SPRINGER BRIEFS IN MATHEMATICS

Jan S. Hesthaven
Gianluigi Rozza
Benjamin Stamm

Certified Reduced Basis Methods for Parametrized Partial Differential Equations

bcam basque center for applied mathematics

Springer

SpringerBriefs in Mathematics

SpringerBriefs in Mathematics showcases expositions in all areas of mathematics and applied mathematics. Manuscripts presenting new results or a single new result in a classical field, new field, or an emerging topic, applications, or bridges between new results and already published works, are encouraged. The series is intended for mathematicians and applied mathematicians.

BCAM SpringerBriefs

BCAM *SpringerBriefs* aims to publish contributions in the following disciplines: Applied Mathematics, Finance, Statistics and Computer Science. BCAM has appointed an Editorial Board, who evaluate and review proposals.

Typical topics include: a timely report of state-of-the-art analytical techniques, bridge between new research results published in journal articles and a contextual literature review, a snapshot of a hot or emerging topic, a presentation of core concepts that students must understand in order to make independent contributions.

Please submit your proposal to the Editorial Board or to Francesca Bonadei, Executive Editor Mathematics, Statistics, and Engineering: francesca.bonadei@springer.com

More information about this series at http://www.springer.com/series/10030

Jan S. Hesthaven · Gianluigi Rozza
Benjamin Stamm

Certified Reduced Basis Methods for Parametrized Partial Differential Equations

basque center for applied mathematics

Springer

Jan S. Hesthaven
EPFL-SB-MATHICSE-MCSS
Ecole Polytechnique Fédérale de Lausanne
EPFL
Lausanne
Switzerland

Benjamin Stamm
Laboratoire Jacques-Louis Lions
Sorbonne Universités, UPMC Univ Paris 06,
CNRS
Paris Cedex 05
France

Gianluigi Rozza
SISSA MathLab
International School for Advanced Studies
Trieste
Italy

ISSN 2191-8198 ISSN 2191-8201 (electronic)
SpringerBriefs in Mathematics
ISBN 978-3-319-22469-5 ISBN 978-3-319-22470-1 (eBook)
DOI 10.1007/978-3-319-22470-1

Library of Congress Control Number: 2015946080

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

To our families and friends

Preface

During the past decade, reduced order modeling has attracted growing interest in computational science and engineering. It now plays an important role in delivering high-performance computing (and bridging applications) across industrial domains, from mechanical to electronic engineering, and in the basic and applied sciences, including neurosciences, medicine, biology, chemistry, etc. Such methods are also becoming increasingly important in emerging application domains dominated by multi-physics, multi-scale problems as well as uncertainty quantification.

This book seeks to introduce graduate students, professional scientists, and engineers to a particular branch in the development of reduced order modeling, characterized by the provision of reduced models of guaranteed fidelity. This is a fundamental development that enables the user to trust the output of the model and balance the needs for computational efficiency and model fidelity. The text develops these ideas by presenting the fundamentals with a gradually increasing complexity; comparisons are made with more traditional techniques and the performance illustrated by means of a few carefully chosen examples. The book does not seek to replace review articles on the topics (such as [1–5]) but aims to widen the perspectives on reduced basis methods and to provide an integrated presentation. The text begins with a basic setting to introduce the general elements of certified reduced basis methods for elliptic affine coercive problems with linear compliant outputs and then gradually widens the field, with extensions to non-affine, non-compliant, non-coercive operators, geometrical parametrization and time-dependent problems.

We would like to point out some original ingredients of the text. Chapter 3 guides the reader through different sampling strategies, providing a comparison between classic techniques based on singular value decomposition (SVD), proper orthogonal decomposition (POD), and greedy algorithms. In this context it also discusses recent results on a priori convergence in the context of the concept of the Kolmogorov N-width [6]. Chapter 4 contains a thorough discussion of the computation of lower bounds for stability factors and a comparative discussion of the various techniques. Chapter 5 focuses on the empirical interpolation method (EIM) [7], which is emerging as a standard element to address problems exhibiting non-affine

parametrizations and nonlinearities. It is our hope that these last two chapters will provide a useful overview of more recent material, allowing readers who wish to address more advanced problems to pursue the development of reduced basis methods for applications of interest to them. Chapter 6 offers an overview of a number of more advanced developments and is intended more as an appetizer than as a solution manual.

Throughout the text we provide some illustrative examples of applications in computational mechanics to guide readers through the various topics. All of the main algorithmic elements are outlined by graphical boxes to assist the reader in his or her efforts to implement the algorithms, emphasizing a matrix notation. An appendix with mathematical preliminaries is also included.

This book is loosely based on a Reduced Basis handbook available online [8], and we thank the co-author of this handbook, our colleague Anthony T. Patera (MIT), for his encouragement, support, and advice during the writing of the book. It benefits from our long-lasting collaboration with him and his many co-workers. We would like to acknowledge all those colleagues who contributed at various levels in the preparation of this manuscript and the related research. In particular, we would like to thank Francesco Ballarin and Alberto Sartori for preparing representative tutorials and the new open-source software library available as a companion to this book at http://mathlab.sissa.it/rbnics. An important role, including the provision of useful feedback, was also played by our very talented and motivated students attending regular doctoral and master classes at EPFL and SISSA (and ICTP), tutorials in Minneapolis and Savannah, and several summer/winter schools on the topic in Paris, Cortona, Hamburg, Udine (CISM), Munich, Sevilla, Pamplona, Barcelona, Torino, and Bilbao.

Lausanne, Switzerland — Jan S. Hesthaven
Trieste, Italy — Gianluigi Rozza
Paris, France — Benjamin Stamm
June 2015

References

1. C. Prudhomme, D.V. Rovas, K. Veroy, L. Machiels, Y. Maday, A.T. Patera, G. Turinici, Reliable real-time solution of parametrized partial differential equations: reduced-basis output bound methods. J. Fluids Eng. **124**, 70–80 (2002)
2. A. Quarteroni, G. Rozza, A. Manzoni, Certified reduced basis approximation for parametrized PDE and applications. J. Math Ind. **3** (2011)
3. G. Rozza, Fundamentals of reduced basis method for problems governed by parametrized PDEs and applications, in *CISM Lectures notes "Separated Representation and PGD based model reduction: fundamentals and applications"* (Springer Vienna, 2014)

4. G. Rozza, P. Huynh, N.C. Nguyen, A.T. Patera, Real-Time Reliable Simulation of Heat Transfer Phenomena, in *ASME, Heat Transfer Summer Conference collocated with the InterPACK09 and 3rd Energy Sustainability Conferences, American Society of Mechanical Engineers* (2009), pp. 851–860
5. G. Rozza, P. Huynh, A.T. Patera, Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: Application to transport and continuum mechanics. Arch. Comput. Methods Eng. **15**, 229–275 (2008)
6. P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, P. Wojtaszczyk, Convergence rates for greedy algorithms in reduced basis methods. SIAM J. Math. Anal. **43**, 1457–1472 (2011)
7. M. Barrault, Y. Maday, N.C. Nguyen, A.T. Patera, An empirical interpolation method: application to efficient reduced-basis discretization of partial differential equations. C.R. Math. **339**, 667–672 (2004)
8. A.T. Patera, G. Rozza, *Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations*, Copyright MIT 2007, MIT Pappalardo Graduate Monographs in Mechanical Engineering, http://www.augustine.mit.edu, 2007

Contents

Chapter 1
Introduction

Models expressed as parametrized partial differential equations are ubiquitous throughout engineering and the applied sciences as models for unsteady and steady heat and mass transfer, acoustics, solid and fluid mechanics, electromagnetics or problems of finance. In such models a number of input-parameters are used to characterize a particular problem and possible variations in its geometric configuration, physical properties, boundary conditions or source terms. The parametrized model implicitly connects these input parameters to outputs of interest of the model, e.g., a maximum system temperature, an added mass coefficient, a crack stress intensity factor, an effective constitutive property, a waveguide transmission loss, a channel flowrate or a pressure drop, etc.

While the development of accurate computational tools to allow the solution of such problems clearly is of broad interest, we focus on problems in which the solution is sought for a large number of different parameter values. Examples of typical applications of relevance are optimization, control, design, uncertainty quantification, real time query and others. It is not only the accuracy of the model that matters, but the computational efficiency of the model is likewise critical. Similar constraints emerge when real-time or near real-time responses are needed for rapid prototyping or computer animations relying on models of increasing physical accuracy, for instance.

In such situations, we need the ability to accurately and efficiently evaluate an output of interest when the input parameters are being varied. However, the complexity and computational cost associated with solving the full partial differential equation for each new parameter value rules out a direct approach. We must therefore seek a different approach that allows us to evaluate the desired output at minimal cost, yet without sacrificing the predictive accuracy of the complex model.

The goal of this text is develop the basic foundation for a class of methods, known as reduced basis (RB) methods, to accomplish this. As a convenient expository vehicle for the introduction of the methodology, we primarily consider the case of linear functional outputs of parametrized linear elliptic coercive partial differential equations. This class of problems, while relatively simple, is relevant to many important applications in transport (e.g., conduction and convection-diffusion,

J.S. Hesthaven et al., *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, SpringerBriefs in Mathematics,
DOI 10.1007/978-3-319-22470-1_1

but also reaction) and continuum mechanics (e.g., linear elasticity, fluid mechanics). Furthermore, they serve as examples to follow when considering more complex applications.

As we shall see soon, it is not the goal of the reduced basis methods to replace the expensive computational model but rather to build upon it. Indeed, the accuracy of the reduced model will be measured directly against the precision of the expensive computational model. This direct comparison against the expensive model that allows us to verify the accuracy of the reduced model and, thus, certify the validity of the predicted output. In other words, the goal is to pursue an algorithmic collaboration rather than an algorithmic competition with the expensive direct solution method.

It should be emphasized that although the focus is on the affine linear elliptic coercive case, the reduced basis approximation and the error estimation methodology provides a much more general methodology which allows extensions to non-affine, non-coercive, time-dependent, and, in some cases, even nonlinear problems. Towards the end of this text we shall discuss some of the extensions in some detail.

Before diving into the details of these methods, their derivation and analysis, let us in the following offer a selective historical overview of the development of certified reduced methods as well as a brief discussion of the content of the text and its use.

1.1 Historical Background and Perspectives

The central idea of the reduced basis approach is the identification of a suitable problem-dependent basis to effectively represent parametrized solutions to partial differential equations. In this simple interpretation, this is certainly not a recent idea and initial work grew out of two related lines of inquiry: one focusing on the need for effective, many-query design evaluation [1], and one from the desire for efficient parameter continuation methods for nonlinear problems [2–5].

These early approaches were soon extended to general finite-dimensional systems as well as certain classes of differential equations [6–12]. Furthermore, a number of different techniques for the identification of different reduced basis approximation spaces, exemplified by Taylor and Lagrange expansions [13] and more recently also Hermite [14] expansions, have emerged [15]. Further early work focused on different applications and specific classes of equations, e.g., fluid dynamics and the incompressible Navier-Stokes equations [14, 16–19].

In this early work, the approximation spaces were local and typically low-dimensional in the number of parameters. Furthermore, the absence of a posteriori error estimators left open questions of the accuracy of the reduced model. This is problematic since ad hoc predictions using a reduced basis, based on sample points far away from the point of interest, is likely to result in an inaccurate model. An a posteriori error estimator is crucial to determine the reliability of the output. Furthermore, sophisticated sampling strategies across the parameter space are crucial to ensure convergence and computational efficiency. This, again, relies on the availability of techniques for effective a posteriori error estimation.

Substantial recent efforts have been devoted to the development of techniques to formulate a posteriori error estimation procedures and rigorous error bounds for outputs of interest [20]. These a posteriori error bounds subsequently allow for the certification of the output of the reduced basis model for any parameter value.

However, the development of effective sampling strategies, in particular in the case of many parameters [21–24], can also be aided by the error estimators. These can play an important role in the development of efficient and effective sampling procedures by utilizing inexpensive error bounds to explore much larger subsets of the parameter domain in search of the most representative snapshots, and to determine when the basis is sufficiently rich.

We note here that such sampling techniques, often of a greedy nature, are similar in objective to, but very different in approach from, the more well-known methods of proper orthogonal decomposition (POD) [25–30]. While the former is directly applicable in the multi-dimensional parameter domain, the latter is most often applied only in the one-dimensional space. Furthermore, the POD approach, relying on the identification of the suitable reduced model by a singular value decomposition of a large number of snapshots, is often prohibitively expensive in the offline phase. However, POD techniques can be combined with the parametric RB approach [31–33]. A brief comparison of greedy and POD approaches for reduced basis constructions can be found in [34, 35].

Early developments of reduced basis methods did not fully decouple the underlying finite element approximation of the parametrized partial differential equation from the subsequent reduced basis projection and its evaluation. As a result, the computational savings offered by the reduced basis approach were typically rather modest [3, 10, 13] despite the small size of the resulting reduced basis problem.

Recent work has focused on achieving a full decoupling of the finite element scheme and the reduced order model through an offline-online procedure. In this approach, the complexity of the offline stage depends on the complexity of the finite element approximation of the parametrized partial differential equation, while the complexity of the online stage depends solely on the complexity of the reduced order model. When combined with the a posteriori error estimation, the online stage guarantees the accuracy of the high-fidelity finite element approximation at the low cost of a reduced order model.

For the case of an affine parameter dependence, in which case the spatial and the parametric dependence in the operator is separable, this offline-online decomposition is natural and has been re-invented repeatedly [17, 19, 36]. However, the combination of this with the rigor of the a posteriori error estimate is more involved and more recent [20, 37]. In the case of nonaffine parameter dependence, the development of offline-online strategies is much less transparent, and has only recently been addressed by the development of the empirical interpolation method [38, 39]. This last development, essential for the overall efficiency of the offline-online decomposition, has opened up for the development of reduced basis methods and their use in real-time and many-query contexts for complex applications, including nonlinear problems.

We note that historically reduced basis methods have been quantified relative to the underlying finite element discretizations [7, 41–45]. However, there are good reasons

to consider alternatives, e.g., a systematic finite volume framework for reduced basis approximation and a posteriori error estimation has been developed in [46].

The reduced basis approach can be extended to the more general case of non-coercive problems. Extensions to Maxwell equations have demonstrate the potential for reduced basis techniques within such a context [47–49]. The development of reduced basis models for problems described by boundary and integral equations is more complicated since the operators typically are non-affine. However, recent work has demonstrated its potential rapid evaluation of electromagnetic scattering applications [50–52].

The special issues associated with saddle-point problems [53, 54], in particular the Stokes equations of incompressible flow, are addressed for divergence-free spaces in [14, 16, 19] and non-divergence-free spaces in [40, 55–57]. A reduced basis optimal control saddle-point framework is introduced in [58, 59].

A recent development, initiated in [60], is the use of reduced basis methods for homogenization [61] and to accelerate multi-scale methods, including heterogeneous multi-scale methods [62–65] reduced order multi-scale finite element methods [66].

The exploration of a 'parameter + time' framework in the context of affine linear parabolic partial differential equations, e.g., the heat equation and the convection-diffusion equation, is discussed at length in [67, 68].

Reduced basis methods can be effectively applied also to nonlinear problems [69–71], although this typically introduces both numerical and theoretical complications, and many questions remain open. For classical problems with a quadratic nonlinearity, there has been substantial progress, e.g., Navier-Stokes/Boussinesq and Burgers' equations in fluid mechanics [19, 21, 23, 72–75] and nonlinear elasticity in solid mechanics.

A posteriori error bounds introduced for linear problems can be effectively extended to steady nonlinear problems (see e.g. [76] for steady incompressible Navier-Stokes equations). However, the most important challenge deals with the reliability and/or the certification of the methodology in the unsteady parabolic problems [77–79]. In such cases, the exponential growth of the estimate seriously compromises a priori and a posteriori error estimates, yielding bounds which are limited to modest (final) times and modest Reynolds numbers [80].

Efforts dealing with (homogeneous or even heterogeneous) couplings in a multiphysics setting, based on domain decomposition techniques, is an area of recent activity. A domain decomposition approach [81, 82], combined with a reduced basis method, has been successfully applied in [43, 83, 84] and further extensions discussed in [85–87]. A coupled multiphysics setting has been proposed for simple fluid-structure interaction problems in [88–90], and for Stokes-Darcy problem in [91].

Optimal control [92–99] as many-query applications continues to be a subject of extensive research and is often of interest also in an industrial context. A main area is the study of efficient techniques to deal with geometric parameters in order to keep the number of parameters manageable while guaranteeing versatility in the parametrization to enable representation of complex shapes. Recent works [100–103] deal with free-form deformation techniques combined with empirical interpolation in bio-medical and aerodynamic problems.

Another active field relates to the development and application of the reduced basis methodology in the context of the quantification of uncertainty, offing another example of application where many-query problems arise naturally [104–106]. Such problems are often characterized by having a high-dimensional parameter space and recent work has focused on the development of efficient ways to explore the parameters space, e.g., modified greedy algorithms and combined adaptive techniques [107–109], and *hp*-reduced basis method [110, 111]. At the same time, improvements in a posteriori error bounds for non-affine problems [112], the reduction of the computational complexity for high-dimensional problems and more efficient estimation of lower bounds of stability factors for complex non-affine problems [113, 114] are under investigation.

In parallel with many of these more fundamental and algorithmic developments, there are substantial activities seeking to improve the computational performance for complex problems by performing the offline work on large scale computational platforms and allow the use of the reduced models on deployed platforms [115].

This brief overview does not pretend to be complete and there are numerous other activities in this fast growing research area in which many advances can be expected in the coming years.

1.2 About this Book

The main target audience of this brief introduction to certified reduced basis methods are researchers with an interest and a need to understand the foundation of such techniques. While it is not intended to be a textbook, it can certainly be used as part of an advanced class and is intended to provide enough material to enable self study and further exploration of more advanced reference material.

To fully benefit from the text, a solid background in finite elements methods for solving linear partial differential equations is needed and an elementary understanding of linear partial differential equations is clearly beneficial. Many ideas and concepts will be introduced throughout the text but rather than providing a complete treatment, we strive to offer references that allows the reader to dive deeper into these as needed.

What remains of the text is organized into 5 additional chapters. In Chap. 2 we describe the basic setting for the affine linear elliptic coercive setting and introduce two illustrative examples that we shall revisit throughout the text to illustrate the performance of the reduced basis methods on successively more complex aspects of the two problems.

Chapter 3 is a key chapter which is strictly devoted to a discussion of the reduced basis methodology. In particular we discuss the reduced basis Galerkin projection and optimality, greedy sampling procedures for the construction of the basis in an optimal manner and recall central elements of the convergence theory.

In Chap. 4 we present the central ideas that allow for the development of rigorous and relatively sharp a posteriori output error bounds for reduced basis approximations. This also includes a discussion of methods for the accurate and efficient estimation of the lower bound of the coercivity-constant, required as part of the a posteriori error estimation procedure.

This sets the stage for Chap. 5 where we pursue the first extension of the basic methodology and discuss the formulation of reduced basis methods for non-affine problems. As we shall realize, the assumption of an affine operator is critically related to the efficiency of the reduced basis method and we discuss a strategy that reduces non-affine operators and data to an approximate affine form. This reduction must, however, be done efficiently to avoid a proliferation of parametric functions and a corresponding degradation of the online response time. This extension, presented in detail in Chap. 5 is based on the Empirical Interpolation Method which we discuss in detail. We shall also demonstrate how this approach allows for the treatment on nonlinear problems.

With the central elements of the reduced basis techniques having been developed in Chaps. 3–5, the final Chap. 6 is devoted to a discussion of a few more advanced developments. In particular, we discuss the development of reduced basis methods for time-dependent problems, problems with a non-compliant output function, non-coercive problems and problems with a parametrization of the geometry.

Throughout the text we emphasize the algorithmic aspects of the reduced basis methods. In particular, for all central elements of the approach we provide an algorithmic breakdown as well as an illustration of the algorithm in a matrix-centric approach. It is the hope that this will assist readers in the implementation of the ideas and easy adoption to problems of their own interest.

1.3 Software Libraries with Support for Reduced Basis Algorithms and Applications

During recent years, software libraries have been developed or extended to include reduced basis algorithms and their application. We provide here a list that, to the best of our knowledge, accounts for available resources at this time. Note that this is provided as a resource only and no guarantees are offered. Questions should be addressed to the authors of the individual software.

- `rbMIT`:
  http://augustine.mit.edu/methodology/methodology_rbMIT_System.htm
  This is a `MATLAB` based library, provided as a companion to the book [34], available at the same link. This library emphasizes geometric affine parametrization of domains and includes primal-dual formulation, offline-online computational steps, error bounds, parametrized stability factor approximation by the Successive Constraint Method, as well as greedy and POD-greedy sampling procedures for basis assembly. Numerous examples are available in heat and mass transfer, linear

elasticity and potential flows. Truth solutions are provided by the finite element method.

- RBMATLAB:
  http://www.ians.uni-stuttgart.de/MoRePaS/software/rbmatlab/1.13.10/doc/index.html
  This is a MATLAB library containing reduced simulation methods for linear and nonlinear, affine or arbitrarily, parameter dependent evolution problems with finite element, finite volume or local discontinuous Galerkin discretizations.
- rb00mit:
  http://augustine.mit.edu/methodology/methodology_rbAPPmit_Client_Software.htm
  This is a package for libMesh (http://libmesh.github.io/), a C++ library for parallel adaptive mesh refinement/coarsening simulations, containing an open source implementation of the certified Reduced Basis method for Android smartphones.
- pyMOR:
  http://www.pymor.org/
  This is a software library for building model order reduction applications using Python.
- Feel++:
  http://www.feelpp.org/
  A C++ library for partial differential equation, solved using generalized Galerkin methods, such as the finite element method, the h/p finite element method, the spectral element method or reduced basis methods.
- DUNE-RB:
  http://users.dune-project.org/projects/dune-rb/wiki
  This is a module for the Dune library (www.dune-project.org) with C++ template classes for use in snapshot generation and the reduced basis offline phases for various discretizations. The focus is on efficient parallel snapshot generation.
- FreeFem++:
  http://www.freefem.org/ff++/
  This is a partial differential equation solver based on its own language. FreeFem scripts can be used to solve multiphysics non linear systems in 2D and 3D, including some support for reduced methods. Tutorials on POD and reduced basis methods are available.
- RBniCS:
  http://mathlab.sissa.it/rbnics
  This software is developed for the construction of the Examples in this book and has been used throughout. The package is based on Python. The high order finite element solver, providing the truth approximation, is based on the FEniCS project (http://fenicsproject.org/).

References

1. R. Fox, H. Miura, An approximate analysis technique for design calculations. AIAA J. **9**, 177–179 (1971)
2. B. Almroth, P. Stern, F. Brogan, Automatic choice of global shape functions in structural analysis. AIAA J. **16**, 525–528 (1978)
3. A.K. Noor, Recent advances in reduction methods for nonlinear problems. Comput. Struct. **13**, 31–44 (1981)
4. A.K. Noor, On making large nonlinear problems small. Comput. Methods Appl. Mech. Eng. **34**, 955–985 (1982)
5. A.K. Noor, J.M. Peters, Reduced basis technique for nonlinear analysis of structures. AIAA J. **18**, 455–462 (1980)
6. A. Barrett, G. Reddien, On the reduced basis method. ZAMM-J. Appl. Math. Mech. Z. Angew. Math. Mechanik **75**, 543–549 (1995)
7. J. Fink, W. Rheinboldt, On the error behavior of the reduced basis technique for nonlinear finite element approximations. ZAMM-J. Appl. Math. Mech. Z. Angew. Math. Mechanik **63**, 21–28 (1983)
8. M.-Y.L. Lee, Estimation of the error in the reduced basis method solution of differential algebraic equation systems. SIAM J. Numer. Anal. **28**, 512–528 (1991)
9. A.K. Noor, C.D. Balch, M.A. Shibut, Reduction methods for nonlinear steady-state thermal analysis. Int. J. Numer. Methods Eng. **20**, 1323–1348 (1984)
10. T. Porsching, M.L. Lee, The reduced basis method for initial value problems. SIAM J. Numer. Anal. **24**, 1277–1287 (1987)
11. W.C. Rheinboldt, Numerical analysis of continuation methods for nonlinear structural problems. Comput. Struct. **13**, 103–113 (1981)
12. W.C. Rheinboldt, On the theory and error estimation of the reduced basis method for multi-parameter problems. Nonlinear Anal. Theor. Meth. Appl. **21**, 849–858 (1993)
13. T. Porsching, Estimation of the error in the reduced basis method solution of nonlinear equations. Math. Comput. **45**, 487–496 (1985)
14. K. Ito, S. Ravindran, A reduced-order method for simulation and control of fluid flows. J. Comput. Phys. **143**, 403–425 (1998)
15. T. Lassila, A. Manzoni, A. Quarteroni, G. Rozza, in *Generalized Reduced Basis Methods and N-width Estimates for the Approximation of the Solution Manifold of Parametric PDEs*, eds. by F. Brezzi, P. Colli Franzone, U. Gianazza, G. Gilardi. Analysis and Numerics of Partial Differential Equations. Springer INdAM Series, vol. 4 (Springer Milan, 2013), pp. 307–329
16. M.D. Gunzburger, *Finite Element Methods for Viscous Incompressible Flows: A guide to theory, practice, and algorithms* (Elsevier, 2012)
17. K. Ito, S. Ravindran, A reduced basis method for control problems governed by PDEs, in *Control and Estimation of Distributed Parameter Systems* (Springer, 1998), pp. 153–168
18. K. Ito, S. Ravindran, Reduced basis method for optimal control of unsteady viscous flows. Int. J. Comput. Fluid Dyn. **15**, 97–113 (2001)
19. J.S. Peterson, The reduced basis method for incompressible viscous flow calculations. SIAM J. Sci. Stat. Comput. **10**, 777–786 (1989)
20. C. Prudhomme, D.V. Rovas, K. Veroy, L. Machiels, Y. Maday, A.T. Patera, G. Turinici, Reliable real-time solution of parametrized partial differential equations: reduced-basis output bound methods. J. Fluids Eng. **124**, 70–80 (2002)
21. N.N. Cuong, K. Veroy, A.T. Patera, Certified real-time solution of parametrized partial differential equations, in *Handbook of Materials Modeling* (Springer, 2005), pp. 1529–1564
22. G. Rozza, Reduced-basis methods for elliptic equations in sub-domains with a posteriori error bounds and adaptivity. Appl. Numer. Math. **55**, 403–424 (2005)
23. K. Veroy, C. Prudhomme, D. Rovas, A. Patera, A posteriori error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations, in *Proceedings of the 16th AIAA Computational Fluid Dynamics Conference*, vol. 3847 (2003)

24. A. Manzoni, A. Quarteroni, G. Rozza, Computational reduction for parametrized PDEs: strategies and applications. Milan J. Math. **80**, 283–309 (2012)
25. M.D. Gunzburger, Perspectives in flow control and optimization, vol. 5, Siam, 2003
26. K. Kunisch, S. Volkwein, Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. SIAM J. Numer. Anal. **40**, 492–515 (2002)
27. P.A. LeGresley, J.J. Alonso, Airfoil design optimization using reduced order models based on proper orthogonal decomposition. AIAA Pap. **2000**, 2545 (2000)
28. S. Ravindran, A reduced-order approach for optimal control of fluids using proper orthogonal decomposition. Int. J. Numer Methods Fluids **34**, 425–448 (2000)
29. L. Sirovich, Turbulence and the dynamics of coherent structures. i-coherent structures. ii-symmetries and transformations. iii-dynamics and scaling. Q. Appl. Math. **45**, 561–571 (1987)
30. K. Willcox, J. Peraire, Balanced model reduction via the proper orthogonal decomposition. AIAA J. **40**, 2323–2330 (2002)
31. T. Bui-Thanh, M. Damodaran, K. Willcox, Proper orthogonal decomposition extensions for parametric applications in compressible aerodynamics. AIAA Pap. **4213** (2003)
32. E.A. Christensen, M. Brøns, J.N. Sørensen, Evaluation of proper orthogonal decomposition-based decomposition techniques applied to parameter-dependent nonturbulent flows. SIAM J. Sci. Comput. **21**, 1419–1434 (1999)
33. M.D. Gunzburger, J.S. Peterson, J.N. Shadid, Reduced-order modeling of time-dependent pdes with multiple parameters in the boundary data. Comput. Methods Appl. Mech. Eng. **196**, 1030–1047 (2007)
34. A.T. Patera, G. Rozza, *Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations*, Copyright MIT 2007, MIT Pappalardo Graduate Monographs in Mechanical Engineering, http://www.augustine.mit.edu, 2007
35. G. Rozza, P. Huynh, A.T. Patera, Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: Application to transport and continuum mechanics. Arch. Comput. Methods Eng. **15**, 229–275 (2008)
36. E. Balmès, Parametric families of reduced finite element models. Theory and applications. Mech. Syst. Signal Process. **10**, 381–394 (1996)
37. P. Huynh, G. Rozza, S. Sen, A.T. Patera, A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants. C.R. Math. **345**, 473–478 (2007)
38. M. Barrault, Y. Maday, N.C. Nguyen, A.T. Patera, An empirical interpolation method: application to efficient reduced-basis discretization of partial differential equations. C.R. Math. **339**, 667–672 (2004)
39. M.A. Grepl, Y. Maday, N.C. Nguyen, A.T. Patera, Efficient reduced-basis treatment of non-affine and nonlinear partial differential equations. ESAIM. Math. Model. Numer. Anal. **41**, 575–605 (2007)
40. G. Rozza, Reduced basis methods for Stokes equations in domains with non-affine parameter dependence. Comput. Vis. Sci. **12**, 23–35 (2009)
41. A.E. Løvgren, Y. Maday, E.M. Rønquist, in *A Reduced Basis Element Method for Complex Flow Systems*, eds. by P. Wesseling, E. Onate, J. Periaux. European Conference on Computational Fluid Dynamics (TU Delft, The Netherlands, 2006)
42. A.E. Løvgren, Y. Maday, E.M. Rønquist, The reduced basis element method for fluid flows, in *Analysis and Simulation of Fluid Dynamics* (Springer, 2007), pp. 129–154
43. A.E. Løvgren, Y. Maday, E.M. Rønquist, A reduced basis element method for the steady Stokes problem. ESAIM Math. Model. Numer. Anal. **40**, 529–552 (2006)
44. A.E. Løvgren, Y. Maday, E.M. Rønquist, A reduced basis element method for the steady Stokes problem: application to hierarchical flow system. Model. Identif. Control **27**, 79–94 (2006)
45. A.T. Patera, E.M. Rønquist, Reduced basis approximation and a posteriori error estimation for a boltzmann model. Comput. Methods Appl. Mech. Eng. **196**, 2925–2942 (2007)
46. M. Drohmann, B. Haasdonk, M. Ohlberger, Reduced basis method for finite volume approximation of evolution equations on parametrized geometries, in *Proceedings of ALGORITMY* (2009), pp. 1–10

47. Y. Chen, J.S. Hesthaven, Y. Maday, J. Rodríguez, Improved successive constraint method based a posteriori error estimate for reduced basis approximation of 2d maxwell's problem. ESAIM Math. Model. Numer. Anal. **43**, 1099–1116 (2009)
48. Y. Chen, J.S. Hesthaven, Y. Maday, J. Rodríguez, Certified reduced basis methods and output bounds for the harmonic Maxwell's equations. SIAM J. Sci. Comput. **32**, 970–996 (2010)
49. Y. Chen, J.S. Hesthaven, Y. Maday, J. Rodríguez, X. Zhu, Certified reduced basis method for electromagnetic scattering and radar cross section estimation. Comput. Methods Appl. Mech. Eng. **233**, 92–108 (2012)
50. M. Fares, J.S. Hesthaven, Y. Maday, B. Stamm, The reduced basis method for the electric field integral equation. J. Comput. Phys. **230**, 5532–5555 (2011)
51. M. Ganesh, J.S. Hesthaven, B. Stamm, A reduced basis method for electromagnetic scattering by multiple particles in three dimensions. J. Comput. Phy. **231**, 7756–7779 (2012)
52. J.S. Hesthaven, B. Stamm, S. Zhang, Certified reduced basis method for the electric field integral equation. SIAM J. Sci. Comput. **34**, A1777–A1799 (2012)
53. D. Boffi, F. Brezzi, M. Fortin, *Mixed Finite Element Methods and Applications*. Springer Series in Computational Mathematics, vol. 44 (Springer, Heidelberg, 2013)
54. F. Brezzi, M. Fortin, *Mixed and Hybrid Finite Element Methods*. Springer Series in Computational Mathematics, vol. 15 (Springer, New York, 1991)
55. G. Rozza, Real time reduced basis techniques for arterial bypass geometries, in *Computational Fluid and Solid Mechanics-Third MIT Conference on Computational Fluid and Solid Mechanics* (Elsevier, 2005), pp. 1283–1287
56. G. Rozza, K. Veroy, On the stability of the reduced basis method for Stokes equations in parametrized domains. Comput. Methods Appl. Mech. Eng. **196**, 1244–1260 (2007)
57. A.-L. Gerner, K. Veroy, Certified reduced basis methods for parametrized saddle point problems. SIAM J. Sci. Comput. **34**, A2812–A2836 (2012)
58. F. Negri, A. Manzoni, G. Rozza, Reduced basis approximation of parametrized optimal flow control problems for the Stokes equations. Comput. Math. Appl. **69**, 319–336 (2015)
59. P. Chen, A. Quarteroni, G. Rozza, Multilevel and weighted reduced basis method for stochastic optimal control problems constrained by steady Stokes equations. Numer. Math. (2015)
60. N.C. Nguyen, A multiscale reduced-basis method for parametrized elliptic partial differential equations with multiple scales. J. Comput. Phys. **227**, 9807–9822 (2008)
61. S. Boyaval, Reduced-basis approach for homogenization beyond the periodic setting. Multiscale Model. Simul. **7**, 466–494 (2008)
62. S. Kaulmann, M. Ohlberger, B. Haasdonk, A new local reduced basis discontinuous Galerkin approach for heterogeneous multiscale problems. C.R. Math. **349**, 1233–1238 (2011)
63. A. Abdulle, Y. Bai, Reduced basis finite element heterogeneous multiscale method for high-order discretizations of elliptic homogenization problems. J. Comput. Phys. **231**, 7014–7036 (2012)
64. A. Abdulle, Y. Bai, Adaptive reduced basis finite element heterogeneous multiscale method. Comput. Methods Appl. Mech. Eng. **257**, 203–220 (2013)
65. A. Abdulle, Y. Bai, G. Vilmart, Reduced basis finite element heterogeneous multiscale method for quasilinear elliptic homogenization problems. Discrete Continuous Dyn. Syst.-Ser. S **8**, 91–118 (2014)
66. J.S. Hesthaven, S. Zhang, X. Zhu, Reduced basis multiscale finite element methods for elliptic problems. Multiscale Model. Simul. **13**, 316–337 (2015)
67. M.A. Grepl, *Reduced-basis approximation and a posteriori error estimation for parabolic partial differential equations*, Ph.D. thesis, Massachusetts Institute of Technology, 2005
68. M.A. Grepl, A.T. Patera, A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations. ESAIM. Math. Model. Numer. Anal. **39**, 157–181 (2005)
69. M.A. Grepl, Y. Maday, N.C. Nguyen, A.T. Patera, Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. ESAIM. Math. Model. Numer. Anal. **41**, 575–605 (2007)

