate the behavior of the discrete inf-sup parameter, the accuracy of which is critical for both the accuracy and bounding properties of our output prediction. In Figures 5-1 and 5-2 we plot the discrete inf-sup parameter β_N^i, $i = 1, \ldots, 4$, and the ratio β_N^i/β, $i = 1, \ldots, 4$, respectively, as a function of k_2 for fixed k_1 (see Section 5.5.1); recall that the index i refers to the method under consideration. We first confirm those aspects of the behavior that we have previously demonstrated. First, β_N^1 and β_N^2 are never less than β_N, as shown in Lemma 5.3.7 and Section 5.3.5, respectively; and $\beta_N^2 \geq \beta_N^1$, as must be the case since the inf space is smaller. The choice $V_N = Y$ ensures stability. Second, we see that $\beta_N^1 \geq \beta_N^3$ and $\beta_N^2 \geq \beta_N^4$, as demonstrated in (5.46) and Section 5.3.7 respectively; the methods with smaller supremizing spaces are perforce less stable. Third,

we see (by closer inspection of the numerical values) that β_N^3 is never greater than β at the sample points, consistent with (5.42); in fact, we observe that equality obtains at the sample points, (5.44), and hence at least in this particular case the conjecture (5.43) appears valid. Fourth, we notice that β_N^4 can be either below or above β, and is clearly the least "controlled" of the four approximations. (Indeed, for other parameter values we observe near zero values of β_N^4 at points quite far away from the true resonances of the system.)

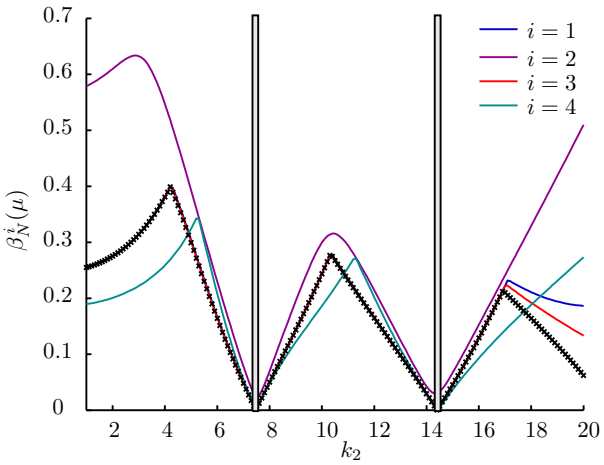

Figure 5-1: The discrete inf-sup parameter for Methods 1, 2, 3, and 4 as a function of k_2 (see text for legend). The symbol $\times$ denotes the exact value of β.

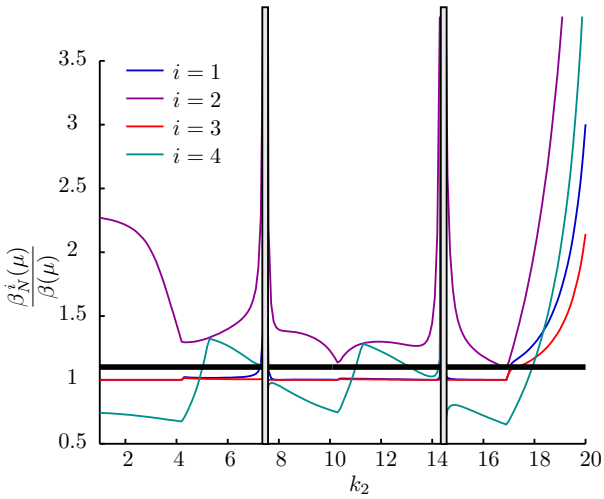

Figure 5-2: The ratio of the discrete inf-sup parameter to the exact inf–sup parameter for Methods 1, 2, 3, and 4, as a function of k_2 (see text for legend). The thick line denotes the "sufficient" limit: if $\beta_N < 1.1\beta$, bounds are guaranteed.

It is clear from Figures 5-1 and 5-2 that β_N^i is indeed a very accurate predictor of β over most of $\mathcal{D}$ for Methods 1 and 3; we anticipated this result in Proposition 6 and the discussion of Section 5.3.6. We now study the convergence of β_N^i to β as N increases. For this test we consider a sample $\mathcal{S}_M = \{\mu_m,\ m = 1, \ldots, M\}$, with the μ_m randomly drawn from $\mathcal{D}$; the particular parameter points selected are given in the second column of Table 5.1. (Note that for a given M, indicated in the first column of Table 5.1, $\mathcal{S}_M$ consists of all $\mu_m, m = 1, \ldots, M$.) We present in Table 5.1 the values of $\beta_N^i - \beta$ for Methods i =1, 2, 3, and 4 for $k_2 = 11$ (and hence $\mu = (11, 11)$). We note that, indeed, β_N^i converges very rapidly for $i = 1$ and $i = 3$ — the two methods in which we include the infimizers in V_N — whereas for $i = 2$ and $i = 4$ we do not obtain convergence — not surprising given the

		$\beta_N^i - \beta$			
M	μ_M	$i=1$	$i=2$	$i=3$	$i=4$
1	(11, 4.7351)	$1.81e-01$	$2.92e+00$	$-1.88e-01$	$1.08e+00$
2	(11,19.0928)	$1.70e-01$	$4.28e-01$	$-2.03e-01$	$4.14e-01$
3	(11,11.4848)	$3.52e-04$	$2.76e-01$	$-1.24e-04$	$-9.85e-02$
4	(11,13.6038)	$9.25e-06$	$5.76e-02$	$3.99e-06$	$5.31e-02$
5	(11, 1.4975)	$6.57e-09$	$4.91e-02$	$2.43e-09$	$4.13e-02$
6	(11, 2.6998)	$1.90e-11$	$4.81e-02$	$4.49e-08$	$3.83e-02$

Table 5.1: The error $\beta_N^i - \beta$ for Methods i =1, 2, 3, and 4, for $k_2 = 11$, as a function of M.

discussion of Section 5.3.3. Note also that whereas the convergence of Method 1 is (and must be) monotonic, this is not necessarily the case for Method 3.

We conclude that Method 2 and in particular Method 4 are not very reliable: we can certainly not guarantee asymptotic bounds for any given fixed $\sigma < 1$; for this reason we do not recommend these techniques, and we focus primarily on Methods 1 and 3 in the remainder of this section. However, in practice, all four methods may perform reasonably well for some smaller σ, in particular since the accuracy of the inf-sup parameter is only a sufficient and not a necessary condition for bounds. Indeed, for our $M = 3$ case of Figures 5-1 and 5-2, Methods 1, 2, and 4 produce bounds for all k_2 less than approximately 18 and Method 3 in fact produces bounds for all k_2 in $\mathcal{D}$; consistent with Proposition 5, bounds are always obtained for all methods so long as $\sigma\beta_N \leq \beta$. The breakdown of bounds for Method 1 (which in fact directly correlates with $\sigma\beta_N^1 > \beta$) is due to the poor infimizer approximation properties of W_N^1 for larger k_2; if we include an additional sample point, $\mu_4 = (11, 20)$, we recover bounds for all $\mathcal{D}$. (In fact, even for lower k_2 the infimizer approximation is not overly good; but thanks to the quadratic convergence proven in Proposition 6 the inf-sup parameter remains quite accurate.)

The fact that Method 3 produces bounds over the entire range is consistent with the "less stable" arguments of Section 5.3.6. However, by these same arguments, in particular Proposition 7, we expect that the bound gap — the controllable error in the output prediction — will be larger for Method 3 than for Method 1. To demonstrate this empirically, we plot in Figure 5-3 $\Delta_N^i/|s|, i = 1$ and $i = 3$, as a function of k_2, for the $M = 3$ case of Figures 5-1 and 5-2. We observe that, indeed, the bound gap is significantly smaller for Method 1 than

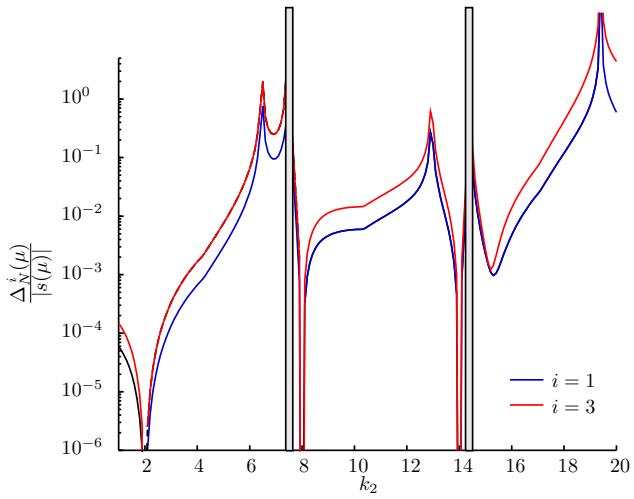

Figure 5-3: The normalized bound gap $\Delta_N^i/|s|$ for Methods i =1 and i =3 as a function of k_2.

		Δ_N^i	
M	μ_M	$i = 1$	$i = 3$
1	(11, 4.7351)	$2.23e-04$	$1.73e-01$
2	(11, 19.0928)	$2.13e-04$	$6.21e-01$
3	(11, 11.4848)	$5.37e-06$	$2.68e-05$
4	(11, 13.6038)	$4.01e-08$	$5.80e-08$
5	(11, 1.4975)	$6.43e-11$	$6.50e-11$
6	(11, 2.6998)	$1.65e-14$	$1.62e-14$

Table 5.2: The bound gap for Methods $i = 1$ and $i = 3$, for $k_2 = 11$, as a function of M.

for Method 3. Note also that the normalized bound gap is quite large for the k_2 at which we no longer obtain bounds for Method 1; no doubt these predictions would be rejected as overly inaccurate and requiring further expansion of the reduced–basis space (thus also recovering the inf-sup parameter accuracy and hence bounds).

As regards the convergence of the bound gap, we present in Table 5.2 convergence results for $\Delta_N^i, i = 1$ and $i = 3$, for $k_2 = 11$ (and hence $\mu = (11, 11)$), as a function of M (analogous to Table 5.1 for the inf-sup parameter). (Note for Methods 2 and 4 the convergence is slower, with the bound gap typically an order of magnitude *larger* than for Methods 1 and 3; this suggests that the inclusion of the infimizers can, in fact, reduce the approximation error — as might be anticipated from (5.41).) We observe that the differences between

M	μ_M	Δ_N^1	η_N^1	Δ_N^3	η_N^3
1	(7.5388, 14.2564)	$1.30e-04$	24.20	$6.09e-04$	12.18
2	(2.9571, 7.1526)	$1.22e-04$	3.51	$3.05e-04$	11.88
3	(4.2387, 9.0533)	$5.82e-05$	1.08	$8.77e-05$	1.49
4	(17.7486, 15.9503)	$1.10e-05$	2.56	$1.24e-05$	3.42
5	(9.7456, 14.0523)	$2.99e-08$	3.47	$3.07e-08$	3.71
6	(3.8279, 16.3388)	$2.81e-08$	3.71	$2.88e-08$	3.97
7	(11.2113, 8.0970)	$5.40e-12$	3.78	$5.44e-12$	3.85

Table 5.3: The bound gap and effectivity at $\mu = (11, 17)$, as a function of M, for Methods $i = 1$ and $i = 3$, for the two-dimensional parameter space $\mathcal{D} =]1, 20[\times]1, 20[$.

Methods 1 and 3 become smaller as M increases; however it is precisely for smaller M that reduced-basis methods are most interesting. We conclude — given that the two methods are of comparable cost – that Method 1 is perhaps preferred, in particular because we can also better guarantee the behavior of the inf-sup parameter. Note that the difference in the true error for Methods 1 and 3 is much smaller than the difference in the error bound for the two methods; this is expected, since the inf-sup parameters do not differ appreciably. It follows that the effectivity (defined in (5.36)) of Method 1 is lower (and hence better) than the effectivity of Method 3; this is not surprising, since for Method 1 the approximation is *designed* to minimize the bound gap.

We close by considering a second set of numerical results included to demonstrate the rapid convergence of the reduced–basis prediction as N increases even in higher dimensional parameter spaces: we now consider $\mathcal{D} =]1, 20[\times]1, 20[$ (without excision of resonances, and hence satisfying our inf-sup condition only "in practice"). In particular we repeat, the convergence scenario of Table 5.2, but now choose our random sample from the two–dimensional space $\mathcal{D} =]1, 20[\times]1, 20[$; we present, in Table 5.5.1, the bound gap and effectivity (defined in (5.36)) for Methods 1 and 3 for a particular "new" parameter point $\mu = (11, 17)$. We observe, first, that we obtain bounds in all cases ($\eta_N \geq 1$) — indicative of an accurate inf-sup parameter prediction; second, that the error (true and estimated) tends to zero very rapidly with increasing M, even in this two–dimensional parameter space; and third, that Method 1 again provides smaller bound gaps (and lower effectivities) than Method 3, consistent with Proposition 7 — though the difference is only significant for very

small M. Note that u and the output s are order 10^{-3}, so the relative errors are roughly 1000 times larger than the absolute errors in the table. Results similar to those reported in Table 5.5.1 are also obtained if we consider the error over a random ensemble of test points μ rather than a single test point.

5.5.2 2-d Example

Formulation

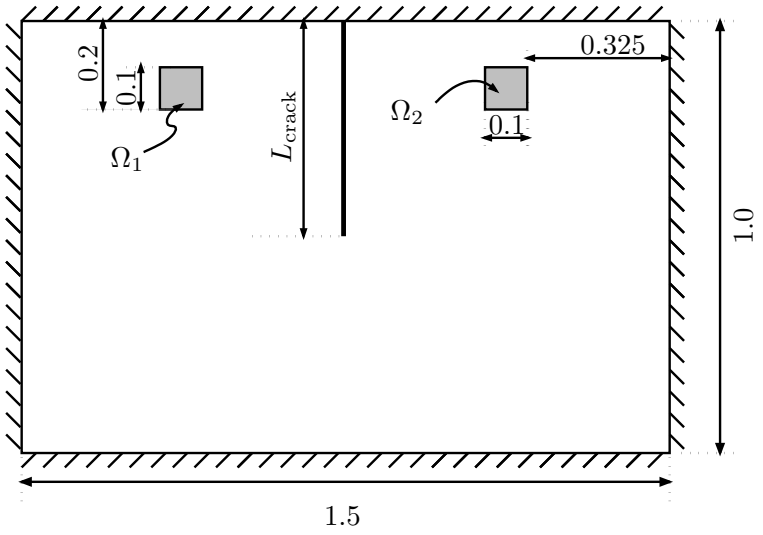

Figure 5-4: Geometrical configuration

To further test our methods we consider here a more realistic two-dimensional example. We restrict in this case our attention only to Method 1 (or the equivalent Method 5). To start, consider the domain Ω shown in Figure 5-4. As before we take $Y = H_0^1(\Omega)$ with inner product $(\cdot,\cdot)_Y$ and norm $\|\cdot\|_Y$. The problem we are interested in solving is the reduced-wave (Helmholtz) equation:

$$\int_\Omega \nabla u \cdot \nabla v \, d\Omega - \omega^2 \int_\Omega uv \, d\Omega = \int_{\Omega_1} v \, d\Omega, \ \forall v \in Y. \tag{5.63}$$

The right-hand side can be understood as an excitation of frequency ω over the domain Ω_1. The resulting solution gives the amplitude, for the given frequency ω, at each point of the domain Ω. In addition, we assume that there is a crack of length L_{crack} which disrupts the

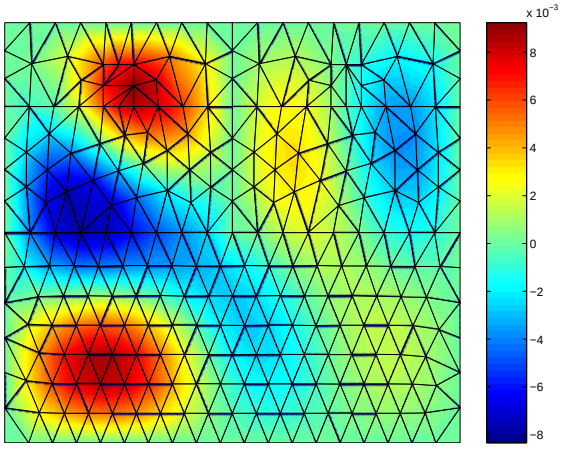

Figure 5-5: $L_{\text{crack}} = 0.5$ and $\omega = 10.0$

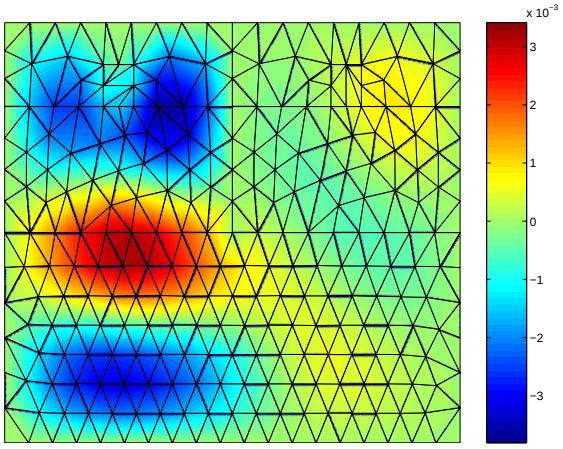

Figure 5-6: $L_{\text{crack}} = 0.5$ and $\omega = 11.0$

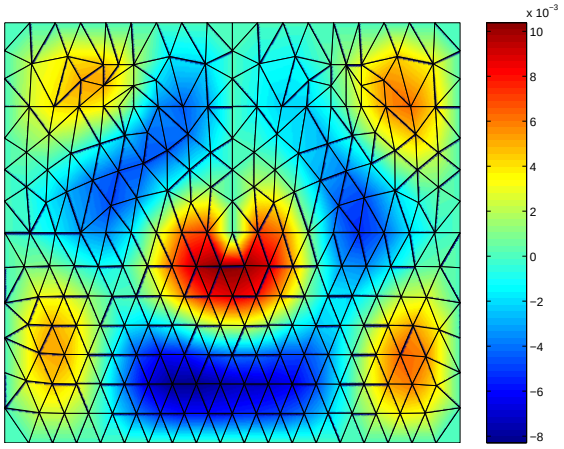

Figure 5-7: $L_{\text{crack}} = 0.5$ and $\omega = 12.0$

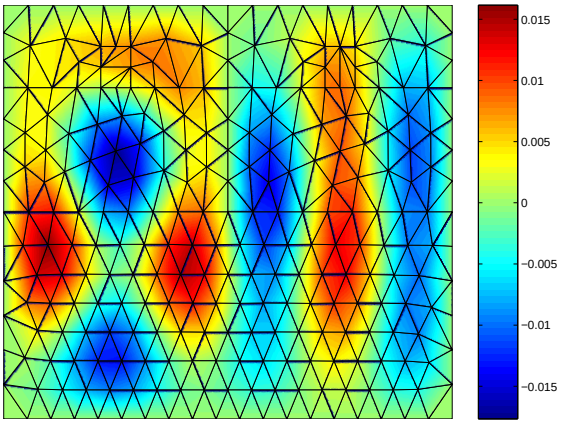

Figure 5-8: $L_{\text{crack}} = 0.5$ and $\omega = 13.0$

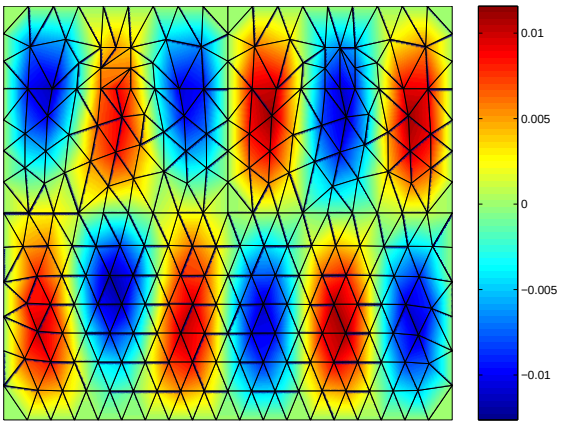

Figure 5-9: $L_{\text{crack}} = 0.5$ and $\omega = 14.0$

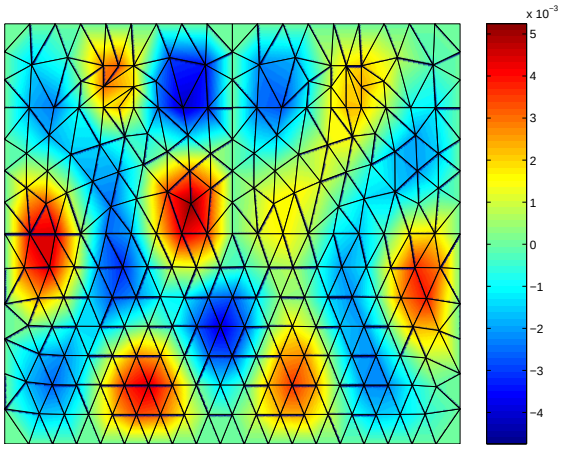

Figure 5-10: $L_{\text{crack}} = 0.5$ and $\omega = 15.0$

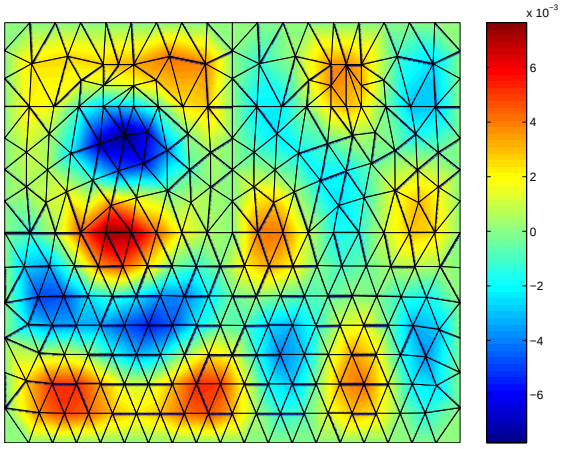

Figure 5-11: $L_{\text{crack}} = 0.5$ and $\omega = 16.0$

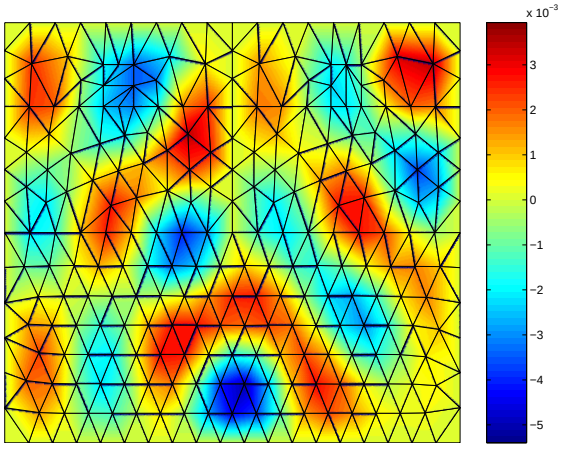

Figure 5-12: $L_{\text{crack}} = 0.5$ and $\omega = 17.0$

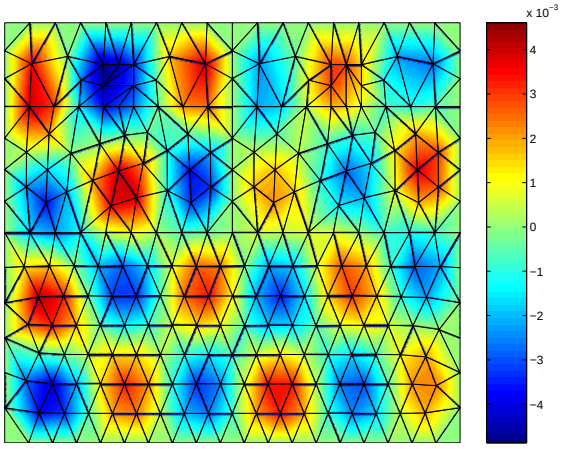

Figure 5-13: $L_{\text{crack}} = 0.5$ and $\omega = 18.0$

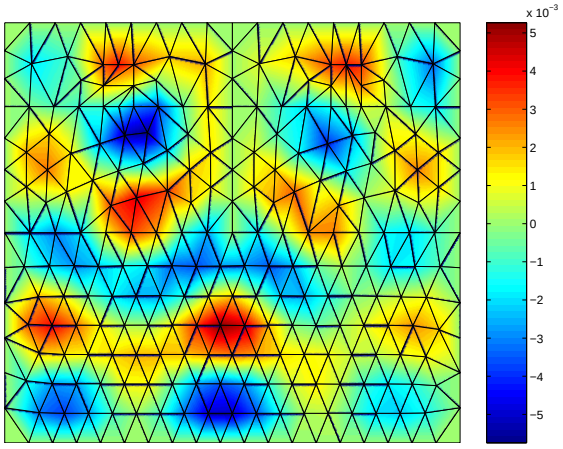

Figure 5-14: $L_{\text{crack}} = 0.5$ and $\omega = 19.0$

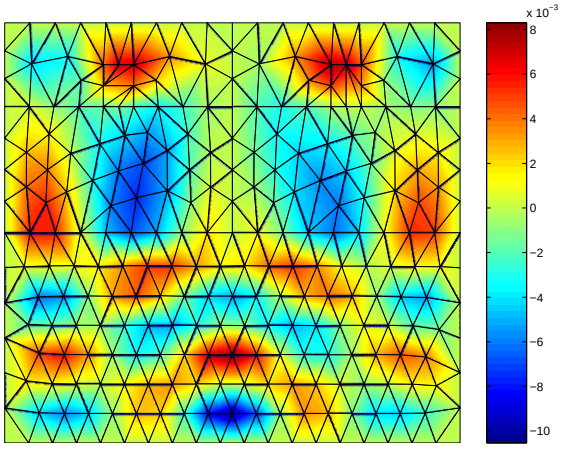

Figure 5-15: $L_{\text{crack}} = 0.3$ and $\omega = 19.0$

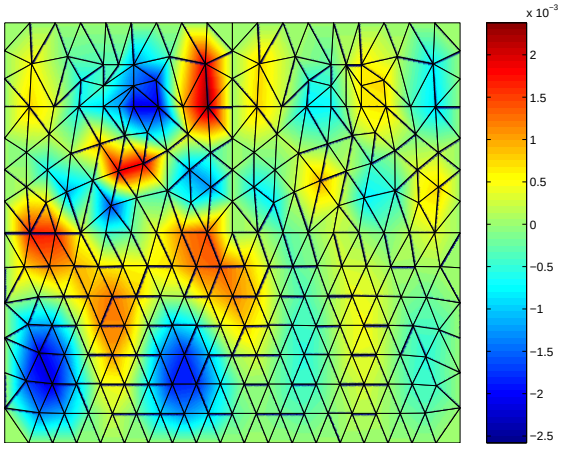

Figure 5-16: $L_{\text{crack}} = 0.7$ and $\omega = 19.0$

continuity of the domain Ω. For the parametrization of the problem we therefore choose $P = 2$ and $\mu = \{\omega, L_{\text{crack}}\}$. For the output we are interested in measuring the response over the small patch Ω_2, and therefore $s(\mu)$ is:

$$s(\mu) = \ell^O(u(\mu)) = \int_{\Omega_2} u(\mu)\, d\Omega. \tag{5.64}$$

The problem above although rather simple in terms of the geometric configuration, is rather interesting in the case of non-destructive evaluation. For example, we could place a sound source over the domain Ω_1 and a sensor on the domain Ω_2. In practice, the signature of the crack is measured (for varying frequencies). Comparing with a database of signatures, we can identify the size of the crack L_{crack} (or more generally even the position of the crack). Here, a reduced-basis model can be used instead of the database (which is rather costly to build), to efficiently and accurately match the measured signature.