70. C. Canuto, T. Tonn, K. Urban, A posteriori error analysis of the reduced basis method for nonaffine parametrized nonlinear PDEs. SIAM J. Numer. Anal. **47**, 2001–2022 (2009)
71. N. Jung, B. Haasdonk, D. Kroner, Reduced basis method for quadratically nonlinear transport equations. Int. J. Comput. Sci. Math. **2**, 334–353 (2009)
72. K. Veroy, C. Prud'homme, A.T. Patera, Reduced-basis approximation of the viscous burgers equation: rigorous a posteriori error bounds. C.R. Math. **337**, 619–624 (2003)
73. A. Quarteroni, G. Rozza, Numerical solution of parametrized Navier-Stokes equations by reduced basis methods. Numer. Methods Partial Differ. Equ. **23**, 923–948 (2007)
74. S. Deparis, G. Rozza, Reduced basis method for multi-parameter-dependent steady Navier-Stokes equations: applications to natural convection in a cavity. J. Comput. Phys. **228**, break4359–4378 (2009)
75. G. Rozza, N.C. Nguyen, A.T. Patera, S. Deparis, Reduced basis methods and a posteriori error estimators for heat transfer problems, in *ASME, Heat Transfer Summer Conference collocated with the InterPACK09 and 3rd Energy Sustainability Conferences* (American Society of Mechanical Engineers, 2009)
76. K. Veroy, A. Patera, Certified real-time solution of the parametrized steady incompressible navier-stokes equations: rigorous reduced-basis a posteriori error bounds. Int. J. Numer. Meth. Fluids **47**, 773–788 (2005)
77. N.C. Nguyen, G. Rozza, A.T. Patera, Reduced basis approximation and a posteriori error estimation for the time-dependent viscous Burgers equation. Calcolo **46**, 157–185 (2009)
78. D.J. Knezevic, N.C. Nguyen, A.T. Patera, Reduced basis approximation and a posteriori error estimation for the parametrized unsteady boussinesq equations. Math. Models Methods Appl. Sci. **21**, 1415–1442 (2011)
79. T. Lassila, A. Manzoni, A. Quarteroni, G. Rozza, in *Model Order Reduction in Fluid Dynamics: Challenges and Perspectives*, eds. by A. Quarteroni, G. Rozza. Reduced Order Methods for Modeling and Computational Reduction. MS & A—Modeling, Simulation and Applications, vol. 9 (Springer International Publishing, 2014), pp. 235–273
80. C. Johnson, R. Rannacher, M. Boman, Numerics and hydrodynamic stability: toward error control in computational fluid dynamics. SIAM J. Numer. Anal. **32**, 1058–1079 (1995)
81. A. Quarteroni, A. Valli, *Numerical Approximation of Partial Differential Equations*. Springer Series in Computational Mathematics, vol. 23 (Springer, Berlin, 1994)
82. A. Quarteroni, A. Valli, Domain decomposition methods for partial differential equations, in *Numerical Mathematics and Scientific Computation* (The Clarendon Press, Oxford University Press, Oxford Science Publications, New York, 1999)
83. A.E. Løvgren, Y. Maday, E.M. Rønquist, The reduced basis element method for fluid flows, in *Analysis and Simulation of Fluid Dynamics* (Springer, 2007), pp. 129–154
84. A. Løvgren, Y. Maday, E. Rønquist, The Reduced basis element method: offline-online decomposition in the nonconforming, nonaffine case, in *Spectral and High Order Methods for Partial Differential Equations* (Springer, 2011), pp. 247–254
85. L. Iapichino, A. Quarteroni, G. Rozza, A reduced basis hybrid method for the coupling of parametrized domains represented by fluidic networks. Comput. Methods Appl. Mech. Eng. **221**, 63–82 (2012)
86. C. Jaeggli, L. Iapichino, G. Rozza, An improvement on geometrical parameterizations by transfinite maps. C.R. Math. **352**, 263–268 (2014)
87. P. Huynh, D.J. Knezevic, A.T. Patera, A static condensation reduced basis element method: approximation and a posteriori error estimation. ESAIM: M2AN, **47**, 213–251 (2013)
88. T. Lassila, G. Rozza, Model reduction of steady fluid-structure interaction problems with free-form deformations and reduced basis methods, in *Proceedings of 10th Finnish Mechanics Days* (Jyvaskyla, Finland, 2009), pp. 454–465
89. T. Lassila, A. Quarteroni, G. Rozza, A reduced basis model with parametric coupling for fluid-structure interaction problems. SIAM J. Sci. Comput. **34**, A1187–A1213 (2012)
90. T. Lassila, A. Manzoni, A. Quarteroni, G. Rozza, A reduced computational and geometrical framework for inverse problems in hemodynamics. Int. J. Numer. Methods Biomed. Eng. **29**, 741–776 (2013)

91. I. Martini, G. Rozza, B. Haasdonk, Reduced basis approximation and a-posteriori error estimation for the coupled Stokes-Darcy system. Adv. Comput. Math. 1–27 (2014)
92. A. Quarteroni, G. Rozza, A. Quaini, in *Reduced Basis Methods for Optimal Control of Advection-diffusion Problems*, eds. by W. Fitzgibbon, R. Hoppe, J. Periaux, O. Pironneau, Y. Vassilevski. Advances in Numerical Mathematics (Institute of Numerical Mathematics, Russian Academy of Sciences and Houston, Department of Mathematics, University of Houston, Moscow, 2007), pp. 193–216
93. T. Tonn, K. Urban, S. Volkwein, Optimal control of parameter-dependent convection-diffusion problems around rigid bodies. SIAM J. Sci. Comput. **32**, 1237–1260 (2010)
94. L. Dedè, Reduced basis method and a posteriori error estimation for parametrized linear-quadratic optimal control problems. SIAM J. Sci. Comput. **32**, 997–1019 (2010)
95. T. Tonn, K. Urban, S. Volkwein, Comparison of the reduced-basis and pod a posteriori error estimators for an elliptic linear-quadratic optimal control problem. Math. Comput. Model. Dyn. Syst. **17**, 355–369 (2011)
96. G. Rozza, *Shape design by optimal flow control and reduced basis techniques: Applications to bypass configurations in haemodynamics*, Ph.D. thesis, EPFL Lausanne, Switzerland, 2005
97. G. Rozza, On optimization, control and shape design of an arterial bypass. Int. J. Numer. Meth. Fluids **47**, 1411–1419 (2005)
98. F. Negri, G. Rozza, A. Manzoni, A. Quarteroni, Reduced basis method for parametrized elliptic optimal control problems. SIAM J. Sci. Comput. **35**, A2316–A2340 (2013)
99. M. Karcher, M. Grepl, A posteriori error estimation for reduced order solutions of parametrized parabolic optimal control problems. ESAIM: M2AN, **48**, 1615–1638 (2014)
100. T. Lassila, G. Rozza, Parametric free-form shape design with pde models and reduced basis method. Comput. Methods Appl. Mech. Eng. **199**, 1583–1592 (2010)
101. G. Rozza, T. Lassila, A. Manzoni, Reduced basis approximation for shape optimization in thermal flows with a parametrized polynomial geometric map, in *Spectral and High Order Methods for Partial Differential Equations* (Springer, 2011), pp. 307–315
102. A. Manzoni, A. Quarteroni, G. Rozza, Shape optimization for viscous flows by reduced basis methods and free-form deformation. Int. J. Numer. Methods Fluids **70**, 646–670 (2012)
103. G. Rozza, A. Manzoni, Model order reduction by geometrical parametrization for shape optimization in computational fluid dynamics, in *Proceedings of the ECCOMAS CFD 2010, V European Conference on Computational Fluid Dynamics* (2010)
104. S. Boyaval, C.L. Bris, Y. Maday, N.C. Nguyen, A.T. Patera, A reduced basis approach for variational problems with stochastic parameters: Application to heat conduction with variable Robin coefficient. Comput. Methods Appl. Mech. Eng. **198**, 3187–3206 (2009)
105. N. Nguyen, G. Rozza, D. Huynh, A.T. Patera, *Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Parabolic PDEs; Application to Real-time Bayesian Parameter Estimation*, eds. by I. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, L. Tenorio, B. van Bloemen Waanders, K. Willcox. Computational Methods for Large Scale Inverse Problems and Uncertainty Quantification (Wiley, UK, 2009)
106. P. Huynh, D. Knezevic, A. Patera, Certified reduced basis model validation: a frequentistic uncertainty framework. Comput. Methods Appl. Mech. Eng. **201**, 13–24 (2012)
107. B. Haasdonk, M. Dihlmann, M. Ohlberger, A training set and multiple bases generation approach for parameterized model reduction based on adaptive grids in parameter space. Math. Comput. Model. Dyn. Syst. **17**, 423–442 (2011)
108. P. Chen, A. Quarteroni, G. Rozza, A weighted reduced basis method for elliptic partial differential equations with random input data. SIAM J. Numer. Anal. **51**, 3163–3185 (2013)
109. P. Chen, A. Quarteroni, G. Rozza, Comparison between reduced basis and stochastic collocation methods for elliptic problems. J. Sci. Comput. **59**, 187–216 (2014)
110. J.L. Eftang, A.T. Patera, E.M. Rønquist, An 'hp' certified reduced basis method for parametrized elliptic partial differential equations. SIAM J. Sci. Comput. **32**, 3170–3200 (2010)
111. J.L. Eftang, D.J. Knezevic, A.T. Patera, An 'hp' certified reduced basis method for parametrized parabolic partial differential equations. Math. Comput. Model. Dyn. Syst. **17**, 395–422 (2011)

112. J.L. Eftang, M.A. Grepl, A.T. Patera, A posteriori error bounds for the empirical interpolation method. C.R. Math. **348**, 575–579 (2010)
113. T. Lassila, G. Rozza, Model reduction of semiaffinely parameterized partial differential equations by two-level affine approximation. C.R. Math. **349**, 61–66 (2011)
114. T. Lassila, A. Manzoni, G. Rozza, On the approximation of stability factors for general parametrized partial differential equations with a two-level affine decomposition. ESAIM: M2AN **46**, 1555–1576 (2012)
115. P. Huynh, D. Knezevic, J. Peterson, A. Patera, High-fidelity real-time simulation on deployed platforms. Comput. Fluids **43**, 74–81 (2011)

Chapter 2
Parametrized Differential Equations

2.1 Parametrized Variational Problems

Let us first introduce a (suitably regular) physical domain $\Omega \in \mathbb{R}^d$ with boundary $\partial\Omega$, where $d = 1, 2$, or 3 is the spatial dimension. We shall consider only real-valued field variables. However, both scalar-valued (e.g., temperature in a Poisson conduction problems) and vector-valued (e.g., displacement in a linear elasticity problem) field variables $w : \Omega \rightarrow \mathbb{R}^{d_v}$ may be considered: here d_v denotes the dimension of the field variable; for scalar-valued fields, $d_v = 1$, while for vector-valued fields, $d_v = d$. We also introduce (boundary measurable) segments of $\partial\Omega$, Γ_i^D, $1 \leq i \leq d_v$, over which we will impose Dirichlet boundary conditions on the components of the field variable.

Let us also introduce the scalar spaces $\mathbb{V}_i$, $1 \leq i \leq d_v$,

$$\mathbb{V}_i = \mathbb{V}_i(\Omega) = \{v \in H^1(\Omega) \mid v|_{\Gamma_i^D} = 0\}, \qquad 1 \leq i \leq d_v.$$

In general $H_0^1(\Omega) \subset \mathbb{V}_i \subset H^1(\Omega)$, and for $\Gamma_i^D = \partial\Omega$, $\mathbb{V}_i = H_0^1(\Omega)$. We construct the space in which our vector-valued field variable shall reside as the Cartesian product $\mathbb{V} = \mathbb{V}_1 \times \cdots \mathbb{V}_{d_v}$; a typical element of $\mathbb{V}$ is denoted $w = (w_1, \ldots, w_{d_v})$. We equip $\mathbb{V}$ with an inner product $(w, v)_{\mathbb{V}}$, $\forall\ w, v, \in \mathbb{V}$, and the induced norm $\|w\|_{\mathbb{V}} = \sqrt{(w, w)_{\mathbb{V}}}$, $\forall\ w \in \mathbb{V}$: any inner product which induces a norm equivalent to the $(H^1(\Omega))^{d_v}$ norm is admissible. Therefore, $\mathbb{V}$ is a Hilbert space.

We finally introduce a (suitably regular) closed parameter domain $\mathbb{P} \in \mathbb{R}^P$, a typical parameter (or input) point, or vector, or P-tuple, denoted as $\mu = (\mu_{[1]}, \mu_{[2]}, \ldots, \mu_{[P]})$. We may thus define our parametric field variable as $u \equiv (u_1, \ldots, u_{d_v}) : \mathbb{P} \rightarrow \mathbb{V}$; here, $u(\mu)$ denotes the field for parameter value $\mu \in \mathbb{P}$.

J.S. Hesthaven et al., *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, SpringerBriefs in Mathematics,
DOI 10.1007/978-3-319-22470-1_2

2.1.1 Parametric Weak Formulation

Let us briefly introduce the general stationary problem in an abstract form. All of the working examples in this text can be cast in this framework. We are given parametrized linear forms $f : \mathbb{V} \times \mathbb{P} \to \mathbb{R}$ and $\ell : \mathbb{V} \times \mathbb{P} \to \mathbb{R}$ where the linearity is with respect to the first variable, and a parametrized bilinear form $a : \mathbb{V} \times \mathbb{V} \times \mathbb{P} \to \mathbb{R}$ where the bilinearity is with respect to the first two variables. Examples of such parametrized linear forms are given in the two examples of Sect. 2.3. The abstract formulation reads: given $\mu \in \mathbb{P}$, we seek $u(\mu) \in \mathbb{V}$ such that

$$a(u(\mu), v; \mu) = f(v; \mu), \qquad \forall\, v \in \mathbb{V}, \tag{2.1}$$

and evaluate

$$s(\mu) = \ell(u(\mu); \mu). \tag{2.2}$$

Here s is an output of interest, $s: \mathbb{P} \to \mathbb{R}$ is the input (parameter)-output relationship, and ℓ takes the role of a linear "output" functional which links the input to the output through the field variable $u(\mu)$.

In this initial part of the text we assume that problems of interest are compliant. A compliant problem of the form (2.1)–(2.2) satisfies two conditions:

(i) $\ell(\cdot\,; \mu) = f(\cdot\,; \mu)$, $\forall\, \mu \in \mathbb{P}$—the output functional and load/source functional are identical.
(ii) The bilinear form $a(\cdot\,, \cdot\,; \mu)$ is symmetric for any parameter value $\mu \in \mathbb{P}$.

Together, these two assumptions greatly simplify the formulation, the a priori convergence theory for the output, and the a posteriori error estimation for the output. Though quite restrictive, there are many interesting problems fulfilling this requirement across mechanics and physics, e.g., material properties, geometrical parametrization, etc. However, we return to the more general non-compliant case in the final Chap. 6.

2.1.2 Inner Products, Norms and Well-Posedness of the Parametric Weak Formulation

The Hilbert space $\mathbb{V}$ is equipped with an intrinsic norm $\|\cdot\|_{\mathbb{V}}$. In many cases this norm coincides with, or is equivalent to, the norm induced by the bilinear form a for a fixed parameter $\bar{\mu} \in \mathbb{P}$:

$$\begin{aligned} (w, v)_{\mathbb{V}} &= a(w, v; \bar{\mu}), \qquad \forall\, w, v \in \mathbb{V}, \\ \|v\|_{\mathbb{V}} &= \sqrt{a(w, w; \bar{\mu})}, \quad \forall\, w \in \mathbb{V}. \end{aligned} \tag{2.3}$$

The well-posedness of the abstract problem formulation (2.1) can be established by the Lax-Milgram theorem [1], see also the Appendix. In order to state a well-posed problem for all parameter values $\mu \in \mathbb{P}$, we assume in addition to the bilinearity and the linearity of the parametrized forms $a(\cdot, \cdot; \mu)$ and $f(\cdot; \mu)$, that

(i) $a(\cdot\,, \cdot\,; \mu)$ is coercive and continuous for all $\mu \in \mathbb{P}$ with respect to the norm $\|\cdot\|_{\mathbb{V}}$, i.e., for every $\mu \in \mathbb{P}$, there exists a positive constant $\alpha(\mu) \geq \alpha > 0$ and a finite constant $\gamma(\mu) \leq \gamma < \infty$ such that

$$a(v, v; \mu) \geq \alpha(\mu)\, \|v\|_{\mathbb{V}}^2 \qquad \text{and} \qquad a(w, v; \mu) \leq \gamma(\mu)\, \|w\|_{\mathbb{V}}\, \|v\|_{\mathbb{V}}, \tag{2.4}$$

for all $w, v \in \mathbb{V}$.

(ii) $f(\cdot\,; \mu)$ is continuous for all $\mu \in \mathbb{P}$ with respect to the norm $\|\cdot\|_{\mathbb{V}}$, i.e., for every $\mu \in \mathbb{P}$, there exists a constant $\delta(\mu) \leq \delta < \infty$ such that

$$f(v; \mu) \leq \delta(\mu)\, \|v\|_{\mathbb{V}}, \qquad \forall\, v \in \mathbb{V}.$$

The coercivity and continuity constants of $a(\cdot\,, \cdot\,; \mu)$ over $\mathbb{V}$ are, respectively, defined as

$$\alpha(\mu) = \inf_{v \in \mathbb{V}} \frac{a(v, v; \mu)}{\|v\|_{\mathbb{V}}^2}, \qquad \text{and} \qquad \gamma(\mu) = \sup_{w \in \mathbb{V}} \sup_{v \in \mathbb{V}} \frac{a(w, v; \mu)}{\|w\|_{\mathbb{V}} \|v\|_{\mathbb{V}}}, \tag{2.5}$$

for every $\mu \in \mathbb{P}$.

Finally, we also may introduce the usual energy inner product and the induced energy norm as

$$(w, v)_\mu = a(w, v; \mu), \qquad \forall\, w, v \in \mathbb{V}\,, \tag{2.6}$$

$$\|w\|_\mu = \sqrt{a(w, w; \mu)}, \qquad \forall\, w \in \mathbb{V}\,, \tag{2.7}$$

respectively; note that these quantities are parameter-dependent. Thanks to the coercivity and continuity assumptions on a, it is clear that (2.6) constitutes a well-defined inner product and (2.7) an induced norm equivalent to the $\|\cdot\|_{\mathbb{V}}$-norm.

2.2 Discretization Techniques

This section supplies an abstract framework of the discrete approximations of the parametric weak formulation (2.1) for conforming approximations, i.e., there is a discrete approximation space $\mathbb{V}_\delta$ in which the approximate solution is sought. This is a subset of $\mathbb{V}$, i.e., $\mathbb{V}_\delta \subset \mathbb{V}$. The conforming nature of the approximation space $\mathbb{V}_\delta$ is an essential assumption in the upcoming presentation of the method in Chap. 3, and for the error estimation discussed in Chap. 4.

As an example, the approximation space $\mathbb{V}_\delta$ can be constructed as a standard finite element method based on a triangulation and using piece-wise linear basis functions. Other examples include spectral methods or higher order finite elements, provided only that the formulation is based on a variational approach.

We denote the dimension of the discrete space $\mathbb{V}_\delta$ by $N_\delta = \dim(\mathbb{V}_\delta)$ and equip $\mathbb{V}_\delta$ with a basis $\{\varphi_i\}_{i=1}^{N_\delta}$. For each $\mu \in \mathbb{P}$, the discrete problem consists of finding $u_\delta(\mu) \in \mathbb{V}_\delta$ such that

$$a(u_\delta(\mu), v_\delta; \mu) = f(v_\delta; \mu), \quad \forall\, v_\delta \in \mathbb{V}_\delta, \tag{2.8}$$

and evaluate

$$s_\delta(\mu) = \ell(u_\delta(\mu); \mu).$$

This problem is denoted as the truth problem. It is a solver of choice in the case where the solution needs to be computed for one parameter value only and it is assumed that this solution, called the truth approximation, can be achieved with as high accuracy as desired.

The computation of the truth solution is, however, potentially very expensive since the space $\mathbb{V}_\delta$ may involved many degrees of freedom N_δ to achieve the desired accuracy level. On the other hand, it provides an accurate approximation $u_\delta(\mu)$ in the sense that the error $\|u(\mu) - u_\delta(\mu)\|_{\mathbb{V}}$ is acceptably small. This model is sometimes also referred to as high fidelity model.

We note that due to the coercivity and continuity of the bilinear form, and the conformity of the approximation space, we ensure the Galerkin orthogonality

$$a\big(u(\mu) - u_\delta(\mu), v_\delta; \mu\big) = 0, \quad \forall\, v_\delta \in \mathbb{V}_\delta,$$

to recover Cea's lemma. Indeed, let $v_\delta \in \mathbb{V}_\delta$ be arbitrary and observe that

$$\|u(\mu) - u_\delta(\mu)\|_{\mathbb{V}} \leq \|u(\mu) - v_\delta\|_{\mathbb{V}} + \|v_\delta - u_\delta(\mu)\|_{\mathbb{V}},$$

by the triangle inequality. Furthermore, it holds that

$$\begin{aligned}\alpha(\mu)\|v_\delta - u_\delta(\mu)\|_{\mathbb{V}}^2 &\leq a\big(v_\delta - u_\delta(\mu), v_\delta - u_\delta(\mu); \mu\big) = a\big(v_\delta - u(\mu), v_\delta - u_\delta(\mu); \mu\big) \\ &\leq \gamma(\mu)\, \|v_\delta - u(\mu)\|_{\mathbb{V}}\, \|v_\delta - u_\delta(\mu)\|_{\mathbb{V}}\end{aligned}$$

by applying the coercivity assumption, the Galerkin orthogonality and the continuity assumption to obtain

$$\|u(\mu) - u_\delta(\mu)\|_{\mathbb{V}} \leq \left(1 + \frac{\gamma(\mu)}{\alpha(\mu)}\right) \inf_{v_\delta \in \mathbb{V}_\delta} \|u(\mu) - v_\delta\|_{\mathbb{V}}.$$

Linear algebra box: The truth solver

We denote the stiffness matrix and the right hand side of the truth problem by $\mathbf{A}_\delta^\mu \in \mathbb{R}^{N_\delta \times N_\delta}$ and $\mathrm{f}_\delta^\mu \in \mathbb{R}^{N_\delta}$, respectively. Further, we denote by $\mathbf{M}_\delta \in \mathbb{R}^{N_\delta \times N_\delta}$ the matrix associated with the inner product $(\cdot, \cdot)_{\mathbb{V}}$ of $\mathbb{V}_\delta$, defined as

$$(\mathbf{M}_\delta)_{ij} = (\varphi_j, \varphi_i)_{\mathbb{V}}, \qquad (\mathbf{A}_\delta^\mu)_{ij} = a(\varphi_j, \varphi_i; \mu), \qquad \text{and} \qquad (\mathrm{f}_\delta^\mu)_i = f(\varphi_i; \mu),$$

for all $1 \le i, j \le N_\delta$. We recall that $\{\varphi_i\}_{i=1}^{N_\delta}$ is of a basis of $\mathbb{V}_\delta$. Then, the truth problem reads: for each $\mu \in \mathbb{P}$, find $\mathrm{u}_\delta^\mu \in \mathbb{R}^{N_\delta}$ s.t.

$$\mathbf{A}_\delta^\mu \, \mathrm{u}_\delta^\mu = \mathrm{f}_\delta^\mu.$$

Then, evaluate the output functional (in the compliant case)

$$s_\delta(\mu) = (\mathrm{u}_\delta^\mu)^T \, \mathrm{f}_\delta^\mu.$$

The field approximation $u_\delta(\mu)$ is obtained by $u_\delta(\mu) = \sum_{i=1}^{N_\delta} (\mathrm{u}_\delta^\mu)_i \, \varphi_i$ where $(\mathrm{u}_\delta^\mu)_i$ denotes the i-th coefficient of the vector u_δ^μ.

This implies that the approximation error $\|u(\mu) - u_\delta(\mu)\|_{\mathbb{V}}$ is closely related to the best approximation error of $u(\mu)$ in the approximation space $\mathbb{V}_\delta$ through the constants $\alpha(\mu), \gamma(\mu)$. More details on the numerical analysis of unparametrized problems can be found in the Appendix.

The linear algebra box The truth solver illustrates the implementation of the truth solver on the level of linear algebra. The size of the unknown vector is N_δ and the size of the stiffness matrix is $N_\delta \times N_\delta$. Depending on the solver of choice to invert the linear system and the properties of the stiffness matrix, the operation count of the map $\mu \to s_\delta(\mu)$ is $\mathcal{O}({N_\delta}^\alpha)$, for $\alpha \ge 1$, but in any case dependent of N_δ.

2.3 Toy Problems

We want to consider simple parametrized examples, intended to be representative of larger classes of problems, to motivate the reader. We consider two model problems: a (steady) heat conduction problem with conductivity and heat flux as parameters; and a linear elasticity problem with load traction conditions as parameters.

We will present generalizations of these examples later in this text in Chaps. 5 and 6. We thus limit ourselves to the following problems only for the introduction of the topic and present some advanced examples later.

2.3.1 Illustrative Example 1: Heat Conduction Part 1

We consider a steady heat conduction problem (we refer the reader in need of more information about thermal problems to [2]) in a two-dimensional domain $\Omega = (-1, 1) \times (-1, 1)$ with outward pointing unit normal n on $\partial\Omega$. The boundary $\partial\Omega$ is split into three parts: the bottom $\Gamma_{\text{base}} = (-1, 1) \times \{-1\}$, the top $\Gamma_{\text{top}} = (-1, 1) \times \{1\}$ and the sides $\Gamma_{\text{side}} = \{\pm 1\} \times (-1, 1)$. The normalized thermal conductivity is denoted by κ. Let Ω_0 be a disk centered at the origin of radius $r_0 = 0.5$ and define $\Omega_1 = \Omega \backslash \overline{\Omega_0}$. Consider the conductivity κ to be constant on Ω_0 and Ω_1, i.e.

$$\kappa|_{\Omega_0} = \kappa_0 \qquad \text{and} \qquad \kappa|_{\Omega_1} = 1.$$

The geometrical set-up is illustrated in Fig. 2.1.

We consider $P = 2$ parameters and $\mathbb{P} = [\mu_{[1]}^{\min}, \mu_{[1]}^{\max}] \times [\mu_{[2]}^{\min}, \mu_{[2]}^{\max}]$ in this model problem. The first one is related to the conductivity in Ω_0, i.e. $\mu_{[1]} = \kappa_0$. We can write $\kappa_\mu = \mathbf{1}_{\Omega_1} + \mu_{[1]}\mathbf{1}_{\Omega_0}$, where $\mathbf{1}$ is the characteristic function of the corresponding set in the sub-script. The second parameter $\mu_{[2]}$ reflects the constant heat flux over Γ_{base}. Our parameter vector is thus given by $\mu = (\mu_{[1]}, \mu_{[2]})$.

The scalar field variable $u(\mu)$ is the temperature that satisfies Poisson's equation in Ω; homogeneous Neumann (zero flux, or insulated) conditions on the side boundaries Γ_{side}; homogeneous Dirichlet (temperature) conditions on the top boundary Γ_{top}; and parametrized Neumann conditions along the bottom boundary Γ_{base}.

The output of interest is the average temperature over the base made up by Γ_{base}. Note that we consider a non-dimensional formulation in which the number of physical parameters has been kept to a minimum.

The strong formulation of this parametrized problem is stated as: for some parameter value $\mu \in \mathbb{P}$, find $u(\mu)$ such that

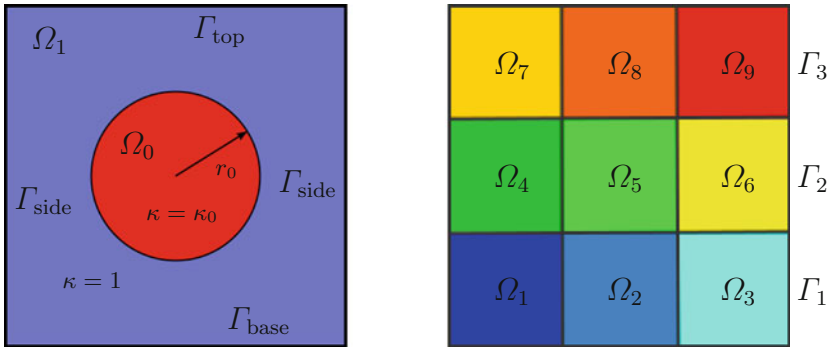

Fig. 2.1 Geometrical set-up (*left*) of the heat conductivity problem, illustrative Example 1, and (*right*) the elasticity problem, illustrative Example 2

$$\begin{cases} \nabla \cdot \kappa_\mu \nabla u(\mu) = 0 & \text{in } \Omega, \\ u(\mu) = 0 & \text{on } \Gamma_{\text{top}}, \\ \kappa_\mu \nabla u(\mu) \cdot n = 0 & \text{on } \Gamma_{\text{side}}, \\ \kappa_\mu \nabla u(\mu) \cdot n = \mu_{[2]} & \text{on } \Gamma_{\text{base}}. \end{cases}$$

The output of interest is given as

$$s(\mu) = \ell(u(\mu); \mu) = \mu_{[2]} \int_{\Gamma_{\text{base}}} u(\mu).$$

We recall that the function space associated with this set of boundary conditions is given by $\mathbb{V} = \{v \in H^1(\Omega) \mid v|_{\Gamma_{\text{top}}} = 0\}$: the Dirichlet boundary conditions are essential; the Neumann boundary conditions are natural.

The weak parametrized formulation then reads: for some parameter $\mu \in \mathbb{P}$, find $u(\mu) \in \mathbb{V}$ such that

$$a(u(\mu), v; \mu) = f(v; \mu) \qquad \forall\, v \in \mathbb{V},$$

with

$$a(w, v; \mu) = \int_\Omega \kappa_\mu \nabla w \cdot \nabla v \qquad \text{and} \qquad f(v; \mu) = \mu_{[2]} \int_{\Gamma_{\text{base}}} v,$$

for all $v, w \in \mathbb{V}$. We endow the space $\mathbb{V}$ with the scalar product

$$(v, w)_{\mathbb{V}} = a(v, w; \bar{\mu}) = \int_\Omega \nabla w \cdot \nabla v, \qquad \forall\, w, v \in \mathbb{V},$$

for $\bar{\mu} = (\bar{\mu}_{[1]}, \bar{\mu}_{[2]})$ such that $\bar{\mu}_{[1]} = 1$. For the problem to be well-posed, we assume that $\mu_{[1]}^{\min} > 0$ so that $\kappa_\mu \geq \min(1, \mu_{[1]}^{\min}) > 0$ and coercivity of the bilinear form a follows. Further, continuity of the forms a and f can be easily obtained using the Cauchy-Schwarz inequality; and linearity and bilinearity can be easily verified as well. We can therefore apply the Lax-Milgram theorem to guarantee existence and uniqueness of the solution $u(\mu) \in \mathbb{V}$ for any parameter value $\mu \in \mathbb{P}$.

A conforming discretization introduces a finite-dimensional subspace $\mathbb{V}_\delta \subset \mathbb{V}$, for instance a standard finite element space. Following the Galerkin approach we obtain the following discrete problem: for some parameter $\mu \in \mathbb{P}$, find $u_\delta(\mu) \in \mathbb{V}_\delta$ such that

$$a(u_\delta(\mu), v_\delta; \mu) = f(v_\delta : \mu) \qquad \forall\, v_\delta \in \mathbb{V}_\delta.$$

In the following illustration, the finite element method, employing piece-wise linear elements, has been chosen as the truth model. The mesh is illustrated in Fig. 2.2 (left) featuring 812 elements. The chosen ranges for the parameters are

$$\mu = (\mu_{[1]}, \mu_{[2]}) \in \mathbb{P} = [0.1, 10] \times [-1, 1].$$

In Fig. 2.3, four representative solutions—snapshots—are depicted for different values of the parameters.

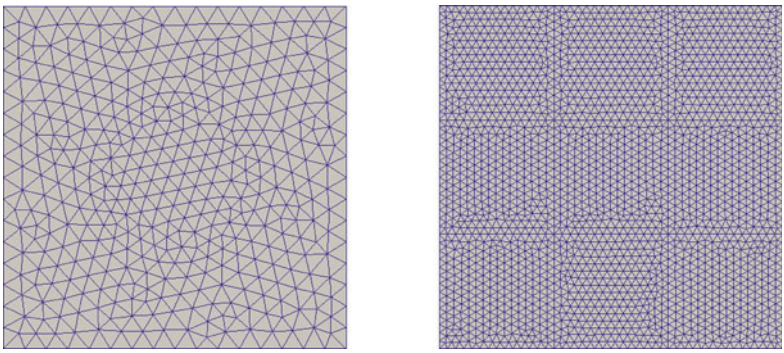

Fig. 2.2 Finite element mesh for Example 1 (*left*) and Example 2 (*right*)

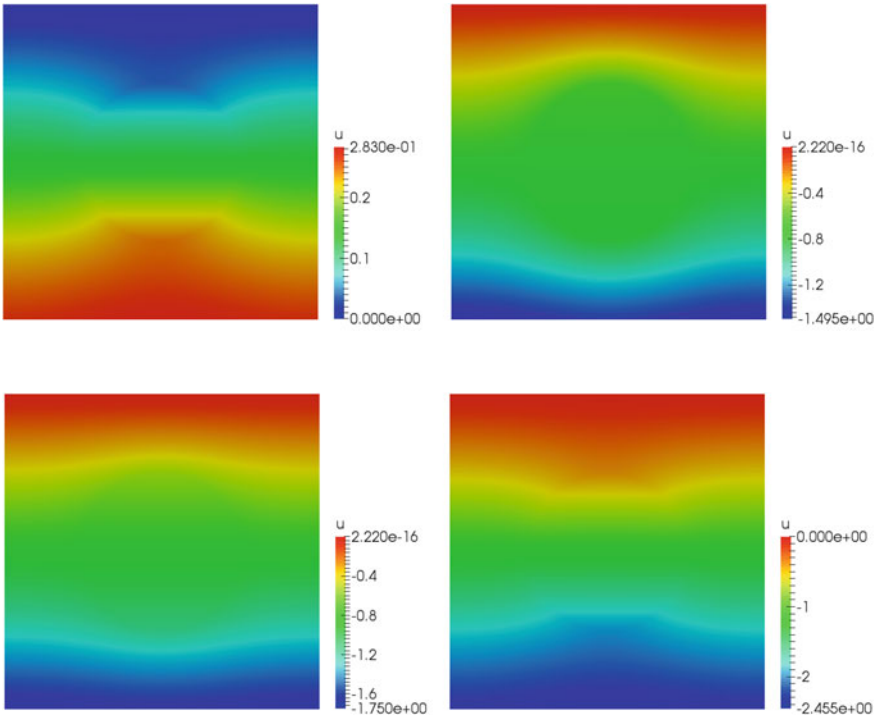

Fig. 2.3 Four different representative solutions for the parametrized conductivity problem (Example 1)

2.3.2 Illustrative Example 2: Linear Elasticity Part 1

We consider a linear elasticity example [3, 4] in the two-dimensional domain $\Omega = (0, 1) \times (0, 1)$, shown in Fig. 2.1 with 9 mini-blocks Ω_i, where the Young's modulus

on each mini-block is denoted by E_i and the Poisson's ratio is set to $\nu = 0.30$. The outward pointing unit normal on $\partial\Omega$ is denoted by n.

We consider $P = 11$ parameters: the 8 Young's moduli with respect to the reference value $E = E_9 = 10$ set in Ω_9 and the 3 horizontal traction/compression load conditions at the right border of the elastic block. Our parameter vector is thus given by $\mu = (\mu_{[1]}, \ldots, \mu_{[P]})$ and we choose for our parameter domain $\mathbb{P} = [\mu_{[1]}^{\min}, \mu_{[1]}^{\max}] \times \cdots \times [\mu_{[P]}^{\min}, \mu_{[P]}^{\max}]$ where

$$\begin{aligned} [\mu_{[p]}^{\min}, \mu_{[p]}^{\max}] &= [1, 100], \qquad p = 1, \ldots, 8, \\ [\mu_{[p]}^{\min}, \mu_{[p]}^{\max}] &= [-1, 1], \qquad p = 9, \ldots, 11. \end{aligned}$$

The local Young's moduli are given by $E_i = \mu_{[i]} E$.

Our vector field variable $u(\mu) = (u_1(\mu), u_2(\mu))$ is the displacement of the elastic block under the applied load: the displacement satisfies the plane-strain linear elasticity equations in Ω in combination with the following boundary conditions: homogeneous Neumann (load-free) conditions are imposed on the top and bottom boundaries Γ_{top} and Γ_{base} of the block; homogeneous Dirichlet (displacement) conditions on the left boundary Γ_{left} (the structure is clamped); and parametrized inhomogeneous Neumann conditions on the right boundary $\Gamma_{\text{right}} = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ with zero shear. The non-trivial (inhomogeneous) boundary conditions are summarized as follows

$$\begin{cases} n \cdot u = \mu_{[9]} & \text{on } \Gamma_1, \\ n \cdot u = \mu_{[10]} & \text{on } \Gamma_2, \\ n \cdot u = \mu_{[11]} & \text{on } \Gamma_3, \end{cases}$$

representing traction loads. The output of interest $s(\mu)$ is the integrated horizontal (traction/compression) displacement over the full loaded boundary Γ_{right}, given by

$$s(\mu) = \int_{\Gamma_1 \cup \Gamma_2 \cup \Gamma_3} u_1(\mu).$$

This corresponds to the compliant situation as we will see later.

The function space associated with this set of boundary conditions is given by

$$\mathbb{V} = \left\{ v \in (H^1(\Omega))^2 \;\middle|\; v|_{\Gamma_{\text{left}}} = 0 \right\}.$$

Hence, the Dirichlet interface and boundary conditions are essential and the Neumann interface and boundary conditions are natural. We then define the load (and also output) functional

$$f_i(v; \mu) = \int_{\Gamma_i} v_1, \qquad \forall\, v = (v_1, v_2) \in \mathbb{V} \quad \text{and} \quad i = 1, 2, 3,$$

such that

$$f(v;\mu) = \sum_{i=1}^{3} \mu_{[i+8]} f_i(v;\mu).$$

The bilinear form associated with the left-hand-side of the problem is given by:

$$a(w,v;\mu) = \sum_{p=1}^{8} \mu_{[p]} E \int_{\Omega_p} \frac{\partial v_i}{\partial x_j} C_{ijkl} \frac{\partial w_k}{\partial x_l} + E \int_{\Omega_9} \frac{\partial v_i}{\partial x_j} C_{ijkl} \frac{\partial w_k}{\partial x_l},$$

where $\mu_{[p]}$ is the ratio between the Young modulus in Ω_p and Ω_9 and Einstein's summation convention is used for the indices i, j, k and l.

For our isotropic material, the elasticity tensor is given by

$$C_{ijkl} = \lambda^1 \delta_{ij}\delta_{kl} + \lambda^2 \left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right),$$

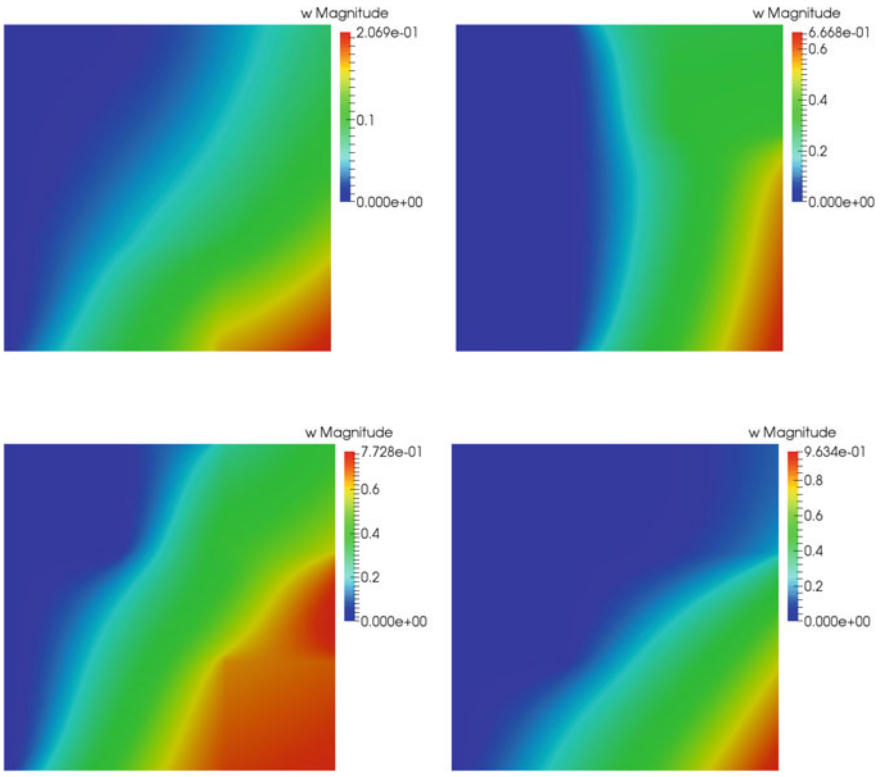

Fig. 2.4 Four representative solutions for the laterally loaded (traction/compression) elastic block (Example 2)

where

$$\lambda^1 = \frac{\nu}{(1+\nu)(1-2\nu)} \quad \text{and} \quad \lambda^2 = \frac{1}{2(1+\nu)} ,$$

are the Lamé constants for plane strain. We recall that the Poisson's ratio is set to $\nu = 0.30$. The weak form is then given by (2.1)–(2.2). The inner product is specified by (2.3),

$$(v, w)_{\mathbb{V}} = a(v, w; \bar{\mu}) = E \int_{\Omega} \frac{\partial v_i}{\partial x_j} C_{ijkl} \frac{\partial w_k}{\partial x_l}, \qquad \forall\, w, v \in \mathbb{V},$$

for some $\bar{\mu} \in \mathbb{P}$ satisfying $\bar{\mu}_{[p]} = 1$, for all $1 \le p \le 8$.

We can now readily verify our hypotheses. First, it is standard to confirm that f is indeed bounded. Second, we readily confirm by inspection that a is symmetric, and we further verify by application of the Korn inequality [5] and the Cauchy-Schwarz inequality that a is coercive and continuous, respectively.

The finite element method, employing piece-wise elements, has been chosen as the truth model. In Fig. 2.2 (right) the mesh is represented, featuring 4,152 elements. In Fig. 2.4, four representative solutions are illustrated.

References

1. A. Quarteroni, A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer Series in Computational Mathematics, vol. 23 (Springer, Berlin, 1994)
2. A. Mills, *Basic Heat Transfer* (Prentice-Hall, Upper Saddle River, 1999)
3. R. Dautray, J. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology*, vol. 2 (Springer Verlag, 1988)
4. R. Leis, *Initial Boundary Value Problems in Mathematical Physics* (Courier Corporation, 2013)
5. G.D.J. Lions, *Inequalities in Mechanics and Physics* (Springer, 1976)

Chapter 3
Reduced Basis Methods

With target applications characterized by computationally intensive parametrized problems that require repeated evaluation, it is clear that we need to seek alternatives to simply solving the full problem many times. This is exactly where reduced models have its place and we are now ready to dive deeper into a discussion of central elements of the certified reduced basis method.

While this initial discussion is the main topic of this chapter we will quickly observe that several elements have to come together to fully realize the method. Some of these elements are not discussed in detail until in later chapters, and are simply stated as assumptions within this chapter.

When introducing reduced models it is inevitable to familiarize the reader with the notion of a solution manifold, that is the set of all solutions to the parametrized problem under variation of the parameter. The final goal of RB methods is to approximate any member of this solution manifold with a low number of, say N, basis functions. This set of basis functions is denoted as the reduced basis.

The reduced basis method is based on a two stage procedure, comprising an offline and an online stage. During the potentially very costly offline stage, one empirically explores the solution manifold to construct a reduced basis that approximates any member of the solution manifold to within a prescribed accuracy. As this involves the solution of at least N truth problems, each with N_δ degrees of freedom, the cost can be high. This results in the identification of an linear N-dimensional reduced basis. The online stage consists of a Galerkin projection, using the parametrized bilinear form $a(\cdot, \cdot; \mu)$ with a varying parameter value $\mu \in \mathbb{P}$, onto the space spanned by the reduced basis. During this stage, one can explore the parameter space at a substantially reduced cost, ideally at a cost independent of N_δ.

This offline/online separation is beneficial in at least two different scenarios. First, if the reduced basis approximation needs to be evaluated for many parameter values, a direct and repeated evaluation of the truth could be prohibitive. Typical examples can be found in areas of optimization, design, uncertainty quantification, query of

J.S. Hesthaven et al., *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, SpringerBriefs in Mathematics,
DOI 10.1007/978-3-319-22470-1_3

simulation based databases etc. Secondly, the online procedure can be embedded in a computer environment that has only limited computational power and memory to allow rapid online query of the response of an otherwise complex system for control, visualization and analysis using a deployed device.

3.1 The Solution Manifold and the Reduced Basis Approximation

Our primary interest is the solution of the parametric exact problem (2.1) given as: find $u(\mu) \in \mathbb{V}$ such that

$$a(u(\mu), v; \mu) = f(v; \mu), \quad \forall v \in \mathbb{V}.$$

We shall refer to this as the exact solution. Let us introduce the notion of solution manifold comprising of all solutions of the parametric problem under variation of the parameters, i.e.,

$$\mathcal{M} = \{u(\mu) \mid \mu \in \mathbb{P}\} \subset \mathbb{V},$$

where each $u(\mu) \in \mathbb{V}$ corresponds to the solution of the exact problem.

In many cases of interest, the exact solution is not available in an analytic or otherwise simple manner, and we seek an approximate solution by seeking $u_\delta(\mu) \in \mathbb{V}_\delta$ such that

$$a(u_\delta(\mu), v_\delta; \mu) = f(v_\delta; \mu), \quad \forall v_\delta \in \mathbb{V}_\delta, \tag{3.1}$$

referred to as the truth. Throughout the subsequent discussion we assume that $\|u(\mu) - u_\delta(\mu)\|_{\mathbb{V}}$ can be made arbitrarily small for any given parameter value, $\mu \in \mathbb{P}$. This simply states that we assume that a computational model is available to solve the truth problem, thus approximate the exact solution at any required accuracy. However, we shall not specify the details of this other than we will require it to be based on variational principles. This accuracy requirement also implies that the computational cost of evaluating the truth model may be very high and depend directly on $N_\delta = \dim(\mathbb{V}_\delta)$.

Following the definition for the continuous problem, we also define the discrete version of the solution manifold

$$\mathcal{M}_\delta = \{u_\delta(\mu) \mid \mu \in \mathbb{P}\} \subset \mathbb{V}_\delta, \tag{3.2}$$

where each $u_\delta(\mu) \in \mathbb{V}_\delta$ corresponds to the solution of the parametric truth problem (3.1).

A central assumption in the development of any reduced model is that the solution manifold is of low dimension, i.e., that the span of a low number of appropriately chosen basis functions represents the solution manifold with a small error. We shall call these basis functions the reduced basis and it will allow us to represent the truth

solution, $u_\delta(\mu)$ based on an N-dimensional subspace $\mathbb{V}_{\mathtt{rb}}$ of $\mathbb{V}_\delta$. Let us initially assume that an N-dimensional reduced basis, denoted as $\{\xi_n\}_{n=1}^N \subset \mathbb{V}_\delta$, is available, then, the associated reduced basis space is given by

$$\mathbb{V}_{\mathtt{rb}} = \text{span}\{\xi_1, \ldots, \xi_N\} \subset \mathbb{V}_\delta.$$

The assumption of the low dimensionality of the solution manifold implies that $N \ll N_\delta$. Given the N-dimensional reduced basis space $\mathbb{V}_{\mathtt{rb}}$, the reduced basis approximation is sought as: for any given $\mu \in \mathbb{P}$, find $u_{\mathtt{rb}}(\mu) \in \mathbb{V}_{\mathtt{rb}}$ s.t.

$$a(u_{\mathtt{rb}}(\mu), v_{\mathtt{rb}}; \mu) = f(v_{\mathtt{rb}}; \mu), \quad \forall v_{\mathtt{rb}} \in \mathbb{V}_{\mathtt{rb}}, \tag{3.3}$$

and evaluate

$$s_{\mathtt{rb}}(\mu) = f(u_{\mathtt{rb}}(\mu); \mu), \tag{3.4}$$

since we assume to be in the compliant case. Otherwise it would be $s_{\mathtt{rb}}(\mu) = \ell(u_{\mathtt{rb}}(\mu); \mu)$. Since the basis functions of $\mathbb{V}_{\mathtt{rb}}$ are given by $\xi_1, \ldots, \xi_N$, we can represent $u_{\mathtt{rb}}(\mu)$ by $u_{\mathtt{rb}}(\mu) = \sum_{n=1}^N (\mathtt{u}_{\mathtt{rb}}^\mu)_n \, \xi_n$ where $\{(\mathtt{u}_{\mathtt{rb}}^\mu)_n\}_{n=1}^N$ denote the coefficients of the reduced basis approximation. In the linear algebra box The reduced basis approximation we further explain the reduced basis approximation at the level of the involved linear algebra.

Linear algebra box: The reduced basis approximation

Let $\{\xi_n\}_{n=1}^N$ denote the reduced basis and define the matrix $\mathbf{B} \in \mathbb{R}^{N_\delta \times N}$ such that

$$\xi_n = \sum_{i=1}^{N_\delta} \mathbf{B}_{in} \, \varphi_i,$$

i.e., the n-th column of $\mathbf{B}$ denotes the coefficients when the n-th basis function ξ_n is expressed in terms of the basis functions $\{\varphi_i\}_{i=1}^{N_\delta}$. Then, the reduced basis solution matrix $\mathbf{A}_{\mathtt{rb}}^\mu \in \mathbb{R}^{N \times N}$ and right hand side $\mathtt{f}_{\mathtt{rb}}^\mu \in \mathbb{R}^N$ defined by

$$(\mathbf{A}_{\mathtt{rb}}^\mu)_{mn} = a(\xi_n, \xi_m; \mu), \qquad \text{and} \qquad (\mathtt{f}_{\mathtt{rb}}^\mu)_m = f(\xi_m; \mu), \qquad 1 \le n, m \le N,$$

can be computed by

$$\mathbf{A}_{\mathtt{rb}}^\mu = \mathbf{B}^T \, \mathbf{A}_\delta^\mu \, \mathbf{B}, \qquad \text{and} \qquad \mathtt{f}_{\mathtt{rb}}^\mu = \mathbf{B}^T \, \mathtt{f}_\delta^\mu.$$

The reduced basis approximation $u_{\mathtt{rb}}(\mu) = \sum_{n=1}^N (\mathtt{u}_{\mathtt{rb}}^\mu)_n \, \xi_n$ is obtained by solving the linear system

$$\mathbf{A}_{\mathtt{rb}}^\mu \, \mathtt{u}_{\mathtt{rb}}^\mu = \mathtt{f}_{\mathtt{rb}}^\mu,$$

and the output of interest evaluated as $s_{\mathtt{rb}}(\mu) = (\mathtt{u}_{\mathtt{rb}}^\mu)^T \, \mathtt{f}_{\mathtt{rb}}^\mu$.

Within this setting, we can now begin to consider a number of central questions, related to the computational efficiency and accuracy of the reduced model as a representation of the truth approximation across the parameter space. If we begin by assuming that the reduced basis approximation is available, questions of accuracy can be addressed by considering the simple statement

$$\|u(\mu) - u_{\mathrm{rb}}(\mu)\|_{\mathbb{V}} \leq \|u(\mu) - u_{\delta}(\mu)\|_{\mathbb{V}} + \|u_{\delta}(\mu) - u_{\mathrm{rb}}(\mu)\|_{\mathbb{V}}.$$

For a given parameter value, $\mu \in \mathbb{P}$, we assume that the accuracy of the first part on the right hand side can be controlled by the accuracy of the truth approximation as assumed above. This assumption also extends to the solution manifold for which we assume that $\mathcal{M}_{\delta}$ approximates $\mathcal{M}$ arbitrarily well.

Hence, accuracy of the reduced basis approximation is guaranteed if we can estimate the accuracy by which the reduced basis approximation approximates the truth for a given parameter value. This error estimation is a key element of the process and will be discussed in detail in Chap. 4. The computational cost on the other hand is dominated by the cost of effectively evaluating (3.3) for a new parameter value and, naturally, the compactness of the reduced basis, N.

While the former challenge is primarily an algorithmic challenge, the latter is a question that clearly has a problem specific answer, i.e., some problems will allow a very efficient reduced basis representation while other problems will escape this entirely. To get a handle on this, it is instructive to introduce the notion of the Kolmogorov N-width. Let us first define

$$E(\mathcal{M}_{\delta}, \mathbb{V}_{\mathrm{rb}}) = \sup_{u_{\delta} \in \mathcal{M}_{\delta}} \inf_{v_{\mathrm{rb}} \in \mathbb{V}_{\mathrm{rb}}} \|u_{\delta} - v_{\mathrm{rb}}\|_{\mathbb{V}}.$$

The Kolmogorov N-width of $\mathcal{M}_{\delta}$ in $\mathbb{V}_{\mathrm{rb}}$ is then defined as

$$d_N(\mathcal{M}_{\delta}) = \inf_{\mathbb{V}_{\mathrm{rb}}} \sup_{u_{\delta} \in \mathcal{M}_{\delta}} \inf_{v_{\mathrm{rb}} \in \mathbb{V}_{\mathrm{rb}}} \|u_{\delta} - v_{\mathrm{rb}}\|_{\mathbb{V}}, \tag{3.5}$$

where the first infimum is taken over all N-dimensional subspaces $\mathbb{V}_{\mathrm{rb}}$ of $\mathbb{V}$. Hence, the N-width measures how well $\mathcal{M}_{\delta}$ can be approximated by some N-dimensional subspace $\mathbb{V}_{\mathrm{rb}}$. If, indeed, the N-width decays rapidly as N increases, it suggests that the solution manifold can be well approximated by a small reduced basis, yielding a compact and efficient approximation across the entire parameter space. In our case of interest, we can think of the regularity of the solution in parameter space, where in some case, the Kolmogorov N-width may even decay exponentially, i.e., $d_N(\mathcal{M}_{\delta}, \mathbb{V}_{\mathrm{rb}}) \leq Ce^{-cN}$. A challenge we shall discuss later in this chapter is how to find this low-dimensional subspace and how compact we can expect it to be provided we have some information about the Kolmogorov N-width for a particular problem.

Following the discussion just above, we assume that the reduced basis is constructed to ensure good approximation properties across the parameter space. This is measured through the Kolmogorov N-width by its capability to approximate any member of $\mathcal{M}_{\delta}$,

$$\sup_{u_\delta \in \mathcal{M}_\delta} \inf_{v_{\mathrm{rb}} \in \mathbb{V}_{\mathrm{rb}}} \|u_\delta - v_{\mathrm{rb}}\|_{\mathbb{V}}.$$

One can relax the supremum-norm over $\mathcal{M}_\delta$ to obtain a different (least-square) notion of optimality as

$$\sqrt{\int_{\mu \in \mathbb{P}} \inf_{v_{\mathrm{rb}} \in \mathbb{V}_{\mathrm{rb}}} \|u_\delta(\mu) - v_{\mathrm{rb}}\|_{\mathbb{V}}^2 \, d\mu}. \tag{3.6}$$

In addition to these considerations about approximating a solution manifold by an N-dimensional space in the best approximation, observe that applying Cea's lemma for a given approximation space $\mathbb{V}_{\mathrm{rb}}$ and a given parameter value $\mu \in \mathbb{P}$ connects the best approximation error with the reduced approximation:

$$\|u(\mu) - u_{\mathrm{rb}}(\mu)\|_{\mathbb{V}} \leq \left(1 + \frac{\gamma(\mu)}{\alpha(\mu)}\right) \inf_{v_{\mathrm{rb}} \in \mathbb{V}_{\mathrm{rb}}} \|u(\mu) - v_{\mathrm{rb}}\|_{\mathbb{V}}. \tag{3.7}$$

Thus, the quality of the reduced basis approximation, based on a Galerkin projection, depends, as the truth approximation, on the coercivity and continuity constants of the bilinear form, both of which are problem-dependent. Furthermore, the quality of also depends on the ability of the reduced basis space to approximate any member of the solution manifold.

In the following, we are going to address the following two fundamental questions:

- How do we generate accurate reduced basis spaces during the offline stage?
- How do we recover the reduced basis solution efficiently during the online stage?

3.2 Reduced Basis Space Generation

While there are several strategies for generating reduced basis spaces, we shall focus on the proper orthogonal decomposition (POD) and the greedy construction in the following. In both cases, one begins by introducing a discrete and finite-dimensional point-set $\mathbb{P}_h \subset \mathbb{P}$ in parameter domain, e.g., it can consist of a regular lattice or a randomly generated point-set intersecting with $\mathbb{P}$. We can then introduce the following set

$$\mathcal{M}_\delta(\mathbb{P}_h) = \{u_\delta(\mu) \mid \mu \in \mathbb{P}_h\}$$

of cardinality $M = |\mathbb{P}_h|$. Of course, it holds that $\mathcal{M}_\delta(\mathbb{P}_h) \subset \mathcal{M}_\delta$ as $\mathbb{P}_h \subset \mathbb{P}$ but if $\mathbb{P}_h$ is fine enough, $\mathcal{M}_\delta(\mathbb{P}_h)$ is also a good representation of $\mathcal{M}_\delta$.

3.2.1 Proper Orthogonal Decomposition (POD)

Proper Orthogonal Decomposition (POD) is an explore-and-compress strategy in which one samples the parameter space, compute the corresponding truth solutions at all sample points and, following compression, retains only the essential information. The N-dimensional POD-space is the space that minimizes the quantity

$$\sqrt{\frac{1}{M}\sum_{\mu\in\mathbb{P}_h}\inf_{v_{\mathtt{rb}}\in\mathbb{V}_{\mathtt{rb}}}\|u_\delta(\mu)-v_{\mathtt{rb}}\|_{\mathbb{V}}^2} \tag{3.8}$$

over all N-dimensional subspaces $\mathbb{V}_{\mathtt{rb}}$ of the span $\mathbb{V}_{\mathcal{M}} = \operatorname{span}\{u_\delta(\mu)\,|\,\mu\in\mathbb{P}_h\}$ of the elements of $\mathcal{M}_\delta(\mathbb{P}_h)$. It is a discrete version of the measure considered in (3.6) using $\mathcal{M}_\delta(\mathbb{P}_h)$ instead of $\mathcal{M}_\delta$.

We introduce an ordering $\mu_1,\dots,\mu_M$ of the values in $\mathbb{P}_h$, hence inducing an ordering $u_\delta(\mu_1),\dots,u_\delta(\mu_M)$ of the elements of $\mathcal{M}_\delta(\mathbb{P}_h)$. For the sake of a short notation, we denote $\psi_m = u_\delta(\mu_m)$ for all $m=1,\dots,M$ in the following. To construct the POD-space, let us define the symmetric and linear operator $\mathtt{C}:\mathbb{V}_{\mathcal{M}}\to\mathbb{V}_{\mathcal{M}}$ defined by

$$\mathtt{C}(v_\delta) = \frac{1}{M}\sum_{m=1}^{M}(v_\delta,\psi_m)_{\mathbb{V}}\,\psi_m,\qquad v_\delta\in\mathbb{V}_{\mathcal{M}},$$

and consider the eigenvalue-eigenfunction pairs $(\lambda_n,\xi_n)\in\mathbb{R}\times\mathbb{V}_{\mathcal{M}}$ of the operator $\mathtt{C}$ with normalization constraint $\|\xi_n\|_{\mathbb{V}}=1$ satisfying

$$(\mathtt{C}(\xi_n),\psi_m)_{\mathbb{V}} = \lambda_n(\xi_n,\psi_m)_{\mathbb{V}},\quad 1\le m\le M. \tag{3.9}$$

Here we assume that the eigenvalues are sorted in descending order $\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_M$. The orthogonal POD basis functions are given by the eigenfunctions $\xi_1,\dots,\xi_M$ and they span $\mathbb{V}_{\mathcal{M}}$. If one truncates the basis and only considers the first N functions $\xi_1,\dots,\xi_N$, they span the N-dimensional space $\mathbb{V}_{\mathtt{POD}}$ that satisfies the optimality criterion (3.8). Further, the projection $P_N:\mathbb{V}\to\mathbb{V}_{\mathtt{POD}}$ for arbitrary functions in $\mathbb{V}$ onto $\mathbb{V}_{\mathtt{POD}}$, defined as

$$(P_N[f],\xi_n)_{\mathbb{V}} = (f,\xi_n)_{\mathbb{V}},\quad 1\le n\le N,$$

is given as

$$P_N[f] = \sum_{n=1}^{N}(f,\xi_n)_{\mathbb{V}}\,\xi_n.$$

In particular, if the projection is applied to all elements in $\mathcal{M}_\delta(\mathbb{P}_h)$ it satisfies the following error estimate

$$\sqrt{\frac{1}{M}\sum_{m=1}^{M}\|\psi_m - P_N[\psi_m]\|_{\mathbb{V}}^2} = \sqrt{\sum_{m=N+1}^{M}\lambda_m}.$$

The computational aspects of the POD procedure is discussed in the linear algebra box Proper Orthogonal Decomposition (POD), highlighting the reliance on basic tools of linear algebra.

Linear algebra box: Proper Orthogonal Decomposition (POD)

Denote $\psi_m = u_\delta(\mu_m)$ for $m = 1, \ldots, M$ and let us construct the correlation matrix $\mathbf{C} \in \mathbb{R}^{M\times M}$ by

$$\mathbf{C}_{mq} = \tfrac{1}{M}(\psi_m, \psi_q)_{\mathbb{V}}, \quad 1 \le m, q \le M.$$

Then, solve for the N largest eigenvalue-eigenvector pairs $(\lambda_n, \mathrm{v}_n)$ of the matrix $\mathbf{C}$ with $\|\mathrm{v}_n\|_{\ell^2(\mathbb{R}^M)} = 1$ such that

$$\mathbf{C}\,\mathrm{v}_n = \lambda_n \mathrm{v}_n, \quad 1 \le n \le N,$$

which is equivalent to (3.9). With the eigenvalues sorted in descending order $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_N$ the orthogonal POD basis functions $\{\xi_1, \ldots, \xi_N\}$ span the POD-space $\mathbb{V}_{\mathrm{POD}} = \mathrm{span}\{\xi_1, \ldots, \xi_N\}$ and are given by the linear combinations

$$\xi_n(x) = \tfrac{1}{\sqrt{M}}\sum_{m=1}^{M}(\mathrm{v}_n)_m\,\psi_m(x), \quad 1 \le n \le N,$$

where $(\mathrm{v}_n)_m$ denotes the m-th coefficient of the eigenvector $\mathrm{v}_n \in \mathbb{R}^M$.

Remark 3.1 (*Relation with singular value decomposition (SVD)*) In the simplified case where the scalar product $(\psi_m, \psi_q)_{\mathbb{V}}$ is replaced by the simple ℓ^2-scalar product of the degrees of freedom of $\psi_m = u_\delta(\mu_m) = \sum_{i=1}^{N_\delta}(\mathrm{u}_\delta^{\mu_m})_i\,\varphi_i$ and $\psi_q = u_\delta(\mu_q) = \sum_{i=1}^{N_\delta}(\mathrm{u}_\delta^{\mu_q})_i\,\varphi_i$, i.e., by

$$\sum_{i=1}^{N_\delta}(\mathrm{u}_\delta^{\mu_m})_i\,(\mathrm{u}_\delta^{\mu_q})_i,$$

the correlation matrix becomes $\mathbf{C} = \frac{1}{M}\mathbf{U}_\delta^T\mathbf{U}_\delta$ where $\mathbf{U}_\delta \in \mathbb{R}^{N_\delta\times M}$ denotes the matrix of the column-wise vectors $\mathrm{u}_\delta^{\mu_m}$ for $m = 1, \ldots, M$. In this case, the eigenvalues of $\mathbf{C}$ correspond to the square of the singular values of $\frac{1}{\sqrt{M}}\mathbf{U}_\delta$.

We note that the orthonormality of the eigenvectors in the sense of $\ell^2(\mathbb{R}^M)$ implies the following orthogonality relationship of the eigenfunctions

$$(\xi_m, \xi_q)_{\mathbb{V}} = M\lambda_n\delta_{mq}, \quad 1 \le m, q \le M$$

where δ_{mq} denotes the Kronecker delta.

While the construction of the POD-basis results in a basis that is optimal in an ℓ^2-sense over the parameter space, the cost of the procedure is potentiality very high. To ensure a reduced basis of sufficient accuracy, a potentially large number M of

truth solutions may be required and, worse yet, a proper choice of M is not known or predictable for a general problem. This implies that one often has to choose $M \gg N$, leading to a very substantial computational overhead by having to compute a large number of truth solutions only to discover that the majority of these solutions do not contribute to the reduced basis. Furthermore, for M and N_δ being large, the cost of computing the reduced basis itself, requiring the solution of a large dense eigenvalue problem, scales like $\mathcal{O}(N N_\delta{}^2)$. While this construction is straightforward and display desirable optimality properties, the lack of an error estimator beyond the magnitude of the largest ignored eigenvalue and, most importantly, the high computational offline cost suggests that we seek an alternative approach.

3.2.2 Greedy Basis Generation

In contrast to the generation of the reduce basis using the proper orthogonal decomposition (POD), the greedy generation of the reduced basis space is an iterative procedure where at each iteration one new basis function is added and the overall precision of the basis set is improved. It requires one truth solution to be computed per iteration and a total of N truth solutions to generate the N-dimensional reduced basis space.

An essential ingredient of the greedy algorithm is the availability of an error $\eta(\mu)$ which predicts the error due to the model order reduction, i.e., it provides an estimate of the error induced by replacing $\mathbb{V}_\delta$ by the reduced basis space $\mathbb{V}_{\mathtt{rb}}$ in the variational formulation. We shall postpone the discussion of how to develop such estimators to Chap. 4 and simply assume here that one is available satisfying

$$\|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_\mu \leq \eta(\mu),$$

for all $\mu \in \mathbb{P}$. Here $u_\delta(\mu)$ is a solution of (3.1) and $u_{\mathtt{rb}}(\mu)$ is solution of (3.3) for a certain reduced basis space $\mathbb{V}_{\mathtt{rb}}$. Alternatively, a different norm, i.e., the parameter-independent norm $\|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}$, or even the measure of the output functional $|s_\delta(\mu) - s_{\mathtt{rb}}(\mu)|$ can be chosen. But in all cases, $\eta(\mu)$ consist of a strict upper bound of the corresponding error-quantity. As already mentioned, details are postponed to Chap. 4.

During this iterative basis selection process and if at the n-th step a n-dimensional reduced basis space $\mathbb{V}_{\mathtt{rb}}$ is given, the next basis function is the one that maximizes the estimated model order reduction error given the n-dimensional space $\mathbb{V}_{\mathtt{rb}}$ over $\mathbb{P}$. That is, we select

$$\mu_{n+1} = \arg\max_{\mu \in \mathbb{P}} \eta(\mu), \tag{3.10}$$

and compute $u_\delta(\mu_{n+1})$ to enrich the reduced basis space as $\mathbb{V}_{\mathtt{rb}} = \mathrm{span}\{u_\delta\}(\mu_1), \ldots, u_\delta(\mu_{n+1})$. This is repeated until the maximal estimated error is below a required error

tolerance. The greedy algorithm always selects the next parameter sample point as the one for which the model error is the maximum as estimated by $\eta(\mu)$. This yields a basis that aims to be optimal in the maximum norm over $\mathbb{P}$ rather than L^2 for the POD basis.

Computing the maximum in (3.10) over the entire parameter space $\mathbb{P}$ is impossible and, as for the POD approach, we introduce a finite point-set $\mathbb{P}_h$. However, since a point in $\mathbb{P}_h$ only requires the evaluation of the error estimator and not a truth solution, the cost per point is small and $\mathbb{P}_h$ can therefore be considerably denser than the one used in the construction of the POD basis, provided the error estimator can be evaluated efficiently. Furthermore, one can utilize that the evaluation of the error estimator is embarrassingly parallel to further accelerate this part of the offline computation. The computational aspects of the greedy basis generation is discussed in the algorithm box The greedy algorithm, highlighting the importance of the error estimator in this development.

Provided this approach leads to a basis of sufficient accuracy and compactness commensurate with the Kolmogorov N-width of the problem, its advantages over the POD basis generation are clear. Not only is the need for solving a potentially very large eigenproblem eliminated but we also dramatically reduce the total cost of the offline computation by only computing the N truth solutions in contrast to the M solutions needed for the POD basis generation, where $M \gg N$ in almost all cases. Note also that the sequence of approximation spaces is hierarchical. Hence, if the N-dimensional reduced basis space is not sufficiently accurate, one can enrich it by adding n additional modes. This which results in exactly the same space as having build the reduced basis space with $N + n$ basis functions.

Algorithm: The greedy algorithm

Input: $\texttt{tol}$, μ_1 and $n = 1$.
Output: A reduced basis space $\mathbb{V}_{\texttt{rb}}$.

1. Compute $u_\delta(\mu_n)$ solution to (3.1) for μ_n and set $\mathbb{V}_{\texttt{rb}} = \text{span}\{u_\delta(\mu_1), \ldots, u_\delta(\mu_n)\}$.
2. For each $\mu \in \mathbb{P}_h$
   a. Compute the reduced basis approximation $u_{\texttt{rb}}(\mu) \in \mathbb{V}_{\texttt{rb}}$ defined by (3.3) for μ.
   b. Evaluate the error estimator $\eta(\mu)$.
3. Choose $\mu_{n+1} = \arg\max_{\mu \in \mathbb{P}_h} \eta(\mu)$.
4. If $\eta(\mu_{n+1}) > \texttt{tol}$, then set $n := n + 1$ and **go to** 1., otherwise **terminate**.

Let us elaborate a little further on the greedy algorithm. We consider a general family $F = \{f(\mu) \mid \mu \in \mathbb{P}\}$ of parametrized functions, $f(\mu) : \Omega \to \mathbb{R}$, for which we find an approximation space using a greedy approach to iteratively select the basis functions as

$$f_{n+1} = \arg\max_{\mu \in \mathbb{P}} \| f(\mu) - P_n f(\mu) \|_{\mathbb{V}},$$

and where $P_n f$ is the orthogonal projection onto $F_n = \text{span}\{f_1, \ldots, f_n\}$, we have the following convergence result for the basic greedy approximation [1] (see also [2] for a generalization to Banach spaces and [3] for the first but less sharp estimates):

Theorem 3.2 *Assume that F has an exponentially small Kolmogorov N-width, $d_N(F) \leq ce^{-aN}$ with $a > \log 2$. Then there exists a constant $\beta > 0$ such that the set F_N, obtained by the greedy algorithm is exponentially accurate in the sense that*

$$\|f - P_N f\|_{\mathbb{V}} \leq Ce^{-\beta N}.$$

In other words, if the underlying problem allows an efficient and compact reduced basis, the greedy approximation will find an exponentially convergent approximation to it.

Recall that the parameter-independent coercivity and continuity constants α and γ, introduced in (2.4), satisfy

$$\begin{aligned} &\forall \mu \in \mathbb{P}: \ a(u, v; \mu) \leq \gamma \|u\|_{\mathbb{V}} \|v\|_{\mathbb{V}}, && \forall u, v \in \mathbb{V}, \\ &\forall \mu \in \mathbb{P}: \ a(u, u; \mu) \leq \alpha \|u\|_{\mathbb{V}}^2, && \forall u \in \mathbb{V}. \end{aligned}$$

Then this convergence behavior can be extended to the reduced basis approximation as [1]:

Theorem 3.3 *Assume that the set of all solutions $\mathcal{M}$ (approximated by $\mathcal{M}_\delta$ in all computations) has an exponentially small Kolmogorov N-width $d_N(\mathcal{M}) \leq ce^{-aN}$, $a > \log\left(1 + \sqrt{\frac{\gamma}{\alpha}}\right)$, then the reduced basis approximation converges exponentially fast in the sense that there exists a $\beta > 0$ such that*

$$\forall \mu \in \mathbb{P}: \|u_\delta(\mu) - u_{\mathrm{rb}}(\mu)\|_{\mathbb{V}} \leq Ce^{-\beta N}.$$

Hence, the reduced basis approximation converges exponentially fast to the truth approximation. It is worth reiterating that the error between the truth approximation and the exact solution is assumed to be very small. If this is violated, the reduced basis approximation will still display exponential convergence to the truth approximation but this will possibly be a poor representation of the exact solution, i.e., the reduced basis approximation would be polluted by the lack of approximability of the truth solution.