To account for geometry variations (due to the varying crack length), we apply a continuous piecewise-affine transformation from Ω to a fixed reference domain $\hat{\Omega}$. The abstract problem statement of (5.1) is thus recovered. It is readily verified that the affine decomposition is obtained for $Q = 8$. We set allowable ranges for the input parameters, $1.0 \leq \omega \leq 20.0$ for the frequency and $0.3 \leq L_{\text{crack}} \leq 0.7$ for the crack size; therefore $\mathcal{D} = [1.0, 20.0] \times [0.3, 0.7]$ (as before we do not excise resonances, and our inf-sup condition is satisfied only "in practice"). We give in Figures 5-5–5-16, solutions for different choices of the input parameters. As we can see even for small variations in the input parameters the solution (and therefore the output) changes appreciably.

Numerical Results

The current example exercises all aspects of our framework. Notice that, since $\ell(v) \neq \ell^O(v)$, we are no longer in compliance and therefore we need both the solution of the primal and the dual problem. For the construction of the reduced-basis spaces, we choose $M_1 = M_3$ and $\mathcal{S}_{M_1} = \mathcal{S}_{M_3}$; also we choose M_2 points different from the previous ones to form $\mathcal{S}_{M_2}$. We then construct the reduced-basis space $W_N = W_N^1$ defined in (5.15).

First, we present in Figures 5-17 and 5-18, the error in the reduced-basis approximation

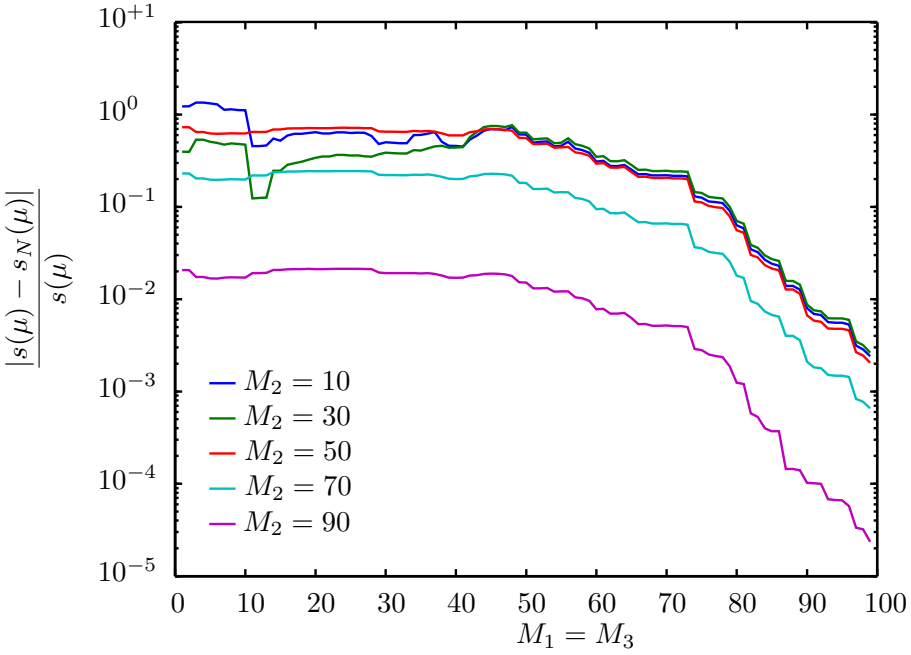

Figure 5-17: Output convergence for $L_{\text{crack}} = 0.4$ and $\omega = 13.5$.

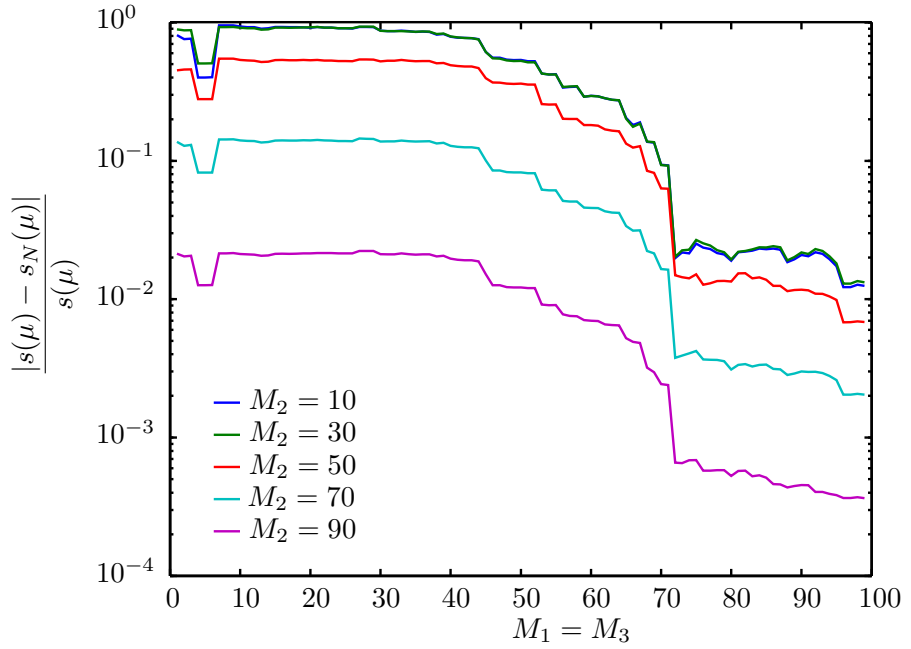

Figure 5-18: Output convergence for $L_{\text{crack}} = 0.4$ and $\omega = 18.0$.

of the output, for increasing values of M_1 and M_2. In the following numerical tests, we will present the results for two different test points $\mu_t^1 = \{0.4, 13.5\}$ and $\mu_t^2 = \{0.4, 18.0\}$. We see that initially, for small values of M_1 or M_2, the error is significant and is not reduced by increasing the number of basis functions included in the reduced-basis space. This can be attributed to the sensitivity of the solution in variations of the input parameters. Initially, for small values of M_1 and M_2 the basis functions included in the definition of the reduced-basis space, have no approximation properties for the solution at the test point. As we further increase M_1 (or M_2) we see that the output approximation converges to the exact value. In fact, for $M_1 = M_2 = 90$ and for the test point μ_t^1, we see that the relative error is less that 10^{-4} — quite acceptable for all practical purposes. For the two different test points we see different convergence rates. Again, this depends on the construction of the reduced-basis space and its ability to approximate a solution at the particular test point. As we can not *a priori* predict the error, the importance of the *a posteriori* error estimator should be clear.

Turning now to the error estimator we give in Figures 5-19 and 5-20 the *a posteriori* effectivity, for the two test points μ_t^1 and μ_t^2. In the computation of the bound gap, we choose $\sigma = 0.5$ and thus a sufficient condition for bounds is that β_N is within 100% of β. We first notice that in both cases the effectivities are always larger than one, and therefore bounds are always obtained. Furthermore, we see that the effectivity is usually between 10 and 100, which is relatively large given also the convergence of the true error. These effectivities can be further improved by developing more appropriate bound conditioners — this development will be considered in a future paper.

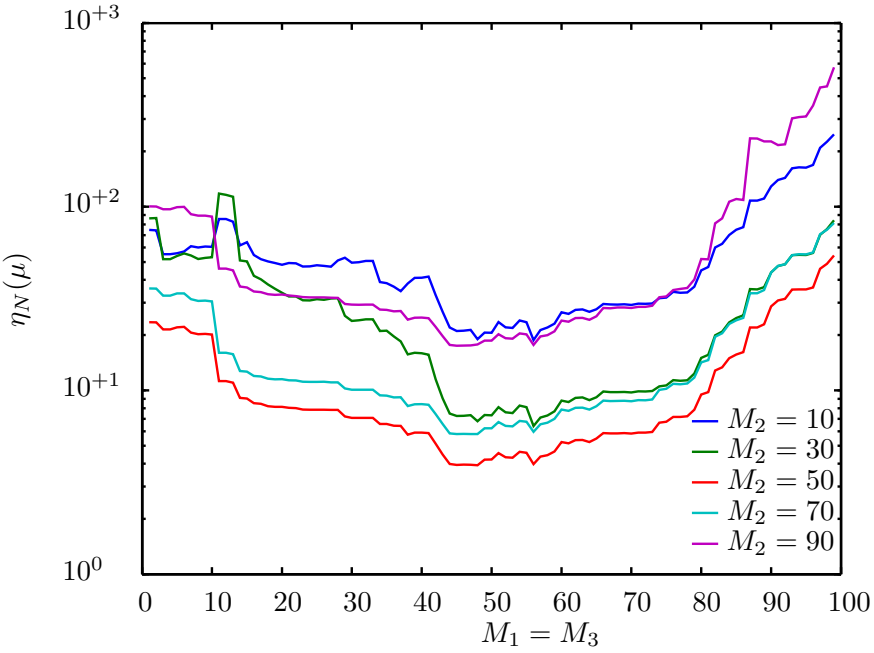

Figure 5-19: Effectivity for $L_{\text{crack}} = 0.4$ and $\omega = 13.5$.

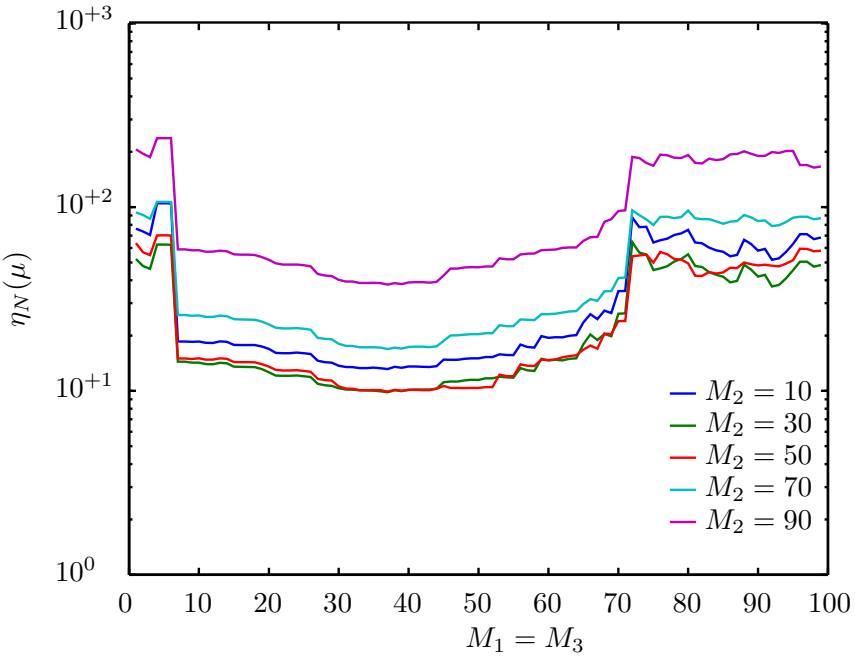

Figure 5-20: Effectivity for $L_{\text{crack}} = 0.4$ and $\omega = 18.0$.

Chapter 6

Stokes Problem

We develop in this Chapter the reduced-basis output bound method for the Stokes problem, considered as a representative example for the class of constrained minimization problems (with equality constraints). The essential new ingredient is the presence of the incompressibility constraint, and relatedly of the pressure which plays the role of the Lagrange multiplier. In addition the fact that the solution variable is a vector will slightly complicate the notation and the treatment.

Our presentation here follows as in the previous chapters: in Section 6.1 we state the problem and provide with some general definitions; then in Section 6.2 we develop the reduced-basis method, and consider issues like stability and accuracy; and, finally, in Section 6.3 we develop an *a posteriori* error estimation framework. The underlying ideas here are similar to the ones for the non-coercive problems, described in the previous Chapter, so we will refer to these as appropriate.

6.1 Problem Description

The system of Stokes equations are of special interest as they model the incompressible flow of highly-viscous fluids. From the numerical point of view, the presence of the incompressibility constraint poses significant problems in stability and special study is required. Moreover, although the Stokes problem is a means by itself, it is also the first (main) step for the

solution of the more general Navier-Stokes equations.

6.1.1 Introduction

To start, consider a Lipschitz-continuous polygonal domain $\Omega \subset \mathbb{R}^d$, with a boundary $\partial\Omega$. Let $u \in X$ a vector with components $u = \{u_1, \ldots, u_d\}$, where the u_i are functions defined on Ω. The non-dimensional strong form for the Stokes equations is:

$$\begin{aligned} -\frac{\partial}{\partial x_j}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) + \frac{\partial p}{\partial x_i} &= f_i, \\ \frac{\partial u_i}{\partial x_i} &= 0; \end{aligned} \tag{6.1}$$

where summation over repeated indices is assumed. Here, p is the pressure which plays the role of a Langrange multiplier in enforcing the incompressibility constraint, and u is the velocity vector. We need to augment the system of equations above with appropriate boundary conditions.

Using the incompressibility constraint we can obtain the simpler — and more familiar — form of the equations:

$$\begin{aligned} -\frac{\partial u_i}{\partial x_j\, \partial x_j} + \frac{\partial p}{\partial x_i} &= f_i, \\ \frac{\partial u_i}{\partial x_i} &= 0. \end{aligned} \tag{6.2}$$

Using either of these forms we can develop a variational statement, which will be the point of departure for the finite-element method. The reason why we mention both approaches is that the first formulation is more general, allowing to include in the variational formulation complex Neumann boundary conditions like, for example, stress boundary conditions, surface tension, etc. On the other hand, the second can only be applied with simple boundary conditions but is appealing due to its simplicity. In general the two formulations will yield discrete problems which, due to the weak imposition of the incompressibility constraint, will give different solutions. Our abstract problem statement given below, encompasses both of these formulations. For the simple example that we will consider in Section 6.5, we prefer

(6.2).

6.1.2 Abstract Problem Statement

To start, let V a closed linear subspace of $H^1(\Omega)$ such that $H^1_0(\Omega) \subset V \subset H^1(\Omega)$, and define $X = (V)^d$; X will be used for the velocity components. Also, for the pressure, we will require $M = L^2(\Omega)$ or $L^2_0(\Omega)$, where

$$L^2_0(\Omega) = \left\{ q \in L^2(\Omega) | \int_\Omega q \, d\Omega = 0 \right\};$$

the second choice is mandated in the case of all-Dirichlet boundary conditions, since then the pressure is defined uniquely up to an additive constant. Finally we will require the product space $Y = X \times M$. The norms and inner products for these spaces are defined in the usual way; for example if $v \in X$ then $\|v\|^2_{H^1(\Omega)} = \|v_1\|^2_{H^1(\Omega)} + \ldots + \|v_d\|^2_{H^1(\Omega)}$.

Suppose also that we are given a parameter μ chosen from a set $\mathcal{D} \subset \mathbb{R}^P$. The parameters of interest for the Stokes problem reflect geometry variations (cf. example in Section 6.5), which makes all the forms parameter-dependent — our presentation below should reflect this. We then look for $[u(\mu), p(\mu)] \in Y$ such that

$$\begin{aligned} a(u(\mu), v; \mu) + b(v, p(\mu); \mu) &= \ell(v; \mu), \ \forall v \in X, \\ b(u(\mu), q; \mu) &= 0, \ \forall q \in M; \end{aligned} \tag{6.3}$$

where $a : X \times X \times \mathcal{D} \to \mathbb{R}$, $b : X \times M \times \mathcal{D} \to \mathbb{R}$ are bilinear forms, and b is non-square. Furthermore, $\ell(\cdot) \in X'$ is a bounded linear functional. We can write the equation above more succinctly as

$$a(u(\mu), v; \mu) + b(v, p(\mu); \mu) + b(u(\mu), q; \mu) = \ell(v; \mu), \ \forall [v, q] \in Y.$$

Finally, as it typical in engineering practice, we assume that we are not interested in calculating the solution or abstract norms of it. Rather we are interested in obtaining performance measures that characterize the particular configuration $\mu \in \mathcal{D}$ and have physical importance like, for example, the flow-rate, lifts or drags. Given the solution $[u(\mu), p(\mu)]$ to

(6.3), the output of interest is obtained from

$$s(\mu) = \ell^O([u(\mu), p(\mu)]; \mu) = \ell_u^O(u(\mu); \mu) + \ell_p^O(p(\mu); \mu); \tag{6.4}$$

with $\ell_u^O(\cdot; \mu) \in X',\ \forall \mu \in \mathcal{D}$ and $\ell_p^O(\cdot; \mu) \in M',\ \forall \mu \in \mathcal{D}$ (notice that $M' = M$ here), which implies that $\ell^O(\cdot; \mu) \in Y',\ \forall \mu \in \mathcal{D}$ is a bounded linear functional. We also require in the following a dual, or adjoint, problem associated with $\ell^O(\cdot; \mu)$: find $[\psi(\mu), \lambda(\mu)] \in Y$ such that

$$\begin{aligned} a(v, \psi(\mu); \mu) + b(v, \lambda(\mu); \mu) &= -\ell_u^O(v; \mu),\ \forall v \in X, \\ b(\psi(\mu), q; \mu) &= -\ell_p^O(q; \mu),\ \forall q \in M. \end{aligned} \tag{6.5}$$

We note that a is symmetric

$$a(w, v; \mu) = a(v, w; \mu),\ \forall w, v \in X^2,\ \forall \mu \in \mathcal{D}.$$

Also we assume that the bilinear forms a and b are:

i) *Continuous*: there exist $\gamma_a(\mu) > 0$ and $\gamma_b(\mu) > 0$ such that

$$\begin{aligned} a(w, v; \mu) &\leq \gamma_a(\mu) \|w\|_X \|v\|_X,\ \forall w, v \in X^2,\ \forall \mu \in \mathcal{D}, \\ b(w, q; \mu) &\leq \gamma_b(\mu) \|w\|_X \|q\|_M,\ \forall w \in X,\ \forall q \in M,\ \forall \mu \in \mathcal{D}. \end{aligned} \tag{6.6}$$

ii) *Stable*: there exist $\alpha(\mu) \geq \tilde{\alpha}_0 > 0$ and $\beta(\mu) > \tilde{\beta}_0 > 0$, such that

$$\begin{aligned} 0 < \tilde{\alpha}_0 \leq \alpha(\mu) &= \inf_{v \in X} \frac{a(v, v; \mu)}{\|v\|_X^2},\ \forall \mu \in \mathcal{D}, \\ 0 < \tilde{\beta}_0 \leq \beta(\mu) &= \inf_{q \in M} \sup_{v \in X} \frac{b(v, q; \mu)}{\|v\|_X \|q\|_M} = \inf_{q \in M} \frac{\|b(\cdot, q; \mu)\|_{X'}}{\|q\|_M}, \forall \mu \in \mathcal{D}. \end{aligned} \tag{6.7}$$

The conditions above are sufficient to ensure existence and uniqueness [97] of the solutions to problems (6.3) and (6.5). Finally, we make the assumption of affine parameter dependence

for all the linear and bilinear forms:

$$a(w,v;\mu) = \sum_{q=1}^{Q_a} \sigma_a^q(\mu) a^q(w,v), \qquad b(w,q) = \sum_{q=1}^{Q_b} \sigma_b^q(\mu) b^q(w,q),$$
$$\ell(w;\mu) = \sum_{q=1}^{Q_f} \sigma_f^q(\mu) \ell_f^q(w), \qquad \ell^O([w,q];\mu) = \sum_{q=1}^{Q_O} \sigma_O^q(\mu) \ell_O^q([w,q]), \tag{6.8}$$

$\forall w, v \in X,\ \forall q \in M$; where $Q_.$ are integers that depend on the problem in consideration.

For the finite-element solution, $Y = X \times M$ is usually replaced by $Y_h = X_h \times M_h$, where X_h and M_h are appropriately defined finite-element spaces. To ensure stability we need to verify that the constants $\alpha_h(\mu)$ and $\beta_h(\mu)$ — obtained from (6.7) by replacing X with X_h, and M with M_h — are positive. In fact, even when they are close to zero, the *a priori* theory [97] suggests that we should expect large approximation errors. For conforming velocity approximation spaces it is easy to verify that $\alpha_h(\mu) \geq \alpha(\mu)$. Ensuring that $\beta_h(\mu)$ is non-zero requires careful selection both of the velocity and pressure approximation spaces. Here we choose to approximate velocity and pressure using $\mathbb{P}^2$ and $\mathbb{P}^1$ triangular elements, respectively. These elements belong to the Taylor-Hood family of elements and satisfy all the requirements above.

6.1.3 Inf-sup supremizers and infimizers

Similar, to Section 5.1.2 we define the supremizer $T_\mu q \in X$ associated with $\|b(\cdot, q; \mu)\|_{X'}$. This supremizer can be calculated from

$$(T_\mu q, v)_X = b(v, q; \mu),\ \forall v \in X. \tag{6.9}$$

We can then express our inf-sup parameter as:

$$\beta(\mu) = \inf_{q \in M} \sup_{v \in X} \frac{b(v,q;\mu)}{\|v\|_X \|q\|_M} = \inf_{q \in M} \frac{\|T_\mu q\|_X}{\|q\|_M} = \frac{\|T_\mu \chi\|_X}{\|\chi\|_M}, \tag{6.10}$$

where

$$\chi(\mu) = \arg\inf_{q \in M} \frac{\|T_\mu q\|_X}{\|q\|_M}. \tag{6.11}$$

Unfortunately, unlike the non-coercive case, it is no longer true that the infimizer $\chi(\mu)$ and the supremizer $T_\mu \chi$ will be collinear. It is useful to recognize that similar to (5.7) and (5.8), $\beta(\mu)$ and $\chi(\mu)$ can be related to the minimum eigenvalue and associated eigenfunction, of an appropriately defined symmetric positive-definite eigenproblem; see Section 5.1.2 for more details.

6.2 Reduced-Basis Approximation

6.2.1 Approximation Spaces

We next define our primal and dual reduced-basis approximation spaces. To wit, for the primal (resp. dual) problem we choose N (resp. M) points $\mu_i,\ i = 1, \ldots, N$ (resp. $\mu_i,\ i = 1, \ldots, M$ not necessarily the same as for the primal problem) in our parameter set $\mathcal{D}$, the collection of which we denote:

$$\mathcal{S}_N^{\text{pr}} = \{\mu_1, \ldots, \mu_N\}\,, \ (\text{resp. } \mathcal{S}_M^{\text{du}} = \{\mu_1, \ldots, \mu_M\}).$$

We then compute $[u(\mu_i), p(\mu_i)] \in Y$ (resp. $[\psi(\mu_i), \lambda(\mu_i)] \in Y$), the solutions of (6.3) (resp. (6.5)), for all $\mu_i \in \mathcal{S}_N^{\text{pr}}$ (resp. $\mu_i \in \mathcal{S}_M^{\text{du}}$), and also $z^{\text{pr}\,q,n} \in X$, (resp. $z^{\text{du}\,q,n} \in X$) $q = 1, \ldots, Q_b$, and $n = 1, \ldots N$ (resp. $n = 1, \ldots, M$) which satisfy

$$\begin{aligned} \left(z^{\text{pr}\,q,n}, v\right)_X &= b^q(v, p(\mu_n)),\ \forall v \in X,\ q = 1, \ldots, Q_b,\ n = 1, \ldots, N \\ (\text{resp. } \left(z^{\text{du}\,q,n}, v\right)_X &= b^q(v, \lambda(\mu_n)),\ \forall v \in X,\ q = 1, \ldots, Q_b,\ n = 1, \ldots, M). \end{aligned} \tag{6.12}$$

We then define the primal and dual "pressure" approximation spaces M_N^{pr} and M_M^{du}

$$\begin{aligned} M_N^{\text{pr}} &= \text{span}\,\{p(\mu_i),\ i = 1, \ldots, N\} \equiv \text{span}\,\{\xi_i^{\text{pr}},\ i = 1, \ldots, N\}\,, \\ M_M^{\text{du}} &= \text{span}\,\{\lambda(\mu_i),\ i = 1, \ldots, M\} \equiv \text{span}\,\{\xi_i^{\text{du}},\ i = 1, \ldots, M\}\,; \end{aligned} \tag{6.13}$$

and the "velocity" approximation spaces $X_N^{\text{pr}}(\mu)$ and $X_M^{\text{du}}(\mu)$

$$
\begin{aligned}
X_N^{\text{pr}}(\mu) &= \text{span}\left\{ u(\mu_i), \sum_{q=1}^{Q_b} \sigma_b^q(\mu) z^{\text{pr}\,q,i},\ i = 1,\dots,N \right\} \equiv \text{span}\left\{ \zeta_i^{\text{pr}},\ i = 1,\dots,2N \right\}, \\
X_M^{\text{du}}(\mu) &= \text{span}\left\{ \psi(\mu_i), \sum_{q=1}^{Q_b} \sigma_b^q(\mu) z^{\text{du}\,q,i},\ i = 1,\dots,M \right\} \equiv \text{span}\left\{ \zeta_i^{\text{du}},\ i = 1,\dots,2M \right\};
\end{aligned}
\tag{6.14}
$$

with dimensions $\dim M_N^{\text{pr}} = N$, $\dim X_N^{\text{pr}}(\mu) = 2N$, $\dim M_M^{\text{du}} = M$, and $\dim X_M^{\text{du}}(\mu) = 2M$. The product spaces $Y_N^{\text{pr}}(\mu) = X_N^{\text{pr}}(\mu) \times M_N^{\text{pr}}$ and $Y_M^{\text{du}}(\mu) = X_M^{\text{du}}(\mu) \times M_M^{\text{du}}$ will also be useful in the following.

If an approximation to the inf-sup parameter is required, we form $\mathcal{S}_K^\chi$ by choosing K points $\mu_i \in \mathcal{D}$, and compute the infimizers $\chi(\mu_i)$, by solving the implied eigenvalue problem of (6.11), for all $\mu_i \in \mathcal{S}_K^\chi$. In addition, we compute $z^{\chi\,q,n} \in X$ for $q = 1,\dots,Q_b$, and $n = 1,\dots K$, which satisfy:

$$
\left(z^{\chi\,q,n}, v\right)_X = b^q(v, \chi(\mu_n)),\ \forall v \in X,\ q = 1,\dots,Q_b,\ n = 1,\dots,K; \tag{6.15}
$$

and define M_K^χ and $X_K^\chi(\mu)$

$$
\begin{aligned}
M_K^\chi &= \text{span}\left\{ \chi(\mu_i),\ i = 1,\dots,K \right\} \equiv \text{span}\left\{ \xi_i^\chi,\ i = 1,\dots,K \right\}, \\
X_K^\chi(\mu) &= \text{span}\left\{ \sum_{q=1}^{Q_b} \sigma_b^q(\mu) z^{\chi\,q,i},\ i = 1,\dots,K \right\} \equiv \text{span}\left\{ \zeta_i^\chi,\ i = 1,\dots,2K \right\};
\end{aligned}
\tag{6.16}
$$

with dimension $\dim M_K^\chi = \dim X_K^\chi(\mu) = K$.

6.2.2 Reduced-Basis Problems

Output Approximation

Using the problem-specific approximation spaces of Section 6.2.1, we can define the reduced-basis problems. We look for $[u_N(\mu), p_N(\mu)] \in Y_N^{\text{pr}}(\mu)$ and $[\psi_M(\mu), \lambda_M(\mu)] \in Y_M^{\text{du}}(\mu)$, such

that:

$$\begin{aligned} a(u_N(\mu);\mu) + b(v, p_N(\mu);\mu) &= \ell(v;\mu),\ \forall v \in X_N^{\mathrm{pr}}(\mu), \\ b(u_N(\mu), q;\mu) &= 0,\ \forall q \in M_N^{\mathrm{pr}}, \end{aligned} \tag{6.17}$$

and,

$$\begin{aligned} a(v, \psi_M(\mu);\mu) + b(v, \lambda_M(\mu);\mu) &= -\ell_u^O(v;\mu),\ \forall v \in X_M^{\mathrm{du}}(\mu), \\ b(\psi_M(\mu), q;\mu) &= -\ell_p^O(q;\mu),\ \forall q \in M_M^{\mathrm{du}}, \end{aligned} \tag{6.18}$$

respectively. If $[u_N, p_N] \in Y_N^{\mathrm{pr}}(\mu)$ and, $[e_u^{\mathrm{pr}}, e_p^{\mathrm{pr}}](\mu) \equiv [u - u_N, p - p_M](\mu)$ is the error, the residual $R_u^{\mathrm{pr}}(\cdot;[u_N, p_N];\mu) \in X'$ is defined

$$\begin{aligned} R_u^{\mathrm{pr}}(v;[u_N, p_N];\mu) &= \ell(v;\mu) - a(u_N(\mu), v;\mu) - b(v, p_N(\mu);\mu), \\ &= a(e_u^{\mathrm{pr}}(\mu), v;\mu) + b(v, e_p^{\mathrm{pr}}(\mu)); \end{aligned} \tag{6.19}$$

where the second line follows from equation (6.3). Similarly the residual related to the incompressibility constraint $R_p^{\mathrm{pr}}(\cdot;[u_N, p_N];\mu) \in M'$ is

$$\begin{aligned} R_p^{\mathrm{pr}}(q;[u_N, p_N];\mu) &= -b(u_N(\mu), q;\mu) \\ &= b(e_u^{\mathrm{pr}}(\mu), q;\mu). \end{aligned} \tag{6.20}$$

We can the define the primal residual $R^{\mathrm{pr}}(\cdot;[u_N, p_N];\mu) \in Y'$, from

$$\begin{aligned} R^{\mathrm{pr}}([w, q];[u_N, p_N];\mu) &= R_u^{\mathrm{pr}}(w;[u_N, p_N];\mu) + R_p^{\mathrm{pr}}(q;[u_N, p_N];\mu) \\ &= a(e_u^{\mathrm{pr}}(\mu), v;\mu) + b(v, e_p^{\mathrm{pr}}(\mu)) + b(e_u^{\mathrm{pr}}(\mu), q;\mu). \end{aligned} \tag{6.21}$$

For the dual problem, if $[\psi_M, \lambda_M] \in Y_M^{\mathrm{du}}(\mu)$ and, $[e_u^{\mathrm{du}}, e_p^{\mathrm{du}}](\mu) \equiv [\psi - \psi_M, \lambda - \lambda_M](\mu)$ is the error, we define in a similar way the residuals $R_u^{\mathrm{du}}(\cdot;[\psi_M, \lambda_M];\mu) \in X'$ and $R_p^{\mathrm{du}}(\cdot;[\psi_M, \lambda_M];\mu) \in M'$:

$$\begin{aligned} R_u^{\mathrm{du}}(v;[\psi_M, \lambda_M];\mu) &= -\ell_u^O(v;\mu) - a(v, \psi_M(\mu);\mu) - b(v, \lambda_M(\mu);\mu), \\ &= a(e_u^{\mathrm{du}}(\mu), v;\mu) + b(v, e_p^{\mathrm{du}}(\mu)); \end{aligned} \tag{6.22}$$

and

$$\begin{aligned} R_p^{\mathrm{du}}(q;[\psi_M,\lambda_M];\mu) &= -\ell_p^O(q;\mu) - b(\psi_M(\mu),q;\mu) \\ &= b(e_u^{\mathrm{du}}(\mu),q;\mu). \end{aligned} \tag{6.23}$$

The dual residual is then $R^{\mathrm{du}}(\cdot;[w_M,q_M];\mu) \in Y'$ is then

$$\begin{aligned} R^{\mathrm{du}}([w,q];[\psi_M,\lambda_M];\mu) &= R_u^{\mathrm{du}}(w;[\psi_M,\lambda_M];\mu) + R_p^{\mathrm{du}}(q;[\psi_M,\lambda_M];\mu) \\ &= a(v,e_u^{\mathrm{du}}(\mu);\mu) + b(v,e_p^{\mathrm{du}}(\mu)) + b(e_u^{\mathrm{du}}(\mu),q;\mu). \end{aligned} \tag{6.24}$$

Regarding the stability of the discrete problems (6.17) and (6.18), we have the coercivity constant $\alpha_{N\,\{M\}}^{\mathrm{pr}\,\{\mathrm{du}\}}(\mu)$

$$\alpha_{N\,\{M\}}^{\mathrm{pr}\,\{\mathrm{du}\}}(\mu) = \inf_{w_{N\,\{M\}} \in X_{N\,\{M\}}^{\mathrm{pr}\,\{\mathrm{du}\}}} \frac{a(w_{N\,\{M\}}, w_{N\,\{M\}};\mu)}{\|w_{N\,\{M\}}\|_X^2}; \tag{6.25}$$

and the inf-sup parameter $\beta_{N\,\{M\}}^{\mathrm{pr}\,\{\mathrm{du}\}}(\mu)$:

$$\beta_{N\,\{M\}}^{\mathrm{pr}\,\{\mathrm{du}\}}(\mu) = \inf_{q_{N\,\{M\}} \in M_{N\,\{M\}}^{\mathrm{pr}\,\{\mathrm{du}\}}} \sup_{w_{N\,\{M\}} \in X_{N\,\{M\}}^{\mathrm{pr}\,\{\mathrm{du}\}}} \frac{b(w_{N\,\{M\}}, q_{N\,\{M\}};\mu)}{\|w_{N\,\{M\}}\|_X \|q_{N\,\{M\}}\|_M}; \tag{6.26}$$

where inside the braces are the modifications of these definitions for the dual problem. For stability of the reduced-basis problems, it is required that these constants are strictly positive; we further discuss stability in Lemma 6.2.1.

The output approximation is then obtained from

$$s_N(\mu) = \ell^O([u_N,p_N](\mu);\mu) - R^{\mathrm{pr}}([\psi_M,\lambda_M](\mu);[u_N,p_N](\mu);\mu); \tag{6.27}$$

the adjoint correction helps improve the accuracy.

Regarding the stability of the reduced-basis problems, we have the following result:

Lemma 6.2.1. *For the discrete coercivity constant $\alpha_N^{\mathrm{pr}}(\mu)$, defined in (6.25) we have:*

$$\alpha_N^{\mathrm{pr}}(\mu) \geq \alpha(\mu), \ \forall \mu \in \mathcal{D}, \tag{6.28}$$

and for the inf-sup parameter $\beta_N^{\rm pr}(\mu)$, defined in (6.26) we have:

$$\beta_N^{\rm pr}(\mu) \geq \beta(\mu), \ \forall \mu \in \mathcal{D}. \tag{6.29}$$

Similar results apply for the dual problem, and also for the inf-sup parameter approximation described in the following section.

Proof. We discuss here only the primal problem. Regarding the coercivity constant since, by definition $X_N^{\rm pr}(\mu) \subset X$, we have that

$$\alpha_N^{\rm pr}(\mu) = \inf_{w_N \in X_N^{\rm pr}(\mu)} \frac{a(w_N, w_N; \mu)}{\|w_N\|_X^2} \geq \inf_{w \in X} \frac{a(w, w; \mu)}{\|w\|_X^2} = \alpha(\mu), \ \forall \mu \in \mathcal{D}.$$

For the inf-sup condition, we notice that any member $q_N \in M_N^{\rm pr}$ can be written $q_N = \sum_{i=1}^N q_{N\,i} \xi_i^{\rm pr}$. Therefore the supremizer, defined in (6.9), can be computed:

$$(T_\mu q_N, v)_X = b(v, q_N; \mu), \ \forall v \in X. \tag{6.30}$$

Using now the affine decomposition assumption we notice that:

$$\begin{aligned}
(T_\mu q_N, v)_X &= \sum_{q=1}^{Q_b} \sigma_b^q(\mu) b^q(v, q_N) \\
&= \sum_{q=1}^{Q_b} \sum_{i=1}^{N} \sigma_b^q(\mu) q_{N\,i} b^q(v, \xi_i^{\rm pr}) \\
&= \sum_{i=1}^{N} q_{N\,i} \left(\sum_{q=1}^{Q_b} \sigma_b^q(\mu) z^{{\rm pr}\ q,i}, v \right)_X ;
\end{aligned}$$

the definition of $X_N^{\rm pr}(\mu)$, (6.14), implies then that $T_\mu q_N(\mu) \in X_N^{\rm pr}(\mu)$. Now notice that if we define $T_\mu^N q_N(\mu) \in X_N^{\rm pr}(\mu)$ the supremizer over $X_N^{\rm pr}(\mu)$:

$$(T_\mu^N q_N, v)_X = b(v, q_N; \mu), \ \forall v \in X_N^{\rm pr}(\mu),$$

and subtracting from the equation above Equation (6.30), we have that for $q_N \in M_N^{\text{pr}}$,

$$(T_\mu^N q_N - T_\mu q_N, v)_X = 0, \ \forall v \in X_N^{\text{pr}}.$$

Therefore choosing $v = T_\mu^N q_N - T_\mu q_N \in X_N^{\text{pr}}$ (from the argument above), we have that $\|T_\mu^N q_N - T_\mu q_N\|_X = 0$ and therefore $T_\mu^N q_N = T_\mu q_N, \ \forall q_N \in M_N^{\text{pr}}$. Therefore we conclude that:

$$\beta(\mu) = \inf_{q_N \in M} \frac{\|T_\mu q_N\|_X}{\|q_N\|_M} \leq \inf_{q_N \in M_N^{\text{pr}}} \frac{\|T_\mu q_N\|_X}{\|q_N\|_M} = \inf_{q_N \in M_N^{\text{pr}}} \frac{\|T_\mu^N q_N\|_X}{\|q_N\|_M} = \beta_N^{\text{pr}}(\mu),$$

as desired. □

Remark 6.2.2. *In the construction of the reduced-basis spaces, we do not necessarily need to choose an equal number of pressure and velocity modes. We can choose N_u^{pr} velocity basis functions for $X_N^{\text{pr}}(\mu)$, and N_p^{pr} basis functions for the for M_p^{pr}. Following the previous Lemma, for stability, we need to augment $X_N^{\text{pr}}(\mu)$ with N_p^{pr} basis functions — and therefore $\dim X_N^{\text{pr}}(\mu) = N_u^{\text{pr}} + N_p^{\text{pr}}$. We discuss how different possible choices affect the accuracy of our predictions in Section 6.5.*

Inf-Sup Parameter Approximation

If also an approximation $\beta_K(\mu)$ to the exact inf-sup parameter $\beta(\mu)$ is also required, we use the reduced-basis spaces $X_K^\chi(\mu)$ and M_K^χ. The inf-sup parameter approximation is then obtained from

$$\beta_K(\mu) = \inf_{w_K \in X_K^\chi(\mu)} \sup_{q_K \in M_K^\chi} \frac{b(w_k, q_k; \mu)}{\|w_K\|_X \|q_K\|_M} = \inf_{w_K \in X_K^\chi(\mu)} \frac{\|T_\mu^K q_K\|_X}{\|q_K\|_M}; \tag{6.31}$$

where for $q_k \in M_K^\chi$, $T_\mu^K q_k \in X_K^\chi(\mu)$, is the solution of

$$(T_\mu^K q_k, v) = b(v, q_k; \mu), \ \forall v \in X_K^\chi(\mu).$$

The infimizer for (6.31), can be obtained by solution of an appropriately defined symmetric positive-definite eigenvalue problem

$$(T_\mu^K \theta(\mu), T_\mu^K q)_X = \rho(\mu)(\theta(\mu), q)_M, \ \forall q \in M_K^\chi; \tag{6.32}$$

the inf-sup parameter is then $\beta_K(\mu) = \sqrt{\rho^{\min}(\mu)}$, where $\rho^{\min}(\mu)$ is the minimum eigenvalue of (6.32). The discussion in Section 5.3.3, regarding the convergence of $\beta_K(\mu)$ to $\beta(\mu)$ also applies here.

6.3 Computational Procedure

The parametric dependence assumed in (6.8) permits us to decouple the computation in two stages: the *off-line* stage, in which (i) the reduced basis is constructed and (ii), some preprocessing is performed; and the *on-line* stage, in which for each new desired value $\mu \in \mathcal{D}$, we compute $s_N(\mu)$. The details of the blackbox technique follow.

6.3.1 Output Prediction

The presentation follows the procedure and the notation introduced in Section 5.4.2.

Off-line Stage

1) Choose $\mathcal{S}_N^{\rm pr}$ and $\mathcal{S}_M^{\rm du}$. For all $\mu_i \in \mathcal{S}_N^{\rm pr}$, calculate $[u(\mu_i), p(\mu_i)] \equiv [\zeta_i^{\rm pr}, \xi_i^{\rm pr}] \in Y$, $i = 1, \ldots, N$, the solution of (6.3). Similarly for the dual, for all $\mu_i \in \mathcal{S}_M^{\rm du}$, calculate $[\psi(\mu_i), \lambda(\mu_i)] \equiv [\zeta_i^{\rm du}, \xi_i^{\rm du}] \in Y$, $i = 1, \ldots, M$, the solution of (6.5).
2) Compute $z^{{\rm pr}\,q,i} \in X$, $q = 1, \ldots, Q_b$, $i = 1, \ldots, N$ and $z^{{\rm du}\,q,n} \in X$, $q = 1, \ldots, Q_b$ $j = 1, \ldots, M$, as in (6.12).
3) Compute $\underline{A}_q^{{\rm pr}\,11} \in \mathbb{R}^{N\times N}$, $\underline{A}_q^{{\rm du}\,11} \in \mathbb{R}^{M\times M}$ and $\underline{A}_q^{{\rm prdu}\,11} \in \mathbb{R}^{M\times N}$ for $q = 1, \ldots, Q_a$, where

$$A_{q\,i,j}^{{\rm pr}\,11} = a^q(\zeta_j^{\rm pr}, \zeta_i^{\rm pr}), \ 1 \le i,j \le N, \ A_{q\,i,j}^{{\rm du}\,11} = a^q(\zeta_j^{\rm du}, \zeta_i^{\rm du}), \ 1 \le i,j \le M,$$
$$A_{q\,i,j}^{{\rm prdu}\,11} = a^q(\zeta_j^{\rm pr}, \zeta_i^{\rm du}), \ 1 \le i \le M, \ 1 \le j \le N;$$

also, $\underline{A}^{\mathrm{pr}\,12}_{q,q'} \in \mathbb{R}^{N\times N}$, $\underline{A}^{\mathrm{du}\,12}_{q,q'} \in \mathbb{R}^{M\times M}$, $\underline{A}^{\mathrm{prdu}\,12}_{q,q'} \in \mathbb{R}^{M\times N}$ and $\underline{A}^{\mathrm{du}\,21}_{q,q'} \in \mathbb{R}^{M\times N}$, $q = 1,\ldots,Q_a$, $q' = 1,\ldots,Q_b$, as

$$A^{\mathrm{pr}\,12}_{q,q'\,i,j} = a^q(z^{\mathrm{pr}\,q',j},\zeta^{\mathrm{pr}}_i),\ 1\le i,j\le N,\ A^{\mathrm{du}\,12}_{q,q'\,i,j} = a^q(z^{\mathrm{du}\,q',j},\zeta^{\mathrm{du}}_i),\ 1\le i,j\le M,$$
$$A^{\mathrm{prdu}\,12}_{q,q'\,i,j} = a^q(z^{\mathrm{pr}\,q',j},\zeta^{\mathrm{du}}_i),\ A^{\mathrm{prdu}\,21}_{q,q'\,i,j} = a^q(\zeta^{\mathrm{pr}}_j,z^{\mathrm{du}\,q',i}),\ 1\le i\le M,\ 1\le j\le N;$$

and $\underline{A}^{\mathrm{pr}\,22}_{q,q',q''} \in \mathbb{R}^{N\times N}$, $\underline{A}^{\mathrm{du}\,22}_{q,q',q''} \in \mathbb{R}^{M\times M}$ and $\underline{A}^{\mathrm{prdu}\,22}_{q,q',q''} \in \mathbb{R}^{M\times N}$, for $q = 1,\ldots,Q_a$ and $q',q'' = 1,\ldots,Q_b$, such that

$$A^{\mathrm{pr}\,22}_{q,q',q''\,i,j} = a^q(z^{\mathrm{pr}\,q',j},z^{\mathrm{pr}\,q'',i}),\ 1\le i,j\le N,\ A^{\mathrm{du}\,22}_{q,q',q''\,i,j} = a^q(z^{\mathrm{du}\,q',j},z^{\mathrm{du}\,q'',i}),\ 1\le i,j\le M,$$
$$A^{\mathrm{prdu}\,22}_{q,q',q''\,i,j} = a^q(z^{\mathrm{pr}\,q',j},z^{\mathrm{du}\,q'',i}),\ 1\le i\le M,\ 1\le j\le N.$$

4) Compute $\underline{B}^{\mathrm{pr}\,1}_q \in \mathbb{R}^{N\times N}$, $\underline{B}^{\mathrm{du}\,1}_q \in \mathbb{R}^{M\times M}$, ${}^1\underline{B}^{\mathrm{prdu}\,1}_q \in \mathbb{R}^{M\times N}$ and ${}^2\underline{B}^{\mathrm{prdu}\,1}_q \in \mathbb{R}^{N\times M}$, for $q = 1,\ldots,Q_b$

$$B^{\mathrm{pr}\,1}_{q\,i,j} = b^q(\zeta^{\mathrm{pr}}_j,\xi^{\mathrm{pr}}_i),\ 1\le i,j\le N,\ B^{\mathrm{du}\,1}_{q\,i,j} = b^q(\zeta^{\mathrm{du}}_j,\xi^{\mathrm{du}}_i),\ 1\le i,j\le M,$$
$${}^1B^{\mathrm{prdu}\,1}_{q\,i,j} = b^q(\zeta^{\mathrm{pr}}_j,\xi^{\mathrm{du}}_i),\ {}^2B^{\mathrm{prdu}\,1}_{q\,i,j} = b^q(\zeta^{\mathrm{du}}_j,\xi^{\mathrm{pr}}_i),\ 1\le i\le M,\ 1\le j\le N;$$

and also $\underline{B}^{\mathrm{pr}\,2}_{q,q'} \in \mathbb{R}^{N\times N}$, $\underline{B}^{\mathrm{du}\,2}_{q,q'} \in \mathbb{R}^{M\times M}$, ${}^1\underline{B}^{\mathrm{prdu}\,2}_{q,q'} \in \mathbb{R}^{M\times N}$ and ${}^2\underline{B}^{\mathrm{prdu}\,2}_{q,q'} \in \mathbb{R}^{M\times N}$, for $q,q' = 1,\ldots,Q_b$, as

$$B^{\mathrm{pr}\,2}_{q,q'\,i,j} = b^q(z^{\mathrm{pr}\,q',j},\xi^{\mathrm{pr}}_i),\ 1\le i,j\le N,\ B^{\mathrm{du}\,2}_{q,q'\,i,j} = b^q(z^{\mathrm{du}\,q',j},\xi^{\mathrm{du}}_i),\ 1\le i,j\le N,$$
$${}^1B^{\mathrm{prdu}\,2}_{q,q'\,i,j} = b^q(z^{\mathrm{pr}\,q',j},\xi^{\mathrm{du}}_i),\ {}^2B^{\mathrm{prdu}\,2}_{q,q'\,i,j} = b^q(z^{\mathrm{du}\,q',j},\xi^{\mathrm{pr}}_i),\ 1\le i\le M,\ 1\le j\le N.$$

5) Compute $\underline{F}^{\mathrm{pr}\,1}_q \in \mathbb{R}^N$, and $\underline{F}^{\mathrm{du}\,1}_q \in \mathbb{R}^M$, for $q = 1,\ldots,Q_f$ as

$$F^{\mathrm{pr}\,1}_{q\,i} = \ell^q_f(\zeta^{\mathrm{pr}}_i),\ 1\le i\le N,\ F^{\mathrm{du}\,1}_{q\,j} = \ell^q_f(\zeta^{\mathrm{du}}_j),\ 1\le j\le M;$$

and also $\underline{F}^{\mathrm{pr}\,2}_{q,q'} \in \mathbb{R}^N$ and $\underline{F}^{\mathrm{du}\,2}_{q,q'} \in \mathbb{R}^M$, for $q = 1,\ldots,Q_f$, $q' = 1,\ldots Q_b$, as

$$F^{\mathrm{pr}\,2}_{q,q'\,i} = \ell^q_f(z^{\mathrm{pr}\,q',i}),\ 1\le i\le N,\ F^{\mathrm{du}\,2}_{q,q'\,j} = \ell^q_f(z^{\mathrm{du}\,q',j}),\ 1\le j\le M.$$

6) Compute $\underline{L}_q^{\mathrm{pr}\,1} \in \mathbb{R}^N$, and $\underline{L}_q^{\mathrm{du}\,1} \in \mathbb{R}^M$, for $q = 1, \ldots, Q_O$ as

$$L_{q\,i}^{\mathrm{pr}\,1} = \ell_O^q([\zeta_i^{\mathrm{pr}}, 0]),\ 1 \le i \le N,\ F_{q\,j}^{\mathrm{du}\,1} = \ell_O^q([\zeta_j^{\mathrm{du}}, 0]),\ 1 \le j \le M;$$

also $\underline{L}_{q,q'}^{\mathrm{pr}\,2} \in \mathbb{R}^N$ and $\underline{L}_{q,q'}^{\mathrm{du}\,2} \in \mathbb{R}^M$, for $q = 1, \ldots, Q_O$, $q' = 1, \ldots Q_b$, as

$$L_{q,q'\,i}^{\mathrm{pr}\,2} = \ell_O^q([z^{\mathrm{pr}\,q',i}, 0]),\ 1 \le i \le N,\ L_{q,q'\,j}^{\mathrm{du}\,2} = \ell_O^q([z^{\mathrm{du}\,q',j}, 0]),\ 1 \le j \le M;$$

and also $\underline{L}_q^{\mathrm{pr}\,3} \in \mathbb{R}^N$, and $\underline{L}_q^{\mathrm{du}\,3} \in \mathbb{R}^M$, for $q = 1, \ldots, Q_O$ as

$$L_{q\,i}^{\mathrm{pr}\,3} = \ell_O^q([0, \xi_i^{\mathrm{pr}}]),\ 1 \le i \le N,\ F_{q\,j}^{\mathrm{du}\,1} = \ell_O^q([0, \xi_j^{\mathrm{du}}]),\ 1 \le j \le M.$$

On-line Stage

Given a new value of the parameter $\mu \in \mathcal{D}$:

1) We form the matrices $\underline{A}^{\mathrm{pr}}(\mu) \in \mathbb{R}^{2N \times 2N}$, $\underline{A}^{\mathrm{du}}(\mu) \in \mathbb{R}^{2M \times 2M}$ and $\underline{A}^{\mathrm{prdu}}(\mu) \in \mathbb{R}^{2M \times 2N}$

$$\underline{A}^{\mathrm{pr}}(\mu) = \begin{pmatrix} \underbrace{\sum_{q=1}^{Q_a} \sigma_a^q(\mu) \underline{A}_q^{\mathrm{pr}\,11}}_{\underline{A}^{\mathrm{pr}\,11}(\mu)} & \underbrace{\sum_{q=1}^{Q_a} \sum_{q'=1}^{Q_b} \sigma_a^q(\mu) \sigma_b^{q'}(\mu) \underline{A}_{q,q'}^{\mathrm{pr}\,12}}_{\underline{A}^{\mathrm{pr}\,12}(\mu)} \\ \left(\underline{A}^{\mathrm{pr}\,12}(\mu)\right)^T & \underbrace{\sum_{q=1}^{Q_a} \sum_{q',q''=1}^{Q_b} \sigma_a^q(\mu) \sigma_b^{q'}(\mu) \sigma_b^{q''}(\mu) \underline{A}_{q,q',q''}^{\mathrm{pr}\,22}}_{\underline{A}^{\mathrm{pr}\,22}(\mu)} \end{pmatrix}$$

$$\underline{A}^{\mathrm{du}}(\mu) = \begin{pmatrix} \underbrace{\sum_{q=1}^{Q_a} \sigma_a^q(\mu) \underline{A}_q^{\mathrm{du}\,11}}_{\underline{A}^{\mathrm{du}\,11}(\mu)} & \underbrace{\sum_{q=1}^{Q_a} \sum_{q'=1}^{Q_b} \sigma_a^q(\mu) \sigma_b^{q'}(\mu) \underline{A}_{q,q'}^{\mathrm{du}\,12}}_{\underline{A}^{\mathrm{du}\,12}(\mu)} \\ \left(\underline{A}^{\mathrm{du}\,12}(\mu)\right)^T & \underbrace{\sum_{q=1}^{Q_a} \sum_{q',q''=1}^{Q_b} \sigma_a^q(\mu) \sigma_b^{q'}(\mu) \sigma_b^{q''}(\mu) \underline{A}_{q,q',q''}^{\mathrm{du}\,22}}_{\underline{A}^{\mathrm{du}\,22}(\mu)} \end{pmatrix}$$

$$\underline{A}^{\mathrm{prdu}}(\mu) = \begin{pmatrix} \underbrace{\sum_{q=1}^{Q_a} \sigma_a^q(\mu)\underline{A}_q^{\mathrm{prdu}\,11}}_{\underline{A}^{\mathrm{prdu}\,11}(\mu)} & \underbrace{\sum_{q=1}^{Q_a}\sum_{q'=1}^{Q_b} \sigma_a^q(\mu)\sigma_b^{q'}(\mu)\underline{A}_{q,q'}^{\mathrm{prdu}\,12}}_{\underline{A}^{\mathrm{prdu}\,12}(\mu)} \\ \underbrace{\sum_{q=1}^{Q_a}\sum_{q'=1}^{Q_b} \sigma_a^q(\mu)\sigma_b^{q'}(\mu)\underline{A}_{q,q'}^{\mathrm{prdu}\,21}}_{\underline{A}^{\mathrm{prdu}\,21}(\mu)} & \underbrace{\sum_{q=1}^{Q_a}\sum_{q',q''=1}^{Q_b} \sigma_a^q(\mu)\sigma_b^{q'}(\mu)\sigma_b^{q''}(\mu)\underline{A}_{q,q',q''}^{\mathrm{prdu}\,22}}_{\underline{A}^{\mathrm{prdu}\,22}(\mu)} \end{pmatrix}$$

2) We form the matrices $\underline{B}^{\mathrm{pr}}(\mu) \in \mathbb{R}^{N\times 2N}$, $\underline{B}^{\mathrm{du}}(\mu) \in \mathbb{R}^{M\times 2M}$, ${}^1\underline{B}^{\mathrm{prdu}} \in \mathbb{R}^{M\times 2N}$ and ${}^2\underline{B}^{\mathrm{prdu}} \in \mathbb{R}^{N\times 2M}$, as

$$\begin{aligned}
\underline{B}^{\mathrm{pr}}(\mu) &= \begin{pmatrix} \sum_{q=1}^{Q_b} \sigma_b^q(\mu)\underline{B}_q^{\mathrm{pr}\,1} & \sum_{q,q'=1}^{Q_b} \sigma_b^q(\mu)\sigma_b^{q'}(\mu)\underline{B}_{q,q'}^{\mathrm{pr}\,2} \end{pmatrix} \\
\underline{B}^{\mathrm{du}}(\mu) &= \begin{pmatrix} \sum_{q=1}^{Q_b} \sigma_b^q(\mu)\underline{B}_q^{\mathrm{du}\,1} & \sum_{q,q'=1}^{Q_b} \sigma_b^q(\mu)\sigma_b^{q'}(\mu)\underline{B}_{q,q'}^{\mathrm{du}\,2} \end{pmatrix} \\
{}^1\underline{B}^{\mathrm{prdu}}(\mu) &= \begin{pmatrix} \sum_{q=1}^{Q_b} \sigma_b^q(\mu){}^1\underline{B}_q^{\mathrm{prdu}\,1} & \sum_{q,q'=1}^{Q_b} \sigma_b^q(\mu)\sigma_b^{q'}(\mu){}^1\underline{B}_{q,q'}^{\mathrm{prdu}\,2} \end{pmatrix} \\
{}^2\underline{B}^{\mathrm{prdu}}(\mu) &= \begin{pmatrix} \sum_{q=1}^{Q_b} \sigma_b^q(\mu){}^2\underline{B}_q^{\mathrm{prdu}\,1} & \sum_{q,q'=1}^{Q_b} \sigma_b^q(\mu)\sigma_b^{q'}(\mu){}^2\underline{B}_{q,q'}^{\mathrm{prdu}\,2} \end{pmatrix}
\end{aligned}$$

3) Form the vectors $\underline{F}^{\mathrm{pr}}(\mu) \in \mathbb{R}^{3N}$ and $\underline{F}^{\mathrm{du}}(\mu) \in \mathbb{R}^{3M}$ as:

$$\underline{F}^{\mathrm{pr}}(\mu) = \begin{pmatrix} \sum_{q=1}^{Q_f} \sigma_f^q(\mu)\underline{F}_q^{\mathrm{pr}\,1} \\ \sum_{q=1}^{Q_f}\sum_{q'=1}^{Q_b} \sigma_f^q(\mu)\sigma_b^{q'}(\mu)\underline{F}_{q,q'}^{\mathrm{pr}\,2} \\ \underline{0}_N \end{pmatrix}, \quad \underline{F}^{\mathrm{du}}(\mu) = \begin{pmatrix} \sum_{q=1}^{Q_f} \sigma_f^q(\mu)\underline{F}_q^{\mathrm{du}\,1} \\ \sum_{q=1}^{Q_f}\sum_{q'=1}^{Q_b} \sigma_f^q(\mu)\sigma_b^{q'}(\mu)\underline{F}_{q,q'}^{\mathrm{du}\,2} \\ \underline{0}_M \end{pmatrix};$$

where $\underline{0}_N$, $\underline{0}_M$ is the N and M-dimensional vector of zeros.