From a practical viewpoint, it is important to observe that the different snapshots $u_\delta(\mu_1), \ldots, u_\delta(\mu_N)$ may be (almost) linearly dependent, resulting in a large condition number of the associated solution matrix. It is therefore advised to orthonormalize the snapshots in order to obtain the basis functions $\xi_1, \ldots, \xi_N$. For instance, one can use the Gram-Schmidt orthonormalization algorithm based on the vector of degrees of freedom of the functions $u_\delta(\mu_n)$ and the discrete scalar product of ℓ^2. Observe that one does not rely on the properties of orthonormality even if this is

Linear algebra box: The affine assumption

The affine assumption is inherited at the linear algebra level to allow the efficient assembly of the reduced basis solution matrix, right hand side and output functional. During the offline stage one precomputes all matrices

$$\begin{aligned} \mathbf{A}_{\mathtt{rb}}^{q} &= \mathbf{B}^{T}\,\mathbf{A}_{\delta}^{q}\,\mathbf{B} \in \mathbb{R}^{N\times N}, && 1 \le q \le Q_{\mathtt{a}}, \\ \mathrm{f}_{\mathtt{rb}}^{q} &= \mathbf{B}^{T}\,\mathrm{f}_{\delta}^{q} \in \mathbb{R}^{N}, && 1 \le q \le Q_{\mathtt{f}}, \\ \mathrm{l}_{\mathtt{rb}}^{q} &= \mathbf{B}^{T}\,\mathrm{l}_{\delta}^{q} \in \mathbb{R}^{N}, && 1 \le q \le Q_{\mathtt{l}}, \end{aligned}$$

where $(\mathbf{A}_{\delta}^{q})_{ij} = a_q(\varphi_j, \varphi_i)$, $(\mathrm{f}_{\delta}^{q})_j = f_q(\varphi_j)$ and $(\mathrm{l}_{\delta}^{q})_j = \ell_q(\varphi_j)$ for $1 \le i, j \le N_\delta$. During the online stage, one can then efficiently build the required operators as

$$\mathbf{A}_{\mathtt{rb}}^{\mu} = \sum_{q=1}^{Q_{\mathtt{a}}} \theta_{\mathtt{a}}^{q}(\mu)\,\mathbf{A}_{\mathtt{rb}}^{q}, \qquad \mathrm{f}_{\mathtt{rb}}^{\mu} = \sum_{q=1}^{Q_{\mathtt{f}}} \theta_{\mathtt{f}}^{q}(\mu)\,\mathrm{f}_{\mathtt{rb}}^{q}, \qquad \mathrm{l}_{\mathtt{rb}}^{\mu} = \sum_{q=1}^{Q_{\mathtt{l}}} \theta_{\mathtt{l}}^{q}(\mu)\,\mathrm{l}_{\mathtt{rb}}^{q}.$$

often desirable for numerical stability. All that is required is a set of basis functions $\xi_1, \ldots, \xi_N$ which spans the same space $\mathbb{V}_{\mathtt{rb}}$ and generates a solution matrix $\mathbf{A}_{\mathtt{rb}}$ with a reasonable condition number.

3.3 Ensuring Efficiency Through the Affine Decomposition

Having addressed the offline computation of the reduced basis, we can now turn to the online stage where the central part is the computation of a solution $u_{\mathtt{rb}}(\mu)$, defined by (3.3). In the ideal setting, the cost of accomplishing this should be independent of the complexity of the truth problem, measured by N_δ, and should depend only on the size $N \ll N_\delta$ of the reduced basis approximation.

To get a handle on this, first note that for each new parameter value $\mu \in \mathbb{P}$, the reduced basis solution matrix $\mathbf{A}_{\mathtt{rb}}^{\mu}$, as defined in the linear algebra box The reduced basis approximation, needs to be assembled. Since the parameter value μ may be in the bilinear form $a(\cdot, \cdot; \mu)$, one generally would need to first assemble the truth matrix $\mathbf{A}_{\delta}^{\mu}$ and then construct $\mathbf{A}_{\mathtt{rb}}^{\mu} = \mathbf{B}^{T}\,\mathbf{A}_{\delta}^{\mu}\,\mathbf{B}$, where $\mathbf{B}$ is the representation of the reduced basis in terms of the basis functions of the truth space $\mathbb{V}_\delta$. This is a computation that depends on N_δ, as $\mathbf{A}_{\delta}^{\mu} \in \mathbb{R}^{N_\delta \times N_\delta}$ and $\mathbf{B} \in \mathbb{R}^{N_\delta \times N}$, and would severely limit the potential for rapid online evaluation of new reduced basis solutions.

Algorithm: The online procedure

Input: A reduced basis model based on the reduced basis space $\mathbb{V}_{\mathtt{rb}}$ and a parameter value $\mu \in \mathbb{P}$.

Output: Fast evaluation of the output functional ($s_{\mathtt{rb}}(\mu)$ and the aposteriori estimate $\eta(\mu)$) that is independent of N_δ.

1. Assemble the reduced basis solution matrix and right hand side:

$$\mathbf{A}^{\mu}_{\mathtt{rb}} = \sum_{q=1}^{Q_{\mathtt{a}}} \theta^q_{\mathtt{a}}(\mu)\, \mathbf{A}^q_{\mathtt{rb}}, \qquad \mathrm{f}^{\mu}_{\mathtt{rb}} = \sum_{q=1}^{Q_{\mathtt{f}}} \theta^q_{\mathtt{f}}(\mu)\, \mathrm{f}^q_{\mathtt{rb}}, \qquad \text{and} \qquad \mathrm{l}^{\mu}_{\mathtt{rb}} = \sum_{q=1}^{Q_{\mathtt{l}}} \theta^q_{\mathtt{l}}(\mu)\, \mathrm{l}^q_{\mathtt{rb}}.$$

2. Solve the linear system

$$\mathbf{A}^{\mu}_{\mathtt{rb}} \mathrm{u}^{\mu}_{\mathtt{rb}} = \mathrm{f}^{\mu}_{\mathtt{rb}},$$

in order to obtain the degrees of freedom $(\mathrm{u}^{\mu}_{\mathtt{rb}})_n$ of the reduced basis solution $u_{\mathtt{rb}}(\mu)$.

3. Computate the output functional $s_{\mathtt{rb}}(\mu) = (\mathrm{u}^{\mu}_{\mathtt{rb}})^T\, \mathrm{l}^{\mu}_{\mathtt{rb}}$.

4. Computate the error estimate $\eta(\mu)$, see the upcoming Chap. 4 for details.

However, this restriction can be overcome if we assume that the forms $a(\cdot, \cdot; \mu)$, $f(\cdot; \mu)$ and $\ell(\cdot; \mu)$ allow the affine decomposition

$$a(w, v; \mu) = \sum_{q=1}^{Q_{\mathtt{a}}} \theta^q_{\mathtt{a}}(\mu)\, a_q(w, v), \tag{3.11}$$

$$f(v; \mu) = \sum_{q=1}^{Q_{\mathtt{f}}} \theta^q_{\mathtt{f}}(\mu)\, f_q(v), \tag{3.12}$$

$$\ell(v; \mu) = \sum_{q=1}^{Q_{\mathtt{l}}} \theta^q_{\mathtt{l}}(\mu)\, \ell_q(v), \tag{3.13}$$

where each form

$$a_q : \mathbb{V} \times \mathbb{V} \to \mathbb{R}, \qquad f_q : \mathbb{V} \to \mathbb{R}, \qquad \ell_q : \mathbb{V} \to \mathbb{R},$$

is independent of the parameter value μ and the coefficients

$$\theta^q_{\mathtt{a}} : \mathbb{P} \to \mathbb{R}, \qquad \theta^q_{\mathtt{f}} : \mathbb{P} \to \mathbb{R}, \qquad \theta^q_{\mathtt{l}} : \mathbb{P} \to \mathbb{R},$$

are scalar quantities which are independent of w and v. Note that we consider the abstract form of a general non-compliant problem for sake of completeness in this section and therefore also illustrate the affine decomposition of the output functional ℓ.

Illustrated by the example of the bilinear form $a(\cdot, \cdot, \mu)$, a series of $Q_{\mathtt{a}}$ $N \times N$-dimensional matrices $\mathbf{A}^q_{\mathtt{rb}}$ (each associated to $a_q(\cdot, \cdot)$) can be precomputed at the offline stage once the reduced basis space is known since the forms $a_q(\cdot, \cdot)$ are

independent of the parameter value. Then, during the online stage, when a new parameter value μ is given, one builds the new solution matrix as

$$\mathbf{A}_{\mathtt{rb}}^{\mu} = \sum_{q=1}^{Q_{\mathtt{a}}} \theta_{\mathtt{a}}^{q}(\mu)\, \mathbf{A}_{\mathtt{rb}}^{q},$$

by weighting the different matrices $\mathbf{A}_{\mathtt{rb}}^{q}$ by the factors parameter dependent $\theta_{\mathtt{a}}^{q}(\mu)$. This operation is independent of N_δ and scales proportionally to $Q_{\mathtt{a}} \cdot N^2$. The treatment of the linear forms $f(\cdot, \mu)$ and $\ell(\cdot, \mu)$ is similar and an account of the steps needed is provided in the algorithm box The affine assumption.

For cases where an affine decomposition does not hold naturally for the operator or the linear forms, one can often find an approximate form that satisfies this property using a technique known as Empirical Interpolation. We shall discuss this technique in detail in Chap. 5.

The evaluation of (3.3) and (3.4) on an algebraic level can then be done as outlined in algorithm box The online procedure, involving only operations that are independent of N_δ.

3.4 Illustrative Examples

We illustrate below how the affine assumption can be satisfied on our toy problems defined in Sect. 2.3 and provide some numerical results illustrating the convergence of the reduced model for the basis being assembled using the greedy- and the POD-algorithms during the offline procedure.

3.4.1 Illustrative Example 1: Heat Conduction Part 2

For Example 1 introduced in Sect. 2.3.1, recall that the bilinear and linear forms are given by

$$a(w, v; \mu) = \int_{\Omega} \kappa_\mu \nabla w \cdot \nabla v \qquad \text{and} \qquad f(v; \mu) = \mu_{[2]} \int_{\Gamma_{\text{base}}} v.$$

Since $\kappa_\mu = \mathbf{1}_{\Omega_1} + \mu_{[1]} \mathbf{1}_{\Omega_0}$ we can decompose the bilinear form into two parts to recover the affine decomposition with $Q_{\mathtt{a}} = 2$, $Q_{\mathtt{f}} = 1$, $\theta_{\mathtt{a}}^{1}(\mu) = 1$, $\theta_{\mathtt{a}}^{2}(\mu) = \mu_{[1]}$, $\theta_{\mathtt{f}}^{1}(\mu) = \mu_{[2]}$ and

$$a_1(w, v) = \int_{\Omega_1} \nabla w \cdot \nabla v, \qquad a_2(w, v) = \int_{\Omega_0} \nabla w \cdot \nabla v, \qquad \text{and} \qquad f_1(v; \mu) = \int_{\Gamma_{\text{base}}} v.$$

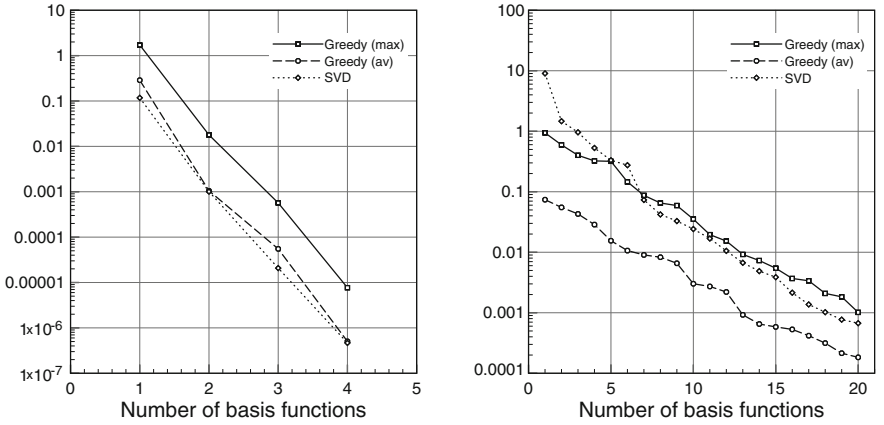

Fig. 3.1 Maximum and average error bound with respect to the number of selected basis functions in comparison with a SVD for the illustrative Example 1 (*left*) and Example 2 (*right*)

Note that the way the affine assumption is satisfied is not unique. Considering for example the bilinear form a, one could alternatively define $Q_{\mathtt{a}} = 2$, $\theta^1_{\mathtt{a}}(\mu) = 1$, $\theta^2_{\mathtt{a}}(\mu) = \mu_{[1]} - 1$ and

$$a_1(w, v) = \int_{\Omega} \nabla w \cdot \nabla v, \qquad a_2(w, v) = \int_{\Omega_0} \nabla w \cdot \nabla v.$$

We now briefly illustrate the reduced basis construction for this problem with basis functions that have been obtained by orthogonalization, through the Gram-Schmidt procedure, of snapshots computed for selected parameters provided by the greedy algorithm based on $\mathbb{P}_h$ of cardinality 1,000. In Fig. 3.1 (left), the maximum and average absolute errors, i.e.,

$$\max_{\mu \in \mathbb{P}_h} \|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_\mu, \qquad \text{and} \qquad \frac{1}{|\mathbb{P}_h|} \sum_{\mu \in \mathbb{P}_h} \|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_\mu, \tag{3.14}$$

with respect to the number of selected basis functions is reported. In addition, we also plot the decay of the relative singular values of the solution matrix over the training set $\mathbb{P}_h$ as a measure of optimality.

In Fig. 3.2, the outcomes provided by the finite element and the reduced basis approximations, for a randomly chosen $\mu = (6.68, 0.94)$ and $N = 5$, are compared. The difference between the two solutions, thus the error function, is also shown, confirming the high accuracy of the model.

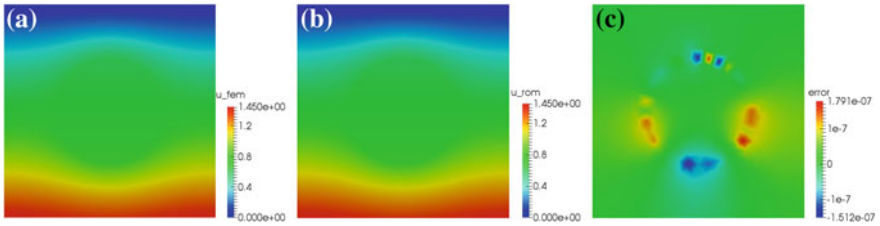

Fig. 3.2 Visual comparison between the truth (finite element) solution (**a**), the reduced basis approximation (**b**) for $\mu = (6.68, 0.94)$. The difference (error) between the two solutions is illustrated in (**c**)

3.4.2 Illustrative Example 2: Linear Elasticity Part 2

Concerning the elastic block of the illustrative Example 2, the form f is clearly affine in the parameter dependency and hence $Q_{\mathrm{f}} = 3$ with $\theta^q_{\mathrm{f}}(\mu) = \mu_{[8+q]}$ for $q = 1, 2, 3$. The bilinear form a is affine for $Q_{\mathrm{a}} = 9$ with $\theta^q_{\mathrm{a}}(\mu) = \mu_{[q]}$, $1 \le q \le 8$, $\theta^9_{\mathrm{a}}(\mu) = 1$, and

$$a_q(w, v) = E \int_{\Omega_q} \frac{\partial v_i}{\partial x_j} C_{ijkl} \frac{\partial w_k}{\partial x_l}, \quad 1 \le q \le 9,$$

where we recall that we use Einstein's summation convention for the indices i, j, k and l. The basis functions that have been obtained by orthogonalization, always through a Gram-Schimdt procedure, of snapshots obtained by the greedy algorithm based on $\mathbb{P}_h$ of cardinality 7,500. The maximum and average errors, given by (3.14), with respect to the number of basis functions employed is reported in Fig. 3.1 (right). In addition, we also plot the decay of the relative singular values of the solution matrix over the training set $\mathbb{P}_h$ as a measure of optimality.

In Fig. 3.3, the outcomes provided by the finite element and the reduced basis solution for a randomly chosen system configuration with reduced basis dimension $N = 53$ are compared, and the difference (error) between the two solutions is shown, confirming the expected accuracy.

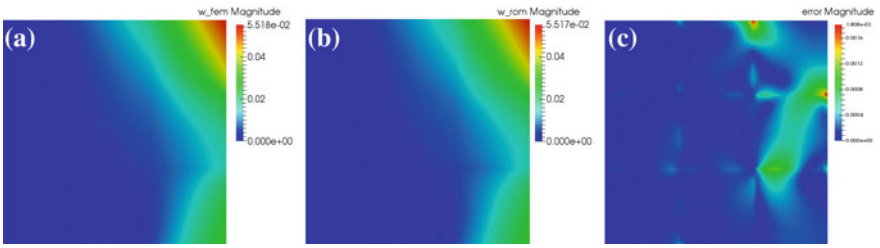

Fig. 3.3 Comparison between truth (finite element) solution (**a**), reduced basis approximation (**b**) for $\mu = (7.737, 7.124, 0.729, 4.620, 3.072, 6.314, 3.590, 7.687, -0.804, 0.129, -0.232)$. The difference (error) between the two solutions is illustrated in (**c**)

3.5 Summary of the Method

The elements of the method can be summarized as follows. In an offline procedure, one seeks to identify a reduced basis space $\mathbb{V}_{\mathtt{rb}} \subset \mathbb{V}_\delta$ that has a dimension as small as possible such that any element of the solution manifold $\mathcal{M}_\delta$ can be approximated with an element of $\mathbb{V}_{\mathtt{rb}}$ to within a desired accuracy. This step can be seen as discarding any degrees of freedom of $\mathbb{V}_\delta$ that are not needed to approximate functions of $\mathcal{M}_\delta$. This is an empirical procedure which depends on the parametrized problem and thus also on the parameter space $\mathbb{P}$ considered in a specific application. Hence, the resulting approximation space $\mathbb{V}_{\mathtt{rb}}$ and its basis are not multi-purpose, but rather tailor made for a specific problem. Most commonly, the construction of $\mathbb{V}_{\mathtt{rb}}$ is built on the greedy algorithm which is based on the existence of an a posteriori error estimator $\eta(\mu)$ which will be developed in the upcoming Chap. 4.

The online procedure comprises a fast evaluation of the map $\mu \mapsto s_{\mathtt{rb}}(\mu)$ which can be evaluated independently of N_δ, the dimension of the truth approximation space $\mathbb{V}_\delta$. The efficiency, i.e., the independence on N_δ, is enabled by the affine assumption of the bilinear form $a(\cdot\,,\,\cdot\,,\mu)$ and the linear form $f(\,\cdot\,;\mu)$.

The construction of the reduced basis approximation and the importance of affine assumption can best be illustrated by a brief overview of the different elements discussed so far as well as a discussion of the computational complexity of the entire

Algorithm: The offline procedure

Input: An abstract truth model of the form (2.8) satisfying the affine assumption (3.11)–(3.13).
An aposteriori error estimator $\eta(\mu)$ such that $\|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_\mu \le \eta(\mu)$.
An error tolerance $\mathtt{tol}$.
A discrete training set $\mathbb{P}_h$.

Output: A reduced basis model based on the reduced basis space $\mathbb{V}_{\mathtt{rb}}$ that guarantees that $\max_{\mu\in\mathbb{P}_h} \|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_\mu \le \mathtt{tol}$.

Initialization: Take $\mu_1 \in \mathbb{P}$ arbitrary and set $n = 1$.

Loop:

(i) **Offline-offline:**

a. Compute $u_\delta(\mu_n)$ as solution to (3.1) for μ_n and set $\mathbb{V}_{\mathtt{rb}} = \mathrm{span}\{u_\delta(\mu_1), \ldots, u_\delta(\mu_n)\}$.
b. Based on $\mathbb{V}_{\mathtt{rb}}$, pre-compute all quantities from the affine-decomposition that are parameter-independent, e.g., the n-dimensional matrices $\mathbf{A}^q_{\mathtt{rb}}$ or the n-dimensional vectors $\mathrm{f}^q_{\mathtt{rb}}$.

(ii) **Offline-online:**

a. For each $\mu \in \mathbb{P}_h$, compute the reduced basis approximation $u_{\mathtt{rb}}(\mu) \in \mathbb{V}_{\mathtt{rb}}$ defined by (3.3) for μ and the error estimator $\eta(\mu)$.
b. Choose $\mu_{n+1} = \arg\max_{\mu\in\mathbb{P}_h} \eta(\mu)$.
c. If $\eta(\mu_{n+1}) > \mathtt{tol}$, then set $n := n + 1$ and **go to** (i), otherwise **terminate**.

approach. We can, for illustration only, compare this with the cost of solving a similar problem using the truth solver for all evaluations.

Let us first consider the offline procedure, outlined in the algorithm box The offline procedure. The computation cost of the offline development of the reduced basis approximation is dominated by the need to solve the truth problem N times, yielding an overall computational cost of $\mathcal{O}(N N_\delta{}^p)$ where $p \leq 3$ is determined by details of the solver, i.e., where direct or iterative solvers are used, the quality of the preconditioner etc. However, since $N_\delta \gg N$ the cost is considerable.

Turning to the cost of the online evaluation of the reduced approximation outlined in the algorithm box The online procedure, we have several terms, i.e., $\mathcal{O}(Q_{\mathtt{a}} N^2)$ to assemble the operator, $\mathcal{O}(Q_{\mathtt{f}} N)$ to assemble the right hand side, $\mathcal{O}(N^3)$ to recover the reduced basis solution, and, finally $\mathcal{O}(Q_{\mathtt{l}} N)$ to evaluate the output of interest in the general non-compliant case ($Q_{\mathtt{l}} = Q_{\mathtt{f}}$ otherwise). To furthermore certify the accuracy through the error estimator we need $\mathcal{O}((Q_{\mathtt{f}} + N Q_{\mathtt{a}})^2)$ operations to evaluate the residual (based on results from Chap. 4). The cost of estimating the stability constant, needed in the error estimator, depends on the details of the method used but is independent of N_δ. The key point is that although the online cost can not be neglected for large values of N and the affine parameters, $Q_{\mathtt{a}}$, $Q_{\mathtt{f}}$, $Q_{\mathtt{l}}$, the entire computational process is independent of N_δ. Hence, for applications where $N_\delta \gg N$, this will be substantially faster than for solving the truth problems, in particular for applications where N_δ is very large, e.g., multi-scale and three-dimensional problems, or for problems where a very large number of parameter evaluations are needed.

References

1. P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, P. Wojtaszczyk, Convergence rates for greedy algorithms in reduced basis methods. SIAM J. Math. Anal. **43**, 1457–1472 (2011)
2. R. DeVore, G. Petrova, P. Wojtaszczyk, Greedy algorithms for reduced bases in Banach spaces. Constr. Approx. **37**, 455–466 (2013)
3. A. Buffa, Y. Maday, A.T. Patera, C. Prud'homme, G. Turinici, A priori convergence of the greedy algorithm for the parametrized reduced basis method. ESAIM Math. Model. Numer. Anal. **46**, 595–603 (2012)

Chapter 4
Certified Error Control

4.1 Introduction

The development of effective and reliable a posteriori error estimators for the field variable or an output of interest is crucial to ensure the reliability and efficiency of the reduced basis approximations. Reduced basis approximations are problem dependent since discretizations are problem specific. They are typically pre-asymptotic since we choose N small to control the online computational cost. Furthermore, the reduced basis functions can not be directly related to specific spatial or temporal scales so problem intuition is of limited value and may even be faulty. Finally, the reduced basis approach is often applied in a real-time context where there is no time for offline verification and errors may be manifested immediately and in deleterious ways. It is thus essential that the development of techniques for the efficient and reliable computation of error estimates plays a central role during both the offline and online stages.

In the greedy algorithm, executed during the offline stage, the error estimates are used as surrogates for the actual error to enable large training sets without dramatically increasing the offline cost. This subsequently results in reduced basis approximations of high fidelity as they can be trained on large and finely sampled spaces of parameters. It is clear that the offline sampling procedures can not be exhaustive, particularly for high-dimensional parameter dimensions P where large parts of the parameter set $\mathbb{P}$ may remain unexplored. Hence, we must accept that we can only accurately account for the online quality of the output for parameters in well sampled parts of $\mathbb{P}$ and the error is controlled over $\mathbb{P}_h$ rather than $\mathbb{P}$. However, the a posteriori estimation guarantees that we can rigorously and efficiently bound the output error in the online procedure for any given new $\mu \in \mathbb{P}$.

During the online stage, the error estimators enable the identification of the minimal reduced dimension N ensuring the required accuracy. This guarantees that constraints are satisfied and feasibility conditions are verified. This subsequently

J.S. Hesthaven et al., *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, SpringerBriefs in Mathematics,
DOI 10.1007/978-3-319-22470-1_4

ensures the validity of the prediction, hence enabling online prediction endowed with the certified accuracy of the truth solution.

In the following we present several a posteriori estimates. These are developed both for the field variable in combination with different norms and for the output functional. They can be used either in the greedy selection of the reduced basis space during the offline procedure or during the online procedure to certify the output, depending on the application and criteria of the user.

4.2 Error Control for the Reduced Order Model

The need for efficient and accurate error bounds places additional requirements on the error estimation. The error bounds must be rigorous and valid for all N and for all parameter values in the parameter domain $\mathbb{P}$. Indeed, non-rigorous error indicators may suffice for adaptivity, but not for reliability. Furthermore, the bounds must be reasonably sharp since overly conservative errors will yield inefficient approximations, with N being large in the RB model, or suboptimal engineering results with, e.g., unnecessary safety margins. Finally, the bounds must be computable at low cost, independent of N_δ, due to the critical role these play in both the offline and the online stage.

We recall that our reduced basis error bounds will be defined relative to the underlying accurate discretization method, e.g., the finite element approximation—the truth. Hence, if the truth approximation is poor, the error estimator will still reflect convergence but it will be convergence to a solution that only poorly approximates the solution of the continuous problem.

4.2.1 Discrete Coercivity and Continuity Constants of the Bilinear Form

The error certification we will be based on the discrete coercivity and continuity constants defined by

$$\alpha_\delta(\mu) = \inf_{v_\delta \in \mathbb{V}_\delta} \frac{a(v_\delta, v_\delta; \mu)}{\|v_\delta\|_{\mathbb{V}}^2}, \qquad \text{and} \qquad \gamma_\delta(\mu) = \sup_{w_\delta \in \mathbb{V}_\delta} \sup_{v_\delta \in \mathbb{V}_\delta} \frac{a(w_\delta, v_\delta; \mu)}{\|w_\delta\|_{\mathbb{V}} \|v_\delta\|_{\mathbb{V}}}, \tag{4.1}$$

since the approximation space is conforming, i.e., $\mathbb{V}_\delta \subset \mathbb{V}$, it holds that $\alpha(\mu) \leq \alpha_\delta(\mu)$ and $\gamma_\delta(\mu) \leq \gamma(\mu)$ where $\alpha(\mu)$ and $\gamma(\mu)$ are the continuous coercivity and continuity constants introduced in (2.5).

4.2.2 Error Representation

The central equation in residual-based a posteriori theory is a quantification of the relationship between the error and the residual. It follows from the problem statements for $u_\delta(\mu)$, defined by (2.8), and $u_{\mathtt{rb}}(\mu)$, defined by (3.3), that the error $e(\mu) = u_\delta(\mu) - u_{\mathtt{rb}}(\mu) \in \mathbb{V}_\delta$ satisfies the classic error equation

$$a(e(\mu), v_\delta; \mu) = r(v_\delta; \mu), \qquad \forall\, v_\delta \in \mathbb{V}_\delta, \tag{4.2}$$

where $r(\,\cdot\,; \mu) \in \mathbb{V}_\delta'$ (the dual space to $\mathbb{V}_\delta$) is the residual,

$$r(v_\delta; \mu) = f(v_\delta; \mu) - a(u_{\mathtt{rb}}(\mu), v_\delta; \mu), \qquad \forall\, v_\delta \in \mathbb{V}_\delta. \tag{4.3}$$

Indeed, (4.2) follows from (4.3) by the bilinearity of a and the definition of $e(\mu)$.

It shall prove convenient to introduce the Riesz representation of $r(\,\cdot\,; \mu)$, denoted by $\hat{r}_\delta(\mu) \in \mathbb{V}_\delta$ and defined as the unique $\hat{r}_\delta(\mu) \in \mathbb{V}_\delta$ satisfying

$$(\hat{r}_\delta(\mu), v_\delta)_{\mathbb{V}} = r(v_\delta; \mu), \qquad \forall\, v_\delta \in \mathbb{V}_\delta. \tag{4.4}$$

Consequently, it holds that

$$\|\hat{r}_\delta(\mu)\|_{\mathbb{V}} = \|r(\,\cdot\,, \mu)\|_{\mathbb{V}_\delta'} = \sup_{v_\delta \in \mathbb{V}_\delta} \frac{r(v_\delta; \mu)}{\|v_\delta\|_{\mathbb{V}}}.$$

We can also write the error equation as

$$a(e(\mu), v_\delta; \mu) = (\hat{r}_\delta(\mu), v_\delta)_{\mathbb{V}}, \qquad \forall\, v_\delta \in \mathbb{V}_\delta. \tag{4.5}$$

As we shall see shortly, the evaluation of the dual norm of the residual through the Riesz representation is a central element in the construction of efficient and reliable a posteriori estimators.

In the next proposition we discuss some error relations which can be established due to the Galerkin framework and the nature of the compliant problem that turns out to be useful in the upcoming proofs.

Proposition 4.1 *For a compliant problem it holds*

$$s_\delta(\mu) - s_{\mathtt{rb}}(\mu) = \|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_\mu^2,$$

for all $\mu \in \mathbb{P}$. *Hence* $s_\delta(\mu) \geq s_{\mathtt{rb}}(\mu)$.

Proof Let $\mu \in \mathbb{P}$ be arbitrary. We first observe by the definitions (2.8) or (3.1) and (3.3) of $u_\delta(\mu) \in \mathbb{V}_\delta$ and $u_{\mathtt{rb}}(\mu) \in \mathbb{V}_{\mathtt{rb}}$ that the following Galerkin orthogonality holds

$$a\big(u_\delta(\mu) - u_{\mathtt{rb}}(\mu), v_{\mathtt{rb}}; \mu\big) = 0, \qquad \forall v_{\mathtt{rb}} \in \mathbb{V}_{\mathtt{rb}}.$$

Then, by the linearity of the right hand side $f(\cdot\,,\mu)$, the definition of $u_\delta(\mu)$ and the Galerkin orthogonality we obtain

$$s_\delta(\mu) - s_{\mathtt{rb}}(\mu) = f\big(e(\mu);\mu\big) = a\big(u_\delta(\mu), e(\mu);\mu\big) = a\big(e(\mu), e(\mu);\mu\big) = \|e(\mu)\|_\mu^2.$$

Since we consider a compliant problem. □

In the following Sects. 4.2.3 and 4.2.4 we initially assume that we have access to a lower bound $\alpha_{\mathtt{LB}}(\mu)$ of the coercivity constant $\alpha_\delta(\mu)$, defined by (4.1), for any value of $\mu \in \mathbb{P}$ in a way that is independent of N_δ. In Sect. 4.3 we shall then revisit this assumption and construct such bounds.

4.2.3 Energy and Output Error Bounds

We define computable error estimators for the energy norm, output, and relative output as

$$\eta_{\mathtt{en}}(\mu) = \frac{\|\hat{r}_\delta(\mu)\|_{\mathbb{V}'}}{\alpha_{\mathtt{LB}}^{1/2}(\mu)}, \tag{4.6a}$$

$$\eta_{\mathtt{s}}(\mu) = \frac{\|\hat{r}_\delta(\mu)\|_{\mathbb{V}'}^2}{\alpha_{\mathtt{LB}}(\mu)} = \big(\eta_{\mathtt{en}}(\mu)\big)^2, \tag{4.6b}$$

$$\eta_{\mathtt{s,rel}}(\mu) = \frac{\|\hat{r}_\delta(\mu)\|_{\mathbb{V}'}^2}{\alpha_{\mathtt{LB}}(\mu)\, s_{\mathtt{rb}}(\mu)} = \frac{\eta_{\mathtt{s}}(\mu)}{s_{\mathtt{rb}}(\mu)}. \tag{4.6c}$$

The following proposition ensures that those estimators are rigorous upper bounds of the quantities they estimate.

Proposition 4.2 *For the computable a posteriori error estimators defined by* (4.6a)–(4.6c) *there holds*

$$\|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_\mu \le \eta_{\mathtt{en}}(\mu) \tag{4.7a}$$

$$s_\delta(\mu) - s_{\mathtt{rb}}(\mu) \le \eta_{\mathtt{s}}(\mu), \tag{4.7b}$$

$$\frac{s_\delta(\mu) - s_{\mathtt{rb}}(\mu)}{s_\delta(\mu)} \le \eta_{\mathtt{s,rel}}(\mu), \tag{4.7c}$$

for all $\mu \in \mathbb{P}$.

Proof It follows directly from the error equation, (4.5) with $v_\delta = e(\mu)$, and the Cauchy-Schwarz inequality that

$$\|e(\mu)\|_\mu^2 = a(e(\mu), e(\mu);\mu) \le \|\hat{r}_\delta(\mu)\|_{\mathbb{V}'}\, \|e(\mu)\|_{\mathbb{V}}. \tag{4.8}$$

Since $\alpha_{\mathrm{LB}}(\mu)$ is assumed to be a lower bound of the coercivity constant $\alpha_\delta(\mu)$ we conclude that

$$\alpha_{\mathrm{LB}}(\mu)\,\|e(\mu)\|_{\mathbb{V}}^2 \le a(e(\mu),e(\mu);\mu) = \|e(\mu)\|_\mu^2.$$

Combining thus with (4.8) yields (4.7a).

Next, we know from Proposition 4.1 that $s_\delta(\mu) - s_{\mathrm{rb}}(\mu) = \|e(\mu)\|_\mu^2$, and since $\eta_{\mathrm{s}}(\mu) = (\eta_{\mathrm{en}}(\mu))^2$, (4.7b) follows. Finally, we can easily deduce that $s_{\mathrm{rb}}(\mu) \le s_\delta(\mu)$ (by Proposition 4.1) which, in combination with (4.7b), implies (4.7c). □

We next introduce the effectivity index associated with these error estimators:

$$\mathtt{eff}_{\mathrm{en}}(\mu) = \frac{\eta_{\mathrm{en}}(\mu)}{\|u_\delta(\mu) - u_{\mathrm{rb}}(\mu)\|_\mu}, \tag{4.9a}$$

$$\mathtt{eff}_{\mathrm{s}}(\mu) = \frac{\eta_{\mathrm{s}}(\mu)}{s_\delta(\mu) - s_{\mathrm{rb}}(\mu)}, \tag{4.9b}$$

$$\mathtt{eff}_{\mathrm{s,rel}}(\mu) = \frac{\eta_{\mathrm{s,rel}}(\mu)}{(s_\delta(\mu) - s_{\mathrm{rb}}(\mu))\,/s(\mu)}, \tag{4.9c}$$

as measures of the quality of the proposed estimator. To ensure rigor of the estimates, we require effectivities ≥ 1 as ensured by Proposition 4.2. For sharpness of the error estimates, however, we desire effectivities as close to unity as possible.

Under the assumption that we remain in the coercive, compliant and, hence, symmetric framework, we recover the following proposition.

Proposition 4.3 *The effectivities* (4.9) *satisfy*

$$\mathtt{eff}_{\mathrm{en}}(\mu) \le \sqrt{\gamma_\delta(\mu)/\alpha_{\mathrm{LB}}(\mu)}, \tag{4.10a}$$

$$\mathtt{eff}_{\mathrm{s}}(\mu) \le \gamma_\delta(\mu)/\alpha_{\mathrm{LB}}(\mu), \tag{4.10b}$$

$$\mathtt{eff}_{\mathrm{s,rel}}(\mu) \le (1+\eta_{\mathrm{s,rel}})\,\gamma_\delta(\mu)/\alpha_{\mathrm{LB}}(\mu), \tag{4.10c}$$

for all $\mu \in \mathbb{P}$. *We recall that* $\alpha_{\mathrm{LB}}(\mu)$ *denotes an lower bound of the coercivity constant* $\alpha_\delta(\mu)$ *and that* $\gamma_\delta(\mu)$ *defines the continuity constant defined by* (4.1).

Proof Consider (4.5) with $v_\delta = \hat{r}_\delta(\mu)$. In combination with the Cauchy-Schwarz inequality this yields

$$\|\hat{r}_\delta(\mu)\|_{\mathbb{V}}^2 = a(e(\mu),\hat{r}_\delta(\mu);\mu) \le \|\hat{r}_\delta(\mu)\|_\mu\,\|e(\mu)\|_\mu. \tag{4.11}$$

By continuity (2.5) of the bilinear form we obtain

$$\|\hat{r}_\delta(\mu)\|_\mu^2 = a(\hat{r}_\delta(\mu),\hat{r}_\delta(\mu);\mu) \le \gamma_\delta(\mu)\,\|\hat{r}_\delta(\mu)\|_{\mathbb{V}}^2 \le \gamma_\delta(\mu)\,\|\hat{r}_\delta(\mu)\|_\mu\,\|e(\mu)\|_\mu. \tag{4.12}$$

Combining (4.11) and (4.12) implies

$$\eta_{\mathtt{en}}^2(\mu) = \frac{\|\hat{r}_\delta(\mu)\|_{\mathbb{V}}^2}{\alpha_{\mathtt{LB}}(\mu)} \leq \frac{\|\hat{r}_\delta(\mu)\|_\mu \|e(\mu)\|_\mu}{\alpha_{\mathtt{LB}}(\mu)} \leq \frac{\gamma_\delta(\mu)}{\alpha_{\mathtt{LB}}(\mu)} \|e(\mu)\|_\mu^2,$$

which establishes (4.10a).