4) Form the vectors $\underline{L}^{\mathrm{pr}}(\mu) \in \mathbb{R}^{3N}$ and $\underline{L}^{\mathrm{du}}(\mu) \in \mathbb{R}^{3M}$ as:

$$\underline{L}^{\mathrm{pr}}(\mu) = \begin{pmatrix} \sum_{q=1}^{Q_O} \sigma_O^q(\mu)\underline{L}_q^{\mathrm{pr}\,1} \\ \sum_{q=1}^{Q_O}\sum_{q'=1}^{Q_b} \sigma_O^q(\mu)\sigma_b^{q'}(\mu)\underline{L}_{q,q'}^{\mathrm{pr}\,2} \\ \sum_{q=1}^{Q_O} \sigma_O^q(\mu)\underline{L}_q^{\mathrm{pr}\,3} \end{pmatrix}, \quad \underline{L}^{\mathrm{du}}(\mu) = \begin{pmatrix} \sum_{q=1}^{Q_f} \sigma_O^q(\mu)\underline{L}_q^{\mathrm{du}\,1} \\ \sum_{q=1}^{Q_f}\sum_{q'=1}^{Q_b} \sigma_O^q(\mu)\sigma_b^{q'}(\mu)\underline{L}_{q,q'}^{\mathrm{du}\,2} \\ \sum_{q=1}^{Q_f} \sigma_O^q(\mu)\underline{L}_q^{\mathrm{du}\,3} \end{pmatrix}.$$

5) Compute $[\underline{u}_N, \underline{p}_N](\mu) \in \mathbb{R}^{2N}\times\mathbb{R}^N$ the reduced-basis solution for the primal problem (6.17):

$$\begin{pmatrix} \underline{A}^{\mathrm{pr}}(\mu) & (\underline{B}^{\mathrm{pr}}(\mu))^T \\ \underline{B}^{\mathrm{pr}}(\mu) & \underline{0}_{N\times N} \end{pmatrix} \begin{pmatrix} \underline{u}_N(\mu) \\ \underline{p}_N(\mu) \end{pmatrix} = \underline{F}^{\mathrm{pr}}(\mu);$$

and $[\underline{\psi}_M(\mu), \underline{\lambda}_M](\mu) \in \mathbb{R}^{2M} \times \mathbb{R}^M$ the reduced-basis solution for the dual problem (6.18):

$$\begin{pmatrix} \underline{A}^{\mathrm{du}}(\mu) & \left(\underline{B}^{\mathrm{du}}(\mu)\right)^T \\ \underline{B}^{\mathrm{du}}(\mu) & \underline{0}_{M\times M} \end{pmatrix} \begin{pmatrix} \underline{\psi}_M(\mu) \\ \underline{\lambda}_M(\mu) \end{pmatrix} = -\underline{L}^{\mathrm{du}}(\mu);$$

6) The output can then be calculated from:

$$s_N(\mu) = \left((\underline{u}_N(\mu))^T \quad (\underline{p}_N(\mu))^T\right) \underline{L}^{\mathrm{pr}}(\mu)$$
$$-(\underline{\psi}_M(\mu))^T \left[\underline{F}^{\mathrm{du}}(\mu) - \underline{A}^{\mathrm{prdu}}(\mu)\underline{u}_N(\mu) - ({}^2\underline{B}^{\mathrm{prdu}}(\mu))^T \underline{p}_N(\mu)\right] - (\underline{\lambda}_M(\mu))^T \left({}^1\underline{B}^{\mathrm{prdu}}(\mu)\underline{u}_N(\mu)\right)$$

Computational Complexity

We don't present with details for the computation of the inf-sup parameter. Following the discussion in Section 5.4 and 6.3.1, the development of a similar off-line/on-line procedure, should be straightforward. As in the previous chapters, the off-line step needs to be performed only once. For the computation of the primal and dual basis functions, a total of $N + M$ Stokes problems need to be solved using an iterative method like, for example, the Uzawa algorithm. In addition, $(N + M)Q_b$ Y-solves are required for the calculation of $z^{n,q}$. Finally, a number of matrix-vector and inner products are required for the formation of a number of auxiliary quantities. The important thing to note it that once the expensive and memory intensive off-line part is completed, a database with $\mathcal{O}((N^2 + M^2)Q_aQ_b^2)$ quantities, is created. In the on-line part, for each new $\mu \in \mathcal{D}$, and using this database: first, $\mathcal{O}((N^2 + M^2)Q_aQ_b^2)$ operations are required to form the reduced-basis problems; and second $\mathcal{O}(N^3 + M^3)$ operations are required to invert the resulting linear systems and compute the output approximation. The important thing to note is that no explicit reference is made to the continuous (or, in practice, finite-element) problem. As N and M will typically small, significant computational savings are expected.

6.4 Error Estimation

To start, define λ_X^1 to be the minimum eigenvalue of $a(\theta, v; \mu) = \lambda(\mu)(\theta, v)_X, \ \forall v \in X$. A lower bound for this eigenvalue is required by the output bound procedure: we assume that a $g(\mu)$ and a $c(\mu) > 0$ is known such that

$$g(\mu)(v, v)_X \leq a(v, v; \mu) \leq c(\mu)(v, v)_X, \ \forall v \in X \text{ and } \forall \mu \in \mathcal{D}. \tag{6.33}$$

It is also possible to include approximation of $\lambda_X^1(\mu)$ as part of the reduced basis approximation.

Remark 6.4.1. *In the following, the more general class of bound conditioners, developed in [113], can also be used. In this case condition (6.33), is replaced by a spectral condition*

$$1 \leq \frac{a(v, v; \mu)}{c(v, v)} \leq \rho, \forall v \in X;$$

with $\rho \geq 1$ a positive number — preferably close to 1 — and c is a parameter-independent symmetric positive-definite form. The development of bound-conditioner-based error estimation procedures will be addressed in a future paper.

6.4.1 A Posteriori Error Analysis

Let $[u, p](\mu) \in X$ the exact solution for the primal problem (6.3), and $[u_N, p_N](\mu) \in Y_N$ the reduced-basis approximation obtained by solving (6.17). Subtracting (6.3) and (6.17), the error $[e_u^{\text{pr}}, e_p^{\text{pr}}](\mu) \equiv [u - u_N, p - p_N] \in Y$ to the primal problem satisfies the following equation:

$$\begin{aligned} a(e_u^{\text{pr}}(\mu), v; \mu) + b(v, e_p^{\text{pr}}(\mu)) &= R_u^{\text{pr}}(v; [u_N, p_N](\mu); \mu), \ \forall v \in X, \\ b(e_u^{\text{pr}}(\mu), q; \mu) &= R_p^{\text{pr}}(q; [u_N, p_N](\mu); \mu), \ \forall q \in M; \end{aligned} \tag{6.34}$$

with similar equation valid for the dual error $[e_u^{\text{du}}, e_p^{\text{du}}](\mu) \equiv [\psi - \psi_M, \lambda - p_M](\mu) \in Y$.

We need a few auxiliary results :

Lemma 6.4.2. *Assume that there exists a constant κ such that*

$$b(w,q;\mu) = \sum_{q=1}^{Q_b} \sigma_b^q(\mu) b^q(w,v) \leq \kappa F_b(\mu) \hat{b}(w,q), \ \forall w \in X, \ \forall q \in M; \tag{6.35}$$

with $\hat{b}(w,q) = \int_\Omega \mathrm{div} w \; q \, d\Omega$, and $F_b(\mu) = \max_{q=1,\ldots,Q_b} |\sigma_b^q(\mu)|$. If d is the number of dimensions for the physical domain ($d = 2$ for two-dimensional, and $d = 3$ for three-dimensional domains), then:

$$|b(w,q;\mu)| \leq \kappa \sqrt{d} F_b(\mu) \|w\|_X \|q\|_M, \ \forall w \in X, \ \forall q \in M. \tag{6.36}$$

Proof. Starting from (6.35), we notice that:

$$|b(w,q;\mu)| \leq \kappa F_b(\mu) |\hat{b}(w,q)| \leq \kappa \|\mathrm{div} w\|_M \|q\|_M \leq \kappa \sqrt{d} \|w\|_X \|q\|_M;$$

where we used $\|\mathrm{div} w\|_M \leq \sqrt{d} \|w\|_X$ — see [36] for a proof. □

Remark 6.4.3. *In the case where we separate the domain in many smaller non-overlapping subdomains and apply affine geometric transformations, it is easy to see that the assumption is true for $\kappa = 1$. In the case of overlapping domains and affine geometry transformations, κ is equal to the maximum number of overlapping domains at any point of the computational domain.*

We now construct a bound for $e_p^{\mathrm{pr}}(\mu)$ in terms of the residuals and other computable quantities:

Lemma 6.4.4. *We define $C_p^1(\mu) = \frac{\kappa c(\mu)^2 \sqrt{d}}{\beta(\mu)^2 g(\mu)} F_b(\mu)$ and $C_p^2(\mu) = \frac{c(\mu)^2}{\beta(\mu)^2}$, then a bound for the error in the pressure $\|e_p^{\mathrm{pr}}(\mu)\|_M$ is obtained from:*

$$\|e_p^{\mathrm{pr}}(\mu)\|_M \leq C_p^1(\mu) \|R_u^{\mathrm{pr}}(\cdot; [u_N, p_N](\mu); \mu)\|_{X'} + C_p^2(\mu) \|R_p^{\mathrm{pr}}(\cdot; [u_N, p_N](\mu); \mu)\|_{M'}; \tag{6.37}$$

with an analogous result for the dual error $e_p^{\mathrm{du}}(\mu)$.

Proof. We start by obtaining an equation for $e_p^{\mathrm{pr}}(\mu)$. To this end we define $T_\mu^b e_p^{\mathrm{pr}}(\mu) \in X$,

the solution of

$$a(T_\mu^b e_p^{\rm pr}(\mu), v; \mu) = b(v, e_p^{\rm pr}(\mu); \mu), \ \forall v \in X; \tag{6.38}$$

similarly define $\varphi^{R_u^{\rm pr}}(\mu) \in X$ from

$$a(\varphi^{R_u^{\rm pr}}(\mu), v; \mu) = R_u^{\rm pr}(v; [u_N, p_N](\mu); \mu), \ \forall v \in X; \tag{6.39}$$

from (6.7) a unique solution to both of these problems will exist. From the error equation (6.34) and the coercivity of a, we see that $e_u^{\rm pr}(\mu) = \varphi^{R_u^{\rm pr}}(\mu) - T_\mu^b e_p^{\rm pr}(\mu)$. Replacing this last expression in the second equation of (6.34) we get:

$$b(T_\mu^b e_p^{\rm pr}(\mu), q; \mu) = b(\varphi^{R_u^{\rm pr}}(\mu), q; \mu) - R_p^{\rm pr}(q; [u_N, p_N](\mu); \mu), \ \forall q \in M; \tag{6.40}$$

note that $b(T_\mu^b \cdot, \cdot; \mu)$ is called the Uzawa operator. We now examine each of the terms in (6.40). First, notice that from Lemma 6.4.2,

$$\begin{aligned} |b(\varphi^{R_u^{\rm pr}}(\mu), q; \mu)| &\leq \kappa\sqrt{d} F_b(\mu) \|\varphi^{R_u^{\rm pr}}(\mu)\|_X \|q\|_M \\ &\leq \kappa\sqrt{d} F_b(\mu) \sup_{v \in X} \frac{\|v\|_X}{a(v, v; \mu)^{1/2}} a(\varphi^{R_u^{\rm pr}}(\mu), \varphi^{R_u^{\rm pr}}(\mu); \mu)^{1/2} \|q\|_M. \end{aligned}$$

But from (6.33),

$$\sup_{v \in X} \frac{\|v\|_X}{a(v, v; \mu)^{1/2}} = \frac{1}{\inf\limits_{v \in X} \dfrac{a(v, v; \mu)^{1/2}}{\|v\|_X}} \leq \frac{1}{\sqrt{g(\mu)}};$$

and also from the Riesz theorem, and (6.33):

$$\begin{aligned} a(\varphi^{R_u^{\rm pr}}(\mu), \varphi^{R_u^{\rm pr}}(\mu); \mu)^{1/2} &= \sup_{v \in X} \frac{R_u^{\rm pr}(v; [u_N, p_N](\mu); \mu)}{a(v, v; \mu)^{1/2}} \\ &\leq \frac{1}{\sqrt{g(\mu)}} \sup_{v \in X} \frac{R_u^{\rm pr}(v; [u_N, p_N](\mu); \mu)}{\|v\|_X} \\ &= \frac{1}{\sqrt{g(\mu)}} \|R_u^{\rm pr}(\cdot; [u_N, p_N](\mu); \mu)\|_{X'}. \end{aligned}$$

Therefore

$$|b(\varphi^{R_u^{\rm pr}}(\mu), q; \mu)| \leq \frac{\kappa\sqrt{d}}{g(\mu)} F_b(\mu) \|R_u^{\rm pr}(\cdot; [u_N, p_N](\mu); \mu)\|_{X'} \|q\|_M. \tag{6.41}$$

Also notice that:

$$|R_p^{\rm pr}(q;[u_N,p_N](\mu);\mu)| \le \sup_{q\in M} \frac{R_p^{\rm pr}(q;[u_N,p_N](\mu);\mu)}{\|q\|_M}\|q\|_M \\ = \|R_p^{\rm pr}(\cdot;[u_N,p_N](\mu);\mu)\|_{M'}\|q\|_M \tag{6.42}$$

Combining (6.40), (6.41) and (6.42) we obtain:

$$\frac{|b(T_\mu^b e_p^{\rm pr}(\mu),q;\mu)|}{\|q\|_M} \le \frac{\kappa\sqrt{d}}{g(\mu)}F_b(\mu)\|R_u^{\rm pr}(\cdot;[u_N,p_N](\mu);\mu)\|_{X'}+\|R_p^{\rm pr}(\cdot;[u_N,p_N](\mu);\mu)\|_{M'},\forall q\in M. \tag{6.43}$$

We now choose $q = e_p^{\rm pr}(\mu)$ and notice that from (6.38):

$$b(T_\mu^b e_p^{\rm pr}(\mu), e_p^{\rm pr}(\mu);\mu) = a(T_\mu^b e_p^{\rm pr}(\mu), T_\mu^b e_p^{\rm pr}(\mu);\mu) \\ = \left(\sup_{v\in X}\frac{b(v,e_p^{\rm pr}(\mu);\mu)}{a(v,v;\mu)^{1/2}}\right)^2,$$

where the second line follows from the Riesz representation theorem. We now choose $v = T_\mu e_p^{\rm pr}(\mu)$ which is obtained from:

$$(T_\mu e_p^{\rm pr}(\mu), v)_X = b(v, e^{\rm pr};\mu),\ \forall v\in X.$$

Then

$$b(T_\mu^b e_p^{\rm pr}(\mu), e_p^{\rm pr}(\mu);\mu) \ge \left(\frac{b(T_\mu e_p^{\rm pr}(\mu), e_p^{\rm pr}(\mu);\mu)}{a(T_\mu e_p^{\rm pr}(\mu), T_\mu e_p^{\rm pr}(\mu))^{1/2}}\right)^2 \\ \ge \left(\frac{\beta(\mu)\|T_\mu e_p^{\rm pr}(\mu)\|_X\|e_p^{\rm pr}(\mu)\|_M}{a(T_\mu e_p^{\rm pr}(\mu), T_\mu e_p^{\rm pr}(\mu))^{1/2}}\right)^2 \\ \ge \left(\beta(\mu)\|e_p^{\rm pr}(\mu)\|_M \inf_{v\in X}\frac{\|v\|_X}{a(v,v;\mu)^{1/2}}\right)^2 \\ \ge \left(\frac{\beta(\mu)}{c(\mu)}\right)^2 \|e_p^{\rm pr}(\mu)\|_M^2; \tag{6.44}$$

where the second line follows from the definition of the inf-sup parameter, and the last line from the right-hand side of (6.33). Combining now, (6.44) and (6.43) we obtain the desired result. A bound for the dual error $e_p^{\rm du}(\mu)$ can be obtained similarly. □

We now develop a similar bound for the error in the velocity $e_u^{\text{pr}}(\mu)$

Lemma 6.4.5. *We define the μ-dependent constants $C_u^1(\mu) = (1 + \kappa\sqrt{d}F_b(\mu)C_p^1(\mu))/g(\mu)$ and $C_u^2(\mu) = \kappa\sqrt{d}F_b(\mu)C_p^2(\mu)/g(\mu)$, then a bound for the error in the velocity $\|e_u^{\text{pr}}(\mu)\|_X$ is obtained from:*

$$\|e_u^{\text{pr}}(\mu)\|_X \leq C_u^1(\mu)\|R_u^{\text{pr}}(\cdot;[u_N,p_N](\mu);\mu)\|_{X'} + C_u^2(\mu)\|R_p^{\text{pr}}(\cdot;[u_N,p_N](\mu);\mu)\|_{M'}; \tag{6.45}$$

with an analogous result for the dual error $e_u^{\text{du}}(\mu) \in X$.

Proof. We start from the first equation of (6.34)

$$a(e_u^{\text{pr}}(\mu),v;\mu) + b(v,e_p^{\text{pr}}(\mu)) = R_u^{\text{pr}}(v;[u_N,p_N](\mu);\mu),\ \forall v \in X.$$

Choosing $v = e_u^{\text{pr}}(\mu)$ in the equation above, we have:

$$\begin{aligned}
g(\mu)\|e_u^{\text{pr}}\|_X^2 &\leq a(e_u^{\text{pr}}(\mu),e_u^{\text{pr}}(\mu);\mu) \\
&\leq |R_u^{\text{pr}}(e_u^{\text{pr}}(\mu);[u_N,p_N](\mu);\mu)| + |b(e_u^{\text{pr}}(\mu),e_p^{\text{pr}}(\mu))| \\
&\leq \sup_{v\in X}\frac{R_u^{\text{pr}}(v;[u_N,p_N](\mu);\mu)}{\|v\|_X}\|e_u^{\text{pr}}(\mu)\|_X + \kappa\sqrt{d}F_b(\mu)\|e_u^{\text{pr}}(\mu)\|_X\|e_p^{\text{pr}}(\mu)\|_M \\
&\leq \left(\|R_u^{\text{pr}}(\cdot;[u_N,p_N](\mu);\mu)\|_{X'} + \kappa\sqrt{d}F_b(\mu)\|e_p^{\text{pr}}(\mu)\|_M\right)\|e_u^{\text{pr}}(\mu)\|_X;
\end{aligned}$$

where we used (6.33), Lemma 6.4.2 and the definition of the dual residual. The desired result follows directly from this last expression, replacing $\|e_u^{\text{pr}}(\mu)\|_M$ with the results from Lemma 6.4.4. □

Using now the two previous Lemmas, we give the *a posteriori* error estimator for the output:

Proposition 10. *Defining:*

$$\begin{aligned}
\underline{\delta}^{\text{pr}}(\mu) &= \left(\|R_u^{\text{pr}}(\cdot;[\psi_M,\lambda_M](\mu);\mu)\|_{X'} \quad \|R_p^{\text{pr}}(\cdot;[\psi_M,\lambda_M](\mu);\mu)\|_{M'}\right)^T, \\
\underline{\delta}^{\text{du}}(\mu) &= \left(\|R_u^{\text{du}}(\cdot;[\psi_M,\lambda_M](\mu);\mu)\|_{X'} \quad \|R_p^{\text{du}}(\cdot;[\psi_M,\lambda_M](\mu);\mu)\|_{M'}\right)^T
\end{aligned} \tag{6.46}$$

and

$$\underline{C}(\mu) = \begin{pmatrix} 1 + \frac{\kappa^2 c(\mu)^2 d F_b(\mu)^2}{\sigma \beta_K(\mu)^2 g(\mu)^2} & \frac{\kappa \sqrt{d} F_b(\mu) c(\mu)^2}{\sigma \beta_K(\mu)^2 g(\mu)} \\ \frac{\kappa c(\mu)^2 \sqrt{d} F_b(\mu)}{\sigma \beta_K(\mu)^2 g(\mu)} & \frac{c(\mu)^2}{\sigma \beta_K(\mu)^2}. \end{pmatrix} \tag{6.47}$$

If the reduced basis approximation $\beta_K(\mu) \to \beta(\mu)$ as $K \to \infty$ then there exists a $K^*(\mu)$ such that $\forall K \geq K^*(\mu)$,

$$|s(\mu) - s_N(\mu)| \leq \underline{\delta}^{\text{du}}(\mu)^T \underline{C}(\mu) \underline{\delta}^{\text{pr}}(\mu). \tag{6.48}$$

Proof. From (6.4) and (6.27) the error in the output is given by

$$\begin{aligned} s(\mu) - s_N(\mu) &= \ell_u^O(u(\mu);\mu) + \ell_p^O(p(\mu);\mu) \\ &\quad - \ell_u^O(u(\mu);\mu) - \ell_p^O(p(\mu);\mu) + R^{\text{pr}}([\psi_M, \lambda_M](\mu); [u_N, p_N](\mu); \mu) \\ &= \ell_u^O(e_u^{\text{pr}}(\mu);\mu) + \ell_p^O(e_p^{\text{pr}}(\mu);\mu) + R^{\text{pr}}([\psi_M, \lambda_M](\mu); [u_N, p_N](\mu); \mu). \end{aligned}$$

Which from the definition of the adjoint problem (6.5) and the primal residual (6.21) can be written as

$$\begin{aligned} s(\mu) - s_N(\mu) &= -a(e_u^{\text{pr}}(\mu), \psi(\mu);\mu) - b(e_u^{\text{pr}}(\mu), \lambda(\mu);\mu) - b(\psi(\mu), e_p^{\text{pr}}(\mu);\mu) \\ &\quad + a(e_u^{\text{pr}}(\mu), \psi_M(\mu);\mu) + b(e_u^{\text{pr}}(\mu), \lambda_M(\mu);\mu) + b(\psi_M(\mu), e_u^{\text{pr}}(\mu);\mu) \\ &= -a(e_u^{\text{pr}}(\mu), e_u^{\text{du}}(\mu);\mu) - b(e_u^{\text{pr}}(\mu), e_p^{\text{du}}(\mu);\mu) - b(e_u^{\text{du}}(\mu), e_u^{\text{pr}}(\mu);\mu) \\ &= -R_u^{\text{du}}(e_u^{\text{pr}}; [\psi_M, \lambda_M](\mu);\mu) - R_p^{\text{du}}(e_p^{\text{pr}}(\mu); [\psi_M, \lambda_M](\mu);\mu); \end{aligned}$$

here the definitions for the primal and dual residuals have been used, equations (6.21) and (6.24), respectively. We then have that:

$$\begin{aligned} |s(\mu) - s_N(\mu)| &\leq \sup_{v \in X} \frac{R_u^{\text{du}}(v; [\psi_M, \lambda_M](\mu);\mu)}{\|v\|_X} \|e_u^{\text{pr}}(\mu)\|_X + \sup_{q \in M} \frac{R_p^{\text{du}}(q; [\psi_M, \lambda_M];\mu)}{\|q\|_M} \|e_p^{\text{pr}}(\mu)\|_M \\ &= \|R_u^{\text{du}}(\cdot; [\psi_M, \lambda_M](\mu);\mu)\|_{X'} \|e_u^{\text{pr}}(\mu)\|_X + \|R_p^{\text{du}}(\cdot; [\psi_M, \lambda_M];\mu)\|_{M'} \|e_p^{\text{pr}}(\mu)\|_M. \end{aligned}$$

which using Lemmas 6.4.4, 6.4.5, and the definitions in (6.46) can be written as

$$|s(\mu) - s_N(\mu)| \leq \underline{\delta}^{\text{du}}(\mu)^T \begin{pmatrix} C_u^1(\mu) & C_u^2(\mu) \\ C_p^1(\mu) & C_p^2(\mu) \end{pmatrix} \underline{\delta}^{\text{pr}}(\mu). \tag{6.49}$$

For $\sigma < 1$, we have from our hypothesis on $\beta_K(\mu)$ that $\sigma\beta_K(\mu) \leq \beta(\mu)$, for K sufficiently large, say $K \geq K^*(\mu)$. From $\sigma\beta_K(\mu) \leq \beta(\mu)$ and (6.49) the desired result, equation (6.48), directly follows. □

The bound obtained in Proposition 10 is easily computable. First, for $C(\mu)$, we need $\beta_K(\mu)$ which is obtained by solving a reduced-basis problem. Second, for $\underline{\delta}^{\text{du}}(\mu)$ and $\underline{\delta}^{\text{pr}}(\mu)$, the dual norms for the primal and dual residuals are required. Following the discussion in Section 6.3 and 5.4.2, an off-line/on-line decomposition can be developed for the efficient calculation of the relevant norms. In the following Section, we will not present numerical results for the *a posteriori* error estimation procedure developed here — these along with the development of bound conditioners, will be presented in a future paper.