Next recall that from Proposition 4.1 we have $s_\delta(\mu) - s_{\mathtt{rb}}(\mu) = \|e(\mu)\|_\mu^2$, and hence

$$\mathtt{eff}_{\mathtt{s}}(\mu) = \frac{\eta_{\mathtt{s}}(\mu)}{s_\delta(\mu) - s_{\mathtt{rb}}(\mu)} = \frac{(\eta_{\mathtt{en}}(\mu))^2}{\|e(\mu)\|_\mu^2} = (\mathtt{eff}_{\mathtt{en}})^2 \leq \frac{\gamma_\delta(\mu)}{\alpha_{\mathtt{LB}}(\mu)}$$

through (4.10a).

Finally, since $\eta_{\mathtt{s,rel}}(\mu) = \eta_{\mathtt{s}}(\mu)/s_{\mathtt{rb}}(\mu)$, we obtain

$$\mathtt{eff}_{\mathtt{s,rel}}(\mu) = \big(s_\delta(\mu)/s_{\mathtt{rb}}(\mu)\big)\, \mathtt{eff}_{\mathtt{s}}(\mu). \tag{4.13}$$

Observe that $s_{\mathtt{rb}}(\mu) \leq s_\delta(\mu)$ as consequence of Proposition 4.1. Applying (4.7b) yields

$$\frac{s_\delta(\mu)}{s_{\mathtt{rb}}(\mu)} = 1 + \frac{s_\delta(\mu) - s_{\mathtt{rb}}(\mu)}{s_{\mathtt{rb}}(\mu)} \leq 1 + \frac{\eta_{\mathtt{s}}(\mu)}{s_\delta(\mu)} = 1 + \eta_{\mathtt{s,rel}}(\mu),$$

which, in combination with (4.13) and (4.10b), proves (4.10c). □

Proposition 4.2 establishes that the estimators (4.9a)–(4.9c) are rigorous upper bounds for the reduced basis error in the energy norm, the reduced basis output error, and the reduced basis relative output error, respectively. Furthermore, the effectivity of the energy-norm and output error estimators is bounded from above independent of N by Proposition 4.3 while being rigorously bounded from below as

$$\gamma_\delta(\mu)/\alpha_{\mathtt{LB}}(\mu) \geq 1,$$

by the definition of the constants.

4.2.4 $\mathbb{V}$-Norm Error Bounds

Although the bounds on the accuracy of the output are arguably the most relevant, it also proves useful (e.g., in a visualization context) to provide a certificate of fidelity for the full field error $u_\delta(\mu) - u_{\mathtt{rb}}(\mu)$ in a norm which is independent of μ. For this, we introduce error estimators in the $\mathbb{V}$-norm and the relative $\mathbb{V}$-norm. We note here that the choice of the $\mathbb{V}$-norm (and hence of $\bar{\mu}$) does not affect the reduced basis output prediction $s_{\mathtt{rb}}(\mu)$ but may affect the sharpness of the a posteriori output error

bounds. In what follows, the $\mathbb{V}$-norm can be replace by any norm on $\mathbb{V}_\delta$ without impacting the validity of the results adapting however the definition of $\alpha_{\mathrm{LB}}(\mu)$.

Let us introduce the following error estimators

$$\eta_{\mathbb{V}}(\mu) = \frac{\|\hat{r}_\delta(\mu)\|_{\mathbb{V}}}{\alpha_{\mathrm{LB}}(\mu)}, \tag{4.14a}$$

$$\eta_{\mathbb{V},\mathtt{rel}}(\mu) = \frac{2\,\|\hat{r}_\delta(\mu)\|_{\mathbb{V}}}{\alpha_{\mathrm{LB}}(\mu)\,\|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}}.$$

Proposition 4.4 *It holds that*

$$\|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}} \le \eta_{\mathbb{V}}(\mu), \qquad \forall \mu \in \mathbb{P}. \tag{4.15}$$

Furthermore, if $\eta_{\mathbb{V},\mathtt{rel}}(\mu) \le 1$ *for* $\mu \in \mathbb{P}$, *then*

$$\frac{\|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}}{\|u_\delta(\mu)\|_{\mathbb{V}}} \le \eta_{\mathbb{V},\mathtt{rel}}(\mu). \tag{4.16}$$

Proof Inequality (4.15) follows from (4.7a) of Proposition 4.2, $\alpha_{\mathrm{LB}}(\mu)\,\|e(\mu)\|_{\mathbb{V}}^2 \le \|e(\mu)\|_\mu^2$, and the definition of $\eta_{\mathbb{V}}(\mu)$, (4.14a).

For (4.16) we first observe that

$$\eta_{\mathbb{V},\mathtt{rel}}(\mu) = 2\,\frac{\|\hat{r}_\delta(\mu)\|_{\mathbb{V}}}{\alpha_{\mathrm{LB}}(\mu)\|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}} = 2\,\frac{\|u_\delta(\mu)\|_{\mathbb{V}}}{\|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}}\,\frac{\eta_{\mathbb{V}}(\mu)}{\|u_\delta(\mu)\|_{\mathbb{V}}}. \tag{4.17}$$

By (4.15) and the assumption that $\eta_{\mathbb{V},\mathtt{rel}}(\mu) \le 1$ we further recover

$$\begin{aligned}\|u_\delta(\mu)\|_{\mathbb{V}} &= \|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}} + \|u_\delta(\mu)\|_{\mathbb{V}} - \|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}} \ge \|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}} - \|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}} \\ &\ge \|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}} - \eta_{\mathbb{V}}(\mu) = (1 - \tfrac{1}{2}\eta_{\mathbb{V},\mathtt{rel}})\,\|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}} \ge \tfrac{1}{2}\|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}.\end{aligned}$$

Combined with (4.15) and (4.17), this yields

$$\eta_{\mathbb{V},\mathtt{rel}}(\mu) = 2\,\frac{\|u_\delta(\mu)\|_{\mathbb{V}}}{\|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}}\,\frac{\eta_{\mathbb{V}}(\mu)}{\|u_\delta(\mu)\|_{\mathbb{V}}} \ge \frac{\eta_{\mathbb{V}}(\mu)}{\|u_\delta(\mu)\|_{\mathbb{V}}} \ge \frac{\|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}}{\|u_\delta(\mu)\|_{\mathbb{V}}}.$$

□

We define the associated effectivities,

$$\mathtt{eff}_{\mathbb{V}}(\mu) = \frac{\eta_{\mathbb{V}}(\mu)}{\|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}},$$

$$\mathtt{eff}_{\mathbb{V},\mathtt{rel}}(\mu) = \frac{\eta_{\mathbb{V},\mathtt{rel}}(\mu)}{\|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}/\|u_\delta(\mu)\|_{\mathbb{V}}}.$$

Proposition 4.5 *It holds that*

$$\mathtt{eff}_{\mathbb{V}}(\mu) \leq \frac{\gamma_\delta(\mu)}{\alpha_{\mathrm{LB}}(\mu)}, \qquad \forall \mu \in \mathbb{P}. \tag{4.19}$$

Furthermore, if $\eta_{\mathbb{V},\mathtt{rel}}(\mu) \leq 1$ *for* $\mu \in \mathbb{P}$*, then*

$$\mathtt{eff}_{\mathbb{V},\mathtt{rel}}(\mu) \leq 3\frac{\gamma_\delta(\mu)}{\alpha_{\mathrm{LB}}(\mu)}. \tag{4.20}$$

Proof Indeed, inequality (4.19) follows directly from (4.10a) and $\|e(\mu)\|_\mu \leq \gamma_\delta(\mu)\,\|e(\mu)\|_{\mathbb{V}}$ as

$$\mathtt{eff}_{\mathbb{V}}(\mu) = \frac{\mathtt{eff}_{\mathtt{en}}(\mu)}{\alpha_{\mathrm{LB}}^{1/2}(\mu)} \leq \frac{\gamma_\delta^{1/2}(\mu)}{\alpha_{\mathrm{LB}}(\mu)}\,\frac{\|e(\mu)\|_\mu}{\|e(\mu)\|_{\mathbb{V}}} \leq \frac{\gamma_\delta(\mu)}{\alpha_{\mathrm{LB}}(\mu)}.$$

To demonstrate (4.20), we first note that

$$\mathtt{eff}_{\mathbb{V},\mathtt{rel}}(\mu) = 2\frac{\|u_\delta(\mu)\|_{\mathbb{V}}}{\|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}}\,\mathtt{eff}_{\mathbb{V}}(\mu) = 2\left(1 + \frac{\|u_\delta(\mu)\|_{\mathbb{V}} - \|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}}{\|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}}\right)\mathtt{eff}_{\mathbb{V}}(\mu)$$

and observe that by (4.15) and the assumption $\mathtt{eff}_{\mathbb{V},\mathtt{rel}}(\mu) \leq 1$ we recover

$$\begin{aligned}\|u_\delta(\mu)\|_{\mathbb{V}} - \|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}} &\leq \|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}} \\ &\leq \alpha_{\mathrm{LB}}^{-1}(\mu)\|\hat{r}_\delta(\mu)\|_{\mathbb{V}} = \tfrac{1}{2}\|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}\,\eta_{\mathbb{V},\mathtt{rel}}(\mu) \leq \tfrac{1}{2}\|u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}}.\end{aligned}$$

From (4.19) we finally get

$$\mathtt{eff}_{\mathbb{V},\mathtt{rel}}(\mu) \leq 3\frac{\gamma_\delta(\mu)}{\alpha_{\mathrm{LB}}(\mu)}. \qquad \square$$

4.2.5 *Efficient Computation of the a Posteriori Estimators*

The error estimators (4.6) and (4.14) all require the evaluation of the residual in the dual norm, computed through the Riesz representation $\|\hat{r}_\delta(\mu)\|_{\mathbb{V}}$ of the residual. Since this evaluation is needed during the online stage of the reduced basis method to certify the output, the computation of $\|\hat{r}_\delta(\mu)\|_{\mathbb{V}}$ must be efficient and at a cost independent of N_δ (the dimension of $\mathbb{V}_\delta$).

To achieve this, we take advantage of the affine decomposition (3.11) and (3.12) and expand the residual as

$$r(v_\delta;\mu) = \sum_{q=1}^{Q_{\mathtt{f}}} \theta_{\mathtt{f}}^q(\mu)\, f_q(v_\delta) - \sum_{q=1}^{Q_{\mathtt{a}}}\sum_{n=1}^{N} \theta_{\mathtt{a}}^q(\mu)\,(\mathtt{u}_{\mathtt{rb}}^\mu)_n\, a_q(\xi_n, v_\delta), \tag{4.21}$$

recalling that $u_{\mathrm{rb}}(\mu) = \sum_{n=1}^{N} (\mathrm{u}_{\mathrm{rb}}^{\mu})_n \, \xi_n$. Let us next introduce the coefficient vector $\mathrm{r}(\mu) \in \mathbb{R}^{Q_{\mathrm{r}}}$, with $Q_{\mathrm{r}} = Q_{\mathrm{f}} + Q_{\mathrm{a}} N$ terms, as

$$\mathrm{r}(\mu) = \left(\theta_{\mathrm{f}}^{1}(\mu), \ldots, \theta_{\mathrm{f}}^{Q_{\mathrm{f}}}(\mu), \, -(\mathrm{u}_{\mathrm{rb}}^{\mu})^T \, \theta_{\mathrm{a}}^{1}(\mu), \ldots, \, -(\mathrm{u}_{\mathrm{rb}}^{\mu})^T \, \theta_{\mathrm{a}}^{Q_{\mathrm{a}}}(\mu) \right)^T .$$

With a similar ordering, we define the vectors of forms $F \in (\mathbb{V}_{\delta}')^{Q_{\mathrm{f}}}$ and $A_q \in (\mathbb{V}_{\delta}')^{N}$ for $1 \le q \le Q_{\mathrm{a}}$ as

$$F = \big(f_1, \, \ldots \, , f_{Q_{\mathrm{f}}} \big), \qquad \text{and} \qquad A_q = \big(a_q(\xi_1, \, \cdot \,), \, \ldots \, , a_q(\xi_N, \, \cdot \,) \big),$$

and the vector of forms $R \in (\mathbb{V}_{\delta}')^{Q_{\mathrm{r}}}$ as

$$R = \big(F, A_1, \, \ldots \, , A_{Q_{\mathrm{a}}} \big)^T .$$

Combining (4.4) and (4.21) we obtain

$$(\hat{r}_{\delta}(\mu), v_{\delta})_{\mathbb{V}} = r(v_{\delta}; \mu) = \sum_{q=1}^{Q_{\mathrm{r}}} \mathrm{r}_q(\mu) \, R_q(v_{\delta}), \qquad \forall v_{\delta} \in \mathbb{V}_{\delta}.$$

Linear algebra box: Efficient computation of the residual norm

We show how the computation of $\|\hat{r}_{\delta}(\mu)\|_{\mathbb{V}}$ in (4.22) can be implemented.
We start with recalling the affine decomposition

$$\mathbf{A}_{\delta}^{\mu} = \sum_{q=1}^{Q_{\mathrm{a}}} \theta_{\mathrm{a}}^{q}(\mu) \, \mathbf{A}_{\delta}^{q}, \qquad \text{and} \qquad \mathrm{f}_{\delta}^{\mu}(v) = \sum_{q=1}^{Q_{\mathrm{f}}} \theta_{\mathrm{f}}^{q}(\mu) \, \mathrm{f}_{\delta}^{q}.$$

Define the matrix $\mathbf{R} \in \mathbb{R}^{N_{\delta} \times Q_{\mathrm{r}}}$ by $\mathbf{R}_{iq} = R_q(\varphi_i)$ for $1 \le i \le N_{\delta}$, $1 \le q \le Q_{\mathrm{r}}$, $Q_{\mathrm{r}} = Q_{\mathrm{f}} + Q_{\mathrm{a}} N$. This can be formed directly by

$$\mathbf{R} = (\mathrm{f}_{\delta}^{1}, \, \ldots \, , \mathrm{f}_{\delta}^{Q_{\mathrm{f}}}, \mathbf{A}_{\delta}^{1} \, \mathbf{B}, \, \ldots \, , \mathbf{A}_{\delta}^{Q_{\mathrm{a}}} \, \mathbf{B})^T,$$

where we recall that $\mathbf{B}_{:n}$ denotes the coefficient column vector of ξ_n in the basis $\{\varphi_i\}_{i=1}^{N_{\delta}}$ of $\mathbb{V}_{\delta}$. Then,

$$\mathbf{G} = \mathbf{R}^T \, \mathbf{M}_{\delta}^{-1} \, \mathbf{R} \in \mathbb{R}^{Q_{\mathrm{r}} \times Q_{\mathrm{r}}} \qquad \text{and} \qquad \|\hat{r}_{\delta}(\mu)\|_{\mathbb{V}} = \sqrt{\mathrm{r}(\mu)^T \, \mathbf{G} \, \mathrm{r}(\mu)},$$

with

$$\mathrm{r}(\mu) = \left(\theta_{\mathrm{f}}^{1}(\mu), \, \ldots \, , \theta_{\mathrm{f}}^{Q_{\mathrm{f}}}(\mu), \, -(\mathrm{u}_{\mathrm{rb}}^{\mu})^T \, \theta_{\mathrm{a}}^{1}(\mu), \, \ldots \, , \, -(\mathrm{u}_{\mathrm{rb}}^{\mu})^T \, \theta_{\mathrm{a}}^{Q_{\mathrm{a}}}(\mu) \right)^T .$$

Here $(\mathbf{M}_{\delta})_{ij} = (\varphi_j, \varphi_i)_{\mathbb{V}}$ is the mass matrix associated with the basis $\{\varphi_i\}_{i=1}^{N_{\delta}}$.

Denoting by $\hat{r}_\delta^q$ the Riesz representation of the form $R_q \in \mathbb{V}_\delta'$, i.e., $(\hat{r}_\delta^q, v_\delta)_\mathbb{V} = R_q(v_\delta)$ for all $v_\delta \in \mathbb{V}_\delta$ and $1 \le q \le Q_\mathrm{r}$, we recover

$$\hat{r}_\delta(\mu) = \sum_{q=1}^{Q_\mathrm{r}} \mathrm{r}_q(\mu)\, \hat{r}_\delta^q,$$

and

$$\|\hat{r}_\delta(\mu)\|_\mathbb{V}^2 = \sum_{q,q'=1}^{Q_\mathrm{r}} \mathrm{r}_q(\mu)\, \mathrm{r}_{q'}(\mu)\, (\hat{r}_\delta^q, \hat{r}_\delta^{q'})_\mathbb{V}. \tag{4.22}$$

In this final expression, the terms $(\hat{r}_\delta^q, \hat{r}_\delta^{q'})_\mathbb{V}$ can be precomputed once and for all in the offline stage of the method. Thus, given any $\mu \in \mathbb{P}$ one can compute $\|\hat{r}_\delta(\mu)\|_\mathbb{V}$ independently of N_δ by directly evaluating (4.22).

4.2.6 *Illustrative Examples 1 and 2: Heat Conduction and Linear Elasticity Part 3*

We present here the effectivities of the a posteriori estimators employed in the numerical tests that were presented in the illustrative examples in Sect. 3.4.

In these tests, we are considering a validation set $\mathbb{P}_h^\mathrm{v} \subset \mathbb{P}$ of 1,000 samples and we are tracking the average value of the error estimator

$$\eta_{\mathtt{en,av}} = \frac{1}{|\mathbb{P}_h^\mathrm{v}|} \sum_{\mu \in \mathbb{P}_h^\mathrm{v}} \eta_\mathtt{en}(\mu)$$

as a function of an increasing number of basis functions N. The maximal and average effectivities are measured by

$$\mathtt{eff}_{\mathtt{en,max}} = \max_{\mu \in \mathbb{P}_h^\mathrm{v}} \frac{\eta_\mathtt{en}(\mu)}{\|u_\delta(\mu) - u_\mathtt{rb}(\mu)\|_\mu} \quad \text{and} \quad \mathtt{eff}_{\mathtt{en,av}} = \frac{1}{|\mathbb{P}_h^\mathrm{v}|} \sum_{\mu \in \mathbb{P}_h^\mathrm{v}} \frac{\eta_\mathtt{en}(\mu)}{\|u_\delta(\mu) - u_\mathtt{rb}(\mu)\|_\mu}. \tag{4.23}$$

The error bounds and effectivity metrics as function of the number of basis functions N employed are then reported in Table 4.1 for both illustrative examples.

Table 4.1 Error bounds and effectivity metrics as function of N for the illustrative example 1 (*top*) and 2 (*bottom*)

N	$\eta_{\text{en,av}}$	$\texttt{eff}_{\text{en,max}}$	$\texttt{eff}_{\text{en,av}}$
1	3.57e−1	2.40	1.24
2	1.72e−3	2.30	1.64
3	8.91e−5	2.34	1.62
4	8.47e−7	20.24	1.68
N	$\eta_{\text{en,av}}$	$\texttt{eff}_{\text{en,max}}$	$\texttt{eff}_{\text{en,av}}$
5	2.08e−2	6.22	1.35
10	3.54e−3	5.06	1.17
15	8.16e−4	5.39	1.40
20	2.43e−4	5.26	1.33

4.3 The Stability Constant

A central feature of the certified reduced basis method is its ability to rigorously bound the error associated with the model reduction. However, as already discussed in Sect. 4.2.3, this step requires the computation of the (discrete) residual and an estimation of a lower bound for the parameter dependent stability constant, e.g., the coercivity or inf-sup constant. Since this must be done during the online phase, the computation of the stability constant must be accomplished with a computational complexity that is independent of the dimension of the underlying truth approximation space, $N_\delta = \dim(\mathbb{V}_\delta)$, to obtain a rigorous error certification in an efficient way. Accomplishing this is a central element of the entire approach and we shall discuss different strategies for this in the following

Let us start by recalling that the definition of the (discrete) coercivity constant is given by

$$\alpha_\delta(\mu) = \inf_{v_\delta \in \mathbb{V}_\delta} \frac{a(v_\delta, v_\delta; \mu)}{\|v_\delta\|_{\mathbb{V}}^2}.$$

The coercivity constant $\alpha_\delta(\mu)$ is thus defined as the smallest eigenvalue of a generalized eigenvalue-problem: Find $(\lambda, w_\delta) \in \mathbb{R}^+ \times \mathbb{V}_\delta$ such that

$$a(w_\delta, v_\delta; \mu) = \lambda(w_\delta, v_\delta)_{\mathbb{V}}, \quad \forall v_\delta \in \mathbb{V}_\delta. \tag{4.24}$$

We recall that if a Galerkin approximation is used for approximating a coercive problem, the underlying solution matrix $\mathbf{A}_\delta^\mu$ is symmetric positive definite.

In the following, we discuss three approaches of increasing generality and complexity for the computation of lower bounds for the coercivity constant $\alpha_\delta(\mu)$. The simplest one, the Min-θ-approach, applies only to a restricted family of problems: so called parametrically coercive problems. Still within this family of problems, the Min-θ-approach can be refined to provide sharper lower bounds, leading to the

multi-parameter Min-θ-approach. Finally, the general coercive case is handled by the Successive Constraint Method (SCM), which can be generalized to non-symmetric, possibly complex, matrices although this is not discussed in this text. We refer to [1–5] for further reading on the extension to general saddle-point problems.

Linear algebra box: The computation of the coercivity constant

Equation (4.24) can be stated on matrix form as: find $(\lambda, \mathrm{w}_\delta) \in \mathbb{R}^+ \times \mathbb{R}^{N_\delta}$ such that

$$\mathbf{A}_\delta^\mu \, \mathrm{w}_\delta = \lambda \, \mathbf{M}_\delta \, \mathrm{w}_\delta$$

where $(\mathbf{A}_\delta^\mu)_{ij} = a(\varphi_j, \varphi_i; \mu)$, $(\mathbf{M}_\delta)_{ij} = (\varphi_j, \varphi_i)_{\mathbb{V}}$ and $\mathrm{w}_\delta \in \mathbb{R}^{N_\delta}$ is the representation vector of the eigenfunctions w_δ in terms of the basis $\{\varphi_i\}_{i=1}^{N_\delta}$ of $\mathbb{V}_\delta$.

4.3.1 Min-θ-approach

We first recall the affine decomposition

$$a(u, v; \mu) = \sum_{q=1}^{Q_\mathrm{a}} \theta_\mathrm{a}^q(\mu) \, a_q(u, v).$$

Parametrically coercive problems are then characterized by

1. $\theta_\mathrm{a}^q(\mu) > 0$, $\forall \mu \in \mathbb{P}$, $q = 1, \ldots, Q_\mathrm{a}$.
2. $a_q(\cdot, \cdot) : \mathbb{V}_\delta \times \mathbb{V}_\delta \to \mathbb{R}$ is semi-positive definite for all $q = 1, \ldots, Q_\mathrm{a}$,

i.e., the bilinear form is a convex combination of semi-positive bilinear forms. Under this assumption and further assuming that the stability $\alpha_\delta(\mu')$ has been computed for a single parameter value $\mu' \in \mathbb{P}$, we observe the following identity

$$\begin{aligned}\alpha_\delta(\mu) &= \inf_{v_\delta \in \mathbb{V}_\delta} \frac{a(v_\delta, v_\delta; \mu)}{\|v_\delta\|_{\mathbb{V}_\delta}^2} = \inf_{v_\delta \in \mathbb{V}_\delta} \sum_{q=1}^{Q_\mathrm{a}} \theta_\mathrm{a}^q(\mu) \frac{a_q(v_\delta, v_\delta)}{\|v_\delta\|_{\mathbb{V}}^2} \\ &= \inf_{v_\delta \in \mathbb{V}_\delta} \sum_{q=1}^{Q_\mathrm{a}} \frac{\theta_\mathrm{a}^q(\mu)}{\theta_\mathrm{a}^q(\mu')} \theta_\mathrm{a}^q(\mu') \frac{a_q(v_\delta, v_\delta)}{\|v_\delta\|_{\mathbb{V}}^2}.\end{aligned}$$

We can derive a lower bound by

$$\alpha_\delta(\mu) \geq \inf_{v_\delta \in \mathbb{V}_\delta} \min_{q=1,\ldots,Q_\mathrm{a}} \frac{\theta_\mathrm{a}^q(\mu)}{\theta_\mathrm{a}^q(\mu')} \sum_{q=1}^{Q_\mathrm{a}} \theta_\mathrm{a}^q(\mu') \frac{a_q(v_\delta, v_\delta)}{\|v_\delta\|_{\mathbb{V}}^2}$$

$$= \min_{q=1,\dots,Q_\mathrm{a}} \frac{\theta_\mathrm{a}^q(\mu)}{\theta_\mathrm{a}^q(\mu')} \underbrace{\inf_{v_\delta \in \mathbb{V}_\delta} \sum_{q=1}^{Q_\mathrm{a}} \theta_\mathrm{a}^q(\mu') \frac{a_q(v_\delta, v_\delta)}{\|v_\delta\|_{\mathbb{V}}^2}}_{=\alpha_\delta(\mu')}$$

$$= \alpha_\delta(\mu') \min_{q=1,\dots,Q_\mathrm{a}} \frac{\theta_\mathrm{a}^q(\mu)}{\theta_\mathrm{a}^q(\mu')} =: \alpha_{\mathrm{LB}}(\mu).$$

While this approach provides a positive lower bound $\alpha_{\mathrm{LB}}(\mu)$ for $\alpha_\delta(\mu)$ it is generally not a sharp bound, possibly resulting in error bounds that are overly conservative. Note also that this lower bound provides the exact coercivity constant if applied at $\mu = \mu'$.

4.3.2 Multi-parameter Min-θ-approach

Remaining in the framework of parametrically coercive problems, the previous approach can be refined by defining a set of M parameter values $\mu_1, \dots, \mu_M$ for which the stability constant $\alpha_\delta(\mu_m)$ is computed from the lowest eigenvalue of (4.24). This is done during the offline phase.

During the online procedure, a lower bound for $\alpha_\delta(\mu)$ for any $\mu \in \mathbb{P}$ is needed. Following the same approach as for the Min-θ-approach for each μ_m, we observe

$$\alpha_\delta(\mu_m) \min_{q=1,\dots,Q_\mathrm{a}} \frac{\theta_\mathrm{a}^q(\mu)}{\theta_\mathrm{a}^q(\mu_m)}$$

is a guaranteed lower bound for all $m = 1, \dots, M$. Thus

$$\alpha_{\mathrm{LB}}(\mu) = \max_{m=1,\dots,M} \left(\alpha_\delta(\mu_m) \min_{q=1,\dots,Q_\mathrm{a}} \frac{\theta_\mathrm{a}^q(\mu)}{\theta_\mathrm{a}^q(\mu_m)} \right)$$

is the sharpest of the lower bounds of $\alpha_\delta(\mu)$ among all candidates. While it is a more accurate approach, it also requires more eigenvalue problems to be solved. Furthermore, this approach remains restricted to parametrically coercive problems. Note also that this lower bound interpolates the exact coercivity constant at the sample points $\mu = \mu_m$.

4.3.3 Illustrative Example 1: Heat Conduction Part 4

Let us illustrate the multi-parameter Min-θ-approach in the case of the Illustrative Example 1: Heat Conduction. We recall that the affine decomposition holds for $Q_\mathrm{a} = 2$ with $\theta_\mathrm{a}^1(\mu) = 1$, $\theta_\mathrm{a}^2(\mu) = \mu_{[1]}$ and

$$a_1(w, v) = \int_{\Omega_1} \nabla w \cdot \nabla v, \qquad a_2(w, v) = \int_{\Omega_0} \nabla w \cdot \nabla v.$$

We therefore recognize that this problem is parametrically coercive and the multi-parameter Min-θ-approach is applicable.

We construct one Min-θ lower bound based on $\mu' = 1$ given by

$$\alpha_{\mathrm{LB}}^{\theta}(\mu) = \alpha_\delta(\mu') \min_{q=1,\ldots,Q_\mathrm{a}} \frac{\theta_\mathrm{a}^q(\mu)}{\theta_\mathrm{a}^q(\mu')} = \alpha_\delta(\mu') \min(1, \mu),$$

and a multi-parameter Min-θ lower bound based on the sample points

$$(\mu_1, \ldots, \mu_5) = (10^{-1}, 10^{-1/2}, 1, 10^{1/2}, 10)$$

as

$$\alpha_{\mathrm{LB}}^{\theta,\mathrm{MP}}(\mu) = \max_{m=1,\ldots,M} \left(\alpha_\delta(\mu_m) \min_{q=1,\ldots,Q_\mathrm{a}} \frac{\theta_\mathrm{a}^q(\mu)}{\theta_\mathrm{a}^q(\mu_m)} \right) = \max_{m=1,\ldots,5} \left(\alpha_\delta(\mu_m) \min\left(1, \frac{\mu}{\mu_m}\right) \right).$$

The results of these lower bounds as well as the value of the coercivity constant itself are illustrated in Fig. 4.1. We note in particular that the exact discrete coercivity constant is recovered by the lower bounds at the sample points.

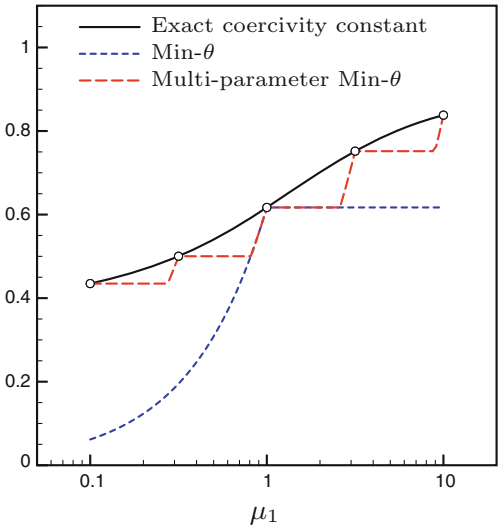

Fig. 4.1 Illustration of the Min-θ and the multi-parameter Min-θ-approach for the Illustrative Example 1

4.3.4 The Successive Constraint Method (SCM)

To address the challenges associated with the need to estimate $\alpha_{\mathrm{LB}}(\mu)$ for more general problems, Huynh et al. [6] proposed a local minimization approach, known as the successive constraint method (SCM). As for the multi-parameter Min-θ-approach, the SCM is an offline/online procedure where generalized eigenvalue problems of size N_δ need to be solved during the offline phase. The online part is then reduced to provide a lower bound $\alpha_{\mathrm{LB}}(\mu)$ of the coercivity constant $\alpha_\delta(\mu)$ for each new parameter value $\mu \in \mathbb{P}$ with an operation count that is independent of the dimension N_δ. The original algorithm [6] was subsequently refined, extended to non-coercive problems and generalized to complex matrices in [1–5].

4.3.4.1 Offline Procedure of SCM

We recall that the coercivity constant can be written as

$$\alpha_\delta(\mu) = \inf_{v_\delta \in \mathbb{V}_\delta} \sum_{q=1}^{Q_{\mathrm{a}}} \theta_{\mathrm{a}}^q(\mu) \frac{a_q(v_\delta, v_\delta)}{\|v_\delta\|_{\mathbb{V}}^2}, \tag{4.25}$$

using the affine decomposition. The key idea of the SCM is to express the right hand side of (4.25) as a minimization problem of the functional

$$\begin{aligned} \mathtt{S} : \mathbb{P} \times \mathbb{R}^{Q_{\mathrm{a}}} &\longrightarrow \mathbb{R} \\ (\mu, \mathrm{y}) &\longmapsto \mathtt{S}(\mu, \mathrm{y}) = \sum_{q=1}^{Q_{\mathrm{a}}} \theta_{\mathrm{a}}^q(\mu)\, y_q \end{aligned}$$

over the set of admissible solutions

$$\mathcal{Y} = \left\{ \mathrm{y} = (y_1, \ldots, y_{Q_{\mathrm{a}}}) \in \mathbb{R}^{Q_{\mathrm{a}}} \;\middle|\; \exists\, v_\delta \in \mathbb{V}_\delta \text{ s.t. } y_q = \frac{a_q(v_\delta, v_\delta)}{\|v_\delta\|_{\mathbb{V}}^2},\ 1 \le q \le Q_{\mathrm{a}} \right\}.$$

Then, we can equivalently write

$$\alpha_\delta(\mu) = \min_{\mathrm{y} \in \mathcal{Y}} \mathtt{S}(\mu, \mathrm{y})$$

and a lower and upper bound can be found by enlarging or restricting the admissible set of solution vectors y. This is done by introducing $\mathcal{Y}_{\mathrm{UB}} \subset \mathcal{Y} \subset \mathcal{Y}_{\mathrm{LB}}$ and defining

$$\alpha_{\mathrm{LB}}(\mu) = \min_{\mathrm{y} \in \mathcal{Y}_{\mathrm{LB}}} \mathtt{S}(\mu, \mathrm{y}), \qquad \text{and} \qquad \alpha_{\mathrm{UB}}(\mu) = \min_{\mathrm{y} \in \mathcal{Y}_{\mathrm{UB}}} \mathtt{S}(\mu, \mathrm{y}).$$

The remaining question is how to efficiently design the spaces $\mathcal{Y}_{\mathrm{UB}}$ and $\mathcal{Y}_{\mathrm{LB}}$ to ensure that any target accuracy for the error quantity

$$1 - \frac{\alpha_{\mathrm{LB}}(\mu)}{\alpha_{\mathrm{UB}}(\mu)},$$

can be achieved.

Denote by $\mathbb{P}_{\mathtt{a}}$ the restriction of $\mathbb{P}$ to the set of actively varying parameters of the bilinear form a. The offline part of the SCM is based on an greedy approach where the n-th iteration of the offline procedure is initiated by assuming that

1. We know the coercivity constants $\alpha_\delta(\mu_j), 1 \le j \le n$, for some parameter values $\mathbb{C}_n = \{\mu_1, \ldots, \mu_n\} \subset \mathbb{P}_{\mathtt{a}}$.
2. Let $\Xi_{\mathtt{a}} \subset \mathbb{P}_{\mathtt{a}}$ be a representative finite point-set discretization of $\mathbb{P}_{\mathtt{a}}$. For each $\mu \in \Xi_{\mathtt{a}}$, we have some lower bound $\alpha_{\mathrm{LB}}^{n-1}(\mu) \ge 0$ of $\alpha_\delta(\mu)$ from the previous iteration. For $n = 1$, set $\alpha_{\mathrm{LB}}^0(\mu) = 0$ for all $\mu \in \Xi_{\mathtt{a}}$.

The eigensolutions $(\alpha_\delta(\mu_j), w_\delta^j) \in \mathbb{R}^+ \times \mathbb{V}_\delta$ are solutions to the generalized eigenvalue problem

$$a(w_\delta^j, v_\delta; \mu_j) = \alpha_\delta(\mu_j)(w_\delta^j, v_\delta)_{\mathbb{V}}, \quad \forall v_\delta \in \mathbb{V}_\delta, \tag{4.26}$$

where $\alpha_\delta(\mu_j)$ is the smallest eigenvalues for each j and w_δ^j the corresponding eigenfunction. The collection of eigenfunctions $\{w_\delta^j\}_{j=1}^n$ provide the corresponding vectors $\{\mathrm{y}^j\}_{j=1}^n$ by

$$(\mathrm{y}^j)_q = \frac{a_q(w_\delta^j, w_\delta^j)}{\|w_\delta^j\|_{\mathbb{V}}^2}, \quad 1 \le q \le Q_{\mathtt{a}},\ 1 \le j \le n,$$

where $(\mathrm{y}^j)_q$ denotes the q-th coefficient of $\mathrm{y}^j \in \mathbb{R}^{Q_{\mathtt{a}}}$. We set

$$\mathcal{Y}_{\mathrm{UB}}^n = \left\{ \mathrm{y}^j \;\middle|\; 1 \le j \le n \right\},$$

which is clearly a subset of $\mathcal{Y}$. For $\mathcal{Y}_{\mathrm{UB}}^n$ we therefore use this finite set of precomputed vectors y^j. Indeed, computing $\alpha_{\mathrm{UB}}^n(\mu) = \min_{\mathrm{y} \in \mathcal{Y}_{\mathrm{UB}}^n} \mathtt{S}(\mu, \boldsymbol{y})$ consists of forming the functional

$$\mathtt{S}(\mu, \mathrm{y}^j) = \sum_{q=1}^{Q_{\mathtt{a}}} \theta_{\mathtt{a}}^q(\mu)\, (\mathrm{y}^j)_q$$

for the different vectors y^j and then choosing the smallest value of the functional. This is clearly independent of N_δ once the vectors y^j have been built.