6.5 Numerical Results

6.5.1 Problem Statement

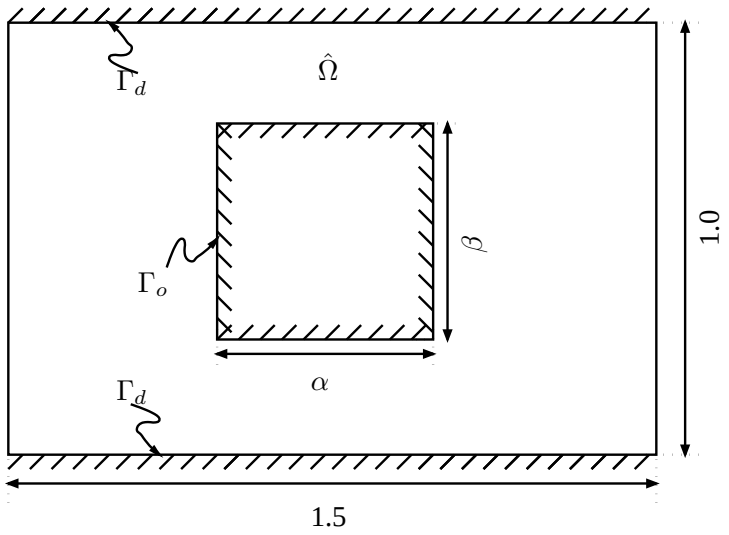

Figure 6-1: Square Obstacle

To illustrate our methods we study the incompressible flow of a highly-viscous fluid in an infinite duct with periodic square obstacles. Assuming a constant pressure gradient applied on the fluid, our interest is to study the effect that the size of the obstacle has to the flow-rate. To compute the velocity and pressure distribution the Stokes equations, (6.1), have to

be solved. For this periodically repeating configuration, we only need consider the domain around one of the obstacles, shown in Figure 6-1, and use periodic boundary conditions for the velocity in the inflow and outflow boundaries. In addition, for the duct Γ_d and obstacle solid walls Γ_o, a no-slip boundary condition is applied.

The basic (non-dimensionalized) geometric dimensions are shown in Figure 6-1. For the parameterization of the problem we are interested in two geometric parameters: the length α and height β of the obstacle — $\mu = (\alpha, \beta) \in \mathbb{R}^{P=2}$. We choose allowable ranges for these parameters $0.1 \leq \alpha \leq 1.0$, and $0.1 \leq \beta \leq 0.8$; that is $\mathcal{D} = [0.1, 1.0] \times [0.1 \times 0.8] \subset \mathbb{R}^2$. Then for a given $\mu \in \mathcal{D}$, we compute the solution $[u(\mu), p(\mu)]$ and from this we obtain the flow-rate — output of interest — $s(\mu)$, from:

$$s(\mu) = \ell^O([u(\mu), p(\mu)]; \mu) = \frac{2}{3} \int_{\hat{\Omega}} u_1 \, d\Omega; \tag{6.50}$$

recall that u_1 is the x-component of the velocity vector.

To account for the geometry variations we map the parameter-dependent domain $\hat{\Omega}$ to a fixed domain Ω, by using affine geometric transformations. Then geometry variations appear as parameter-dependent properties over the fixed domain Ω. We thus obtain an equation of the form (6.3), with $X = V^2$, where $H_0^1(\Omega) \subset V \subset H^1(\Omega)$ satisfies the aforementioned boundary conditions. Also, since there is no interaction with the environment, the pressure is defined up to an additive constant. To eliminate this uncontrollable mode we choose $M = L_0^2(\Omega)$ for the pressure. Under these assumptions a unique solution $[u(\mu), p(\mu)]$ will exist for (6.3). It is easy to verify that a decomposition of the form (6.8), with $Q_a = 6$, $Q_b = 4$ and $Q_c = Q_O = 3$, exists for the linear and bilinear forms. In practice for the solution of (6.3), X and M are replaced by X_h and M_h, suitably chosen finite-dimensional approximation spaces. Here, we use the Taylor-Hood family of elements, where:

$$X_h = \left\{ v \in X \cap C^0(\Omega) | v|_{T_h} \in \mathbb{P}^2(T_h), \ \forall T_h \in \mathcal{T}_h \right\}, \text{ and,}$$
$$M_h = \left\{ q \in M \cap C^0(\Omega) | q|_{T_h} \in \mathbb{P}^1(T_h), \ \forall T_h \in \mathcal{T}_h \right\};$$

with $\mathcal{T}_h$ a suitably fine triangulation of the domain Ω. This choice, ensures discrete stability

[36], and the discrete problems have a unique solution obtained by using the Uzawa algorithm.

6.5.2 Results

The calculations can be simplified by noticing that, for our problem $\ell^0([v,q];\mu) = c\ell([v,q];\mu)$ with c a constant — here, $c = 2/3$. Given the symmetry of the Stokes operator, and choosing $N = M$, and $\mathcal{S}_N^{\rm pr} \equiv \mathcal{S}_M^{\rm du}$ we see that the resulting primal and dual reduced-basis spaces will coincide. As a consequence the primal and dual solutions at each $\mu \in \mathcal{D}$ will be co-linear; for this specific case, denoted as "compliance," the computational procedure can be simplified. More specifically, for the particular case in consideration, the adjoint correction term of (6.27) will vanish by virtue of the Galerkin orthogonality. This suggests that we only need consider the primal problem. The disadvantage over the segregated primal-dual approach (with $\mathcal{S}_N^{\rm pr} \neq \mathcal{S}_M^{\rm du}$) is an increase — for a given accuracy — roughly by a factor of two to four of the off-line and on-line computational complexity. A discussion on how the relative choice of basis functions for the primal and dual problem affects the accuracy and computational cost, was given in the previous chapters — the same conclusions apply here.

Here we focus on a different problem, which is the relative selection of basis functions used for the approximation of the velocity and the pressure. Following Remark 6.2.2, we can choose independently $N_p^{\rm pr}$, the number of pressure basis functions, and $N_u^{\rm pr}$ the number of velocity basis functions. Recall, that to ensure stability of the reduced-basis problems, we need to augment the velocity (supremizing-) space with $N_p^{\rm pr}$ parameter-dependent basis functions. The velocity approximation-space has then total dimension $N_u^{\rm pr} + N_p^{\rm pr}$.

To form the reduced-basis space we define $N^{\rm pr} = \max\{N_u^{\rm pr}, N_p^{\rm pr}\}$, and select $N^{\rm pr}$ points $\mu_i \in \mathcal{D}$ to form the sample set $\mathcal{S}_N^{\rm pr} = \{\mu_i, i = 1, \ldots, N^{\rm pr}\}$. We then compute the solution of (6.3) $[u(\mu), p(\mu)]$ for all $\mu \in \mathcal{S}_N^{\rm pr}$ using the finite-element method; representative solutions are shown in Figure 6-3. To form the pressure reduced-basis space $M_N^{\rm pr}$ we pick the first $N_p^{\rm pr}$ pressure basis-functions. We then compute the related supremizing functions from (6.15). For the velocity reduced-basis space $X_N^{\rm pr}$, we include the first $N_u^{\rm pr}$ velocity basis functions, and in addition the parameter-dependent functions of (6.14). For the efficient calculation of the reduced-basis predictions, the off-line/on-line computational procedure presented in

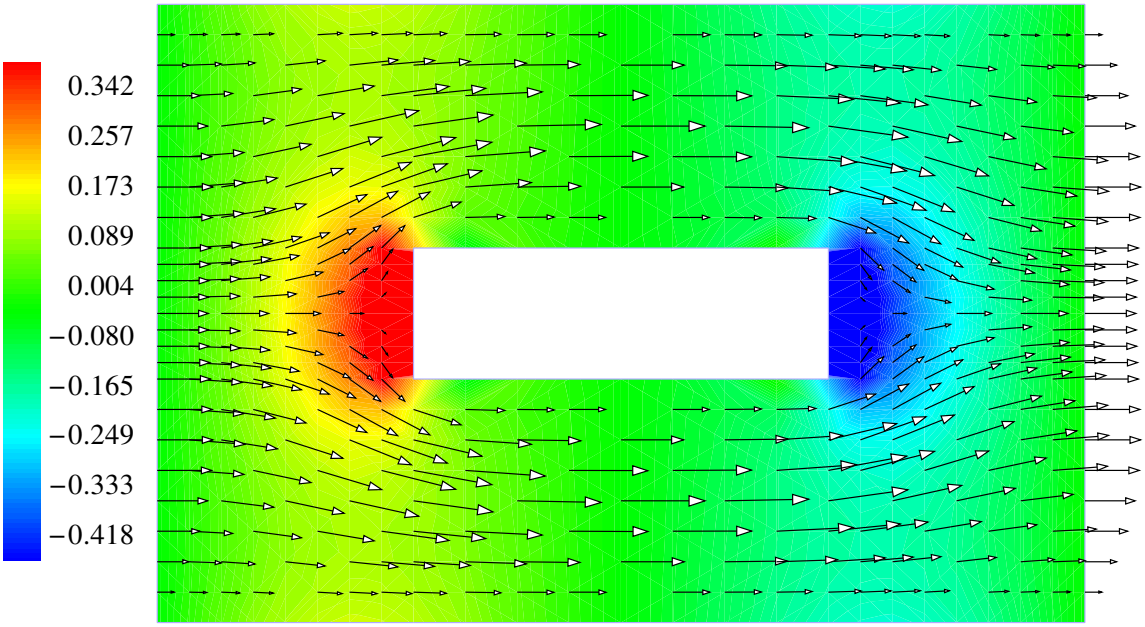

Figure 6-2: FEM Solution for $\alpha = 0.671$ and $\beta = 0.212$

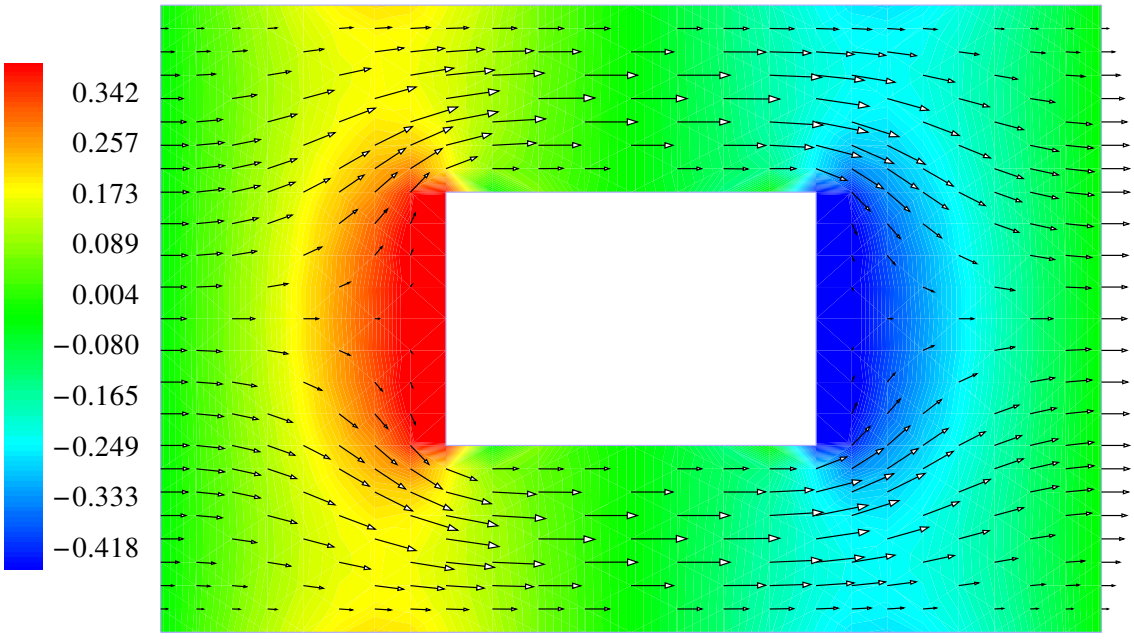

Figure 6-3: FEM Solution for $\alpha = 0.590$ and $\beta = 0.404$

Section 6.3 can be utilized.

Test 1: Output Approximation

As a first test, we investigate the accuracy of the reduced-basis predictions as a function of N_u^{pr} and N_p^{pr}. In Figures 6.5.2 and 6-4 the relative error in the output is plotted as a function of N_u^{pr}, for two test points $\mu_1 = \{0.5, 0.5\}$ and $\mu_2 = \{0.2, 0.1\}$, respectively. It should be clear that the size of N_u^{pr} and N_p^{pr} is directly related to the approximation properties of the velocity and pressure spaces.

As N_u^{pr} is increasing and for a fixed value of N_p^{pr}, say $N_p^{\text{pr}} = 10$, we see that initially the error is decreasing very rapidly, and after some point it remains constant. The *a priori* error analysis states that

$$|s(\mu) - s_N(\mu)| \leq c_1 \inf_{w_N \in X_N^{\text{pr}}} \|u(\mu) - w_N\|_X^2 + c_2 \inf_{q_N \in M_N^{\text{pr}}} \|p(\mu) - q_N\|_M^2; \tag{6.51}$$

where c_1 and c_2, depend on the continuity and stability constants of the bilinear forms a and b. This suggests that the accuracy in the output depends both in the approximation of the pressure as well as the velocity — this is confirmed by these plots.

Initially, for N_u^{pr} small, the velocity approximation error dominates over the pressure approximation error. As we increase N_u^{pr}, the velocity reduced-basis space becomes richer, and therefore the error in the velocity and consequently in the output is reduced. At some point (as determined by the size of N_p^{pr}), the velocity approximation error becomes smaller than the pressure approximation error. Therefore, further increasing N_u^{pr} no longer contributes to the accuracy of the output, as then the dominant error is now due to the inaccurate approximation of the pressure. Even though for the particular output — the flowrate — only the velocity appears explicitly in (6.50), the discussion above suggests that a balancing of the pressure and velocity errors is essential for the accurate approximation of the output. A choice like $N_u^{\text{pr}} = N_p^{\text{pr}}$, suggested in Section 6.2.1 is desirable for good convergence in the output. For this particular choice, the convergence of the relative error in the output, is shown in Figure 6-6.

Regarding the convergence rate we notice, following any of the curves in 6.5.2, that the

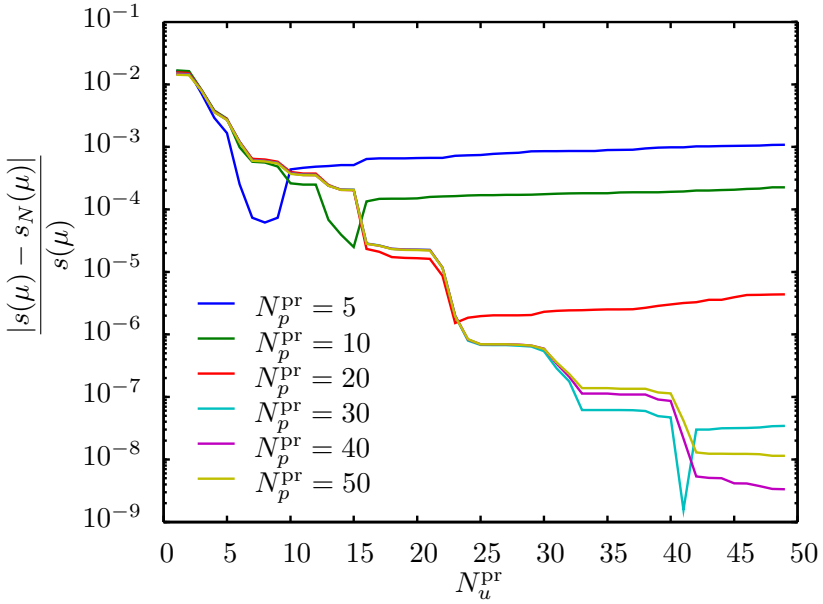

Figure 6-4: Relative error as a function of N_u^{pr}, for different N_p^{pr}, for $\mu = \{0.5, 0.5\}$.

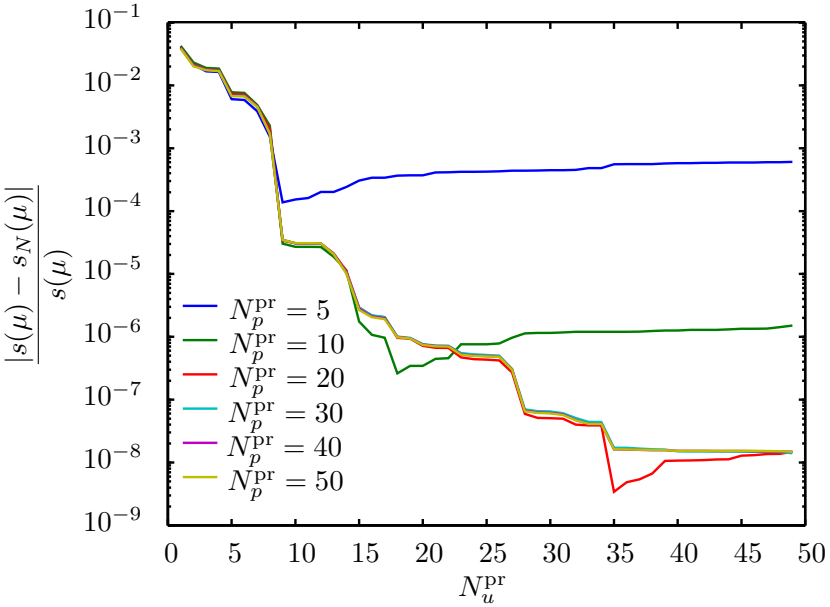

Figure 6-5: Relative error as a function of N_u^{pr}, for different N_p^{pr}, for $\mu = \{0.2, 0.1\}$.

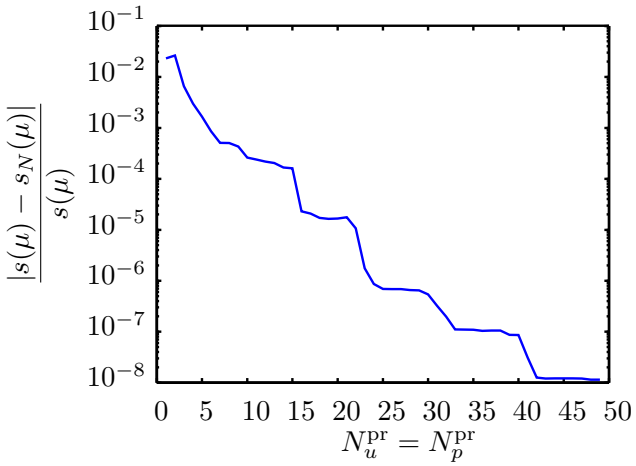

Figure 6-6: Convergence of the relative error in the output as a function of $N_u^{\rm pr} = N_p^{\rm pr}$ ($\alpha = 0.2$,$\beta = 0.1$).

error in the velocity converges to zero at an exponential rate. Indirectly, comparing the different curves in Figure 6.5.2, a similar conclusion can be reached for the error in the pressure. In both cases, for an increase of $N_u^{\rm pr}$ (or $N_p^{\rm pr}$) by ten, the approximation error goes down roughly by an order of magnitude (for the output given the quadratic convergence — expected from (6.51) — it goes down by roughly two orders of magnitude). The deterioration in this convergence rate as $N_u^{\rm pr}$ or $N_p^{\rm pr}$ become large can be attributed to ill-conditioning, as the basis functions become close to linearly dependent. Regarding the computational cost, the evaluation of $s_N(\mu)$ for $N_u^{\rm pr} = N_p^{\rm pr} = 25$ is roughly 1000 times faster, compared to the solution of the finite-element problem for the evaluation of $s(\mu)$; the output is predicted with an error less than 10^{-6} (as we can see from Figure 6-6) which is acceptable for many applications. Of course, these savings are realized only in the limit of many evaluations, after the off-line cost is offset.

Test 2: Inf-Sup Parameter Approximation

If an approximation to the inf-sup parameter is required (for example in the case of *a posteriori* error estimation frameworks), the methodology described in 6.2.2, can be used. To wit, we choose $K = 50$ and form the reduced-basis spaces $X_K^{\chi}$and M_k^{χ}. Using these reduced-basis spaces we can compute an approximation $\beta_K(\mu)$ to the exact inf-sup parameter $\beta(\mu)$. It should be mentioned that for the efficient computation of $\beta_K(\mu)$, an off-line/on-line

$\beta \setminus \alpha$	0.1000	0.3250	0.5500	0.7750	1.0000
0.100	1.13e-06	3.06e-07	1.37e-07	2.97e-07	6.58e-07
0.275	3.73e-08	2.11e-08	3.99e-07	2.20e-07	3.45e-05
0.450	6.92e-07	7.10e-07	4.33e-06	1.94e-05	5.35e-04
0.625	2.34e-07	4.65e-07	1.55e-05	2.36e-04	2.07e-03
0.800	4.18e-04	1.45e-04	4.46e-05	6.68e-04	5.59e-03

Table 6.1: Relative error $\frac{\beta_K(\mu)-\beta(\mu)}{\beta(\mu)}$ for different $\mu \in \mathcal{D}$ ($K = 50$)

procedure can be developed, under the same assumptions as for the output prediction (i.e. affine parameter dependence).

We present in Table 6.5.2, the relative error in the prediction of the inf-sup parameter $\frac{\beta_K(\mu)-\beta(\mu)}{\beta(\mu)}$, for different $\mu = \{\alpha, \beta\} \in \mathcal{D}$. Consistent with Lemma 6.2.1, we notice that $\beta_K(\mu)$ is always larger than the exact inf-sup parameter. In addition, we notice that the prediction is very accurate for all the test points. The relatively larger errors for α and β large can, at least partially, be attributed to the choice of basis functions for the construction of the reduced-basis spaces.

Chapter 7

Eigenvalue Problems

7.1 Introduction

Given two Hilbert spaces X and Y satisfying $Y \subset X$, we consider the symmetric eigenvalue problem : find $(u(\mu), \lambda(\mu)) \in Y \times \mathbb{R}$ such that

$$a(u(\mu), v; \mu) = \lambda(\mu) m(u(\mu), v), \ \forall v \in Y, \tag{7.1}$$

with the normalization condition

$$m(u(\mu), u(\mu)) = 1;$$

where $\mu \in \mathcal{D} \subset \mathbb{R}^P$ is a multi-parameter, and, for any fixed μ in $\mathcal{D}$, $a(v, w; \mu)$ and $m(v, w)$ are symmetric bilinear forms such that $a(\cdot, \cdot; \mu)$ is uniformly continuous in Y, and $m(\cdot, \cdot)$ is continuous and coercive in X. We further require the existence of a positive function $g(\mu)$ and a symmetric coercive continuous bilinear form $\hat{a}(v, w)$ such that, for a positive constant $c > 0$,

$$\underline{c}||v||_Y^2 \leq g(\mu)\hat{a}(v, v) \leq a(v, v; \mu), \ \forall v \in Y, \ \forall \mu \in \mathcal{D}. \tag{7.2}$$

We focus here on the situation, common in engineering design and optimization, in which we wish to evaluate $\lambda(\mu)$ at many points μ in the parameter space $\mathcal{D}$.

7.2 The Reduced-Basis Approximation

In what follows, $0 < \lambda^1(\mu) \leq \lambda^2(\mu) \leq \cdots$ and $u^1(\mu), u^2(\mu), \ldots$ denote respectively the eigenvalues and eigenfunctions of (7.1) at a given point $\mu \in \mathcal{D}$. We suppose that our output of interest is the first eigenvalue $\lambda^1(\mu)$; we will further assume that $\lambda^1(\mu) < \lambda^2(\mu)$. Note that $m(u^j(\mu), u^k(\mu)) = \delta_{jk}$ and hence $a(u^j(\mu), u^k(\mu); \mu) = \lambda^j(\mu)\delta_{jk}$, where δ_{jk} is the Kronecker symbol.

We start by constructing the reduced basis : we select the sample set $S_N = \{\mu_1, \ldots, \mu_{N/2}\}$ (suppose N even), and compute $u^1(\mu_i)$ and $u^2(\mu_i)$, $i = 1, \ldots, N/2$. We then define the reduced-basis space:

$$W_N = \text{span}\{\zeta_1, \ldots, \zeta_N\} = \text{span}\{u^1(\mu_1), u^2(\mu_1), \ldots, u^1(\mu_{N/2}), u^2(\mu_{N/2})\}.$$

We then consider, for any value μ of interest, the approximate solution: find $(u_N(\mu), \lambda_N(\mu)) \in W_N \times \mathbb{R}$ such that

$$\begin{aligned} a(u_N(\mu), v_N; \mu) &= \lambda_N(\mu) m(u_N(\mu), v_N), \ \forall v_N \in W_N, \text{ and} \\ m(u_N(\mu), u_N(\mu)) &= 1. \end{aligned} \tag{7.3}$$

As we recall below in Lemma 7.3.1, $\lambda_N^1(\mu)$, the first discrete eigenvalue, is larger than $\lambda^1(\mu)$; we now construct a lower bound for $\lambda^1(\mu)$. We first introduce a reconstructed error $\hat{e}(\mu)$, in Y, solution of

$$g(\mu)\hat{a}(\hat{e}(\mu), v) = 2[\lambda_N^1 m(u_N^1(\mu), v) - a(u_N^1(\mu), v; \mu)], \ \forall v \in Y. \tag{7.4}$$

For any positive γ such that $\beta(\mu) = 1 - \gamma - \frac{\lambda_N^1(\mu)}{\lambda_N^2(\mu)}$ is positive, the proposed lower bound is then

$$\lambda_N^-(\mu) = \lambda_N^1(\mu) - \frac{g(\mu)}{4\beta(\mu)} \hat{a}(\hat{e}(\mu), \hat{e}(\mu)).$$

We shall explain in Section 7.4 how to compute efficiently the solutions of (7.3) and (7.4): we exploit (i) a decomposition of the bilinear form a, (ii) linear superposition, and (iii) certain *a priori* estimates for the eigenvalue problem. In particular, the reduced-basis

and μ-independent functions are pre-computed (for a similar idea, see [84]); the complexity of the real-time reduced-basis and bound calculations is thus independent of the dimension of the underlying expensive space Y. Before discussing the computational considerations, we derive and analyze a bound error expression, and prove the asymptotic bounding properties and optimal convergence rate of the bound gap.