For $\mathcal{Y}_{\mathrm{LB}}$ we define first a rectangular box $\mathcal{B} = \prod_{q=1}^{Q_{\mathrm{a}}} [\sigma_q^-, \sigma_q^+] \subset \mathbb{R}^{Q_{\mathrm{a}}}$ containing $\mathcal{Y}$ by setting

$$\sigma_q^- = \inf_{v_\delta \in \mathbb{V}_\delta} \frac{a_q(v_\delta, v_\delta)}{\|v_\delta\|_{\mathbb{V}}^2} \quad \text{and} \quad \sigma_q^+ = \sup_{v_\delta \in \mathbb{V}_\delta} \frac{a_q(v_\delta, v_\delta)}{\|v_\delta\|_{\mathbb{V}}^2}.$$

This corresponds to computing the smallest and the largest eigenvalues of a generalized eigenvalue problem for each $a_q(\cdot, \cdot)$ and can be computed once at the beginning of the SCM algorithm. To ensure that the set $\mathcal{Y}_{\mathrm{LB}}$ is as small as possible while containing $\mathcal{Y}$, we impose some additional restrictions, which result in sharper lower bounds. These constraints depend on the value of the actual parameter μ and we distinguish between two types:

1. Constraints based on the exact eigenvalues for some close parameter values among the set $\mathbb{C}_n$.
2. Constraints based on the previous lower bounds $\alpha_{\mathrm{LB}}^{n-1}$ for some neighbor parameter values.

Observe that, in contrast to $\mathcal{Y}_{\mathrm{UB}}$, the space $\mathcal{Y}_{\mathrm{LB}}$ will change with variation of the parameter μ as the constraints change with μ, reflected by the notation $\mathcal{Y}_{\mathrm{LB}}(\mu)$ in the following.

Next, we introduce the function that provides close parameter values

$$\mathbb{P}_M(\mu; E) = \begin{cases} M \text{ closest points to } \mu \text{ in } E & \text{if } \mathrm{card}(E) > M, \\ E & \text{if } \mathrm{card}(E) \leq M, \end{cases}$$

for either $E = \mathbb{C}_n$ or $E = \Xi_{\mathrm{a}}$. For some M_{e} and M_{p}, we define

$$\mathcal{Y}_{\mathrm{LB}}^n(\mu) = \Big\{ \mathrm{y} \in \mathcal{B} \,\Big|\, \mathrm{S}(\mu', \mathrm{y}) \geq \alpha_\delta(\mu'), \quad \forall \mu' \in \mathbb{P}_{M_{\mathrm{e}}}(\mu; \mathbb{C}_n), \\ \mathrm{S}(\mu', \mathrm{y}) \geq \alpha_{\mathrm{LB}}^{n-1}(\mu'), \quad \forall \mu' \in \mathbb{P}_{M_{\mathrm{p}}}(\mu; \Xi_{\mathrm{a}} \backslash \mathbb{C}_n) \Big\}.$$

It can be shown that $\mathcal{Y}_{\mathrm{UB}}^n \subset \mathcal{Y} \subset \mathcal{Y}_{\mathrm{LB}}^n(\mu)$ (see [6] for the proof). Consequently, $\mathcal{Y}_{\mathrm{UB}}^n$, $\mathcal{Y}$ and $\mathcal{Y}_{\mathrm{LB}}^n(\mu)$ are nested as

$$\mathcal{Y}_{\mathrm{UB}}^1 \subset \mathcal{Y}_{\mathrm{UB}}^2 \subset \cdots \subset \mathcal{Y}_{\mathrm{UB}}^n \subset \cdots \subset \mathcal{Y} \subset \cdots \subset \mathcal{Y}_{\mathrm{LB}}^n(\mu) \subset \cdots \subset \mathcal{Y}_{\mathrm{LB}}^2(\mu) \subset \mathcal{Y}_{\mathrm{LB}}^1(\mu).$$

Note that finding $\alpha_{\mathrm{LB}}^n(\mu) = \min_{\boldsymbol{y} \in \mathcal{Y}_{\mathrm{LB}}^n(\mu)} \mathrm{S}(\mu, \boldsymbol{y})$ corresponds to a linear programming problem of Q_{a} design variables and $2Q_{\mathrm{a}} + M_{\mathrm{e}} + M_{\mathrm{p}}$ conditions. The complexity of the linear programming problem is thus independent of N_δ.

Having defined the two sets $\mathcal{Y}_{\mathrm{LB}}^n(\mu)$ and $\mathcal{Y}_{\mathrm{UB}}^n$, we can define a greedy selection to enrich the space $\mathbb{C}_n$ and build $\mathbb{C}_{n+1}$ at all stages of n. The algorithm is outlined in the algorithm box Offline-procedure of the SCM.

Algorithm: Offline-procedure of the SCM

Input: An error tolerance $\texttt{Tol}$, some initial set $\mathbb{C}_1 = \{\mu_1\}$ and $n = 1$.
Output: The sample points $\mathbb{C}_N = \{\mu_1, \ldots, \mu_N\}$, the corresponding coercivity constants $\alpha_\delta(\mu_n)$ and vectors y^n, $n = 1, \ldots, N$, as well as the lower bounds $\alpha_{\mathrm{LB}}^N(\mu)$ for all $\mu \in \Xi_{\mathrm{a}}$.

1. For each $\mu \in \Xi_{\mathrm{a}}$:
   a. Compute the upper bound $\alpha_{\mathrm{UB}}^n(\mu) = \min_{\mathrm{y} \in \mathcal{Y}_{\mathrm{UB}}^n} \mathrm{S}(\mu, \mathrm{y})$.
   b. Compute the lower bound $\alpha_{\mathrm{LB}}^n(\mu) = \min_{\mathrm{y} \in \mathcal{Y}_{\mathrm{LB}}^n(\mu)} \mathrm{S}(\mu, \mathrm{y})$.
   c. Define the error estimate $\eta(\mu; \mathbb{C}_n) = 1 - \frac{\alpha_{\mathrm{LB}}^n(\mu)}{\alpha_{\mathrm{UB}}^n(\mu)}$.
2. Select $\mu_{n+1} = \mathrm{argmax}_{\mu \in \mathbb{P}} \eta(\mu; \mathbb{C}_n)$ and set $\mathbb{C}_{n+1} = \mathbb{C}_n \cup \{\mu_{n+1}\}$.
3. If $\max_{\mu \in \mathbb{P}} \eta(\mu; \mathbb{C}_n) \leq \texttt{Tol}$, **terminate**.
4. Solve the generalized eigenvalue problem (4.26) associated with μ_{n+1}, store $\alpha_\delta(\mu_{n+1})$, y^{n+1}.
5. Set $n := n + 1$ and **goto 1.**

4.3.4.2 Online Procedure of SCM

Once the offline-procedure is completed, we denote $\mathcal{Y}_{\mathrm{LB}}^n(\mu)$ by $\mathcal{Y}_{\mathrm{LB}}(\mu)$ and $\mathcal{Y}_{\mathrm{UB}}^n$ by $\mathcal{Y}_{\mathrm{UB}}$.

For an arbitrary parameter value $\mu \in \mathbb{P}$, we can compute a lower bound $\alpha_{\mathrm{LB}}(\mu)$ by only retaining the information about $\alpha_\delta(\mu)$ for all $\mu \in \mathbb{C}_n$ and $\alpha_{\mathrm{LB}}(\mu)$ for all $\mu \in \Xi_{\mathrm{a}}$: For any new $\mu \in \mathbb{P}$, find the solution of

$$\alpha_{\mathrm{LB}}(\mu) = \min_{\mathrm{y} \in \mathcal{Y}_{\mathrm{LB}}(\mu)} \mathrm{S}(\mu, \mathrm{y}),$$

which consists again of a linear program with Q_{a} design variables and $2Q_{\mathrm{a}} + M_{\mathrm{e}} + M_{\mathrm{p}}$ constraints. Note that we now consider, during the online stage, any parameter $\mu \in \mathbb{P}$ not necessarily contained in Ξ_{a}. This implies that we must add the additional constraint that $\mathrm{S}(\mu, \mathrm{y}) \geq 0$. Further, note that $1 - \frac{\alpha_{\mathrm{LB}}(\mu)}{\alpha_{\mathrm{UB}}(\mu)}$ still provides an indicator of the sharpness of the bounds that can be evaluated a posteriori.

4.3.4.3 Numerical Results

We refer to Sects. 6.3.1 and 6.5 in Chap. 6 where we employ the SCM with complete numerical examples. Within these examples, we illustrate the convergence of the SCM-greedy algorithm with respect to the number of solved eigenvalue problems.

4.3.5 A Comparitive Discussion

In this section we compare the lower bounds obtained by the Min-θ-approach and the multi-parameter Min-θ-approach with the ones obtained by the SCM in the parametrically coercive case. We assume throughout this section that the problem is parametrically coercive.

For a given $\mu \in \mathbb{P}_{\mathtt{a}}$, denote by $\hat{q}$ the (or an) index such that

$$\frac{\theta_{\mathtt{a}}^{\hat{q}}(\mu)}{\theta_{\mathtt{a}}^{\hat{q}}(\mu')} = \min_{q=1,\ldots,Q_{\mathtt{a}}} \frac{\theta_{\mathtt{a}}^{q}(\mu)}{\theta_{\mathtt{a}}^{q}(\mu')},$$

and observe that the lower bound provided by the Min-θ-approach is provided by

$$\alpha_{\mathrm{LB}}^{\theta}(\mu) = \alpha_{\delta}(\mu') \frac{\theta_{\mathtt{a}}^{\hat{q}}(\mu)}{\theta_{\mathtt{a}}^{\hat{q}}(\mu')}.$$

On the other hand, consider the SCM with $\mathbb{C}_1 = \{\mu'\}$, $M_{\mathtt{p}} = 0$ and denote the corresponding minimization space used for the lower bound as

$$\mathcal{Y}_{\mathrm{LB}}^{\mu'}(\mu) = \left\{ \mathrm{y} \in \mathcal{B} \,\middle|\, \mathtt{S}(\mu', \mathrm{y}) \geq \alpha_{\delta}(\mu') \right\}.$$

Then, the following lemma holds.

Lemma 4.6 *For parametrically coercive problems, consider the Min-θ-approach based upon the computation of $\alpha_{\delta}(\mu')$ and the lower bound of the SCM based upon $\mathcal{Y}_{\mathrm{LB}}^{\mu'}(\mu)$. Then, the Min-$\theta$ lower bound $\alpha_{\mathrm{LB}}^{\theta}(\mu)$ is at most as sharp as the lower bound provided by this SCM, i.e.,*

$$\min_{\mathrm{y} \in \mathcal{Y}_{\mathrm{LB}}^{\mu'}(\mu)} \mathtt{S}(\mu, \mathrm{y}) \geq \alpha_{\mathrm{LB}}^{\theta}(\mu).$$

Proof For any $\mathrm{y} \in \mathcal{Y}_{\mathrm{LB}}^{\mu'}(\mu)$ it holds that

$$\mathtt{S}(\mu', \mathrm{y}) \geq \alpha_{\delta}(\mu').$$

Multiplying by $\frac{\theta_{\mathtt{a}}^{\hat{q}}(\mu)}{\theta_{\mathtt{a}}^{\hat{q}}(\mu')}$ yields

$$\frac{\theta_{\mathtt{a}}^{\hat{q}}(\mu)}{\theta_{\mathtt{a}}^{\hat{q}}(\mu')} \mathtt{S}(\mu', \mathrm{y}) \geq \alpha_{\mathrm{LB}}^{\theta}(\mu).$$

Additionally we obtain

$$\begin{aligned}\mathtt{S}(\mu', \mathrm{y}) &= \sum_{q=1}^{Q_\mathtt{a}} \theta_\mathtt{a}^q(\mu) \frac{\theta_\mathtt{a}^q(\mu')}{\theta_\mathtt{a}^q(\mu)} \mathrm{y}_q \leq \max_{q=1,\ldots,Q_\mathtt{a}} \frac{\theta_\mathtt{a}^q(\mu')}{\theta_\mathtt{a}^q(\mu)} \sum_{q=1}^{Q_\mathtt{a}} \theta_\mathtt{a}^q(\mu) \mathrm{y}_q \\ &= \max_{q=1,\ldots,Q_\mathtt{a}} \frac{\theta_\mathtt{a}^q(\mu')}{\theta_\mathtt{a}^q(\mu)} \mathtt{S}(\mu, \mathrm{y}).\end{aligned}$$

Now, since

$$\max_{q=1,\ldots,Q_\mathtt{a}} \frac{\theta_\mathtt{a}^q(\mu')}{\theta_\mathtt{a}^q(\mu)} = \frac{1}{\min_{q=1,\ldots,Q_\mathtt{a}} \frac{\theta_\mathtt{a}^q(\mu)}{\theta_\mathtt{a}^q(\mu')}} = \frac{\theta_\mathtt{a}^{\hat{q}}(\mu')}{\theta_\mathtt{a}^{\hat{q}}(\mu)},$$

the result follows. □

Proposition 4.7 *For parametrically coercive problems, consider the Min-θ-approach based upon the computation of $\alpha_\delta(\mu')$ and the SCM assuming that $\mu' \in \mathbb{C}_n$. Then, the Min-θ lower bound is at most as sharp as the lower bound provided by the SCM based upon $\mathbb{C}_n$ and any $M_\mathtt{p} \geq 0$, i.e.,*

$$\alpha_{\mathrm{LB}}^{\mathtt{SCM}}(\mu) \geq \alpha_{\mathrm{LB}}^{\theta}(\mu).$$

Proof It is easy to see that $\boldsymbol{\mathcal{Y}}_{\mathrm{LB}}(\mu) \subset \boldsymbol{\mathcal{Y}}_{\mathrm{LB}}^{\mu'}(\mu)$ as long as $\mu' \in \mathbb{C}_n$, which is assumed here, and therefore

$$\alpha_{\mathrm{LB}}^{\mathtt{SCM}}(\mu) = \min_{\mathrm{y} \in \boldsymbol{\mathcal{Y}}_{\mathrm{LB}}(\mu)} \mathtt{S}(\mu, \mathrm{y}) \geq \min_{\mathrm{y} \in \boldsymbol{\mathcal{Y}}_{\mathrm{LB}}^{\mu'}(\mu)} \mathtt{S}(\mu, \mathrm{y}) \geq \alpha_{\mathrm{LB}}^{\theta}(\mu),$$

by Lemma 4.6. □

We turn now our attention to the comparison with the multi-parameter Min-θ-approach. For a given $\mu \in \mathbb{P}_\mathtt{a}$, let $\hat{q}$ and μ' denote an index and snapshot parameter on which the multi-parameter Min-θ-approach lower bound is based, i.e.,

$$\alpha_{\mathrm{LB}}^{\theta,\mathtt{mp}}(\mu) = \alpha_\delta(\mu') \frac{\theta_\mathtt{a}^{\hat{q}}(\mu)}{\theta_\mathtt{a}^{\hat{q}}(\mu')} = \max_{m=1,\ldots,M} \left(\alpha_\delta(\mu_m) \min_{q=1,\ldots,Q_\mathtt{a}} \frac{\theta_\mathtt{a}^q(\mu)}{\theta_\mathtt{a}^q(\mu_m)} \right).$$

Denote $\boldsymbol{\mu} = \{\mu_1, \ldots, \mu_M\}$ and define

$$\boldsymbol{\mathcal{Y}}_{\mathrm{LB}}^{\boldsymbol{\mu}}(\mu) = \left\{ \mathrm{y} \in \boldsymbol{\mathcal{B}} \,\middle|\, \mathtt{S}(\mu_m, \mathrm{y}) \geq \alpha_\delta(\mu_m),\ \forall\, 1 \leq m \leq M \right\}.$$

Then, the following statement holds.

Lemma 4.8 *For parametrically coercive problems, consider the multi-parameter Min-θ-approach based upon the computation of $\alpha_\delta(\mu_1), \ldots, \alpha_\delta(\mu_M)$ and the SCM*

based upon $\mathcal{Y}^{\boldsymbol{\mu}}_{\mathrm{LB}}(\mu)$. *Then, the multi-parameter Min-*θ *lower bound* $\alpha^{\theta,\mathrm{mp}}_{\mathrm{LB}}(\mu)$ *is at most as sharp as the lower bound provided by this SCM, i.e.,*

$$\min_{\mathrm{y}\in\mathcal{Y}^{\boldsymbol{\mu}}_{\mathrm{LB}}(\mu)} \mathrm{S}(\mu,\mathrm{y}) \geq \alpha^{\theta,\mathrm{mp}}_{\mathrm{LB}}(\mu).$$

Proof For any $\mathrm{y} \in \mathcal{Y}^{\boldsymbol{\mu}}_{\mathrm{LB}}(\mu)$ it holds that

$$\mathrm{S}(\mu',\mathrm{y}) \geq \alpha_\delta(\mu').$$

Multiplying by $\frac{\theta^{\hat{q}}_{\mathrm{a}}(\mu)}{\theta^{\hat{q}}_{\mathrm{a}}(\mu')}$ yields

$$\frac{\theta^{\hat{q}}_{\mathrm{a}}(\mu)}{\theta^{\hat{q}}_{\mathrm{a}}(\mu')}\mathrm{S}(\mu',\mathrm{y}) \geq \alpha^{\theta,\mathrm{mp}}_{\mathrm{LB}}(\mu).$$

Additionally we easily obtain, as in the proof of Lemma 4.6,

$$\mathrm{S}(\mu',\mathrm{y}) = \sum_{q=1}^{Q_{\mathrm{a}}} \theta^q_{\mathrm{a}}(\mu)\frac{\theta^q_{\mathrm{a}}(\mu')}{\theta^q_{\mathrm{a}}(\mu)}\,\mathrm{y}_q \leq \max_{m=1,\ldots,M}\max_{q=1,\ldots,Q_{\mathrm{a}}} \frac{\theta^q_{\mathrm{a}}(\mu_m)}{\theta^q_{\mathrm{a}}(\mu)}\mathrm{S}(\mu,\mathrm{y}).$$

Now, since

$$\max_{m=1,\ldots,M}\max_{q=1,\ldots,Q_{\mathrm{a}}} \frac{\theta^q_{\mathrm{a}}(\mu')}{\theta^q_{\mathrm{a}}(\mu)} = \frac{1}{\min_{m=1,\ldots,M}\min_{q=1,\ldots,Q_{\mathrm{a}}} \frac{\theta^q_{\mathrm{a}}(\mu)}{\theta^q_{\mathrm{a}}(\mu')}} = \frac{\theta^{\hat{q}}_{\mathrm{a}}(\mu')}{\theta^{\hat{q}}_{\mathrm{a}}(\mu)},$$

the result follows. □

Finally we conclude with the following result.

Proposition 4.9 *For parametrically coercive problems, consider the multi-parameter Min-*θ*-approach based upon the computation of* $\alpha_\delta(\mu_1),\ldots,\alpha_\delta(\mu_M)$ *and the SCM assuming that* $\mu_m \in \mathbb{C}_n$ *for all* $1 \leq m \leq M$. *Then, the multi-parameter Min-*θ *lower bound is at most as sharp as the lower bound provided by the SCM based upon* $\mathbb{C}_n$ *and any* $M_{\mathrm{p}} > 0$, *i.e.,*

$$\alpha^{\mathrm{SCM}}_{\mathrm{LB}}(\mu) \geq \alpha^{\theta}_{\mathrm{LB}}(\mu).$$

Proof It is easy to see that $\mathcal{Y}_{\mathrm{LB}}(\mu) \subset \mathcal{Y}^{\boldsymbol{\mu}}_{\mathrm{LB}}(\mu)$ as long as $\mu_m \in \mathbb{C}_n$ for all $1 \leq m \leq M$, which is assumed here, and therefore

$$\alpha^{\mathrm{SCM}}_{\mathrm{LB}}(\mu) = \min_{\mathrm{y}\in\mathcal{Y}_{\mathrm{LB}}(\mu)} \mathrm{S}(\mu,\mathrm{y}) \geq \min_{\mathrm{y}\in\mathcal{Y}^{\mu'}_{\mathrm{LB}}(\mu)} \mathrm{S}(\mu,\mathrm{y}) \geq \alpha^{\theta}_{\mathrm{LB}}(\mu),$$

by Lemma 4.8. □

We end this comparative discussion by noting that the parameter values for which the coercivity constant is computed is automatically detected in the case of the SCM, while needs to be specified a priori by the user in the case of the (multi-parameter) Min-θ-approach.

References

1. Y. Chen, J.S. Hesthaven, Y. Maday, J. Rodríguez, A monotonic evaluation of lower bounds for inf-sup stability constants in the frame of reduced basis approximations. C.R. Math. **346**, 1295–1300 (2008)
2. Y. Chen, J.S. Hesthaven, Y. Maday, J. Rodríguez, Improved successive constraint method based a posteriori error estimate for reduced basis approximation of 2d Maxwell's problem. ESAIM. Math. Model. Numer. Anal. **43**, 1099–1116 (2009)
3. Y. Chen, J.S. Hesthaven, Y. Maday, J. Rodríguez, Certified reduced basis methods and output bounds for the harmonic Maxwell's equations. SIAM J. Sci. Comput. **32**, 970–996 (2010)
4. J.S. Hesthaven, B. Stamm, S. Zhang, Certified reduced basis method for the electric field integral equation. SIAM J. Sci. Comput. **34**, A1777–A1799 (2012)
5. P. Huynh, D. Knezevic, Y. Chen, J.S. Hesthaven, A. Patera, A natural-norm successive constraint method for inf-sup lower bounds. Comput. Methods Appl. Mech. Eng. **199**, 1963–1975 (2010)
6. P. Huynh, G. Rozza, S. Sen, A.T. Patera, A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants. C.R. Math. **345**, 473–478 (2007)

Chapter 5
The Empirical Interpolation Method

5.1 Motivation and Historical Overview

As discussed previously, the computational efficiency of the reduced basis method relies strongly on the affine assumption, i.e., we generally assume that

$$a(w, v; \mu) = \sum_{q=1}^{Q_a} \theta_a^q(\mu)\, a_q(w, v), \qquad \forall w, v \in \mathbb{V},\ \forall \mu \in \mathbb{P}, \tag{5.1}$$

and similarly for the righthand side and the output of interest. Unfortunately, this assumption fails for the majority of problems one would like to consider and it is essential to look for ways to overcome this assumption by approximating the non-affine elements in a suitable way.

This is the key motivation behind the development of the Empirical Interpolation Method (EIM) which seeks to approximate a general parametrized function by a sum of affine terms, i.e., on the form

$$f(x, y) \approx \sum_{q=1}^{Q} g_q(x)\, h_q(y).$$

Such a decomposition often allows to establish an affine decomposition of the form (5.1). As we shall discuss later, this is a powerful idea that allows for substantial extensions of the applicability of reduced basis methods, including to nonlinear problems.

The central idea of EIM was presented first in [1] and applied in the context of reduced order modeling in [2]. In [3] a broader set of applications of the EIM approach is discussed and an a posteriori error analysis is presented in [2, 4, 5]. Extensions to hp-adpative EIM is discussed [6] with an additional emphasis on high-dimensional parameter spaces being discussed in [7]. In [8, 9], the authors introduce the discrete

J.S. Hesthaven et al., *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, SpringerBriefs in Mathematics,
DOI 10.1007/978-3-319-22470-1_5

EIM (DEIM) as a special case of EIM. In this approach, the function is given in terms of a finite set of vectors to allow for a simple way to build reduced order models for nonlinear problems. A theoretical analysis of EIM is offered in [3] and a more general family of EIM was recently introduced in [5, 10, 11]. A non-intrusive version is developed in [12]. An overview illustrating the connection with other low rank approximation techniques can be found in [13].

5.2 The Empirical Interpolation Method

In a general setting, EIM is designed to approximate functions $g : \Omega \times \mathbb{P}_{\texttt{EIM}} \to \mathbb{R}$ in situations where each function $g_\mu := g(\cdot, \mu)$, $\mu \in \mathbb{P}_{\texttt{EIM}}$, belongs to some Banach space $\mathcal{X}_\Omega$ and $\mathbb{P}_{\texttt{EIM}}$ denotes the corresponding parameter space. The approximation is obtained through an interpolation operator $\texttt{I}_q$ that interpolates the function $g(\cdot, \mu)$ at some particular interpolation points $x_1, \ldots, x_Q \in \Omega$ as a linear combination of some carefully chosen basis functions $\{h_1, \ldots, h_Q\}$. These basis functions do not consist of multi-purpose basis functions such as polynomials or trigonometric functions but belong to the particular family $\{g_\mu\}_{\mu \in \mathbb{P}_{\texttt{EIM}}}$ related to the problem being considered. They are built empirically by means of linear combinations of Q snapshots $g_{\mu_1}, \ldots, g_{\mu_Q}$ where the sample points $\mu_1, \ldots, \mu_Q \in \mathbb{P}_{\texttt{EIM}}$ are chosen using a greedy approach.

Since an interpolation process requires point-wise evaluations of the functions we assume that each function g_μ belongs to $C^0(\overline{\Omega})$, thus $C^0(\overline{\Omega}) \subset \mathcal{X}_\Omega$. The interpolant $\texttt{I}_Q[g_\mu]$ of g_μ for $\mu \in \mathbb{P}_{\texttt{EIM}}$ is expressed as

$$\texttt{I}_Q[g_\mu](x) = \sum_{q=1}^{Q} c_q(\mu)\, h_q(x), \quad x \in \Omega, \tag{5.2}$$

and defined by the interpolation statement

$$\texttt{I}_Q[g_\mu](x_j) = g_\mu(x_j), \quad j = 1, \ldots, Q. \tag{5.3}$$

The interpolation is recovered by solving the following linear system

$$\sum_{q=1}^{Q} c_q(\mu)\, h_j(x_j) = g_\mu(x_j), \quad j = 1, \ldots, Q,$$

expressed as $\mathbf{T}\, \mathrm{c}_\mu = \mathrm{g}_\mu$ with Q unknowns and

$$\mathbf{T}_{ij} = h_j(x_i), \quad (\mathrm{c}_\mu)_j = c_j(\mu), \quad (\mathrm{g}_\mu)_i = g_\mu(x_i), \quad i, j = 1, \ldots, Q. \tag{5.4}$$

The remaining question is how to determine the basis functions $\{h_1, \ldots, h_Q\}$ and the interpolation points $x_1, \ldots, x_Q$ and ensure that the system is uniquely solvable, i.e., that the interpolation matrix $\mathbf{T}_{ij} = h_j(x_i)$ is invertible.

As already mentioned, the basis functions are chosen as linear combinations of some selected snapshots $g_{\mu_1}, \ldots, g_{\mu_q}$. In this manner, one does not rely on the smoothness of g with respect to x (in contrast to polynomial approximations) but on the fact that each function g_μ can be well approximated by similar functions $g_{\mu 1}, \ldots, g_{\mu Q}$. This is related to the fact that the manifold

$$\mathcal{M}_{\texttt{EIM}} = \{g_\mu \,|\, \mu \in \mathbb{P}_{\texttt{EIM}}\}$$

has small Kolmogorov N-width (3.5).

The construction of the basis functions and the interpolation points is based on a greedy algorithm in which we add the particular function g_μ that is least well approximated by the current interpolation operator. In a similar fashion, the interpolation point is chosen as the point in space where the corresponding error function is maximized.

Note that EIM is defined with respect to a given norm on Ω given by $\mathcal{X}_\Omega$ and we generally consider $L^p(\Omega)$-norms for $1 \leq p \leq \infty$. The greedy EIM algorithm is outlined in the algorithm box Empirical Interpolation Method. The output is a Q-term

Algorithm: Empirical Interpolation Method

Input: A family of functions $g_\mu : \Omega \to \mathbb{R}$, parametrized by a parameter $\mu \in \mathbb{P}_{\texttt{EIM}}$ and a target error tolerance `tol`.

Output: A set of Q basis functions $\{h_q\}_{q=1}^Q$ and interpolation points $\{x_q\}_{q=1}^Q$.

Set $q = 1$. Do while `err` $<$ `tol`:

1. Pick the sample point

$$\mu_q = \arg\sup_{\mu \in \mathbb{P}_{\texttt{EIM}}} \left\| g_\mu - \mathtt{I}_{q-1}[g_\mu] \right\|_{\mathcal{X}_\Omega},$$

and the corresponding interpolation point

$$x_q = \arg\sup_{x \in \Omega} \left| g_{\mu_q}(x) - \mathtt{I}_{q-1}[g_{\mu_q}](x) \right|. \tag{5.5}$$

2. Define the next basis function as the scaled error function

$$h_q = \frac{g_{\mu_q} - \mathtt{I}_{q-1}[g_{\mu_q}]}{g_{\mu_q}(x_q) - \mathtt{I}_{q-1}[g_{\mu_q}](x_q)}. \tag{5.6}$$

3. Define the error

$$\texttt{err} = \left\| \texttt{err}_p \right\|_{L^\infty(\mathbb{P}_{\texttt{EIM}})} \quad \text{with} \quad \texttt{err}_p(\mu) = \left\| g_\mu - \mathtt{I}_{q-1}[g_\mu] \right\|_{\mathcal{X}_\Omega},$$

and set $q := q + 1$.

approximation. We note that the basis functions $\{h_1, \ldots, h_Q\}$ and the snapshots $\{g_{\mu_1}, \ldots, g_{\mu_Q}\}$ span the same space by construction, i.e.,

$$\mathbb{V}_Q = \text{span}\{h_1, \ldots, h_Q\} = \text{span}\{g_{\mu_1}, \ldots, g_{\mu_Q}\}.$$

However, the former basis functions are preferred to the latter due to the properties

$$\mathbf{T}_{ii} = h_i(x_i) = 1, \quad 1 \leq i \leq Q \qquad \text{and} \quad \mathbf{T}_{ij} = h_j(x_i) = 0, \quad 1 \leq i < j \leq Q.$$

This construction of the basis functions and interpolation points satisfies the following properties [1]:

- The basis functions $\{h_1, \ldots, h_q\}$ are linearly independent.
- The interpolation matrix $\mathbf{T}_{ij}$ is lower triangular with unity diagonal and hence invertible.
- The empirical interpolation procedure is well-posed in $\mathcal{X}_\Omega$ as long as convergence is not achieved.

Further, one easily shows that the interpolation operator $\mathtt{I}_Q$ is the identity if restricted to the space $\mathbb{V}_Q$, i.e., there holds

$$\mathtt{I}_Q[g_{\mu_i}](x) = g_{\mu_i}(x), \quad i = 1, \ldots, q, \quad \forall x \in \Omega,$$

and

$$\mathtt{I}_Q[g_\mu](x_i) = g_\mu(x_i), \quad i = 1, \ldots, q, \quad \forall \mu \in \mathbb{P}_{\mathtt{EIM}},$$

by the interpolation property.

Remark 5.1 As explained in [3], this algorithm can be applied to the selection of the interpolation points only, in the case where the family of interpolating functions is predefined. This can be the case if a canonical basis and ordering like the set of polynomials is selected beforehand. In this case, the algorithm selects reasonable interpolation points on arbitrary domains Ω. The sequence of sets of generated interpolations points for different numbers of basis functions is hierarchical. Note that the construction of good interpolation points on arbitrary non-tensorized domains is far from trivial. The proposed procedure can provide good, but not optimal, points in such a case, c.f. [3].

If the $L^\infty(\Omega)$-norm is considered, the error analysis of the interpolation procedure involves the Lebesgue constant $\Lambda_q = \sup_{x\in\Omega} \sum_{i=1}^q |L_i(x)|$ where $L_i \in \mathbb{V}_q$ are the Lagrange functions satisfying $L_i(x_j) = \delta_{ij}$. In this case, the following bound holds [1]

$$\left\| g_\mu - \mathtt{I}_q[g_\mu] \right\|_{L^\infty(\Omega)} \leq (1 + \Lambda_q) \inf_{v_q \in \mathbb{V}_q} \left\| g_\mu - v_q \right\|_{L^\infty(\Omega)}.$$

Although in practice a very conservative bound, an upper bound of the Lebesque constant is given by

$$\Lambda_q \leq 2^q - 1,$$

see [3]. Further, assume that $\mathcal{M}_{\texttt{EIM}} \subset \mathcal{X}_\Omega \subset L^\infty(\Omega)$ and that there exists a sequence of finite dimensional spaces

$$\mathbb{Z}_1 \subset \mathbb{Z}_2 \subset \ldots, \quad \dim(\mathbb{Z}_q) = q, \quad \text{and} \quad \mathbb{Z}_q \subset \mathcal{M}_{\texttt{EIM}}, \quad \forall q,$$

such that there exists $c > 0$ and $\alpha > \log(4)$ with

$$\inf_{v_q \in \mathbb{Z}_q} \left\| g_\mu - v_q \right\|_{\mathcal{X}_\Omega} \leq c e^{-\alpha q}, \quad \mu \in \mathbb{P}_{\texttt{EIM}}.$$

Then the upper bound on the Lebesque constant yields the estimate

$$\left\| g_\mu - \mathtt{I}_q[g_\mu] \right\|_{L^\infty(\Omega)} \leq c e^{-(\alpha - \log(4))\, q}.$$

Remark 5.2 The worst-case scenario in which the Lebesgue constant scales like $\Lambda_q \leq 2^q - 1$ is rather artificial. In implementations involving functions belonging to some reasonable set with a small Kolmogorov N-width, one observes a growth of the Lebesgue constant that is much more reasonable. In many cases a linear growth is observed, similar to what is observed for classic interpolation estimates where the growth is typically bounded by the logarithm of the cardinality of the interpolation points. The generation of interpolation points can be seen as a generalization of Leja points for arbitrary basis functions, see [13].

Remark 5.3 Recovering an affine decomposition (5.2) is closely related to the identification of a low rank approximation of a bivariate function $f(x, y)$ of the form

$$f(x, y) \approx \sum_{q=1}^{Q} g_q(x)\, h_q(y).$$

The similarities between EIM and the Adaptive Cross Approximation (ACA) are discussed in detail in [13].

Remark 5.4 A generalization of the EIM, known as Generalized Empirical Interpolation Method (GEIM) [10, 11], consists in replacing the point-wise interpolation (5.3) by the statement

$$\sigma_j(\mathtt{I}_Q[g_\mu]) = \sigma_j(g_\mu), \quad j = 1, \ldots, Q,$$

where the σ_j are well-chosen functionals among a set Σ of functionals. The particular case of $\sigma_j = \delta_{x_j}$ (δ_{x_j} denoting Dirac's delta-function) recovers the EIM. In the general case, the set Σ can consist of functionals requiring less regularity on the

function g_μ than continuity, e.g., L^2-regularity. The functionals can consist of local averages or convolutions with Gaussian's, to mention a few examples.

The algorithm presented in the linear algebra box Empirical Interpolation Method can easily be adapted to this case by replacing (5.5) by

$$\sigma_q = \arg\sup_{\sigma\in\Sigma} \left|\sigma(g_{\mu_q}) - \sigma(\mathtt{I}_{q-1}[g_{\mu_q}])\right|,$$

and (5.6) by

$$h_q = \frac{g_{\mu_q} - \mathtt{I}_{q-1}[g_{\mu_q}]}{\sigma_q(g_{\mu_q}) - \sigma_q(\mathtt{I}_{q-1}[g_{\mu_q}])}.$$

An elaborate convergence analysis of GEIM can be found in [11].

5.3 EIM in the Context of the RBM

As already mentioned, there are several scenarios in which EIM becomes an essential component of the reduced basis method paradigm to ensure computational efficiency. Most of these are related to ensure an affine decomposition of the form (5.1) or the associated problem with force or output functionals, or in the more general case to deal with nonlinear problems. In the following we discuss in more detail how EIM can be used to address these challenges.