7.3 Bound Properties

First, we recall the classical result, where $e(\mu) = u^1(\mu) - u_N^1(\mu)$

Lemma 7.3.1. *We have*

$$a(e(\mu), e(\mu); \mu) - \lambda^1(\mu) m(e(\mu), e(\mu)) = \lambda_N^1(\mu) - \lambda^1(\mu) > 0, \tag{7.5}$$

and also, if $m(e(\mu), e(\mu))$ is small enough,

$$a(e(\mu), e(\mu); \mu) - \lambda_N^1(\mu) m(e(\mu), e(\mu)) = (1 - m(e(\mu), e(\mu)))(\lambda_N^1(\mu) - \lambda^1(\mu)) > 0. \tag{7.6}$$

Proof. For (7.5) see Lemma 9.1 and Equation (8.42) in [9]; (7.6) then immediately follows. □

Second, we prove the following error expression

Lemma 7.3.2. *The bound satisfies*

$$\begin{aligned} \lambda_N^-(\mu) = \lambda^1(\mu) - \frac{g(\mu)}{2} \hat{a} \left(\sqrt{2\beta(\mu)} e(\mu) - \frac{\hat{e}(\mu)}{\sqrt{2\beta(\mu)}}, \sqrt{2\beta(\mu)} e(\mu) - \frac{\hat{e}(\mu)}{\sqrt{2\beta(\mu)}} \right) \\ - \{ [a(e(\mu), e(\mu); \mu) - \beta(\mu) g(\mu) \hat{a}(e(\mu), e(\mu))] - \lambda_N^1(\mu) m(e(\mu), e(\mu)) \}. \end{aligned} \tag{7.7}$$

Proof. We take $v = e(\mu)$ in (7.4) and add two times $a(u^1(\mu), e(\mu); \mu) - \lambda^1(\mu) m(u^1(\mu), e(\mu)) = 0$ to the right-hand side; from $m(u^1(\mu), u^1(\mu)) = m(u_N^1(\mu), u_N^1(\mu)) = 1$ and Lemma 7.3.1 we are then able to derive that

$$g(\mu) \hat{a}(\hat{e}(\mu), e(\mu)) = a(e(\mu), e(\mu); \mu) - \lambda_N^1(\mu) m(e(\mu), e(\mu)) + \lambda_N^1(\mu) - \lambda^1(\mu). \tag{7.8}$$

We complete the proof by expanding the second term of the right-hand side of (7.7), and then evoking the definition of $\lambda_N^-(\mu)$ and the equality (7.8). □

Note that (7.6) and (7.7) states already that, for $\beta(\mu)$ small enough, $\lambda_N^-(\mu)$ is a lower bound for $\lambda^1(\mu)$. The following inequalities make that statement more precise.

Lemma 7.3.3. *We have*

$$\lambda^1(\mu)m(e(\mu),e(\mu)) \leq \frac{1}{4\lambda^1(\mu)}a(e(\mu),e(\mu);\mu)^2\left(1-\frac{\lambda^1(\mu)}{\lambda^2(\mu)}\right)+\frac{\lambda^1(\mu)}{\lambda^2(\mu)}a(e(\mu),e(\mu);\mu); \quad (7.9)$$

furthermore, if we suppose $a(e(\mu),e(\mu);\mu)$ is sufficiently small,

$$a(e(\mu),e(\mu);\mu) \leq (\lambda_N^1(\mu)-\lambda^1(\mu))\left(1-\frac{\lambda^1(\mu)}{\lambda^2(\mu)}\right)^{-1}+O[(\lambda_N^1(\mu)-\lambda^1(\mu))^2]. \quad (7.10)$$

Proof. We expand $e(\mu)=\sum_{j=1}^{\infty}\alpha^j u^j$ and write, as in the proof of Theorem 9.1 in [9],

$$\begin{aligned} a(e(\mu),e(\mu);\mu) &= \lambda^1(\alpha^1)^2+\sum_{j=2}^{\infty}\lambda^j(\alpha^j)^2 = \lambda^1(\alpha^1)^2+\sum_{j=2}^{\infty}\lambda^j(\alpha^j)^2\left(1-\frac{\lambda^1}{\lambda^j}\right)\left(1-\frac{\lambda^1}{\lambda^j}\right)^{-1} \\ &\leq \lambda^1(\alpha^1)^2+[a(e(\mu),e(\mu);\mu)-\lambda^1 m(e,e)]\left(1-\frac{\lambda^1}{\lambda^2}\right)^{-1}, \end{aligned}$$

where we have evoked our normalizations $m(u^j,u^k)=\delta_{jk}$, $a(u^j,u^k)=\lambda^j\delta_{jk}$. We now note that $\alpha^1=\frac{1}{\lambda^1}a(e,u^1;\mu)=1-\frac{1}{\lambda^1}a(u_N^1,u^1;\mu)\leq\frac{1}{2\lambda^1}a(u^1,u^1;\mu)+\frac{1}{2\lambda^1}a(u_N^1,u_N^1;\mu)-\frac{1}{\lambda^1}a(u_N^1,u^1;\mu)$, hence $\alpha^1\leq\frac{1}{2\lambda^1}a(e,e;\mu)$, where we have evoked $\lambda_N^1\geq\lambda^1$ from (7.5). This, together with the fact that $\alpha^1=m(e,u^1)=\frac{m(e,e)}{2}\geq 0$ (again from $m(u^j,u^k)=\delta_{jk}$), directly yields (7.9).

From (7.5) and (7.9) we obtain $\frac{1}{4\lambda^1}a(e,e;\mu)^2-a(e,e;\mu)+(\lambda_N^1-\lambda^1)\left(1-\frac{\lambda^1}{\lambda^2}\right)^{-1}\geq 0$; then

$$a(e,e;\mu)\leq 2\lambda^1\left\{1-\sqrt{1-\frac{(\lambda_N^1-\lambda^1)}{\lambda^1}(1-\lambda^1\lambda^2)^{-1}}\right\},$$

for $a(e,e;\mu)$ sufficiently small; (7.10) follows from expanding the square root. □

Finally, we prove

Proposition 11. *Assume our reduced-basis approximation is convergent in the sense that* $a(e(\mu), e(\mu); \mu) \to 0$, $\lambda_N^1(\mu) \to \lambda^1(\mu)$, *and* $\lambda_N^2(\mu) \to \lambda^2(\mu)$ *as* $N \to \infty$. *Then there exist an* $N^*(\mu)$ *such that, for* $N \geq N^*(\mu)$,

$$\lambda_N^-(\mu) \leq \lambda^1(\mu) - \gamma a(e(\mu), e(\mu); \mu) + O[(\lambda_N^1(\mu) - \lambda^1(\mu))^2 + (\lambda_N^2(\mu) - \lambda^2(\mu))(\lambda_N^1(\mu) - \lambda^1(\mu))], \tag{7.11}$$

and hence an $N^{**}(\mu)$ *such that* $\lambda_N^-(\mu) \leq \lambda^1(\mu)$ *for* $N \geq N^{**}(\mu)$. *Furthermore, the rate of convergence of the bound gap* $\Delta_N = \lambda_N^1(\mu) - \lambda_N^-(\mu)$ *is optimal in the sense that* $\Delta_N \leq C\|e(\mu)\|_Y^2$.

Proof. We write, for $N > N^*(\mu)$ such that (7.10) is satisfied,

$$\begin{aligned}
&[a(e, e; \mu) - \beta g(\mu)\hat{a}(e, e)] - \lambda_N^1 m(e, e) && (7.12)\\
&\quad \geq \left(1 - \beta - \frac{\lambda^1}{\lambda^2}\right) a(e, e; \mu) - \frac{1}{4\lambda^1} a(e, e; \mu)^2 \left(1 - \frac{\lambda_1}{\lambda^2}\right) - (\lambda_N^1 - \lambda^1) m(e, e) && (7.13)\\
&\quad \geq \left(1 - \beta - \frac{\lambda^1}{\lambda_N^2}\right) a(e, e; \mu) + \left(\frac{\lambda^1}{\lambda_N^2} - \frac{\lambda^1}{\lambda^2}\right) a(e, e; \mu) + O[(\lambda_N^1 - \lambda^1)^2] && (7.14)\\
&\quad \geq \gamma a(e, e; \mu) + O[(\lambda_N^1 - \lambda^1)^2 + (\lambda_N^2 - \lambda^2)(\lambda_N^1 - \lambda^1)], && (7.15)
\end{aligned}$$

where we have used (7.2) and (7.9) in the first inequality, (7.5) and (7.10) in the second inequality, and (7.10) and our choice of $\beta(\mu)$ (see Section 7.2) in the final inequality. Finally, (7.7) and (7.2) together with the previous inequality yield (7.11), and (7.5) then ensures the existence of $N^{**}(\mu)$.

To prove the optimality of the bound gap, we add $2[a(u^1, v; \mu) - \lambda^1 m(u^1, v)] = 0$ to the right-hand side of (7.4) to obtain $g(\mu)\hat{a}(\hat{e}, v) = 2[a(e, v; \mu) - \lambda_N^1 m(e, v)] + 2(\lambda_N^1 - \lambda^1) m(u^1, v)$. We then take $v = \hat{e}(\mu)$ and use (7.2) and the continuity of a and m to show $\|\hat{e}\|_Y \leq C[\|e\|_Y + (\lambda_N^1 - \lambda^1)]$. We conclude the proof by noting that $\lambda_N^1 - \lambda^1 \leq C\|e\|_Y^2$, which is a consequence of (7.5), and thus the bound gap $\Delta_N(\mu) = \frac{g(\mu)}{4\beta}\hat{a}(\hat{e}, \hat{e}) \leq C_1\|\hat{e}\|_Y^2 \leq C_2\|e\|_Y^2$. □

Note that our hypothesis that $\lambda_N^2(\mu)$ is a sufficiently good approximation of $\lambda^2(\mu)$ is realistic, since we have included the second eigenfunctions in the reduced-basis. Before illustrating our method with numerical results, we describe, under some realistic hypotheses,

ways to evaluate efficiently both the discrete solution and the bounds.

7.4 Computational approach

We henceforth assume that a can be decomposed as $a(v,w;\mu) = \sum_{q=1}^{Q-1} \sigma^q(\mu) a^q(v,w)$, where the σ^q are mappings from $\mathcal{D}$ into $\mathbb{R}$, and the a^q are bilinear forms. We presume that the dimension of Y, dim$Y$, is large, and that for any (v,w) in Y^2, $a(v,w;\mu)$, $m(v,w)$, and $\hat{a}(v,w)$ require $O(\text{dim}Y)^\alpha$ operations to evaluate, where α is a positive real number (typically 1, and at most 2).

As an example, we consider $Y = \{v \in H_0^1(]0,1[)|\ v|_{T_h} \in \mathbb{P}_1(T_h),\ \forall T_h \in \mathcal{T}_h\}$, where $\mathcal{T}_h$ is a triangulation of $]0,1[$, and $X = L^2(]0,1[)$. We take $a(v,w;\mu) = \nu_1 \int_0^\omega v_x w_x + \nu_2 \int_\omega^1 v_x w_x$ and $m(v,w) = \int_0^1 vw$, where $0 < \omega < 1$; the parameter $\mu = (\nu_1, \nu_2)$ lies in the set $D = [1,10]^2$. For this problem, $Q = 3$, $\sigma^1(\mu) = \nu_1$, $\sigma^2(\mu) = \nu_2$, $a^1(v,w) = \int_0^\omega v_x w_x$, $a^2(v,w) = \int_\omega^1 v_x w_x$, $\hat{a}(v,w) = \int_0^1 v_x w_x$, and $g(\mu) = \min(\nu_1, \nu_2)$.

The procedure has two distinct stages: the pre-processing stage and the real-time model.

Step 1 — Off-line Step

After computing the reduced basis, we compute, for $q = 1, \ldots, Q$ and $n = 1, \ldots, N$, the functions $z_n^q \in Y$, solutions of

$$\begin{aligned} &\hat{a}(z_n^q, v) = -a^q(\zeta_n, v),\ \forall v \in Y,\ 1 \le q \le Q-1,\ \text{and} \\ &\hat{a}(z_n^Q, v) = -m(\zeta_n, v); \end{aligned} \tag{7.16}$$

we then assemble the matrices $A^q \in \mathbb{R}^{N\times N}, q = 1, \ldots, Q-1, M \in \mathbb{R}^{N\times N}$, and $\Gamma \in \mathbb{R}^{N\times N\times Q\times Q}$, defined by $A^q_{mn} = a^q(\zeta_m, \zeta_n)$, $q = 1, \ldots, Q-1$, $M_{mn} = m(\zeta_m, \zeta_n)$, and $\Gamma_{mnpq} = \hat{a}(z_n^p, z_m^q)$.

Step 2 — On-line Step

Given $\mu \in \mathcal{D}$, in order to solve the discrete problem (7.3), we compute $(\eta^i(\mu), \lambda_N^i(\mu)) \in \mathbb{R}^N \times \mathbb{R}$, $i = 1, 2$, the first two eigenpairs of the problem $A_N(\mu)\eta = \lambda_N M_N \eta$, $\eta^T M \eta = 1$.

Here $A_N(\mu) = \sum_{q=1}^{Q-1} \sigma^q(\mu) A^q$, and η^T denotes the transpose of the vector $\eta \in \mathbb{R}^N$; note that $u_N^1(\mu) = \sum_{n=1}^N \eta_n^1 \zeta_n$. Then, by linear superposition, we can evaluate our lower bound as

$$\lambda_N^-(\mu) = \lambda_N^1(\mu) - \frac{1}{\beta g(\mu)} \sum_{m=1}^N \sum_{n=1}^N \sum_{p=1}^Q \sum_{q=1}^Q \eta_m^1(\mu)\eta_n^1(\mu)\sigma^p(\mu)\sigma^q(\mu)\Gamma_{mnpq}, \tag{7.17}$$

where $\sigma^Q(\mu) = -\lambda_N^1(\mu)$; note that $\hat{e}$ defined in (7.4) verifies $\hat{e} = \frac{2}{g(\mu)} \sum_{n=1}^N \sum_{q=1}^Q \sigma^q(\mu)\eta_n^1 z_n^q$.

Computational Complexity

The off-line step requires N eigensolves and the inversion of NQ symmetric positive-definite linear systems (with identical operators) in the expensive space Y; the matrices A^q, M, and Γ are constructed in less than $\mathcal{O}[(N^2Q^2)(\dim Y)^\alpha]$ operations. In contrast, the real-time model in Step 2 is inexpensive — the operation count (and storage) is independent of $\dim Y$: for each new point $\mu \in \mathcal{D}$, $\lambda_N(\mu)$ and $\lambda_N^-(\mu)$ are obtained in less than $\mathcal{O}(N^3 + N^2Q^2)$ operations; the first term accounts for the eigenvalue solve, and the second term for the assembly of $A_N(\mu)$ (in fact, N^2Q operations) and the calculation of the sum in (7.16).

7.5 Numerical Example

We consider the problem defined by $\sigma^1(\mu) = 1$, $\sigma^2(\mu) = \mu$, $a^1(v,w) = \int_{\Omega^1} \nabla v \cdot \nabla w$, $a^2(v,w) = \int_{\Omega^2} \nabla v \cdot \nabla w$, and $m(v,w) = \int_\Omega vw$, where $\Omega =]0,1[\times]0,1[$, $\Omega^2 =]0,0.5[\times]0,0.5[$, $\Omega^1 = \Omega - \overline{\Omega^2}$, and $\mu \in \mathcal{D} = [1,9]$. We take $\hat{a}(v,w) = \int_\Omega \nabla v \cdot \nabla w$, and $g(\mu) = 1$. Our Hilbert space Y is the finite element space $Y = \{v \in H^1(\Omega) \cap C^0(\Omega) |\ v|_{T_h} \in \mathbb{P}_1(T_h),\ \forall T_h \in \mathcal{T}_h, v|_{\Gamma_D} = 0\}$, where $\mathcal{T}_h$ is a fine triangulation of the domain Ω; the homogeneous Dirichlet boundary Γ_D is defined as $\overline{\Gamma}_D = \{(x,1),\ 0 \le x \le 1\} \cup \{(1,y),\ 0 \le y \le 1\}$. Our sample set is defined by $S_N = \{\mu_1, \ldots, \mu_{N/2}\} = \{1, 3, 5, \ldots, N-1\}$ for $N \le 8$ — this is certainly not optimal since our target value is $\mu = 9$ (extrapolation), however it serves well to illustrate the technique. We define $\eta_N = \Delta_N(\mu)/(\lambda_N^1(\mu) - \lambda^1(\mu))$ as the effectivity index. We observe in Table 7.1 exponential convergence of $\lambda_N^1(\mu)$ and $\lambda_N^2(\mu)$ towards $\lambda^1(\mu)$ and $\lambda^2(\mu)$, respectively, as we increase N. The effectivities show that bounds $(\lambda_N^1(\mu) \ge \lambda^1(\mu) \ge \lambda_N^-(\mu))$ are indeed

obtained for each case, $\eta_N(\mu) \geq 1$ (hence $N^{**} = 2$), and also demonstrate the efficiency and optimality of the method, as $\eta_N(\mu)$ is at most 6.3, and thus the error bars are tight.

N	$(\lambda_N^1 - \lambda^1)/\lambda^1$	$(\lambda_N^2 - \lambda^2)/\lambda^2$	$\Delta_N(\mu)/\lambda^1(\mu)$	$\eta_N(\mu)$
2	1.3×10^{-1}	2.1×10^{0}	8.3×10^{-1}	6.3
4	5.2×10^{-3}	2.1×10^{-1}	2.4×10^{-2}	4.7
6	7.3×10^{-6}	3.2×10^{-4}	3.6×10^{-5}	4.9
8	1.0×10^{-9}	1.4×10^{-8}	5.6×10^{-9}	5.2

Table 7.1: Numerical Results

Chapter 8

Concluding Discussion

8.1 Summary

The focus of this thesis has been the development of reduced-basis output bound methods for different classes of parameter-dependent partial differential equations. The essential ingredients are model-order reduction and the development of relevant *a posteriori* error estimators for outputs of interest.

The issue of model-order reduction has received considerable attention in the literature. Much of the earlier work has focused on the reduction of time-dependent non-linear systems, with the goal of minimizing computational complexity. The case of parameter-dependence has been considered in the context of reduced-basis methods but most of the earlier work has been local both in theory and practice. Our choice of global approximation spaces ensures good approximation properties for *wide* ranges of the input parameters. Thus instead of creating a reduced-order model for a particular system, we create a model valid for general parametric *families* of systems — of special interest in the contexts of design, optimization and control.

From the numerical point of view, we studied how different projection methods affect the accuracy and stability of the reduced-basis problems. For coercive-elliptic and parabolic problems, it was found that a Galerkin projection is sufficient for stability. For other problems, like non-coercive elliptic and the Stokes problem, it was found that a Galerkin projec-

tion (on spaces spanned by solution vectors for different parameter points) did not preserve stability. The remedy in the case of non-coercive problems has been the use of minimum-residual instead of Galerkin, or alternatively a Petrov-Galerkin projection method were the supremizing space was augmented by problem-specific functions which help ensure stability. Similarly, for the Stokes problem, to ensure stability we also had to augment the velocity space with pressure-dependent basis functions. Moreover, ensuring optimal convergence rates for the output prediction, required solving a dual problem associated with the output of interest. For this primal-dual procedure, we developed relevant *a priori* error bounds directly for outputs of interest.

More importantly, a critical ingredient for the successful application of these methods, is the development of *a posteriori* error estimation procedures, directly for outputs of interest. It is understood that the error incurred by the model-order reduction depends on a number of factors: the choice of reduced-basis functions, the problem in consideration, even the output of interest — to name a few. Our approach is based on evaluating appropriate dual norms of the residuals to the primal and dual problems. We prove that these estimators are bounds to the true error, and thereby uncertainty in our predictions is greatly reduced.

On a more practical side, integration of the aforementioned components, required a carefully developed computational procedure (specific to each particular class of problems). The assumption of affine parameter dependence for all the linear and bilinear forms, permitted the decoupling of the computation in two stages: an expensive preprocessing step that needs to be performed only once, and an *inexpensive* on-line step which needs to be performed for each new set of input parameters.

Finally, corroborating results were presented for each class of problems. On one hand, their purpose has been to verify the theoretical claims. On the other hand, to better understand practical implementation issues like, for example, the relative choice of dimensions for the primal and dual spaces.

8.2 Suggestions for future work

We conclude this section by giving some suggestions for future work:

- On certain cases the *a posteriori* effectivity index has been rather high, suggesting that the error estimator largely overestimates the true error. To improve the situation, more general bound conditioner procedures have to be developed, following the ideas presented in [113]. The main difficulty there is to construct these conditioners such that a specific spectral condition is satisfied. See [111] for more details.

- In this thesis, the only non-linear problems considered were eigenvalue problems. The extension to the Burgers equation and the Navier-Stokes equation should be possible. The ingredients presented for the Stokes problem, will also be required there. The extension to problems with general non-quadratic nonlinearities, at present, seems difficult.

- The issue of ill-conditioning arising when the basis functions are close to linearly dependent has not been discussed. A proposed way would be to use the proper orthogonal decomposition to compute the "most energetic" modes. Then a bound for the L^2 truncation error can be obtained in terms of the singular values of a correlation matrix. The theory presented in [93] (originally, for L^∞ error bounds) can be adapted.

- The assumption of affine dependence is critical for computational efficiency, but also rather limiting for certain problems (esp. when considering complex geometric variations). Procedures for — at least partially — relaxing this requirement should be possible to develop. See [106] for more details.

- Of importance is also the integration of these methods in optimization, inverse design or control frameworks, and their use for realistic problems.

- If a system comprises of many connected components, and for each of those component a reduced-order model exists, it is interesting to develop error estimation procedures for the whole system. A related problem is the presence of uncertainties on the input parameters. Some of the theory in [93] should be relevant.

- Finally a more theoretical issue is the convergence of the error with the number of basis functions used. In [101] a local result was established for multi-parameter problems;

in [69], the exponential convergence has been proven globally for single-parameter problems. A global theory for multi-parameter problems, does not exist, even though the numerical results presented here suggest that this conjecture might be true.

Appendix A

Parabolic Problem — Computational Procedure

A.1 Discontinuous Galerkin — Case q=0

From the definition of $\mathbb{P}_q(I_l;V)$ for $q=0$, $\mathbb{P}_0(I_l;V) = \{v : I_l \mapsto V |\ v(t) = v_s, v_s \in V\}$ — for each time interval I_l the functions will be constant. Defining $u(t;\mu) = u^l(\mu) \in V,\ t \in I_l, \forall l \in \mathcal{L}$ and $\psi^l(\mu) = \psi(t;\mu) \in V,\ t \in I_l, \forall l \in \mathcal{L}$ equations (4.24), (4.25) simplify for the case $q=0$ to: find $u^l(\mu) \in V,\ \psi^l(\mu) \in V,\ \forall l \in \mathcal{L}$, from:

$$\begin{aligned} b(u^l(\mu), v;\mu) + \Delta\tau^l a(u^l(\mu), v;\mu) &= \Delta\tau^l f(v) + b(u^{l-1}(\mu), v;\mu),\ \forall v \in V,\ \forall l \in \mathcal{L}. \\ b(\psi^l(\mu), v;\mu) + \Delta\tau^l a(v, \psi^l(\mu);\mu) &= -\Delta\tau^l \ell^O(v) + b(\psi^{l+1}(\mu), v;\mu),\ \forall v \in V,\ \forall l \in \mathcal{L}; \end{aligned} \quad \text{(A.1)}$$

with $u^0(\mu) = u_0 \in X$ and $\psi^{L+1}(\mu) = -g^O \in X$; here $b(w,v;\mu) \equiv (w,v)$ as the L^2 inner product will also be assumed to be parameter-dependent. The output is then obtained from,

$$\begin{aligned} s(\mu) &= \sum_{l\in\mathcal{L}} \Delta\tau^l \ell^O(u^l(\mu)) + b(g^O, u^L(\mu);\mu) \\ &= \sum_{l\in\mathcal{L}} \Delta\tau^l f(\psi^l(\mu)) + b(u_0, \psi^1(\mu);\mu). \end{aligned} \quad \text{(A.2)}$$

A.1.1 Reduced-basis

We first form the reduced-basis spaces $W_N^{\rm pr},\ W_M^{\rm du}$ by solution of (A.1). The reduced-basis approximations $u_N^l(\mu),\ \psi_M^l(\mu)$ to $u^l(\mu),\ \psi^l(\mu)$ can be written as

$$u_N^l(\mu) = \sum_{j=1}^N u_{N\,j}^l \zeta_j = (\underline{u}_N^l)^T \underline{\zeta},\ \text{and}\ \psi_M^l(\mu) = \sum_{j=1}^M \psi_{M\,j}^l \xi_j = (\underline{\psi}_M^l)^T \underline{\xi}; \tag{A.3}$$

with $\underline{u}_N^l(\mu) \in \mathbb{R}^N$ and $\underline{\psi}_M^l(\mu) \in \mathbb{R}^M$. Using the expressions above, the reduced-basis problem for the primal variable becomes:

$$\begin{aligned} b(u_N^l(\mu), v; \mu) + \Delta\tau^l a(u_N^l(\mu), v; \mu) &= \Delta\tau^l f(v) + b(u_N^{l-1}(\mu), v; \mu),\ \forall v \in W_N^{\rm pr},\ \forall l \in \mathcal{L} \\ \left(\underline{B}^{\rm pr}(\mu) + \Delta\tau^l \underline{A}^{\rm pr}(\mu)\right) \underline{u}_N^l(\mu) &= \Delta\tau^l \underline{f} + \underline{r}^{{\rm pr},l-1},\ \forall l \in \mathcal{L}; \end{aligned} \tag{A.4}$$

here $\underline{A}^{\rm pr}(\mu) \in \mathbb{R}^{N\times N}$ is the matrix with entries $A_{ij}^{\rm pr} = a(\zeta_j, \zeta_i; \mu),\ 1 \le i, j \le N$; $\underline{B}^{\rm pr}(\mu) \in \mathbb{R}^{N\times N}$ has entries $B_{ij}^{\rm pr} = b(\zeta_j, \zeta_i; \mu)$; $\underline{f} \in \mathbb{R}^N$ is the vector defined by $f_i = f(\zeta_i)$; and $\underline{r}^{{\rm pr},l} \in \mathbb{R}^N$ is: for $\ell = 0$, equal to $r_i^{{\rm pr},0} = b(u_0, \zeta_i; \mu)$, and for $\ell \in \mathcal{L}$, $\underline{r}^{{\rm pr},l} = \underline{B}^{\rm pr}(\mu)\underline{u}_N^l(\mu)$. The reduced-basis problem for the dual variable is:

$$\begin{aligned} b(\psi_M^l(\mu), v; \mu) + \Delta\tau^l a(v, \psi_M^l(\mu); \mu) &= -\Delta\tau^l \ell^O(v) + b(\psi_M^{l+1}(\mu), v; \mu),\ \forall v \in W_M^{\rm du},\ \forall l \in \mathcal{L}, \\ \left(\underline{B}^{\rm du}(\mu) + \Delta\tau^l \underline{A}^{\rm du}(\mu)\right) \underline{\psi}_M^l(\mu) &= -\Delta\tau^l \underline{\ell}^O + \underline{r}^{{\rm du},l+1},\ \forall l \in \mathcal{L}; \end{aligned} \tag{A.5}$$

here $\underline{A}^{\rm du}(\mu) \in \mathbb{R}^{M\times M}$ is the matrix with entries $A_{ij}^{\rm du} = a(\xi_i, \xi_j; \mu),\ 1 \le i, j \le M$; $\underline{B}^{\rm du}(\mu) \in \mathbb{R}^{M\times M}$ has entries $B_{ij}^{\rm du} = b(\xi_j, \xi_i; \mu)$; $\underline{\ell}^O \in \mathbb{R}^M$ is the vector defined by $\ell_i^O = \ell^O(\xi_i)$; and $\underline{r}^{{\rm du},l} \in \mathbb{R}^M$ is: for $\ell = L+1$, equal to $r_i^{{\rm du},L+1} = -b(g^O, \xi_i; \mu)$, and for $\ell \in \mathcal{L}$, $\underline{r}^{{\rm du},l} = \underline{B}^{\rm du}(\mu)\underline{\psi}_M^l(\mu)$. The output can be calculated from,

$$s_N(\mu) = \sum_{l\in\mathcal{L}} \Delta\tau^l \ell^O(u_N^l(\mu)) + b(g^O, u_N^L(\mu); \mu) = \sum_{l\in\mathcal{L}} \Delta\tau^l \underline{L}_N^T \underline{u}_N^l(\mu) + \underline{G}_N(\mu)^T \underline{u}_N^L(\mu); \tag{A.6}$$

with $\underline{L}_N \in \mathbb{R}^N$, with entries $L_{N\,i} = \ell^O(\zeta_i)$, for $1 \le i \le N$; also, $\underline{G}_N(\mu) \in \mathbb{R}^N$ with $G_{N\,i}(\mu) = b(g^O, \zeta_i; \mu)$.