5.3.1 Non-affine Parametric Coefficients

In this first and still abstract example, we assume, as an example, that the bilinear form has the following form

$$a(w, v; \mu) = \int_\Omega g(x; \mu)\, b(w, v; x)\, dx,$$

where $b(w, v; x)$ is bilinear in w and v for any $x \in \Omega$ but the coefficient function g has a non-trivial dependency on μ. If there is no known affine decomposition of the form

$$g(x; \mu) = \sum_{q=1}^{Q_\mathrm{a}} c_q(\mu)\, h_q(x), \qquad \mu \in \mathbb{P},\ x \in \Omega,$$

one can apply EIM to recover an approximate affine decomposition

$$g(x; \mu) \approx \sum_{q=1}^{Q_\mathrm{a}} c_q(\mu)\, h_q(x), \qquad \mu \in \mathbb{P},\ x \in \Omega,$$

Linear algebra box: Empirical Interpolation Method

Given a discrete representation of the spaces Ω and $\mathbb{P}_{\texttt{EIM}}$ by $\Omega_h = \{x_1, \ldots, x_M\}$ and $\mathbb{P}^h_{\texttt{EIM}} = \{\mu_1, \ldots, \mu_N\}$ which are M- resp. N-dimensional point-sets of Ω and $\mathbb{P}_{\texttt{EIM}}$. Consider the representative matrix of g defined by

$$\mathbf{G}_{ij} = g(x_i, \mu_j), \quad 1 \le i \le M, \quad 1 \le j \le N.$$

For the sake of notation we apply the notation $\mathbf{A}_{:j}$ to express the j-th column of any matrix $\mathbf{A}$. Assume that we are given a set of basis vectors $\mathbf{H}_Q = [\mathrm{h}_1, \ldots, \mathrm{h}_Q]$ and the set of interpolation indices $\boldsymbol{i}_Q = (i_1, \ldots, i_Q)$. The discrete interpolation operator $\mathbb{I}_Q : \mathbb{R}^Q \to \mathbb{R}^M$ of some column vector $\mathrm{g} \in \mathbb{R}^Q$ is given as the span of the basis vectors $\{\mathrm{h}_q\}_{q=1}^Q$ through $\mathbb{I}_Q[\mathrm{g}] = \mathbf{H}_Q\, \mathrm{a}_{(\mathrm{g})}$ for the vector $\mathrm{a}_{(\mathrm{g})}$ such that $\mathbf{T}\, \mathrm{a}_{(\mathrm{g})} = \mathrm{g}$ with

$$\mathbf{T}_{kq} = (\mathbf{H}_Q)_{i_k q}, \qquad k, q = 1, \ldots, Q.$$

Then, EIM can be expressed as:
Set $q = 1$. Do while $\texttt{err} < \texttt{tol}$

1 .Pick the sample index

$$j_q = \underset{j=1,\ldots,M}{\arg\max} \left\| \mathbf{G}_{:j} - \mathbb{I}_{q-1}[\mathbf{G}_{\boldsymbol{i}_{q-1} j}] \right\|_{\ell^p},$$

and the corresponding interpolation index

$$i_q = \underset{i=1,\ldots,N}{\arg\max} \left| \mathbf{G}_{i j_q} - (\mathbb{I}_{q-1}[\mathbf{G}_{\boldsymbol{i}_{q-1} j_q}])_i \right|.$$

2 .Define the next approximation column by

$$\mathrm{h}_q = \frac{\mathbf{G}_{:j_q} - \mathbb{I}_{q-1}[\mathbf{G}_{\boldsymbol{i}_{q-1} j_q}]}{\mathbf{G}_{i_q j_q} - (\mathbb{I}_{q-1}[\mathbf{G}_{\boldsymbol{i}_{q-1} j_q}])_{i_q}}.$$

3 .Define the error level by

$$\texttt{err} = \max_{j=1,\ldots,M} \left\| \mathbf{G}_{:j} - \mathbb{I}_{q-1}[\mathbf{G}_{\boldsymbol{i}_{q-1} j}] \right\|_{\ell^p}$$

and set $q := q + 1$.

This procedure allows the definition of an approximation of any coefficient of the matrix $\mathbf{G}$. In some cases, however, one would seek to obtain an approximation of $g(x, \mu)$ for *any* $(x, \mu) \in \Omega \times \mathbb{P}_{\texttt{EIM}}$. This is possible as the interpolation points are provided by $x_{i_1}, \ldots, x_{i_Q}$. The construction of the (continuous) basis functions h_q is subsequently based on mimicking part 2 in a continuous context. During the discrete version outlined above, one saves the following data

$$\begin{aligned} \mathbf{S}_{:q} &= \mathrm{a}_{(\mathbf{G}_{\boldsymbol{i}_{q-1} j_q})}, \qquad \text{from} \qquad \mathbb{I}_{q-1}[\mathbf{G}_{\boldsymbol{i}_{q-1} j_q}] = \mathbf{H}_{q-1}\, \mathrm{a}_{(\mathbf{G}_{\boldsymbol{i}_{q-1} j_q})}, \\ \mathbf{S}_{qq} &= \mathbf{G}_{i_q j_q} - (\mathbb{I}_{q-1}[\mathbf{G}_{\boldsymbol{i}_{q-1} j_q}])_{i_q}. \end{aligned}$$

Then, the continuous basis functions can be recovered by the recursive formula

$$h_q = \frac{g(\cdot, \mu_{i_q}) - \sum_{j=1}^{q-1} \mathbf{S}_{jq}\, h_j}{\mathbf{S}_{qq}}.$$

and therefore

$$a(w, v; \mu) \approx \sum_{q=1}^{Q_a} c_q(\mu)\, a_q(w, v) = \sum_{q=1}^{Q_a} c_q(\mu) \int_\Omega h_q(x)\, b(w, v; x)\, dx.$$

This technique can be applied in an entirely similar manner to obtain approximate affine decompositions of the right hand side $f(v; \mu)$ or output functionals, as needed.

5.3.2 *Illustrative Example 1: Heat Conduction Part 5*

Let us now illustrate how the Illustrative Example 1: Heat Conduction can be extended to account for a geometric parametrization and ensure that the affine decomposition holds by the tools provided by EIM.

Assume that the geometric configuration of the inclusion Ω_0 is parametrized. We still assume that the inclusion Ω_0 is a disk centered at the origin but it has a variable radius $\mu_{[3]} \in [r_{\mathtt{min}}, r_{\mathtt{max}}]$. We denote the disk as $\Omega_0(\mu)$ and the geometrical set-up is shown in Fig. 5.1. Furthermore, let us introduce the reference radius $r_0 = 0.5$—the one introduced in Sect. 2.3.1. Then $\Omega_0 = \Omega(\mu)$ for any μ such that $\mu_{[3]} = r_0$. In addition to the existing diffusion parameter $\mu_{[1]} = \kappa_0$ and the flux intensity $\mu_{[2]}$ we now have introduced a geometric parameter $\mu_{[3]}$. We therefore have that $P = 3$ and we denote the vector of parameters as $\mu = (\mu_{[1]}, \mu_{[2]}, \mu_{[3]})$.

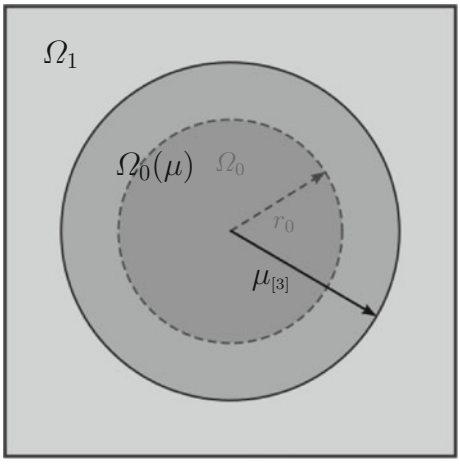

Fig. 5.1 The geometrical set-up for the extended illustrative Example 1

The thermal conductivity, as a function acting on $x \in \overline{\Omega}$, is defined by

$$\kappa_\mu = 1 + (\mu_{[1]} - 1)\, \mathbf{1}_{\Omega_0(\mu)},$$

where $\mathbf{1}_{\Omega_0(\mu)}$ denotes the characteristic function of the parametrized inclusion $\Omega_0(\mu)$.

In terms of approximability of both the induced solution manifold $\mathcal{M}_\delta$ and $\mathcal{G} = \{\kappa_\mu \,|\, \forall \mu \in \mathbb{P}\}$, it would not be a good idea to apply EIM directly to the family of functions κ_μ. The resulting approximation space would not be low dimensional as one would expect errors of order one if approximating the characteristic function $\mathbf{1}_{\Omega_0(\mu)}$ by linear combinations of characteristic functions based on some different sample radii.

However, a reformulation of the problem can result in an approach with a low dimensional approximation space and, thus, a reduced computational cost.

For each $r \in [r_{\min}, r_{\max}]$, we introduce a continuous transformation (more precisely a homeomorphism) $T_r : \Omega \to \Omega$ such that $T_r|_{\partial\Omega} = I_d$ and $|T_r(\hat{x})| = r$ for all $\hat{x} \in \Omega$ such that $|\hat{x}| = r_0$ and thus $\mathrm{Im}(T_r|_{\Omega_0}) = \Omega_0(\mu)$. More precisely, we can realize this by defining r_-, r_+ such that $0 < r_- < r_{\min} < r_{\max} < r_+ < 1$ and $T_r(\hat{x}) = \varphi_r(|\hat{x}|)\,\hat{x}$ for all $\hat{x} \in \Omega$ where

$$\varphi_r(|\hat{x}|) = \begin{cases} 1 & \text{if} \quad 0 \le |\hat{x}| < r_- \\ \frac{1}{r_0(r_0 - r_-)}[r_0(r_0 - |\hat{x}|) + r(|\hat{x}| - r_-)] & \text{if} \quad r_- \le |\hat{x}| < r_0 \\ \frac{1}{r_0(r_0 - r_+)}[r_0(r_0 - |\hat{x}|) + r(|\hat{x}| - r_+)] & \text{if} \quad r_0 \le |\hat{x}| < r_+ \\ 1 & \text{if} \quad r_+ \le |\hat{x}| \end{cases}.$$

Recall that the illustrative example can be expressed as: for any $\mu \in \mathbb{P}$, find $u(\mu) \in \mathbb{V}$ such that

$$a(u(\mu), v; \mu) = f(v; \mu), \qquad \forall v \in \mathbb{V}, \tag{5.7}$$

where

$$a(w, v; \mu) = \int_\Omega \kappa_\mu \, \nabla w \cdot \nabla v, \qquad \text{and} \qquad f(v; \mu) = \mu_{[2]} \int_{\Gamma_{\text{base}}} v.$$

By substitution we observe that

$$a(w, v; \mu) = \int_\Omega \kappa_\mu(T_{\mu_{[3]}}(\hat{x})) \left((\hat{\mathbf{J}}_{\mu_{[3]}}(\hat{x}))^{-1} \nabla \hat{w}(\hat{x})\right) \cdot \left((\hat{\mathbf{J}}_{\mu_{[3]}}(\hat{x}))^{-1} \nabla \hat{v}(\hat{x})\right) \left|\det \hat{\mathbf{J}}_{\mu_{[3]}}(\hat{x})\right| d\hat{x},$$
$$f(v; \mu) = \mu_{[2]} \int_{\Gamma_{\text{base}}} \hat{v}(\hat{x})\, d\hat{x},$$

where $\hat{w}(\hat{x}) = w(T_{\mu_{[3]}}(\hat{x}))$, $\hat{v}(\hat{x}) = v(T_{\mu_{[3]}}(\hat{x}))$ and $(\hat{\mathbf{J}}_{\mu_{[3]}}(\hat{x}))_{kl} = \frac{\partial (T_{\mu_{[3]}}(\hat{x}))_l}{\partial \hat{x}_k}$. In a similar manner, define $\hat{\kappa}_\mu(\hat{x}) = \kappa_\mu(T_{\mu_{[3]}}(\hat{x}))$ and observe that $\hat{\kappa}_\mu|_{\Omega_0} = \mu_{[1]}$ and

$\hat{\kappa}_\mu|_{\Omega_1} = 1$. Based on this, we define new bilinear and linear forms $\hat{a}$ and $\hat{f}$ as

$$\hat{a}(\hat{w}, \hat{v}; \mu) = \int_{\Omega_1} \nabla\hat{w}(\hat{x}) \cdot \left(\mathbf{G}_{\mu_{[3]}}(\hat{x})\, \nabla\hat{v}(\hat{x})\right) d\hat{x} + \mu_{[1]} \int_{\Omega_0} \nabla\hat{w}(\hat{x}) \cdot \left(\mathbf{G}_{\mu_{[3]}}(\hat{x})\, \nabla\hat{v}(\hat{x})\right) d\hat{x},\ \hat{f}(\hat{v}; \mu) = \mu_{[2]} \int_{\Gamma_{\text{base}}} \hat{v}(\hat{x})\, d\hat{x},$$

with

$$\mathbf{G}_{\mu_{[3]}}(\hat{x}) = \left|\det \hat{\mathbf{J}}_{\mu_{[3]}}(\hat{x})\right| \left((\hat{\mathbf{J}}_{\mu_{[3]}}(\hat{x}))^{-1}\right)^T (\hat{\mathbf{J}}_{\mu_{[3]}}(\hat{x}))^{-1}.$$

By the regularity of the homeomorphic mapping T_r, (5.7) is equivalent to the statement: for any $\mu \in \mathbb{P}$, find $\hat{u}(\mu) \in \mathbb{V}$ such that

$$\hat{a}(\hat{u}(\mu), \hat{v}; \mu) = \hat{f}(\hat{v}; \mu), \qquad \forall \hat{v} \in \mathbb{V}.$$

While this resolves the problem of the affine dependency of κ_μ, we have introduced a new dependency of the parameter in the bilinear form $\hat{a}$ through $\hat{\mathbf{J}}_{\mu_{[3]}}$. Introducing the short notation $\varphi_{\mu_{[3]}}$ and $\varphi'_{\mu_{[3]}}$ for $\varphi_{\mu_{[3]}}(|\hat{x}|)$ and $\varphi'_{\mu_{[3]}}(|\hat{x}|)$, respectively, we can write

$$\hat{\mathbf{J}}_{\mu_{[3]}}(\hat{x}) = \varphi_{\mu_{[3]}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\varphi'_{\mu_{[3]}}}{|\hat{x}|} \begin{pmatrix} \hat{x}_1^2 & \hat{x}_1\hat{x}_2 \\ \hat{x}_1\hat{x}_2 & \hat{x}_2^2 \end{pmatrix},$$

so that

$$\det \hat{\mathbf{J}}_{\mu_{[3]}}(\hat{x}) = \varphi_{\mu_{[3]}} \left(\varphi_{\mu_{[3]}} + |\hat{x}|\varphi'_{\mu_{[3]}}\right).$$

Under some mild assumptions on $r_{\texttt{min}}, r_{\texttt{max}}$ with respect to r_-, r_+ and r_0, one can show that this determinant is always positive. This allows us to express the inverse Jacobian as

$$\hat{\mathbf{J}}^{-1}_{\mu_{[3]}}(\hat{x}) = \frac{1}{\det \hat{\mathbf{J}}_{\mu_{[3]}}(\hat{x})} \left[\varphi_{\mu_{[3]}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\varphi'_{\mu_{[3]}}}{|\hat{x}|} \begin{pmatrix} \hat{x}_2^2 & -\hat{x}_1\hat{x}_2 \\ -\hat{x}_1\hat{x}_2 & \hat{x}_1^2 \end{pmatrix}\right]$$

and, consequently, we obtain

$$\mathbf{G}_{\mu_{[3]}}(\hat{x}) = \frac{1}{\det \hat{\mathbf{J}}_{\mu_{[3]}}(\hat{x})} \left[\varphi^2_{\mu_{[3]}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \left(\frac{\varphi_{\mu_{[3]}}\varphi'_{\mu_{[3]}}}{|\hat{x}|} + (\varphi'_{\mu_{[3]}})^2\right) \begin{pmatrix} \hat{x}_2^2 & -\hat{x}_1\hat{x}_2 \\ -\hat{x}_1\hat{x}_2 & \hat{x}_1^2 \end{pmatrix}\right].$$

We can now apply the EIM to get an affine decomposition of the form

$$\mathbf{G}_{\mu_{[3]}}(\hat{x}) \approx \mathbf{I}_Q[\mathbf{G}_{\mu_{[3]}}](\hat{x}) = \sum_{q=1}^{Q} \alpha_q(\mu_{[3]})\, \mathbf{H}_q(\hat{x})$$

for some basis functions $\mathbf{H}_1, \ldots, \mathbf{H}_Q$ chosen by EIM.

The framework of EIM needs to be slightly adapted to the case where the function to be interpolated is a matrix function in $\mathbb{R}^{N\times M}$. An interpolation point at the q-th iteration is specified by a point $\hat{x}_q \in \Omega$ and indices (i_q, j_q) with $1 \le i_q \le N$ and $1 \le j_q \le M$ that are encoded in a functional σ_q defined by

$$\sigma_q\left(\mathbf{G}_{\mu_{[3]}}\right) = (\mathbf{G}_{\mu_{[3]}})_{i_q j_q}(\hat{x}_q).$$

The set of all such functionals shall be denoted by Λ. Given Q functionals $\sigma_q \in \Lambda$ of this form chosen by EIM, the coefficients c_q can be recovered by solving the linear system

$$\sum_{q=1}^{Q} c_q(\mu_{[3]})\,\sigma_i(\mathbf{H}_q) = \sigma_i\left(\mathbf{G}_{\mu_{[3]}}\right), \qquad \forall i = 1,\ldots,Q.$$

An overview of the algorithm in this more general setting is given in the algorithm box Empirical Interpolation Method for matrix functions. Let us finally mention that an alternative would be to apply a separate scalar EIM for each of the components independently. This has the drawback that linear dependencies among different coefficients are not explored, resulting in an affine decomposition with a larger number of terms.

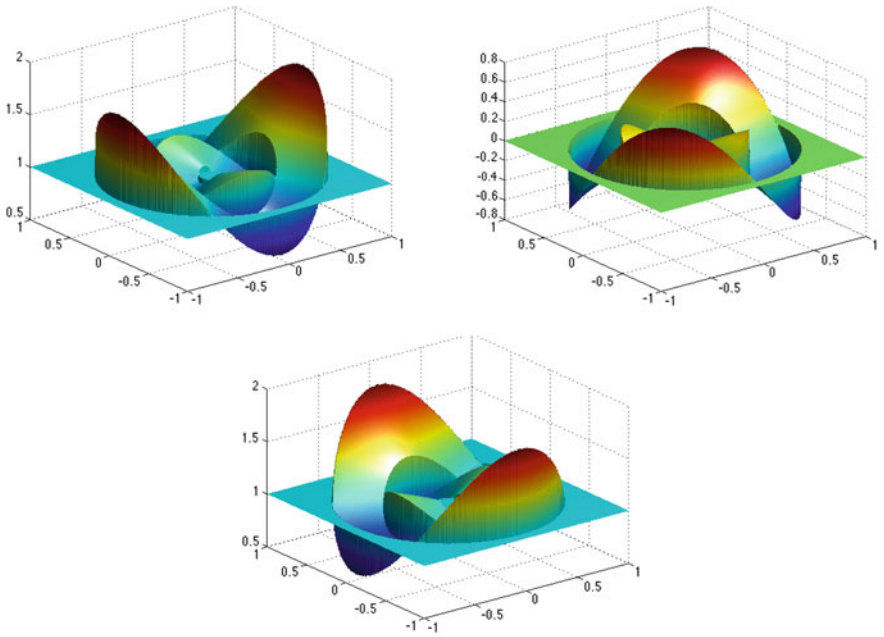

Fig. 5.2 The three components of the symmetric $\mathbf{G}_{\mu_{[3]}}$ for $\mu_{[3]} = 0.6$ with $r_0 = 0.5$

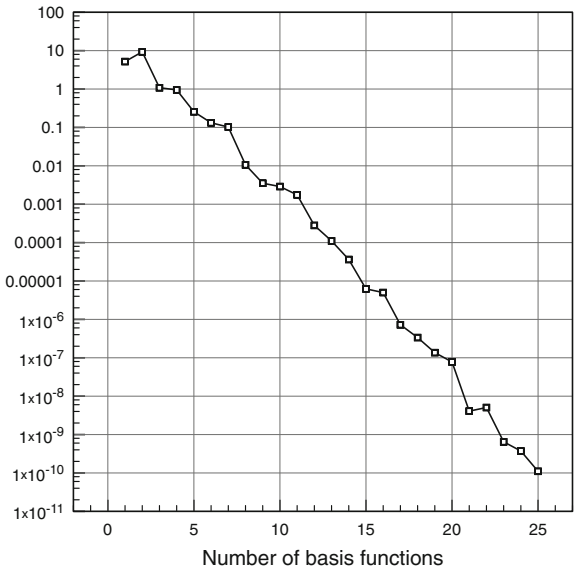

Fig. 5.3 Convergence of the EIM with respect to the number Q of basis functions

In Fig. 5.2 we present the solution, using $r_- = 0.1, r_+ = 0.9$ and $r_0 = 0.5$ for the particular parameter value $\mu_{[3]} = r = 0.6$. Figure 5.3 shows the convergence of the EIM for $\mu_{[3]} = r \in [0.3, 0.7]$ with respect to the number Q of basis functions. One can clearly see an exponential decay.

5.3.3 Illustrative Example 1: Heat Conduction Part 6

5.3.3.1 Problem Formulation

We illustrate the application of the EIM to a variety of further challenges: nonlinear problems. In the context of the Illustrative Example 1, assume that Ω_0 is again a fixed domain as presented in Sect. 2.3.1, but that the problem is made more complex by adding a nonlinearity to the equation. We consider the following nonlinear truth problem: for any $\mu \in \mathbb{P}$, find $u_\delta(\mu) \in \mathbb{V}_\delta$ such that

$$\big(\nabla u_\delta(\mu), \nabla v_\delta\big)_{\Omega_1} + \mu_{[1]}\big(\nabla u_\delta(\mu), \nabla v_\delta\big)_{\Omega_0} + \big(g(u_\delta(\mu)), v_\delta\big)_{\Omega} = f(v_\delta; \mu), \quad \forall v_\delta \in \mathbb{V}_\delta, \tag{5.8}$$

where $g : \mathbb{R} \to \mathbb{R}$ is a nonlinear function, $(\cdot,\cdot)_{\Omega_1}$ and $(\cdot,\cdot)_{\Omega_0}$ denote the L^2-scalar products on Ω_1 and Ω_0 respectively. In this case, the reduced basis approximation reads: find $u_{\mathrm{rb}} \in \mathbb{V}_{\mathrm{rb}}$ such that

$$\big(\nabla u_{\mathrm{rb}}(\mu), \nabla v_{\mathrm{rb}}\big)_{\Omega_1} + \mu_{[1]}\big(\nabla u_{\mathrm{rb}}(\mu), \nabla v_{\mathrm{rb}}\big)_{\Omega_0} + \big(g(u_{\mathrm{rb}}(\mu)), v_{\mathrm{rb}}\big)_{\Omega} = \big(f, v_{\mathrm{rb}}\big)_{\Omega}, \quad \forall v_{\mathrm{rb}} \in \mathbb{V}_{\mathrm{rb}}. \tag{5.9}$$

Unfortunately, this most immediate formulation will not be computationally efficient as the cost of evaluating the solution depends on N_δ through the nonlinear term. Indeed, consider the associated system of nonlinear equations obtained by testing (5.8) with any test-function $\varphi_j \in \mathbb{V}_\delta$: find $\mathrm{u}_\delta^\mu \in \mathbb{R}^{N_\delta}$ such that

$$(\mathbf{A}_\delta^1 + \mu_{[1]}\mathbf{A}_\delta^2)\,\mathrm{u}_\delta^\mu + \mathrm{g}_\delta(\mathrm{u}_\delta^\mu) = \mathrm{f}_\delta, \tag{5.10}$$

with $\mathrm{g}_\delta(\mathrm{u}_\delta^\mu) \in \mathbb{R}^{N_\delta}$ being defined component-wise by $\big(\mathrm{g}_\delta(\mathrm{u}_\delta^\mu)\big)_i = \big(g(u_\delta(\mu)), \varphi_i\big)_\Omega$ and where $\mathbf{A}_\delta^1$, $\mathbf{A}_\delta^2$, f_δ and u_δ are defined following the standard notation used in this text. The reduced basis counter-part reads then: find $\mathrm{u}_{\mathtt{rb}}^\mu \in \mathbb{R}^N$ such that

$$(\mathbf{A}_{\mathtt{rb}}^1 + \mu_{[1]}\mathbf{A}_{\mathtt{rb}}^2)\,\mathrm{u}_{\mathtt{rb}}^\mu + \mathrm{g}_{\mathtt{rb}}(\mathrm{u}_{\mathtt{rb}}^\mu) = \mathrm{f}_{\mathtt{rb}}, \tag{5.11}$$

with $\mathrm{g}_{\mathtt{rb}}(\mathrm{u}_{\mathtt{rb}}^\mu) \in \mathbb{R}^N$ defined component-wise by $\big(\mathrm{g}_{\mathtt{rb}}(\mathrm{u}_{\mathtt{rb}}^\mu)\big)_n = \big(g(u_{\mathtt{rb}}(\mu)), \xi_n\big)_\Omega$ and where again $\mathbf{A}_{\mathtt{rb}}^1$, $\mathbf{A}_{\mathtt{rb}}^2$, $\mathrm{f}_{\mathtt{rb}}$ and $\mathrm{u}_{\mathtt{rb}}$ are defined following the standard notation. It follows therefore that

$$\mathrm{g}_{\mathtt{rb}}(\mathrm{u}_{\mathtt{rb}}^\mu) = \mathbf{B}^T \mathrm{g}_\delta(\mathbf{B}\mathrm{u}_{\mathtt{rb}}^\mu).$$

This assembly process clearly depends on N_δ as the matrix $\mathbf{B} \in \mathbb{R}^{N_\delta \times N}$ is of dimension $N_\delta \times N$.

In the following we will discuss two different strategies which are based on the EIM to obtain efficient and accurate approximations of this term.

5.3.3.2 Using the Empirical Interpolation Method

As proposed in [3], a remedy can be based on the EIM applied to the family of functions

$$\mathcal{M}_{\delta,g} = \Big\{ g_\mu := g(u_\delta(\mu)) \;\Big|\; \mu \in \mathbb{P} \text{ and } u_\delta(\mu) \text{ solution to the truth problem (5.8)} \Big\},$$

where each member is a function of Ω, i.e. $g_\mu : \Omega \to \mathbb{R}$.

During the Offline-stage, one applies the EIM to the family of functions $\mathcal{M}_{\delta,g}$ in order to obtain Q interpolation points $x_1, \ldots, x_Q \in \Omega$, Q sample points $\mu_1, \ldots, \mu_Q \in \mathbb{P}$ and Q basis functions $h_1, \ldots, h_Q$ such that

$$\mathrm{span}\{h_1, \ldots, h_Q\} = \mathrm{span}\{g_{\mu_1}, \ldots, g_{\mu_Q}\}.$$

Then, for any $\mu \in \mathbb{P}$, the interpolant $\mathtt{I}_Q[g_\mu] = \sum_{q=1}^Q c_q(\mu)\, h_q$ is defined by

$$\mathtt{I}_Q[g_\mu](x_q) = g_\mu(x_q) = g(u_\delta(x_q;\mu)), \qquad \forall q = 1, \ldots, Q.$$

Algorithm: Empirical Interpolation Method for matrix functions

Input: A family of vector functions $\mathbf{G}_\mu : \Omega \to \mathbb{R}^{N\times M}$, parametrized by a parameter $\mu \in \mathbb{P}_{\texttt{EIM}}$ and a target error tolerance $\texttt{tol}$.

Output: A set of Q basis functions $\{\mathbf{H}_q\}_{q=1}^{Q}$ and interpolation functionals $\{\sigma_q\}_{q=1}^{Q}$.

Set $q = 1$. Do while $\texttt{err} < \texttt{tol}$:

1. Pick the sample point

$$\mu_q = \arg\sup_{\mu\in\mathbb{P}_{\texttt{EIM}}} \sup_{\substack{1\le i\le N\\ 1\le j\le M}} \left\|(\mathbf{G}_\mu)_{ij} - (\mathbf{I}_{q-1}[\mathbf{G}_\mu])_{ij}\right\|_{\mathcal{X}_\Omega},$$

and the corresponding interpolation functional

$$\sigma_q = \arg\sup_{\sigma\in\Lambda} \left|\sigma\left(\mathbf{G}_{\mu_q} - \mathbf{I}_{q-1}[\mathbf{G}_{\mu_q}]\right)\right|.$$

2. Define the next basis function as

$$h_q = \frac{\mathbf{G}_{\mu_q} - \mathbf{I}_{q-1}[\mathbf{G}_{\mu_q}]}{\sigma_q\left(\mathbf{G}_{\mu_q} - \mathbf{I}_{q-1}[\mathbf{G}_{\mu_q}]\right)}.$$

3. Define the error

$$\texttt{err} = \left\|\texttt{err}_p\right\|_{L^\infty(\mathbb{P}_{\texttt{EIM}})} \quad \text{with} \quad \texttt{err}_p(\mu) = \sup_{\substack{1\le i\le N\\ 1\le j\le M}} \left\|(\mathbf{G}_\mu)_{ij} - (\mathbf{I}_{q-1}[\mathbf{G}_\mu])_{ij}\right\|_{\mathcal{X}_\Omega},$$

and set $q := q + 1$.

To recover an online procedure that is independent on N_δ, the strategy is to replace the term $g(u_{\texttt{rb}}(\mu))$ by its interpolant $\texttt{I}_Q[g(u_{\texttt{rb}}(\mu))] = \sum_{q=1}^{Q} c_q(\mu)\, h_q$ in (5.9) to obtain an approximate nonlinear term

$$\left(g(u_{\texttt{rb}}(\mu)), v_{\texttt{rb}}\right)_\Omega \;\simeq\; \sum_{q=1}^{Q} c_q(\mu)\left(h_q, v_{\texttt{rb}}\right)_\Omega.$$

If the reduced basis approximation is expressed as $u_{\texttt{rb}}(\mu) = \sum_{n=1}(\mathrm{u}_{\texttt{rb}}^{\mu})_n\, \xi_n$, then

$$a_q(\mu) = \sum_{k=1}^{Q}(\mathbf{T}^{-1})_{qk}\, g(u_{\texttt{rb}}(x_k;\mu)) = \sum_{k=1}^{Q}(\mathbf{T}^{-1})_{qk}\, g\left(\sum_{n=1}^{N}(\mathrm{u}_{\texttt{rb}}^{\mu})_n\, \xi_n(x_k)\right),$$

where $\mathbf{T}^{-1}$ denotes the inverse of the interpolation matrix $(\mathbf{T})_{kq} = h_q(x_k)$ provided by the EIM. Therefore the approximate nonlinear part has an affine decomposition

$$\left(g(u_{\texttt{rb}}(\mu)), v_{\texttt{rb}}\right)_\Omega \approx \sum_{q=1}^{Q}\theta_q(\mathrm{u}_{\texttt{rb}}^{\mu})\, b_q(v_{\texttt{rb}}),$$

where

$$b_q(v_{\mathrm{rb}}) = \big(h_q, v_{\mathrm{rb}}\big)_{\Omega},$$
$$\theta_q(\mathrm{u}^{\mu}_{\mathrm{rb}}) = \sum_{k=1}^{Q} (\mathbf{T}^{-1})_{qk}\; g\left(\sum_{n=1}^{N} (\mathrm{u}^{\mu}_{\mathrm{rb}})_n\, \xi_n(x_k)\right).$$

The efficient implementation of the reduced basis approximation then becomes: find $u_{\mathrm{rb}}(\mu) = \sum_{n=1}(\mathrm{u}^{\mu}_{\mathrm{rb}})_n\, \xi_n \in \mathbb{V}_{\mathrm{rb}}$ such that

$$\big(\nabla u_{\mathrm{rb}}, \nabla v_{\mathrm{rb}}\big)_{\Omega_1} + \mu_{[1]}\big(\nabla u_{\mathrm{rb}}, \nabla v_{\mathrm{rb}}\big)_{\Omega_0} + \sum_{q=1}^{Q} \theta_q(\mathrm{u}^{\mu}_{\mathrm{rb}})\, b_q(v_{\mathrm{rb}}) = f(v_{\mathrm{rb}}; \mu), \qquad \forall v_{\mathrm{rb}} \in \mathbb{V}_{\mathrm{rb}}, \tag{5.12}$$

which translates into the N-dimensional system of nonlinear algebraic equations

$$(\mathbf{A}^1_{\mathrm{rb}} + \mu_{[1]}\mathbf{A}^2_{\mathrm{rb}})\, \mathrm{u}^{\mu}_{\mathrm{rb}} + \sum_{q=1}^{Q} \theta_q(\mathrm{u}^{\mu}_{\mathrm{rb}})\, \mathrm{b}^q_{\mathrm{rb}} = \mathrm{f}_{\mathrm{rb}}, \tag{5.13}$$

for the unknown $\mathrm{u}^{\mu}_{\mathrm{rb}}$, where $\mathrm{b}^q_{\mathrm{rb}} \in \mathbb{R}^N$ with $(\mathrm{b}^q_{\mathrm{rb}})_n = b_q(\xi_n)$. This can then be solved with Newton's method for example.

5.3.3.3 The Discrete Empirical Interpolation Method (DEIM)

In the context of nonlinear equations, a slightly different technique, called discrete EIM (DEIM) and introduced in [9] as a variant of the general EIM, has become popular in the POD-community. Indeed, the differences with respect to the strategy outlined above can be summarized in (i) the generation of the collateral basis functions $\{h_1, \ldots, h_Q\}$ and (ii) the representation of the nonlinearity by its finite-dimensional expression $\mathrm{g}_{\mu} = \mathrm{g}_{\delta}(\mathrm{u}^{\mu}_{\delta}) \in \mathbb{R}^{N_\delta}$ rather than as the generic nonlinear function $g_{\mu} = g(u_{\delta}(\mu)) : \Omega \to \mathbb{R}$. Therefore, the focus is naturally on the algebraic representations in the form of (5.10) and (5.11).

Consider the set

$$\mathcal{M}_{N_\delta, g} = \Big\{ \mathrm{g}_{\mu} := \mathrm{g}_{\delta}(\mathrm{u}^{\mu}_{\delta}) \;\Big|\; \mu \in \mathbb{P} \text{ and } u_{\delta}(\mu) = \sum_{i=1}^{N_\delta} (\mathrm{u}^{\mu}_{\delta})_i\, \varphi_i \text{ solution to the truth problem (5.8)} \Big\}.$$

Then, different snapshots $\{\mathrm{g}_{\delta}(\mathrm{u}^{\mu_1}_{\delta}), \ldots, \mathrm{g}_{\delta}(\mathrm{u}^{\mu_M}_{\delta})\} \subset \mathbb{R}^{N_\delta}$ for a large number M of samples are collected in order represent $\mathcal{M}_{N_\delta, g}$ accurately. A representative basis $\{\mathrm{b}^1_{\delta}, \ldots, \mathrm{b}^Q_{\delta}\} \subset \mathbb{R}^{N_\delta}$ of Q terms is obtained by means of a Proper Orthogonal Decomposition (POD) (see Sect. 3.2.1 for a discussion of POD). This means that

any of the snapshots can be obtained as a linear combination of the basis functions $\{\mathrm{b}_\delta^1, \ldots, \mathrm{b}_\delta^Q\}$ up to some precision.