Assuming that all the parameter-dependent operators, depend affinely on the parameter,

we can write:

$$A^{\rm pr}_{ij}(\mu) = a(\zeta_j,\zeta_i;\mu) = \sum_{q=1}^{Q_a} \sigma_a^q(\mu) a^q(\zeta_j,\zeta_i) \quad \rightarrow \quad \underline{A}^{\rm pr}(\mu) = \sum_{q=1}^{Q_a} \sigma_a^q(\mu) \underline{A}^{{\rm pr}\,q}$$

$$B^{\rm pr}_{ij}(\mu) = b(\zeta_j,\zeta_i;\mu) = \sum_{q=1}^{Q_b} \sigma_b^q(\mu) b^q(\zeta_j,\zeta_i) \quad \rightarrow \quad \underline{B}^{\rm pr}(\mu) = \sum_{q=1}^{Q_b} \sigma_b^q(\mu) \underline{B}^{{\rm pr}\,q}$$

$$r^{{\rm pr},0}_i = b(u_0,\zeta_i;\mu) = \sum_{q=1}^{Q_b} \sigma_b^q(\mu) b^q(u_0,\zeta_i) \quad \rightarrow \quad \underline{r}^{{\rm pr},0} = \sum_{q=1}^{Q_b} \sigma_b^q(\mu) \underline{r}^{{\rm pr},0\,q};$$

with $1 \le i,\ j \le N$; for the dual we can similarly form $\underline{A}^{{\rm du}\,q}$, $\underline{B}^{{\rm du}\,q}$, $\underline{r}^{{\rm du},L+1\,q}$; and for the output $G_{N\,i}(\mu) = b(g^O,\zeta_i;\mu) = \sum_{q=1}^{Q_b} \sigma_b^q(\mu) b^q(g^O,\zeta_i) \rightarrow \underline{G}_N(\mu) = \sum_{q=1}^{Q_b} \sigma_b^q(\mu) \underline{G}^q$.

The off-line/on-line decomposition should be clear. In the *off-line* stage, we first compute from (A.1) the reduced-basis vectors $\zeta_i,\ i = 1,\ldots,N$, and $\xi_i = 1,\ldots,M$. We then compute and store the μ-independent quantities $\underline{A}^{\{{\rm pr,du}\}\,q}$, $\underline{B}^{\{{\rm pr,du}\}\,q}$, $\underline{r}^{{\rm pr},0\,q}$, $\underline{r}^{{\rm du},L+1\,q}$, $\underline{f}$, $\underline{\ell}^O$, $\underline{G}^q$. The computational cost is then $(N+M)L$ V-solves, and $O((N^2+M^2)(Q_a+Q_b))$ V-inner products. The storage requirements are: $O((N^2+M^2)(Q_a+Q_b))$ for all the μ-independent quantities. In the *on-line* stage, for each new $\mu \in \mathcal{D}$, we form using the *precomputed* information, all the required vectors and matrices; this requires $O((N^2+M^2)(Q_a+Q_b))$ operations. We then solve (A.4) for $u_N(t;\mu)$, and (A.5) for $\psi_M(t;\mu)$. The systems are dense so a direct solver can be used and the cost is $O((N^3+M^3)L)$; in the special case of constant time-step $\Delta\tau$, we can factor the matrices using LU (or Cholesky) factorization and the cost reduces to $O(N^3+M^3+L(N^2+M^2))$. Finally, from (A.6) we compute the output approximation $s_N(\mu)$.

Thus as required, the incremental or marginal cost to evaluate, $s_N(\mu)$ for any given new μ — as proposed in a design, optimization, or inverse-problem context — is very small: first, because N, M are very small, typically $O(10)$; and second, because the reduced-order problems can be very rapidly assembled and inverted thanks to the off-line/on-line decomposition.

A.1.2 Output Bounds

The second step is the computation of the output bounds; following (4.29) and (4.30), we need to compute the following quantities: $I_4 = (g(\mu))^2 \int_0^T \hat{a}(\hat{e}^{\rm pr}(\mu), \hat{e}^{\rm pr}(\mu))\, dt$, $I_5 = (g(\mu))^2 \int_0^T \hat{a}(\hat{e}^{\rm pr}(\mu), \hat{e}^{\rm du}(\mu))dt$, $I_6 = \sum_{l\in\mathcal{L}} \mathcal{R}_l^{\rm pr}(\psi_M(t;\mu);\mu)$, $I_7 = (g(\mu))^2 \int_0^T \hat{a}(\hat{e}^{\rm du}(\mu), \hat{e}^{\rm du}(\mu))dt$. To efficiently calculate the bounds, we develop next a two-stage computational procedure; the essential enabler, as before, is the affine decomposition assumption.

Since $\hat{e}^{\rm pr}(t;\mu),\ \hat{e}^{\rm du}(t;\mu) \in V^q(\mathcal{I};V)$, we define $\hat{e}^{{\rm pr}\,l}(\mu) \equiv \hat{e}^{\rm pr}(t;\mu),\ t\in I_l,\ \forall l\in\mathcal{L}$ and $\hat{e}^{{\rm du}\,l}(\mu) \equiv \hat{e}^{\rm du}(t;\mu),\ t\in I_l,\ \forall l\in\mathcal{L}$. Following (4.28), the representations of the error $\hat{e}^{\rm pr}(t;\mu),\ \hat{e}^{\rm du}(t;\mu)$ can be obtained from:

$$\begin{aligned} g(\mu)\hat{a}(\hat{e}^{{\rm pr}\,l}(\mu), v) &= f(v) - \frac{1}{\Delta\tau^l} b(u_N^l(\mu) - u_N^{l-1}(\mu), v;\mu) - a(u_N^l(\mu), v;\mu), \text{ and} \\ g(\mu)\hat{a}(\hat{e}^{{\rm du}\,l}(\mu), v) &= -\ell^O(v) - \frac{1}{\Delta\tau^l} b(\psi_M^l(\mu) - \psi_M^{l+1}(\mu), v;\mu) - a(\psi_M^l(\mu), v;\mu); \end{aligned}$$

$\forall v\in V,\ \forall l\in\mathcal{L}$. Using the affine decomposition assumption, we get ($\forall v\in V,\ \forall l\in\mathcal{L}$):

$$\begin{aligned} g(\mu)\hat{a}(\hat{e}^{{\rm pr}\,l}(\mu), v) &= f(v) - \frac{1}{\Delta\tau^l}\sum_{q=1}^{Q_b}\sum_{j=1}^{N} \sigma_b^q(\mu)\left(u_{N\,j}^l(\mu) - (1-\delta_{l\,1})u_{N\,j}^{l-1}(\mu)\right) b^q(\zeta_j, v) \\ &\quad - \frac{\delta_{l\,1}}{\Delta\tau^l}\sum_{q=1}^{Q_b}\sum_{j=1}^{Q_u} \sigma_b^q(\mu)\sigma_u^j(\mu) b^q(u_0^j, v) - \sum_{q=1}^{Q_a}\sum_{j=1}^{N} \sigma_a^q(\mu) u_{N\,j}^l(\mu) a^q(\zeta_j, v); \end{aligned}$$

with δ_{ij} the Kronecker delta which $\delta_{ij}=1$ is $i=j$ and $\delta_{ij}=0$ otherwise. During the off-line stage, we compute: $\hat{z}_0^{\rm pr}\in V$ from $\hat{a}(\hat{z}_0^{\rm pr}, v) = f(v),\ \forall v\in V$; $\hat{z}_{a\,j}^{q\,{\rm pr}}\in V$ for $j=1,\ldots,N$ and $q=1,\ldots,Q_a$ from $\hat{a}(\hat{z}_{a\,j}^{q\,{\rm pr}}, v) = -a^q(\zeta_j, v),\ \forall v\in V$; $\hat{z}_{b\,j}^{q\,{\rm pr}}\in V$ for $j=1,\ldots,N$ and $q=1,\ldots,Q_b$ from $\hat{a}(\hat{z}_{b\,j}^{q\,{\rm pr}}, v) = -b^q(\zeta_j, v),\ \forall v\in V$; and $\hat{z}_{u\,j}^{q\,{\rm pr}}\in V$ for $j=1,\ldots,Q_u$ and $q=1,\ldots,Q_b$ from $\hat{a}(\hat{z}_{u\,j}^{q\,{\rm pr}}, v) = -b^q(u_0^j, v),\ \forall v\in V$. Then $\hat{e}^{{\rm pr}\,l}(\mu)$ can be computed from:

$$\begin{aligned} \hat{e}^{{\rm pr}\,l}(\mu) &= \frac{1}{g(\mu)}\left[\hat{z}_0^{\rm pr} + \frac{1}{\Delta\tau^l}\sum_{q=1}^{Q_b}\sum_{j=1}^{N} \sigma_b^q(\mu)\left(u_{N\,j}^l(\mu) - (1-\delta_{l\,1})u_{N\,j}^{l-1}(\mu)\right)\hat{z}_{b\,j}^{q\,{\rm pr}}\right. \\ &\quad \left. + \frac{\delta_{l\,1}}{\Delta\tau^l}\sum_{q=1}^{Q_b}\sum_{j=1}^{Q_u} \sigma_b^q(\mu)\sigma_u^j(\mu)\hat{z}_{u\,j}^{q\,{\rm pr}} + \sum_{q=1}^{Q_a}\sum_{j=1}^{N} \sigma_a^q(\mu) u_{N\,j}^l(\mu)\hat{z}_{a\,j}^{q\,{\rm pr}}\right], \ \forall l\in\mathcal{L}. \quad \text{(A.7)} \end{aligned}$$

Similarly for the computation of $\hat{e}^{\mathrm{du}\,l}(\mu)$ we have that ($\forall v \in V,\ \forall l \in \mathcal{L}$):

$$
\begin{aligned}
g(\mu)\hat{a}(\hat{e}^{\mathrm{du}\,l}(\mu), v) \;=\; & -\ell^O(v) - \frac{1}{\Delta\tau^l}\sum_{q=1}^{Q_b}\sum_{j=1}^{M}\sigma_b^q(\mu)\left(\psi^l_{M\,j}(\mu) - (1-\delta_{l\,L})\psi^{l+1}_{M\,j}(\mu)\right) b^q(\xi_j, v) \\
& -\frac{\delta_{l\,L}}{\Delta\tau^l}\sum_{q=1}^{Q_b}\sum_{j=1}^{Q_g}\sigma_b^q(\mu)\sigma_g^j(\mu) b^q(g_j^O, v) - \sum_{q=1}^{Q_a}\sum_{j=1}^{M}\sigma_a^q(\mu)\psi^l_{M\,j}(\mu) a^q(\xi_j, v).
\end{aligned}
$$

During the off-line stage, we compute: $\hat{z}_0^{\mathrm{du}} \in V$ from $\hat{a}(\hat{z}_0^{\mathrm{du}}, v) = -\ell^O(v),\ \forall v \in V$; $\hat{z}_{a\,j}^{q\,\mathrm{du}} \in V$ for $j = 1, \ldots, M$ and $q = 1, \ldots, Q_a$ from $\hat{a}(\hat{z}_{a\,j}^{q\,\mathrm{du}}, v) = -a^q(\xi_j, v),\ \forall v \in V$; $\hat{z}_{b\,j}^{q\,\mathrm{du}} \in V$ for $j = 1, \ldots, M$ and $q = 1, \ldots, Q_b$ from $\hat{a}(\hat{z}_{b\,j}^{q\,\mathrm{du}}, v) = -b^q(\xi_j, v),\ \forall v \in V$; and $\hat{z}_{g\,j}^{q\,\mathrm{du}} \in V$ for $j = 1, \ldots, Q_g$ and $q = 1, \ldots, Q_b$ from $\hat{a}(\hat{z}_{g\,j}^{q\,\mathrm{du}}, v) = -b^q(g_j^O, v),\ \forall v \in V$. Then $\hat{e}^{\mathrm{du}\,l}(\mu)$ can be computed from:

$$
\begin{aligned}
\hat{e}^{\mathrm{du}\,l}(\mu) \;=\; & \frac{1}{g(\mu)}\left[\hat{z}_0^{\mathrm{du}} + \frac{1}{\Delta\tau^l}\sum_{q=1}^{Q_b}\sum_{j=1}^{M}\sigma_b^q(\mu)\left(\psi^l_{M\,j}(\mu) - (1-\delta_{l\,L})\psi^{l+1}_{M\,j}(\mu)\right)\hat{z}_{b\,j}^{q\,\mathrm{du}}\right. \\
& \left. + \frac{\delta_{l\,L}}{\Delta\tau^l}\sum_{q=1}^{Q_b}\sum_{j=1}^{Q_g}\sigma_b^q(\mu)\sigma_g^j(\mu)\hat{z}_{g\,j}^{q\,\mathrm{du}} + \sum_{q=1}^{Q_a}\sum_{j=1}^{M}\sigma_a^q(\mu)\psi^l_{M\,j}(\mu)\hat{z}_{a\,j}^{q\,\mathrm{du}}\right],\ \forall l \in \mathcal{L}. \quad \text{(A.8)}
\end{aligned}
$$

In (A.7) and (A.8) the parameter dependence enters only through the coefficients $\sigma^q(\mu)$ and the primal and dual solutions to the reduced-basis problems.

In the equations above, $\hat{z}_0^{\mathrm{pr}}$, $\hat{z}_{b\,j}^{q\,\mathrm{pr}}$, $\hat{z}_{u\,j}^{q\,pr}$, $\hat{z}_{a\,j}^{q\,pr}$ and $\hat{z}_0^{\mathrm{du}}$, $\hat{z}_{b\,j}^{q\,\mathrm{du}}$, $\hat{z}_{g\,j}^{q\,du}$, $\hat{z}_{a\,j}^{q\,du}$, do not depend on the parameter μ or time. They need only be computed once, and then from (A.7) and (A.8), $\hat{e}^{\mathrm{pr}}(\mu)$ or $\hat{e}^{\mathrm{du}}(\mu)$ can be computed for different parameters μ; the parameter dependence enters only through the coefficients and the primal and dual solutions to the reduced-basis problems. We can go one step further; since we are not interested on $\hat{e}^{\mathrm{pr}}(\mu)$ or $\hat{e}^{\mathrm{du}}(\mu)$, but rather on I_4, I_4, I_6, and I_7, we can insert (A.7) and (A.8) in the definition of those quantities. The (quite long) expanded forms are shown in appendix A.2, Equations (A.9), (A.10), (A.11), and (A.12).

In the *off-line* stage, we compute the μ-independent error components $\hat{z}$; this requires $O((Q_a+Q_b)(N+M)+(Q_u+Q_g)Q_b)$ V linear systems solves. Then, using these error components, we compute and store the μ-independent quantities required in (A.9), (A.10), (A.11)

and (A.12); for example in (A.9), we compute and store $\hat{a}(\hat{z}_{b\,j}^{q\,\mathrm{pr}}, \hat{z}_{b\,j'}^{q'\,\mathrm{pr}})$, for $j, j' = 1, \ldots, N$ and $q, q' = 1, \ldots, Q_b$. If $Q = \max\{Q_a,\ Q_b,\ Q_u,\ Q_u\}$, then for the computation of those auxiliary quantities we need $O((N^2 + M^2)Q)$ V inner products. The storage requirements are then $O((N^2+M^2)Q)$ for the storage of auxiliary quantities. In the *on-line* stage, for each new parameter point μ, we compute $I_4(\mu)$, $I_5(\mu)$, $I_6(\mu)$ and $I_7(\mu)$ from (A.9), (A.10), (A.11) and (A.12). The operations required are $O((N^2 + M^2)Q)$ and *independent* of the dimension of space V.

The upper and lower bounds $s^{\pm}(\mu)$ can then be computed from:

$$\begin{aligned} s_B(\mu) &= s_N(\mu) - \frac{1}{2g(\mu)} I_5(\mu) - I_6(\mu), \\ \Delta(\mu) &= \frac{1}{2g(\mu)} \sqrt{I_4(\mu) I_7(\mu)}; \end{aligned}$$

and $s^{\pm}(\mu) = s_B(\mu) \pm \Delta(\mu)$.

A.2 Formulas

We give below explicit expressions for the calculation of the output bounds. Recall that $I_4 = (g(\mu))^2 \int_0^T \hat{a}(\hat{e}^{\rm pr}(\mu), \hat{e}^{\rm pr}(\mu))$, $I_5 = (g(\mu))^2 \int_0^T \hat{a}(\hat{e}^{\rm pr}(\mu), \hat{e}^{\rm du}(\mu))$, $I_6 = \sum_{l\in\mathcal{L}} \mathcal{R}_l^{\rm pr}(\psi_M(t;\mu);\mu)$, and $I_7 = (g(\mu))^2 \int_0^T \hat{a}(\hat{e}^{\rm du}(\mu), \hat{e}^{\rm du}(\mu))$.

$$
\begin{aligned}
I_4 = \quad & (g(\mu))^2 \int_0^T \hat{a}(\hat{e}^{pr}(t;\mu), \hat{e}^{pr}(t;\mu)) = (g(\mu))^2 \sum_{l\in\mathcal{L}} \Delta\tau^l \hat{a}(\hat{e}^{{\rm pr}\,l}(\mu), \hat{e}^{{\rm pr}\,l}(\mu)) \\
= \quad & \sum_{l\in\mathcal{L}} \sum_{q,q'=1}^{Q_b} \sum_{j,j'=1}^{N} \frac{\sigma_b^q(\mu)\sigma_b^{q'}(\mu)}{\Delta\tau^l} \left(u_{N\,j}^l(\mu) - (1-\delta_{l\,1}) u_{N\,j}^{l-1}(\mu)\right) \\
& \left(u_{N\,j'}^l(\mu) - (1-\delta_{l\,1}) u_{N\,j'}^{l-1}(\mu)\right) \hat{a}(\hat{z}_{b\,j}^{q\,{\rm pr}}, \hat{z}_{b\,j'}^{q'\,{\rm pr}}) \\
& + T\hat{a}(\hat{z}_0^{\rm pr}, \hat{z}_0^{\rm pr}) + \sum_{l\in\mathcal{L}} \sum_{q,q'=1}^{Q_b} \sum_{j,j'=1}^{Q_u} \frac{\delta_{l\,1}}{\Delta\tau^l} \sigma_b^q(\mu)\sigma_b^{q'}(\mu)\sigma_u^j(\mu)\sigma_u^{j'}(\mu) \hat{a}(\hat{z}_{u\,j}^{q\,{\rm pr}}, \hat{z}_{u\,j'}^{q'\,{\rm pr}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q,q'=1}^{Q_a} \sum_{j,j'=1}^{N} \Delta\tau^l \sigma_a^q(\mu)\sigma_a^{q'}(\mu) u_{N\,j}^l(\mu) u_{N\,j'}^l(\mu) \hat{a}(\hat{z}_{a\,j}^{q\,{\rm pr}}, \hat{z}_{a\,j'}^{q'\,{\rm pr}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_b} \sum_{j=1}^{Q_u} 2\delta_{l\,1} \sigma_b^q(\mu)\sigma_u^j(\mu) \hat{a}(\hat{z}_0^{\rm pr}, \hat{z}_{u\,j}^{q\,{\rm pr}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_b} \sum_{j=1}^{N} 2\sigma_b^q(\mu) \left(u_{N\,j}^l(\mu) - (1-\delta_{l\,1}) u_{N\,j}^{l-1}(\mu)\right) \hat{a}(\hat{z}_0^{\rm pr}, \hat{z}_{b\,j}^{q\,{\rm pr}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_a} \sum_{j=1}^{N} 2\Delta\tau^l \sigma_a^q(\mu) u_{N\,j}^l(\mu) \hat{a}(\hat{z}_0^{\rm pr}, \hat{z}_{a\,j}^{q\,{\rm pr}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q,q'=1}^{Q_b} \sum_{j=1}^{N} \sum_{j'=1}^{Q_u} \frac{2\delta_{l\,1}}{\Delta\tau^l} \sigma_b^q(\mu)\sigma_b^{q'}(\mu)\sigma_u^{j'}(\mu) \left(u_{N\,j}^l(\mu) - (1-\delta_{l\,1}) u_{N\,j}^{l-1}(\mu)\right) \hat{a}(\hat{z}_{b\,j}^{q\,{\rm pr}}, \hat{z}_{u\,j'}^{q'\,{\rm pr}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_b} \sum_{j,j'=1}^{N} \sum_{q'=1}^{Q_a} 2\sigma_a^{q'}(\mu)\sigma_b^q(\mu) \left(u_{N\,j}^l(\mu) - (1-\delta_{l\,1}) u_{N\,j}^{l-1}(\mu)\right) u_{N\,j'}^l(\mu) \hat{a}(\hat{z}_{b\,j}^{q\,{\rm pr}}, \hat{z}_{a\,j'}^{q'\,{\rm pr}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_b} \sum_{j=1}^{Q_u} \sum_{q'=1}^{Q_a} \sum_{j'=1}^{N} 2\delta_{l\,1} \sigma_b^q(\mu)\sigma_u^j(\mu)\sigma_a^{q'}(\mu) u_{N\,j'}^l(\mu) \hat{a}(\hat{z}_{u\,j}^{q\,{\rm pr}}, \hat{z}_{a\,j'}^{q'\,{\rm pr}}). \qquad \text{(A.9)}
\end{aligned}
$$

$$
\begin{aligned}
I_5= \quad & (g(\mu))^2 \int_0^T \hat{a}(\hat{e}^{pr}(t;\mu),\hat{e}^{\mathrm{du}}(t;\mu)) = (g(\mu))^2 \sum_{l\in\mathcal{L}} \Delta\tau^l \hat{a}(\hat{e}^{\mathrm{pr}\,l}(\mu),\hat{e}^{\mathrm{du}\,l}(\mu)) \\
= \quad & \sum_{l\in\mathcal{L}} \sum_{q,q'=1}^{Q_b} \sum_{j=1}^{N} \sum_{j'=1}^{M} \frac{\sigma_b^q(\mu)\sigma_b^{q'}(\mu)}{\Delta\tau^l} \left(u_{N\,j}^l(\mu) - (1-\delta_{l\,1})u_{N\,j}^{l-1}(\mu)\right) \\
& \left(\psi_{M\,j'}^l(\mu) - (1-\delta_{l\,L})\psi_{M\,j'}^{l+1}(\mu)\right) \hat{a}(\hat{z}_{b\,j}^{q\,\mathrm{pr}}, \hat{z}_{b\,j'}^{q'\,\mathrm{du}}) \\
& + T\hat{a}(\hat{z}_0^{\mathrm{pr}}, \hat{z}_0^{\mathrm{du}}) + \sum_{l\in\mathcal{L}} \sum_{q,q'=1}^{Q_b} \sum_{j=1}^{Q_u} \sum_{j'=1}^{Q_g} \frac{\delta_{l\,1}\delta_{l\,L}}{\Delta\tau^l} \sigma_b^q(\mu)\sigma_b^{q'}(\mu)\sigma_u^j(\mu)\sigma_g^{j'}(\mu)\hat{a}(\hat{z}_{u\,j}^{q\,\mathrm{pr}}, \hat{z}_{g\,j'}^{q'\,\mathrm{du}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q,q'=1}^{Q_a} \sum_{j=1}^{N} \sum_{j'=1}^{M} \Delta\tau^l \sigma_a^q(\mu)\sigma_a^{q'}(\mu) u_{N\,j}^l(\mu)\psi_{M\,j'}^l(\mu)\hat{a}(\hat{z}_{a\,j}^{q\,\mathrm{pr}}, \hat{z}_{a\,j'}^{q'\,\mathrm{du}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_b} \sum_{j=1}^{Q_u} \delta_{l\,1}\sigma_b^q(\mu)\sigma_u^j(\mu)\hat{a}(\hat{z}_0^{\mathrm{du}}, \hat{z}_{u\,j}^{q\,\mathrm{pr}}) + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_b} \sum_{j=1}^{Q_g} \delta_{l\,L}\sigma_b^q(\mu)\sigma_g^j(\mu)\hat{a}(\hat{z}_0^{\mathrm{pr}}, \hat{z}_{g\,j}^{q\,\mathrm{du}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_b} \sum_{j=1}^{N} \sigma_b^q(\mu) \left(u_{N\,j}^l(\mu) - (1-\delta_{l\,1})u_{N\,j}^{l-1}(\mu)\right) \hat{a}(\hat{z}_0^{\mathrm{du}}, \hat{z}_{b\,j}^{q\,\mathrm{pr}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_b} \sum_{j=1}^{M} \sigma_b^q(\mu) \left(\psi_{M\,j}^l(\mu) - (1-\delta_{l\,L})\psi_{M\,j}^{l+1}(\mu)\right) \hat{a}(\hat{z}_0^{\mathrm{pr}}, \hat{z}_{b\,j}^{q\,\mathrm{du}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_a} \sum_{j=1}^{N} \Delta\tau^l \sigma_a^q(\mu) u_{N\,j}^l(\mu)\hat{a}(\hat{z}_0^{\mathrm{du}}, \hat{z}_{a\,j}^{q\,\mathrm{pr}}) + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_a} \sum_{j=1}^{M} \Delta\tau^l \sigma_a^q(\mu)\psi_{M\,j}^l(\mu)\hat{a}(\hat{z}_0^{\mathrm{pr}}, \hat{z}_{a\,j}^{q\,\mathrm{du}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q,q'=1}^{Q_b} \sum_{j=1}^{N} \sum_{j'=1}^{Q_g} \frac{\delta_{l\,L}}{\Delta\tau^l} \sigma_b^q(\mu)\sigma_b^{q'}(\mu)\sigma_g^{j'}(\mu) \left(u_{N\,j}^l(\mu) - (1-\delta_{l\,1})u_{N\,j}^{l-1}(\mu)\right) \hat{a}(\hat{z}_{b\,j}^{q\,\mathrm{pr}}, \hat{z}_{g\,j'}^{q'\,\mathrm{du}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q,q'=1}^{Q_b} \sum_{j=1}^{M} \sum_{j'=1}^{Q_u} \frac{\delta_{l\,1}}{\Delta\tau^l} \sigma_b^q(\mu)\sigma_b^{q'}(\mu)\sigma_u^{j'}(\mu) \left(\psi_{M\,j}^l(\mu) - (1-\delta_{l\,L})\psi_{M\,j}^{l+1}(\mu)\right) \hat{a}(\hat{z}_{b\,j}^{q\,\mathrm{du}}, \hat{z}_{u\,j'}^{q'\,\mathrm{pr}}) \\
& + \sum_{q=1}^{Q_b} \sum_{j=1}^{N} \sum_{j'=1}^{M} \sum_{q'=1}^{Q_a} \sigma_a^{q'}(\mu)\sigma_b^q(\mu) \left(u_{N\,j}^l(\mu) - (1-\delta_{l\,1})u_{N\,j}^{l-1}(\mu)\right) \psi_{M\,j'}^l(\mu)\hat{a}(\hat{z}_{b\,j}^{q\,\mathrm{pr}}, \hat{z}_{a\,j'}^{q'\,\mathrm{du}}) \\
& + \sum_{q=1}^{Q_b} \sum_{j=1}^{M} \sum_{j'=1}^{N} \sum_{q'=1}^{Q_a} \sigma_a^{q'}(\mu)\sigma_b^q(\mu) \left(\psi_{M\,j}^l(\mu) - (1-\delta_{l\,L})\psi_{M\,j}^{l+1}(\mu)\right) u_{N\,j'}^l(\mu)\hat{a}(\hat{z}_{b\,j}^{q\,\mathrm{du}}, \hat{z}_{a\,j'}^{q'\,\mathrm{pr}}) \\
& + \sum_{q=1}^{Q_b} \sum_{j=1}^{Q_u} \sum_{q'=1}^{Q_a} \sum_{j'=1}^{M} \delta_{l\,1}\sigma_b^q(\mu)\sigma_u^j(\mu)\sigma_a^{q'}(\mu)\psi_{M\,j'}^l(\mu)\hat{a}(\hat{z}_{u\,j}^{q\,\mathrm{pr}}, \hat{z}_{a\,j'}^{q'\,\mathrm{du}}) \\
& + \sum_{q=1}^{Q_b} \sum_{j=1}^{Q_g} \sum_{q'=1}^{Q_a} \sum_{j'=1}^{N} \delta_{l\,L}\sigma_b^q(\mu)\sigma_g^j(\mu)\sigma_a^{q'}(\mu)u_{N\,j'}^l(\mu)\hat{a}(\hat{z}_{g\,j}^{q\,\mathrm{du}}, \hat{z}_{a\,j'}^{q'\,\mathrm{pr}}). \qquad \text{(A.10)}
\end{aligned}
$$

$$
\begin{aligned}
I_6 = \quad & \sum_{l\in\mathcal{L}} \mathcal{R}_l^{\mathrm{pr}}(\psi_M(t;\mu);\mu) = g(\mu)\sum_{l\in\mathcal{L}} \hat{a}(\psi_M^l(\mu);\mu) \\
= \quad & g(\mu)\sum_{l\in\mathcal{L}}\Bigg[\sum_{j=1}^{M}\psi_{M\,j}^l(\mu) f(\xi_j) \\
& -\sum_{q=1}^{Q_b}\sum_{j=1}^{N}\sum_{j'=1}^{M}\frac{\sigma_b^q(\mu)}{\Delta\tau^l}\left(u_{N\,j}^l(\mu) - (1-\delta_{l\,1})u_{N\,j}^{l-1}(\mu)\right)\psi_{M\,j'}^l(\mu) b^q(\zeta_j,\xi_{j'}) \\
& -\sum_{q=1}^{Q_b}\sum_{j=1}^{Q_u}\sum_{j'=1}^{M}\frac{\delta_{l\,1}}{\Delta\tau^l}\sigma_b^q(\mu)\sigma_u^j(\mu)\psi_{M\,j'}^l(\mu) b^q(u_0^j,\xi_{j'}) \\
& -\sum_{q=1}^{Q_a}\sum_{j=1}^{N}\sum_{j'=1}^{M}\sigma_a^q(\mu) u_{N\,j}^l(\mu)\psi_{M\,j'}^l(\mu) a^q(\zeta_j,\xi_{j'})\Bigg].
\end{aligned}
\tag{A.11}
$$

$$
\begin{aligned}
I_7 = \quad & (g(\mu))^2 \int_0^T \hat{a}(\hat{e}^{du}(t;\mu), \hat{e}^{du}(t;\mu)) = (g(\mu))^2 \sum_{l\in\mathcal{L}} \Delta\tau^l \hat{a}(\hat{e}^{\mathrm{du}\,l}(\mu), \hat{e}^{\mathrm{du}\,l}(\mu)) \\
= \quad & \sum_{l\in\mathcal{L}} \sum_{q,q'=1}^{Q_b} \sum_{j,j'=1}^{M} \frac{\sigma_b^q(\mu)\sigma_b^{q'}(\mu)}{\Delta\tau^l} \left(\psi_{M\,j}^l(\mu) - (1-\delta_{l\,L})\psi_{M\,j}^{l+1}(\mu)\right) \\
& \left(\psi_{M\,j'}^l(\mu) - (1-\delta_{l\,L})\psi_{M\,j'}^{l+1}(\mu)\right) \hat{a}(\hat{z}_{b\,j}^{q\,\mathrm{du}}, \hat{z}_{b\,j'}^{q'\,\mathrm{du}}) \\
& + T\hat{a}(\hat{z}_0^{\mathrm{du}}, \hat{z}_0^{\mathrm{du}}) + \sum_{l\in\mathcal{L}} \sum_{q,q'=1}^{Q_b} \sum_{j,j'=1}^{Q_g} \frac{\delta_{l\,L}}{\Delta\tau^l} \sigma_b^q(\mu)\sigma_b^{q'}(\mu)\sigma_g^j(\mu)\sigma_g^{j'}(\mu)\hat{a}(\hat{z}_{g\,j}^{q\,\mathrm{du}}, \hat{z}_{g\,j'}^{q'\,\mathrm{du}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q,q'=1}^{Q_a} \sum_{j,j'=1}^{M} \Delta\tau^l \sigma_a^q(\mu)\sigma_a^{q'}(\mu)\psi_{M\,j}^l(\mu)\psi_{M\,j'}^l(\mu)\hat{a}(\hat{z}_{a\,j}^{q\,\mathrm{du}}, \hat{z}_{a\,j'}^{q'\,\mathrm{du}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_b} \sum_{j=1}^{Q_g} 2\delta_{l\,L}\sigma_b^q(\mu)\sigma_g^j(\mu)\hat{a}(\hat{z}_0^{\mathrm{du}}, \hat{z}_{g\,j}^{q\,\mathrm{du}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_b} \sum_{j=1}^{M} 2\sigma_b^q(\mu) \left(\psi_{M\,j}^l(\mu) - (1-\delta_{l\,L})\psi_{M\,j}^{l+1}(\mu)\right) \hat{a}(\hat{z}_0^{\mathrm{du}}, \hat{z}_{b\,j}^{q\,\mathrm{du}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_a} \sum_{j=1}^{M} 2\Delta\tau^l \sigma_a^q(\mu)\psi_{M\,j}^l(\mu)\hat{a}(\hat{z}_0^{\mathrm{du}}, \hat{z}_{a\,j}^{q\,\mathrm{du}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q,q'=1}^{Q_b} \sum_{j=1}^{M} \sum_{j'=1}^{Q_g} \frac{2\delta_{l\,L}}{\Delta\tau^l} \sigma_b^q(\mu)\sigma_b^{q'}(\mu)\sigma_g^{j'}(\mu) \left(\psi_{M\,j}^l(\mu) - (1-\delta_{l\,L})\psi_{M\,j}^{l+1}(\mu)\right) \hat{a}(\hat{z}_{b\,j}^{q\,\mathrm{du}}, \hat{z}_{g\,j'}^{q'\,\mathrm{du}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_b} \sum_{j,j'=1}^{M} \sum_{q'=1}^{Q_a} 2\sigma_a^{q'}(\mu)\sigma_b^q(\mu) \left(\psi_{M\,j}^l(\mu) - (1-\delta_{l\,L})\psi_{M\,j}^{l+1}(\mu)\right) \psi_{M\,j'}^l(\mu)\hat{a}(\hat{z}_{b\,j}^{q\,\mathrm{du}}, \hat{z}_{a\,j'}^{q'\,\mathrm{du}}) \\
& + \sum_{l\in\mathcal{L}} \sum_{q=1}^{Q_b} \sum_{j=1}^{Q_g} \sum_{q'=1}^{Q_a} \sum_{j'=1}^{M} 2\delta_{l\,L}\sigma_b^q(\mu)\sigma_g^j(\mu)\sigma_a^{q'}(\mu)\psi_{M\,j'}^l(\mu)\hat{a}(\hat{z}_{g\,j}^{q\,\mathrm{du}}, \hat{z}_{a\,j'}^{q'\,\mathrm{du}}). \qquad \text{(A.12)}
\end{aligned}
$$

Bibliography

[1] R. A. Adams. *Sobolev Spaces*. Academic Press, 1975.

[2] M. Ainsworth and J. T. Oden. A posteriori error estimation in finite element analysis. *Comp. Meth. Appl. Mech. Engrg.*, 142:1–88, 1997.

[3] M. Ainsworth and J. T. Oden. *A Posteriori Error Estimation in Finite Element Analysis*. Wiley-Interscience, 2000.

[4] M. A. Akgun, J. H. Garcelon, and R. T. Haftka. Fast exact linear and non-linear structural reanalysis and the Sherman-Morrison-Woodbury formulas. *International Journal for Numerical Methods in Engineering*, 50(7):1587–1606, March 2001.

[5] E. Allgower and K. Georg. Simplicial and continuation methods for approximating fixed-points and solutions to systems of equations. *SIAM Review*, 22(1):28–85, 1980.

[6] 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.

[7] A.C Antoulas and D.C. Sorensen. Approximation of large-scale dynamical systems: An overview. Technical report, Rice University, 2001.

[8] J. A. Atwell and B. B. King. Proper orthogonal decomposition for reduced basis feedback controllers for parabolic equations. *Mathematical and Computer Modelling*, 33(1-3):1–19, Jan-Feb 2001.

[9] I. Babuska and J. Osborn. Eigenvalue problems. In *Handbook of numerical analysis*, volume II, pages 641–787. Elsevier, 1991.

[10] I. Babuska and W.C. Rheinboldt. A posteriori error estimates for the finite-element method. *Int. J. Num. Meth. Engrg.*, 18:736–754, 1978.

[11] Gareth A. Baker, James H. Bramble, and Vidar Thomee. Single step galerkin approximations for parabolic problems. *Mathematics of Computation*, 31(140):818–847, October 1977.

[12] E. Balmes. Parametric families of reduced finite element models. theory and applications. *Mechanical Systems and Signal Processing*, 10(4):381–394, 1996.

[13] R. E. Bank and A. Weiser. Some a posteriori error estimators for elliptic partial differential equations. *Math. Comput.*, 44(170):283–301, 1985.

[14] A. Barrett and G. Reddien. On the reduced basis method. *Z. Angew. Math. Mech.*, 75(7):543–549, 1995.

[15] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible navier-stokes equations. *Journal of Computational Physics*, 131:267–279, 1997.

[16] F. Brezzi, G. Manzini, D. Marini, P. Pietra, and A. Russo. Discontinuous galerkin approximations for elliptic problems. *Numerical Methods for Partial Differential Equations*, 16(4):365–378, Jul 2000.

[17] Paul Castillo, Bernardo Cockburn, Ilara Perugia, and Dominik Schötzau. An a priori error analysis of the local discontinuous galerkin method for elliptic problems. *SIAM Journal of Numerical Analysis*, 38(5):1676–1706, 2000.

[18] T. F. Chan and W. L. Wan. Analysis of projection methods for solving linear systems with multiple right-hand sides. *SIAM Journal on Scientific Computing*, 18(6):1698, 1721 1997.

[19] E.A. Christensen, M. Brøns, and J.N. Sørensen. Evaluation of proper orthogonal decomposition-based decomposition techniques applied to parameter-dependent nonturbulent flows. *SIAM J. Scientific Computing*, 21(4):1419–1434, 2000.

[20] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. Classics in Applied Mathematics, 40. SIAM, 2002.

[21] A.E. Deane, I.G. Kevrekidis, G.E. Karniadakis, and S.A. Orszag. Low-dimensional models for complex geometry flows: Application to grooved channels and circular cylinders. *Phys. Fluids. A*, 3(10):2337–2354, October 1991.

[22] Earl H. Dowell and Kenneth C. Hall. Modelling of fluid structure interaction. *Annu. Rev. Fluid. Mech.*, 33:445–490, 2001.

[23] L. Machiels D.V. Rovas and Y. Maday. Reduced basis output bound methods for parabolic problems. *Computer Methods in Applied Mechanics and Engineering*, 2001. Submitted.

[24] Keneth Eriksson, Claes Johnson, and Stig Larsson. Adaptive finite element methods for parabolic problems iv: Analytic semigroups. *SIAM Journal of Numerical Analysis*, 35(4):1315–1325, August 1998.

[25] Kenneth Eriksson and Claes Johnson. Error estimates and automatic time step control for nonlinear parabolic problems. *SIAM Journal of Numerical Analysis*, 24(1):12–23, February 1987.

[26] Kenneth Eriksson and Claes Johnson. Adaptive finite element methods for parabolic problems i: A linear model problem. *SIAM Journal of Numerical Analysis*, 28(1):43–77, February 1991.

[27] Kenneth Eriksson and Claes Johnson. Adaptive finite element methods for parabolic problems ii: Optimal error estimates in $l_\infty l_2$ and $l_\infty l_\infty$. *SIAM Journal of Numerical Analysis*, 32(3):706–740, June 1995.

[28] Kenneth Eriksson and Claes Johnson. Adaptive finite element methods for parabolic problems iv: Nonlinear problems. *SIAM Journal of Numerical Analysis*, 32(6):1729–1749, December 1995.

[29] Kenneth Eriksson and Claes Johnson. Adaptive finite element methods for parabolic problems v: Long-time integration. *SIAM Journal of Numerical Analysis*, 32(6):1750–1763, December 1995.

[30] C. Farhat, L. Crivelli, and F.X. Roux. Extending substructure based iterative solvers to multiple load and repeated analyses. *Computer Methods in Applied Mechanics and Engineering*, 117(1-2):195–209, July 1994.

[31] C. Farhat and F. X. Roux. Implicit parallel processing in structural mechanics. Technical Report CU-CSSC-93-26, Center for Aerospace Structures, University of Colorado, Boulder, CO, 1993.

[32] P. Feldmann and R.W. Freund. Efficient linear circuit analysis by pade-approximation via the lanczos process. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 14(5):639–649, May 1995.

[33] 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.

[34] J. P. Fink and W. C. Rheinboldt. Local error estimates for parametrized non-linear equations. *SIAM J. Numerical Analysis*, 22:729–735, 1985.

[35] M. Fortin and F. Brezzi. *Mixed and Hybrid Finite Element Methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer Verlag, July 1991.

[36] V. Girault and P.A. Raviart. *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, volume 5 of *Springer Series in Computational Mathematics*. Springer Verlag, August 1986.

[37] M. D. Gunzburger. *Finite element methods for viscous incompressible flows*. Academic Press, 1989.

[38] D.J. Higham and N.J. Higham. Componentwise perturbation-theory for linear-systems with multiple right-hand sides. *Linear Algebra and its Applications*, 174:111–129, 1992.

[39] L. Hörmander. *Linear Partial Differential Operators*, volume 1. Springer-Verlag, 1964.

[40] 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.

[41] K. Ito and J.D. Schroeter. Reduced order feedback synthesis for viscous incompressible flows. *Mathematical And Computer Modelling*, 33(1-3):173–192, Jan-Feb 2001.

[42] Pierre Jamet. Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain. *SIAM Journal on Numerical Analysis*, 15(5):912–928, Oct. 1978.

[43] M.L. Joyner, H.T. Banks, B. Wincheski, and W.P. Winfree. Nondestructive evaluation using a reduced order computational methodology. ICASE Report 2000-10, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, March 2000. NASA//CR-2000-209870.

[44] Pearson K. On lines and planes of closest fit to points in space. *Philosophical Magazine*, 2:609–629, 1901.

[45] Matt Kamon, Frank Wang, and Jacob White. Generating nearly optimally compact models from krylov-subspace based reduced-order models. *IEEE Transactions on Circuits and Systems — II: Analog and Digital Processing*, 47(4):239–248, April 2000.

[46] C. Kanzow. Some noninterior continuation methods for linear complementarity problems. *SIAM Journal on Matrix Analysis and Applications*, 17(4):851–868, 1996.

[47] Ohannes Karakashian and Charalambos Makridakis. A space-time finite element method for the nonlinear schrödinger equation: The discontinuous galerkin method. *Mathematics of Computation*, 67(222):479–499, April 1998.

[48] K. Karhunen. Zur spektraltheorie stochastischer prozesse. *Annales Academiae Scientiarum Fennicae*, 37, 1946.

[49] K. Kline. Dynamic analysis using a reduced-basis of exact modes and ritz vectors. *AIAA Journal*, 24(12):2022–2029, 1986.

[50] M. Kojima, N. Megiddo, and T. Noma. Homotopy continuation methods for nonlinear complementarity problems. *Mathematics of Operations Research*, 16(4):754, 774 1991.

[51] P. Krysl, S. Lall, and J.E. Marsden. Dimensional model reduction in non-linear finite element dynamics of solids and structures. *International Journal for Numerical Methods in Engineering*, 51:479–504, 2001.

[52] P. Ladeveze and D. Leguillon. Error estimation procedures in the finite element method and applications. *SIAM J. Numer. Anal.*, 20:485–509, 1983.

[53] S.-H. Lai and B.C. Vermuri. Generalized capacitance matrix theorems and algorithm for solving linear systems. *SIAM Journal on Scientific Computing*, 19(3):1024–1045, May 1998.

[54] Sanjay Lall, Petr Krysl, and Jerrold E. Marsden. Structure-preserving model reduction of mechanical systems. *preprint*, 2001.

[55] Sanjay Lall, Jerrold E. Marsden, and Sonja Glavaski. A subspace approach to balanced trucation for model reduction of nonlinear control systems. *International Journal on Robust and Nonlinear Control*, 2001. to appear.

[56] M. Y. Lin Lee. Estimation of the error in the reduced-basis method solution of differential algebraic equations. *SIAM Journal of Numerical Analysis*, 28:512–528, 1991.

[57] J.L. Lions and E. Magenes. *Non-Homogenous Boundary Value Problems and Applications*. Springer-Verlag, 1972.

[58] J. Liu and W.C. Rheinboldt. A posteriori finite element error estimators for parametrized nonlinear boundary value problems. *Numerical Functional Analysis and Optimization*, 17(6), 1996.

[59] MM. Loeve. *Probablity Theory*. Van Nostrand, 1955.

[60] John Lumley and Peter Blossey. Control of turbulence. *Annu. Rev. Fluid. Mech.*, 30:311–327, 1998.

[61] 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.

[62] L. Machiels, Y. Maday, and A. T. Patera. A "flux-free" nodal Neumann subproblem approach to output bounds for partial differential equations. *C. R. Acad. Sci. Paris, Série I*, 330(3):249–254, February 2000.

[63] L. Machiels, Y. Maday, and A. T. Patera. Output bounds for reduced-order approximations of elliptic partial differential equations. *Comp. Meth. Appl. Mech. Engrg.*, 190(26-27):3413–3426, 2001.

[64] L. Machiels, A. T. Patera, J. Peraire, and Y. Maday. A general framework for finite element a posteriori error control: Application to linear and nonlinear convection-dominated problems. In *ICFD Conference on numerical methods for fluid dynamics*, Oxford, England, 1998.

[65] L. Machiels, J. Peraire, and A. T. Patera. A posteriori finite element output bounds for the incompressible Navier-Stokes equations; Application to a natural convection problem. *Journal of Computational Physics*, 172:401–425, 2001.

[66] Y. Maday, L. Machiels, A. T. Patera, and D. V. Rovas. Blackbox reduced-basis output bound methods for shape optimization. In *Proceedings* 12^{th} *International Domain Decomposition Conference*, pages 429–436, Chiba, Japan, 2000.

[67] Y. Maday, A. T. Patera, and J. Peraire. A general formulation for a posteriori bounds for output functionals of partial differential equations; Application to the eigenvalue problem. *C. R. Acad. Sci. Paris, Série I*, 328:823–828, 1999.

[68] Y. Maday, A.T. Patera, and D.V. Rovas. A blackbox reduced-basis output bound method for noncoercive linear problems. *Nonlinear Partial Differential Equations and Their Applications*, 2001. Proceedings of the College De France Seminars.

[69] Y. Maday, A.T. Patera, and G. Turinici. A priori convergence theory for reduced–basis approximations of single–parameter elliptic partial differential equations. *Journal of Scientific Computing*, 17(1–4):437–446, December 2002.

[70] Y. Maday and E. M. Rønquist. A geometrical reduced-basis element method. *C. R. Acad Sci. Paris, Série I*, 2002.

[71] C. G. Markidakis and I. Babuska. On the stability of the discontinuous galerkin method for the heat equation. *SIAM Journal of Numerical Analysis*, 34(1):389–401, 1997.

[72] B.C. Moore. Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, 26(1):17–32, 1981.

[73] Tamal Mukherjee, Gary K. Fedder, D. Ramaswamy, and J. White. Emerging simulation approaches for micromachined devices. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 19(12):1572–1589, December 2000.

[74] D. A. Nagy. Modal representation of geometrically nonlinear behaviour by the finite element method. *Computers and Structures*, 10:683–688, 1979.

[75] A. K. Noor. Recent advances in reduction methods for nonlinear problems. *Comput. Struct.*, 13:31–44, 1981.

[76] A. K. Noor. On making large nonlinear problems small. *Comp. Meth. Appl. Mech. Engrg.*, 34:955–985, 1982.

[77] A. K. Noor, C. M. Andresen, and J. A. Tanner. Exploiting symmetries in the modeling and analysis of tires. *Comp. Meth. Appl. Mech. Engrg.*, 63:37–81, 1987.

[78] A. K. Noor, C. D. Balch, and M. A. Shibut. Reduction methods for non-linear steady-state thermal analysis. *Int. J. Num. Meth. Engrg.*, 20:1323–1348, 1984.

[79] A. K. Noor and J. M. Peters. Reduced basis technique for nonlinear analysis of structures. *AIAA Journal*, 18(4):455–462, April 1980.

[80] A. K. Noor and J. M. Peters. Multiple-parameter reduced basis technique for bifurcation and post-buckling analysis of composite plates. *Int. J. Num. Meth. Engrg.*, 19:1783–1803, 1983.

[81] A. K. Noor and J. M. Peters. Recent advances in reduction methods for instability analysis of structures. *Comput. Struct.*, 16:67–80, 1983.

[82] A. K. Noor, J. M. Peters, and C. M. Andersen. Mixed models and reduction techniques for large-rotation nonlinear problems. *Comp. Meth. Appl. Mech. Engrg.*, 44:67–89, 1984.

[83] M. Paraschivoiu and A. T. Patera. A hierarchical duality approach to bounds for the outputs of partial differential equations. *Comp. Meth. Appl. Mech. Engrg.*, 158(3-4):389–407, June 1998.

[84] M. Paraschivoiu, J. Peraire, Y. Maday, and A. T. Patera. Fast bounds for outputs of partial differential equations. In J. Borgaard, J. Burns, E. Cliff, and S. Schreck, editors, *Computational methods for optimal design and control*, pages 323–360. Birkhäuser, 1998.

[85] A. T. Patera and E. M. Rønquist. A general output bound result: Application to discretization and iteration error estimation and control. *Math. Models Methods Appl. Sci.*, 11(4):685–712, 2001.

[86] A. T. Patera, D. Rovas, and L. Machiels. Reduced–basis output–bound methods for elliptic partial differential equations. *SIAG/OPT Views-and-News*, 11(1), April 2000.

[87] A.T. Patera and J. Peraire. *Error Estimation and Solution Adaptive Procedures in CFD*, chapter A general Langragian formulation for the computation of a posteriori finite element bounds. Springer-Verlag, 2002.

[88] A.T. Patera and E. M. Rønquist. A general output bound result: Application to discretization and iteration error estimation and control. *Mathematical Models and Methods in Applied Science*, 2000. MIT FML Report 98-12-1.

[89] J. S. Peterson. The reduced basis method for incompressible viscous flow calculations. *SIAM J. Sci. Stat. Comput.*, 10(4):777–786, July 1989.

[90] N.A. Pierce and M. B. Giles. Adjoint recovery of superconvergent functionals from pde approximations. *SIAM Review*, 42(2):247–264, 2000.

[91] 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.

[92] T. A. Porsching and M. Y. Lin Lee. The reduced-basis method for initial value problems. *SIAM Journal of Numerical Analysis*, 24:1277–1287, 1987.

[93] C. Prud'homme and A.T. Patera. Reduced-basis output bounds for approximately parametrized elliptic coercive partial differential equations. *Computing and Visualization in Science*, 2002. Submitted.

[94] 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 bounds methods. *Journal of Fluids Engineering*, 124(1):70–80, March 2002.

[95] C. Prud'homme, D. Rovas, K. Veroy, and A.T. Patera. Mathematical and computational framework for reliable real–time solution of parametrized partial differential equations. *M2AN*, 2002. Submitted.

[96] C. Prud'homme, D.V. Rovas, K. Veroy, L. Machiels, Y. Maday, A.T. Patera, and G. Turnici. Reduced–basis output bound methods for parametrized partial differential equations. In *Proceedings SMA Symposium*, January 2002.

[97] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*. Springer, 2nd edition, 1997.

[98] S. S. Ravindaran. A reduced-order approach for optimal control of fluids using proper orthogonal decomposition. *Int. J. Numer. Meth. Fluids*, 34:425–448, 2000.

[99] W.C. Rheinboldt. Solution fields of non-linear equations and continuation methods. *SIAM Journal on Numerical Analysis*, 17(2):221–237, 1980.

[100] W.C. Rheinboldt. Numerical analysis of continuation methods for nonlinear structural problems. *Computers and Structures*, 13(1-3):103–113, 1981.

[101] 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.

[102] Clarence W. Rowley and Jerrold E. Marsden. Reconstruction equations and the karhunen-loeve expansion for systems with symmetry. *Physica D*, 142(1–2):1–19, Aug 2000.

[103] Dominik Schötzau and Christoph Schwab. Time discretization of parabolic problems by the hp-version of the discontinuous galerkin finite element method. *SIAM Journal of Numerical Analysis*, 38(3):837–875, 2000.

[104] L.Miguel Silveira, Mattan Kamon, Ibrahim Elfadel, and Jacob White. A coordinate-transformed arnoldi algorithm for generating guaranteed stable reduced-order models of rlc circuits. *Comput. Methods Appl. Mech. Engrg.*, 169:377–389, 1999.

[105] 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.

[106] Y. Solodukhov. *(In progress).* PhD thesis, Massachusetts Institute of Technology, 2004.

[107] G. Strang and G. Fix. *An Analysis of the Finite Element Method.* Wellesley Cambridge Pr, 1973.

[108] M. Suarjana. Successive conjugate-gradient methods for structural-analysis with multiple load cases. *Int. J. Numer. Methods Eng.*, 37:4185–4203, 1994.

[109] Vidar Thomee. *Galerkin Finite Element Methods for Parabolic Problems*, chapter The Discontinuous Galerkin Time Stepping Method, pages 181–208. Springer Series in Computational Mathematics. Springer, June 1997.

[110] Vidar Thomee. From finite differences to finite elements. a short history of numerical analysis of partial differential equations. *Journal of Computational and Applied Mathematics*, 128:1–54, 2001.

[111] K. Veroy. *Reduced Basis Methods Applied to Problems in Elasticity: Analysis and Applications*. PhD thesis, Massachusetts Institute of Technology, 2003. In progress.

[112] K. Veroy, T. Leurent, C. Prud'homme, D. Rovas, and A.T. Patera. Reliable real–time solution of parametrized elliptic partial differential equations: Application to elasticity. In *Proceedings SMA Symposium*, January 2002.

[113] 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 JL Lions.

[114] Mary Fanett Wheeler. A priori l_2 error estimates for galerkin approximations to parabolic partial differential equations. *SIAM Journal on Numerical Analysis*, 10(4):723–759, Sep. 1973.

[115] K. Willcox and J. Peraire. Balanced model reduction via the proper orthogonal decomposition. In *15th AIAA Computational Fluid Dynamics Conference*. AIAA, June 2001.

[116] E. L. Yip. A note on the stability of solving a rank-p modification of a linear system by the Sherman-Morrison-Woodbury formula. *SIAM Journal on Scientific and Statistical Computating*, 7(2):507–513, April 1986.