Given any function $u_\delta(\mu) \in \mathbb{V}_\delta$, represented by $\mathrm{u}_\delta^\mu \in \mathbb{R}^{N_\delta}$, the coefficients of the linear combination of $\{\mathrm{b}_\delta^1, \ldots, \mathrm{b}_\delta^Q\}$, needed to approximate $\mathrm{g}_\delta(\mathrm{u}_\delta^\mu)$ are provided through interpolation at some selected indices. Indeed, let $\{i_1, \ldots, i_Q\}$ be Q distinct indices among $\{1, \ldots, N_\delta\}$ so that the matrix $\mathbf{T} \in \mathbb{R}^{Q\times Q}$ with $\mathbf{T}_{kq} = (\mathrm{b}_\delta^q)_{i_k}$ is invertible. Then, the approximation to $\mathrm{g}_\delta(\mathrm{u}_\delta^\mu)$ is given by imposing the interpolation at the specified indices

$$\sum_{q=1}^{Q} \mathbf{T}_{kq}\, c_q(\mu) = \left(\mathrm{g}_\delta(\mathrm{u}_\delta^\mu)\right)_{i_k}, \qquad k = 1, \ldots, Q,$$

to uniquely determine the coefficients $c_q(\mu)$. The approximation is then given by

$$\mathrm{g}_\delta(\mathrm{u}_\delta^\mu) \approx \sum_{q=1}^{Q} c_q(\mu)\, \mathrm{b}_\delta^q.$$

If provided with a reduced basis approximation $u_{\mathtt{rb}}(\mu) \in \mathbb{V}_{\mathtt{rb}}$, represented by $\mathrm{u}_{\mathtt{rb}}^\mu$, the approximation comprises

$$\mathrm{g}_{\mathtt{rb}}(\mathrm{u}_{\mathtt{rb}}^\mu) \approx \sum_{q=1}^{Q} \theta_q(\mathrm{u}_{\mathtt{rb}}^\mu)\, \mathrm{b}_{\mathtt{rb}}^q,$$

with $\mathrm{b}_{\mathtt{rb}}^q = \mathbf{B}^T \mathrm{b}_\delta^q \in \mathbb{R}^N$. Here $\mathbf{B} \in \mathbb{R}^{N_\delta \times N}$ represents the reduced basis functions ξ_n in the truth space $\mathbb{V}_\delta$. Further, the coefficients $\{a_q(\mu)\}_{q=1}^Q$ are obtained through the interpolation

$$\sum_{q=1}^{Q} \mathbf{T}_{kq}\, c_q(\mu) \;= \left(\mathrm{g}_{\mathtt{rb}}(\mathrm{u}_{\mathtt{rb}}^\mu)\right)_{i_k}, \qquad k = 1, \ldots, Q, \tag{5.14}$$

and thus

$$\theta_q(\mathrm{u}_{\mathtt{rb}}^\mu) = \sum_{k=1}^{Q} (\mathbf{T}^{-1})_{qk}\, \left(\mathrm{g}_{\mathtt{rb}}(\mathrm{u}_{\mathtt{rb}}^\mu)\right)_{i_k}, \qquad q = 1, \ldots, Q.$$

With the above definitions the reduced basis approximation is obtained by the solution of the following N-dimensional system of nonlinear equations

$$(\mathbf{A}_{\mathtt{rb}}^1 + \mu_{[1]}\mathbf{A}_{\mathtt{rb}}^2)\, \mathrm{u}_{\mathtt{rb}}^\mu + \sum_{q=1}^{Q} \theta_q(\mathrm{u}_{\mathtt{rb}}^\mu)\, \mathrm{b}_{\mathtt{rb}}^q = \mathrm{f}_{\mathtt{rb}},$$

similar to (5.13) but with a slightly different definition of the coefficients θ_q and the vectors $\mathrm{b}_{\mathtt{rb}}^q$.

There are two remaining questions: (i) How to select the interpolation indices i_k, and (ii) how to access efficiently (independent on N_δ) the coefficients $\left(\mathrm{g}_{\mathtt{rb}}(\mathrm{u}_{\mathtt{rb}}^\mu)\right)_{i_k}$.

The interpolating indices are found by applying an EIM-like procedure where the basis vectors $\mathrm{b}_\delta^1, \ldots, \mathrm{b}_\delta^Q$ (including its specific order) are given. The indices are then defined iteratively by

$$i_q = \underset{i=1,\ldots,N_\delta}{\arg\max} \left|(\mathrm{b}_\delta^q - \mathbb{I}_{q-1}[\mathrm{b}_\delta^q])_i\right|,$$

where $\mathbb{I}_{q-1}$ denotes the interpolant based on the basis functions $\mathrm{b}_\delta^1, \ldots, \mathrm{b}_\delta^{q-1}$ and the interpolation indices $i_1, \ldots, i_{q-1}$ obtained at previous iterations. Such a procedure was already introduced in [3] in order to find good interpolation points for interpolation by arbitrary (possibly polynomial) functions on non-standard domains and is referred to as Magic Points.

Finally, for an efficient implementation, it is important that the procedure

$$\mathrm{u}_{\mathtt{rb}}^\mu \mapsto \left\{\left(\mathrm{g}_\delta(\mathrm{u}_{\mathtt{rb}}^\mu)\right)_{i_k}\right\}_{k=1}^Q,$$

is independent of N_δ, as this is needed to obtain the coefficients in (5.14). This is indeed possible if the corresponding mass matrix $(\varphi_i, \varphi_j)_\Omega$ is sparse and we refer to [9] for the details. Note that the sparsity is essential for an efficient implementation and therefore the approach is not suited for spectral methods for example.

Let us conclude with some remarks. The complete strength of the EIM is not used in the above presentation as the interpolating matrix $\mathbf{T}$ can be ill-conditioned. However, this is avoidable by a simple trick. It is recommended to also update the basis functions by substracting the previous interpolant and scaling the error function by

$$\tilde{\mathrm{b}}_\delta^q = \frac{\mathrm{b}_\delta^q - \tilde{\mathbb{I}}_{q-1}[\mathrm{b}_\delta^q]}{(\mathrm{b}_\delta^q - \tilde{\mathbb{I}}_{q-1}[\mathrm{b}_\delta^q])_{i_q}},$$

where now $\tilde{\mathbb{I}}_{q-1}$ denotes the interpolant based on the modified basis functions $\tilde{\mathrm{b}}_\delta^1, \ldots, \tilde{\mathrm{b}}_\delta^{q-1}$ and interpolation indices $i_1, \ldots, i_{q-1}$ obtained at the previous iterations. This guarantees that the linear system, needed to solve the interpolation problem, is always well-conditioned since the matrix is lower triangular with ones on the diagonal.

5.3.3.4 A Comparitive Discussion

As a general feature, it should be mentioned that the basis generation is handled differently in EIM and in DEIM. While EIM is based on a greedy-procedure, the DEIM is based on a POD-procedure. In both cases however, the basis generation can

Table 5.1 Characteristic differences between the EIM and DEIM to handle nonlinearities

	EIM	DEIM
Input	Interpolation points $x_1, \ldots, x_Q$ Basis functions $h_1, \ldots, h_Q : \Omega \to \mathbb{R}$	Interpolation indices $i_1, \ldots, i_Q$ Basis vectors $\mathrm{b}_\delta^1, \ldots, \mathrm{b}_\delta^Q \in \mathbb{R}^{N_\delta}$
Interpolation matrix Reduced basis vectors	$(\mathbf{T})_{kq} = h_q(x_k)$ $(\mathrm{b}_{\mathtt{rb}}^q)_n = \left(h_q, \xi_n\right)_\Omega, \quad \forall n = 1, \ldots, N$	$\mathbf{T}_{kq} = (\mathrm{b}_\delta^q)_{i_k}$ $\mathrm{b}_{\mathtt{rb}}^q = \mathbf{B}^T \mathrm{b}_\delta^q$
Nonlinearity in $\mathrm{u}_{\mathtt{rb}}^\mu$	$\theta_q(\mathrm{u}_{\mathtt{rb}}^\mu) = \sum_{k=1}^{Q} (\mathbf{T}^{-1})_{qk}\ g\left(\sum_{n=1}^{N} (\mathrm{u}_{\mathtt{rb}}^\mu)_n\, \xi_n(x_k)\right)$	$\theta_q(\mathrm{u}_{\mathtt{rb}}^\mu) = \sum_{k=1}^{Q} (\mathbf{T}^{-1})_{qk}\ \left(\mathrm{g}_{\mathtt{rb}}(\mathrm{u}_{\mathtt{rb}}^\mu)\right)_{i_k}$

be uncoupled from the general framework (working with the nonlinear function g_μ versus its finite-dimensional representation g_μ) so that either basis generation can potentially be applied.

A distinguishing element of the two techniques is the order in which EIM and discretization is applied. In the EIM-based approach, one first applies the function-based EIM which subsequently is discretized. On the contrary, in DEIM the starting point for the EIM is the discrete representation of the nonlinearity. The level of discretization may render the two results different, recovering the same approximation in the continuous limit.

The different aspects of the online phase can be summarized in Table 5.1.

References

1. M. Barrault, Y. Maday, N.C. Nguyen, A.T. Patera, An empirical interpolation method: application to efficient reduced-basis discretization of partial differential equations. C.R. Math. **339**, 667–672 (2004)
2. M.A. Grepl, Y. Maday, N.C. Nguyen, A.T. Patera, Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. ESAIM Math. Model. Numer. Anal. **41**, 575–605 (2007)
3. Y. Maday, N.C. Nguyen, A.T. Patera, G.S. Pau, A general, multipurpose interpolation procedure: the magic points. Commun. Pure Appl. Anal. **8**, 383–404 (2007)
4. J.L. Eftang, M.A. Grepl, A.T. Patera, A posteriori error bounds for the empirical interpolation method. C.R. Math. **348**, 575–579 (2010)
5. P. Chen, A. Quarteroni, G. Rozza, A weighted empirical interpolation method: a priori convergence analysis and applications. ESAIM: M2AN **48**, 943–953 (2014)
6. J.L. Eftang, B. Stamm, Parameter multi-domain hp empirical interpolation. Int. J. Numer. Methods Eng. **90**, 412–428 (2012)
7. J.S. Hesthaven, B. Stamm, S. Zhang, Efficient greedy algorithms for high-dimensional parameter spaces with applications to empirical interpolation and reduced basis methods. ESAIM Math. Model. Numer. Anal. **48**, 259–283 (2014)
8. S. Chaturantabut, D.C. Sorensen, Discrete empirical interpolation for nonlinear model reduction, in *Proceedings of the 48th IEEE Conference on Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference CDC/CCC 2009* (IEEE, 2009), pp. 4316–4321

9. S. Chaturantabut, D.C. Sorensen, Nonlinear model reduction via discrete empirical interpolation. SIAM J. Sci. Comput. **32**, 2737–2764 (2010)
10. Y. Maday, O. Mula, A generalized empirical interpolation method: application of reduced basis techniques to data assimilation, in *Analysis and Numerics of Partial Differential Equations* (Springer, Berlin, 2013), pp. 221–235
11. Y. Maday, O. Mula, G. Turinici, et al., A priori convergence of the generalized empirical interpolation method, in *10th International Conference on Sampling Theory and Applications (SampTA 2013), Breman, Germany* (2013), pp. 168–171
12. F. Casenave, A. Ern, T. Lelievre, A nonintrusive reduced basis method applied to aeroacoustic simulations. Adv. Comput. Math. 1–26 (2014)
13. M. Bebendorf, Y. Maday, B. Stamm, Comparison of some reduced representation approximations, in *Reduced Order Methods for Modeling and Computational Reduction* (Springer, Berlin, 2014), pp. 67–100

Chapter 6
Beyond the Basics

In this final chapter we discuss some important extensions to the methodology presented in details in the previous chapters to motivate the readers to apply the reduced basis methodology to more complex problems of their own interest. The extensions are related to (i) time-dependent problems, (ii) geometrical parametrization of the computational domain and (iii) non-compliant outputs and primal-dual approximations, and (iv) non-coercive problems. Although the discussion of these extensions are accompanied by numerical tests, we conclude with a final numerical example to offer evidence of the potential of the discussed techniques for complex three-dimensional applications.

6.1 Time-Dependent Problems

Consider the following parabolic model problem: Given a parameter value $\mu \in \mathbb{P} \subset \mathbb{R}^P$, evaluate the output of interest

$$s(t;\mu) = \ell(u(t;\mu)), \qquad \forall\, t \in I := [0, T_{\mathrm{f}}],$$

where $u(\mu) \in C^0(I; L^2(\Omega)) \cap L^2(I; \mathbb{V})$ satisfies

$$\big(\partial_t u(t;\mu), v\big)_{L^2(\Omega)} + a\big(u(t;\mu), v;\mu\big) = g(t) f(v), \quad \forall v \in \mathbb{V}, \quad \forall t \in I, \qquad (6.1)$$

subject to initial condition $u(0;\mu) = u_0 \in L^2(\Omega)$. Here $g(t) \in L^2(I)$ is called the control function. To keep the presentation simple we assume that the right hand side f and the output functional ℓ are independent of the parameter although this assumption can be relaxed. As in the elliptic case, we assume that the bilinear form $a(\cdot, \cdot; \mu)$ is coercive and continuous (2.4), satisfies the affine assumption (3.11) and is time-invariant. We denote by $(\cdot, \cdot)_{L^2(\Omega)}$ the $L^2(\Omega)$ scalar product. In the following, we consider a compliant output as

J.S. Hesthaven et al., *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, SpringerBriefs in Mathematics,
DOI 10.1007/978-3-319-22470-1_6

$$s(t;\mu) = f(u(t;\mu)), \qquad \forall\, t \in I := [0, T_{\mathrm{f}}].$$

6.1.1 Discretization

We next introduce a finite difference discretization in time and maintain the discrete approximation space $\mathbb{V}_\delta$ in space to discretize (6.1). We first divide the time interval I into K subintervals of equal length $\Delta t = T_{\mathrm{f}}/K$ and define $t^k = k\Delta t$, $0 \le k \le K$. Hence, given $\mu \in \mathbb{P}$, we seek $u_\delta^k(\mu) \in \mathbb{V}_\delta$, $0 \le k \le K$, such that

$$\tfrac{1}{\Delta t}\big(u_\delta^k(\mu) - u_\delta^{k-1}(\mu), v_\delta\big)_{L^2(\Omega)} + a\big(u_\delta^k(\mu), v_\delta; \mu\big) = g(t^k) f(v_\delta), \qquad \forall v_\delta \in \mathbb{V}_\delta,\ 1 \le k \le K, \tag{6.2}$$

subject to initial condition $(u_\delta^0, v_\delta)_{L^2(\Omega)} = (u_0, v_\delta)_{L^2(\Omega)}$, $\forall v_\delta \in \mathbb{V}_\delta$. Recalling the compliance assumption, we evaluate the output for $0 \le k \le K$,

$$s_\delta^k(\mu) = f(u_\delta^k(\mu)). \tag{6.3}$$

Equation (6.2), comprising a backward Euler-Galerkin discretization of (6.1), shall be our point of departure. We presume that Δt is sufficiently small and N_δ is sufficiently large such that $u_\delta^k(\mu)$ and $s_\delta^k(\mu)$ are effectively indistinguishable from $u(t^k;\mu)$ and $s(t^k;\mu)$, respectively. The development readily extends to Crank-Nicholson or higher order discretization. The solution process for the truth solver is outlined in Linear algebra box: Algebraic formulation of the time-dependent truth problem.

Linear algebra box: Algebraic formulation of the time-dependent truth problem

We develop the algebraic equations associated with (6.2)–(6.3). The truth approximation $u_\delta^k(\mu) \in \mathbb{V}_\delta$ is expressed as

$$u_\delta^k(\mu) = \sum_{i=1}^{N_\delta} (\mathrm{u}_{\delta,k}^\mu)_i\, \varphi_i, \tag{6.4}$$

given the basis $\mathbb{V}_\delta = \mathrm{span}\{\varphi_1, \ldots, \varphi_{N_\delta}\}$ of the truth space $\mathbb{V}_\delta$. By testing with $v_\delta = \varphi_j$ for all $1 \le j \le N$ in (6.2) and using (6.4) we obtain

$$\big(\tfrac{1}{\Delta t}\mathbf{M}_\delta + \mathbf{A}_\delta^\mu\big)\, \mathrm{u}_{\delta,k}^\mu = \tfrac{1}{\Delta t}\mathbf{M}_\delta\, \mathrm{u}_{\delta,k-1}^\mu + g(t^k)\, \mathrm{f}_\delta,$$

which defines the coefficient vector $\mathrm{u}_{\delta,k}^\mu$ at the k-th step. The matrices $\mathbf{A}_\delta^\mu$, $\mathbf{M}_\delta$ and the vector f_δ are defined by

$$(\mathbf{M}_\delta)_{ij} = (\varphi_j, \varphi_i)_{L^2(\Omega)}, \qquad (\mathbf{A}_\delta^\mu)_{ij} = a(\varphi_j, \varphi_i; \mu), \qquad \text{and} \qquad (\mathrm{f}_\delta)_i = f(\varphi_i),$$

for all $1 \le i, j \le N_\delta$. We can evaluate the output as

$$s_\delta^k(\mu) = (\mathrm{u}_{\delta,k}^\mu)^T\, \mathrm{f}_\delta.$$

As in the elliptic case, we consider the discrete solution manifold

$$\mathcal{M}_\delta^K = \left\{ u_\delta^k(\mu) \;\middle|\; 1 \le k \le K,\ \mu \in \mathbb{P} \right\} \subset \mathbb{V}_\delta,$$

as the analogous entity of (3.2) for the parabolic case and seek a small representative basis thereof, i.e., the reduced basis for the parabolic problem. The reduced basis approximation [1, 2] is based on an N-dimensional reduced basis space $\mathbb{V}_{\mathrm{rb}}$, generated by a sampling procedure which combines spatial snapshots in time and parameter space in an optimal fashion. Given $\mu \in \mathbb{P}$, we seek $u_{\mathrm{rb}}^k(\mu) \in \mathbb{V}_{\mathrm{rb}}$, $0 \le k \le K$, such that

$$\tfrac{1}{\Delta t}(u_{\mathrm{rb}}^k(\mu) - u_{\mathrm{rb}}^{k-1}(\mu), v_{\mathrm{rb}})_{L^2(\Omega)} + a\big(u_{\mathrm{rb}}^k(\mu), v_{\mathrm{rb}}; \mu\big) = g(t^k)\, f(v_{\mathrm{rb}}), \quad \forall v_{\mathrm{rb}} \in \mathbb{V}_{\mathrm{rb}},\ 1 \le k \le K, \tag{6.5}$$

Linear algebra box: Algebraic formulation of the time-dependent reduced basis problem

We develop the algebraic equations associated with (6.5)–(6.6). The reduced approximation $u_{\mathrm{rb}}^k(\mu) \in \mathbb{V}_{\mathrm{rb}}$ is expressed as

$$u_{\mathrm{rb}}^k(\mu) = \sum_{n=1}^{N} (\mathrm{u}_{\mathrm{rb},k}^\mu)_n\, \xi_n, \tag{6.7}$$

given the reduced basis $\mathbb{V}_{\mathrm{rb}} = \mathrm{span}\{\xi_1, \ldots, \xi_N\}$. By testing with $v_{\mathrm{rb}} = \xi_n$ for all $1 \le n \le N$ in (6.5) and using (6.7) we obtain:

$$\left(\tfrac{1}{\Delta t}\mathbf{M}_{\mathrm{rb}} + \mathbf{A}_{\mathrm{rb}}^\mu\right) \mathrm{u}_{\mathrm{rb},k}^\mu = \tfrac{1}{\Delta t}\mathbf{M}_{\mathrm{rb}}\, \mathrm{u}_{\mathrm{rb},k-1}^\mu + g(t^k)\, \mathrm{f}_{\mathrm{rb}}, \tag{6.8}$$

for defining the reduced basis coefficients $(\mathrm{u}_{\mathrm{rb},k}^\mu)_n$, $1 \le n \le N$, at the k-th step. The matrices $\mathbf{A}_{\mathrm{rb}}^\mu$, $\mathbf{M}_{\mathrm{rb}}$ and the vector f_{rb} are defined by

$$(\mathbf{M}_{\mathrm{rb}})_{nm} = (\xi_m, \xi_n)_{L^2(\Omega)}, \qquad (\mathbf{A}_{\mathrm{rb}}^\mu)_{nm} = a(\xi_m, \xi_n; \mu), \qquad \text{and} \qquad (\mathrm{f}_{\mathrm{rb}})_n = f(\xi_n),$$

for all $1 \le n, m \le N$. Subsequently we evaluate the reduced basis output as

$$s_{\mathrm{rb}}^k(\mu) = (\mathrm{u}_{\mathrm{rb},k}^\mu)^T\, \mathrm{f}_{\mathrm{rb}}.$$

Using the affine decompositions, (6.8) can be written as

$$\left(\tfrac{1}{\Delta t}\mathbf{M}_{\mathrm{rb}} + \sum_{q=1}^{Q_{\mathrm{a}}} \theta_{\mathrm{a}}^q(\mu)\, \mathbf{A}_{\mathrm{rb}}^q\right) \mathrm{u}_{\mathrm{rb},k}^\mu = \tfrac{1}{\Delta t}\mathbf{M}_{\mathrm{rb}}\, \mathrm{u}_{\mathrm{rb},k-1}^\mu + g(t^k)\, \mathrm{f}_{\mathrm{rb}},$$

where

$$(\mathbf{A}_{\mathrm{rb}}^q)_{nm} = a_q(\xi_m, \xi_n), \qquad 1 \le n, m \le N.$$

Note that the matrices and vectors are related to the truth matrices and vectors by

$$\mathbf{M}_{\mathrm{rb}}^q = \mathbf{B}^T \mathbf{M}_\delta^q\, \mathbf{B}, \qquad \mathbf{A}_{\mathrm{rb}} = \mathbf{B}^T \mathbf{A}_\delta\, \mathbf{B}, \qquad \text{and} \qquad \mathrm{f}_{\mathrm{rb}} = \mathbf{B}^T \mathrm{f}_\delta,$$

where $\mathbf{B}$ is the matrix that represents the reduced basis in terms of the truth basis; see linear algebra box The reduced basis approximation in Chap. 3.

subject to $(u^0_{\mathrm{rb}}(\mu), v_{\mathrm{rb}})_{L^2(\Omega)} = (u^0_\delta, v_{\mathrm{rb}})_{L^2(\Omega)}, \forall v_{\mathrm{rb}} \in \mathbb{V}_{\mathrm{rb}}$ and evaluate the associated output: for $0 \le k \le K$,

$$s^k_{\mathrm{rb}}(\mu) = f(u^k_{\mathrm{rb}}(\mu)). \tag{6.6}$$

The solution process for the reduced basis approximation is outlined in Linear algebra box: Algebraic formulation of the time-dependent reduced basis problem.

The offline-online procedure is now straightforward since the unsteady case is very similar to the steady case discussed before. There are, however, a few critical issues to address. As regards storage, we must now append to the elliptic offline data set the mass matrix $\mathbf{M}_{\mathrm{rb}}$ associated with the unsteady term. Furthermore, we must multiply the elliptic operation counts by K to arrive at $\mathcal{O}(KN^3)$ for the online operation count, where K is the number of time steps. Nevertheless, the cost of the online evaluation of $s^k_{\mathrm{rb}}(\mu)$ remains independent of N_δ even in the unsteady case.

6.1.2 POD-greedy Sampling Algorithm

We now discuss a widely used sampling strategy to construct reduced basis spaces for the time-dependent parabolic case based on combining proper orthogonal decomposition in time with a greedy approach in parameter space. Let us denote by $\mathbb{P}_h$ a finite sample of points in $\mathbb{P}$, serving as a surrogate for $\mathbb{P}$ in the calculation of errors and error bounds across the parameter domain.

A purely greedy approach [1] may encounter difficulties best treated by including elements of the proper orthogonal decomposition selection process [3]. Hence, to capture the causality associated with the evolution, the sampling method combines the proper orthogonal decomposition, see Sect. 3.2.1, for the time-trajectories with the greedy procedure in the parameter space [1, 4, 5] to enable the efficient treatment of the higher dimensions and extensive ranges of parameter variation.

Let us first summarize the basic POD optimality property, already discussed in Sect. 3.2.1, applied to a time-trajectory: given K elements $u^k_\delta(\mu) \in \mathbb{V}_\delta$, $1 \le k \le K$, the procedure POD($\{u^1_\delta(\mu), \ldots, u^K_\delta(\mu)\}, M$), with $M < K$, returns M $\mathbb{V}$-orthonormal functions $\{\xi_m, 1 \le m \le M\}$ for which the space $\mathbb{V}_{\mathrm{POD}} = \mathrm{span}\{\xi_m, 1 \le m \le M\}$ is optimal in the sense of

$$\mathbb{V}_{\mathrm{POD}} = \underset{Y_M \subset \mathrm{span}\{u^k_\delta(\mu), 1 \le j \le K\}}{\arg\inf} \left(\frac{1}{K} \sum_{k=1}^{K} \inf_{v \in Y_M} \|u^k_\delta(\mu) - v\|^2_{\mathbb{V}} \right)^{1/2}.$$

Here Y_M denotes a M-dimensional linear subspace of $\mathbb{V}$. The POD-greedy algorithm, as outlined in the algorithm box The POD-greedy algorithm, comprises an intertwined greedy and POD algorithm. The greedy algorithm provides the outer algorithm where, for each new selected parameter point μ_n, the first N_1 principal

Algorithm: The POD-greedy algorithm

Input: $\texttt{tol}$, μ_1, $n = 1$, N_1 and N_2, $\mathcal{Z} = 0$.
Output: A reduced basis space $\mathbb{V}_{\texttt{rb}}$.

1. Compute $\{u_\delta^1(\mu_n), \ldots, u_\delta^K(\mu_n)\}$ as the sequence of truth solutions to (6.2) for μ_n.
2. Compress this time-trajectory using a POD and retain the relevant modes: $\{\zeta_1, \ldots, \zeta_{N_1}\} = \text{POD}(\{u_\delta^1(\mu_n), \ldots, u_\delta^K(\mu_n)\}, N_1)$.
3. $\mathcal{Z} \leftarrow \{\mathcal{Z}, \{\zeta_1, \ldots, \zeta_{N_1}\}\}$.
4. Set $N \leftarrow N + N_2$ and compute $\{\xi_1, \ldots, \xi_N\} = \text{POD}(\mathcal{Z}, N)$.
5. $\mathbb{V}_{\texttt{rb}} = \text{span}\{\xi_1, \ldots, \xi_N\}$.
6. $\mu_{n+1} = \arg\max_{\mu \in \mathbb{P}_h} \eta(t^K; \mu)$.
7. If $\eta(t^K; \mu_{n+1}) > \texttt{tol}$, then set $n := n + 1$ and **go to** 1., otherwise **terminate**.

components of the time-trajectory $u_\delta^1(\mu_n), \ldots, u_\delta^K(\mu_n)$ are recovered. In a subsequent step, the existing N-dimensional reduced basis space is enriched with those components to build a new $N + N_2$ dimensional basis. Finally, the a posteriori error estimator is used to define a new sample point μ_{n+1} which minimizes the estimated error over the training set $\mathbb{P}_h$.

To initiate the POD-greedy sampling procedure we specify $\mathbb{P}_h$, an initial sample point μ_1 and a tolerance $\texttt{tol}$. The algorithm depends on two suitable integers N_1 and N_2.

We choose N_1 to satisfy an internal POD error criterion based on the usual sum of eigenvalues and $\texttt{tol}$. Furthermore, we choose $N_2 \leq N_1$ to minimize duplication with the existing reduced basis space. It is important to observe that the POD-greedy method is based on successive greedy cycles so that new information will always be retained and redundant information rejected. A purely greedy approach in both t and μ [1], though often generating good spaces, can stall. Furthermore, since the proper orthogonal decomposition is conducted in only one dimension, the computational cost remains relatively low, even for large parameter domains and extensive parameter train samples.

As we discuss in Sect. 6.1.3, $\eta(t^k; \mu)$ is assumed to provide a sharp and inexpensive a posteriori error bound for $\|u_\delta^k(\mu) - u_{\texttt{rb}}^k(\mu)\|_{\mathbb{V}}$. In practice, we exit the POD-greedy sampling procedure at $N = N_{\max}$ for which a prescribed error tolerance

$$\max_{\mu \in \mathbb{P}_h} \eta(t^K; \mu) < \texttt{tol},$$

is satisfied. Note that the POD-greedy generates hierarchical spaces $\mathbb{V}_{\text{POD}}^N$, $1 \leq N \leq N_{\max}$, which is computationally advantageous.

Concerning the computational aspects, a crucial point is that the operation count for the POD-greedy algorithm is additive and not multiplicative in the number of training points in $\mathbb{P}_h$ and N_δ. In a pure proper orthogonal decomposition approach,

we would need to evaluate the "truth" solution for each $\mu \in \mathbb{P}_h$. Consequently, in the POD-greedy approach we can take $\mathbb{P}_h$ relatively large and we can expect that a reduced model will provide rapid convergence uniformly over the parameter domain.

A slightly different version of the POD-greedy algorithm was proposed in [3] where the error trajectories $u_\delta^k(\mu_n) - u_{\mathrm{rb}}^k(\mu_n)$ are added instead of $u_\delta^k(\mu_n)$ to avoid the second POD-compression.

6.1.3 A Posteriori Error Bounds for the Parabolic Case

Let us discuss the a posteriori error estimation for affinely parametrized parabolic coercive partial differential equations. As for the elliptic case, discussed in Chap. 3, we need two ingredients to construct the a posteriori error bounds. The first is the Riesz representation $\hat{r}_\delta^k(\mu)$ of the residual $r^k(\cdot\,;\mu)$ such that

$$\|\hat{r}_\delta^k(\mu)\|_{\mathbb{V}} = \sup_{v_\delta \in \mathbb{V}_\delta} \frac{r^k(v_\delta;\mu)}{\|v_\delta\|_{\mathbb{V}}}, \quad 1 \le k \le K,$$

where $r^k(\cdot\,;\mu)$ is the residual associated with the reduced basis approximation (6.5) defined by

$$r^k(v_\delta;\mu) = g(t^k)f(v_\delta) - \tfrac{1}{\Delta t}\big(u_{\mathrm{rb}}^k(\mu) - u_{\mathrm{rb}}^{k-1}(\mu), v_\delta\big)_{L^2(\Omega)} - a\big(u_{\mathrm{rb}}^k(\mu), v_\delta;\mu\big),$$

for all $v_\delta \in \mathbb{V}_\delta$, $1 \le k \le K$. The second ingredient is a lower bound for the coercivity constant $\alpha_\delta(\mu)$, $0 < \alpha_{\mathrm{LB}}(\mu) \le \alpha_\delta(\mu)$, $\forall \mu \in \mathbb{P}$.

We can now define our error bounds in terms of these two ingredients as it can readily be proven [1, 3] that for all $\mu \in \mathbb{P}$,

$$\|u_\delta^k(\mu) - u_{\mathrm{rb}}^k(\mu)\|_\mu \le \eta_{\mathrm{en}}^k(\mu) \quad \text{and} \quad |s_\delta^k(\mu) - s_{\mathrm{rb}}^k(\mu)| \le \eta_{\mathrm{s}}^k(\mu), \quad 1 \le k \le K,$$

where $\eta_{\mathrm{en}}^k(\mu)$ and $\eta_{\mathrm{s}}^k(\mu)$ are given as

$$\eta_{\mathrm{en}}^k(\mu) = \left(\frac{\Delta t}{\alpha_{\mathrm{LB}}(\mu)} \sum_{k'=1}^{k} \|\hat{r}_\delta^{k'}(\mu)\|_{\mathbb{V}}^2\right)^{\frac{1}{2}} \quad \text{and} \quad \eta_{\mathrm{s}}^k(\mu) = (\eta_{\mathrm{en}}^k(\mu))^2.$$

We assume here for simplicity that $u_\delta^0(\mu) \in \mathbb{V}_{\mathrm{rb}}$; otherwise there will be an additional contribution to $\eta_{\mathrm{en}}^k(\mu)$.

Although based on the same elements as the elliptic case, the offline-online procedure for the error bound is a bit more involved. Following [2], consider the decomposition of the residual using the affine decomposition

$$r^k(v_\delta;\mu) = g(t^k)\, f(v_\delta) - \tfrac{1}{\Delta t}\sum_{n=1}^{N}(\Delta^{\mu}_{\mathrm{rb},k})_n\,(\xi_n, v_\delta)_{L^2(\Omega)} - \sum_{q=1}^{Q_\mathrm{a}}\sum_{n=1}^{N}\theta^q_\mathrm{a}(\mu)\,(\mathrm{u}^{\mu}_{\mathrm{rb},k})_n\, a_q(\xi_n, v_\delta),$$

with $(\Delta^{\mu}_{\mathrm{rb},k})_n = (\mathrm{u}^{\mu}_{\mathrm{rb},k})_n - (\mathrm{u}^{\mu}_{\mathrm{rb},k-1})_n$ and then introduce the coefficient vector $\mathrm{r}^k(\mu) \in \mathbb{R}^{Q_\mathrm{r}}$, with $Q_\mathrm{r} = 1 + N + Q_\mathrm{a}N$ terms, as

$$\mathrm{r}^k(\mu) = \Big(g(t^k),\ -\tfrac{1}{\Delta t}(\Delta^{\mu,k}_{\mathrm{rb}})^T,\ -(\mathrm{u}^{\mu,k}_{\mathrm{rb}})^T\,\theta^1_\mathrm{a}(\mu)\,,\ldots,\ -(\mathrm{u}^{\mu,k}_{\mathrm{rb}})^T\,\theta^{Q_\mathrm{a}}_\mathrm{a}(\mu)\Big)^T.$$

With a similar ordering, we define $M \in (\mathbb{V}'_\delta)^N$ and $A_q \in (\mathbb{V}'_\delta)^N$ for $1 \le q \le Q_\mathrm{a}$ by

$$M = \big((\xi_1,\,\cdot\,)_{L^2(\Omega)},\ldots,(\xi_N,\,\cdot\,)_{L^2(\Omega)}\big), \qquad \text{and} \qquad A_q = \big(a_q(\xi_1,\,\cdot\,),\ldots,a_q(\xi_N,\,\cdot\,)\big),$$

and the vector of forms $R \in (\mathbb{V}'_\delta)^{Q_\mathrm{r}}$ as

$$R = (f, M, A_1, \ldots, A_{Q_\mathrm{a}})^T,$$

to obtain

$$r^k(v_\delta;\mu) = \sum_{q=1}^{Q_\mathrm{r}} \mathrm{r}^k_q(\mu)\, R_q(v_\delta), \qquad \forall v_\delta \in \mathbb{V}_\delta.$$

As in Chap. 4, denoting by $\hat{r}^q_\delta$ the Riesz representation of R_q, i.e. $(\hat{r}^q_\delta, v_\delta)_\mathbb{V} = R_q(v_\delta)$ for all $v_\delta \in \mathbb{V}_\delta$ and $1 \le q \le Q_\mathrm{r}$, we recover

$$\hat{r}^k_\delta(\mu) = \sum_{q=1}^{Q_\mathrm{r}} \mathrm{r}^k_q(\mu)\,\hat{r}^q_\delta,$$

and

$$\|\hat{r}^k_\delta(\mu)\|^2_\mathbb{V} = \sum_{q,q'=1}^{Q_\mathrm{r}} \mathrm{r}^k_q(\mu)\,\mathrm{r}^k_{q'}(\mu)\,\big(\hat{r}^q_\delta, \hat{r}^{q'}_\delta\big)_\mathbb{V}. \tag{6.9}$$

Observe that $(\hat{r}^q_\delta, \hat{r}^{q'}_\delta)_\mathbb{V}$ are time-independent.

The offline-online decomposition is now clear. In the μ-independent construction stage we find the Riesz-representations $\hat{r}^q_\delta$, and the inner products $(\hat{r}^q_\delta, \hat{r}^{q'}_\delta)_\mathbb{V}$ at possibly large computational cost $\mathcal{O}(Q_\mathrm{r}\,{N_\delta}^2 + Q^2_\mathrm{r}\,N_\delta)$. In the μ-dependent online procedure we simply evaluate (6.9) from the stored inner products in $\mathcal{O}(Q^2_\mathrm{r})$ operations per time step and hence $O(Q^2_\mathrm{r}\,K)$ operations in total. The cost and storage in the online phase is again independent of N_δ.

We may also pursue a primal-dual reduced approximations [1, 5, 6], explained in more detail in Sect. 6.3, to ensure an accelerated rate of convergence of the output and a more robust estimation of the output error. However, in cases where many

outputs are of interest, e.g., inverse problems, the primal-only approach described above can be more efficient and also more flexible by being expanded to include additional output functionals.

6.1.4 Illustrative Example 3: Time-Dependent Heat Conduction

In this example we solve a time-dependent heat transfer problem in the two-dimensional domain shown in Fig. 6.1, using the POD-greedy approach.

The bilinear form for the problem (6.1) is given as

$$a(w, v; \mu) = \sum_{p=1}^{8} \mu_{[p]} \int_{\Omega_p} \nabla w \cdot \nabla v + \int_{\Omega_9} \nabla w \cdot \nabla v,$$

where $\mu_{[p]}$ is the ratio between the conductivity of the Ω_p and Ω_9 subdomains, respectively, and

$$\mu_{[p]} \in [0.1, 10] \qquad \text{for } p = 1, \ldots, 8.$$

Homogeneous Dirichlet boundaries conditions have been applied on Γ_{top} and thus

$$w = 0 \quad \text{on } \Gamma_{\text{top}}.$$

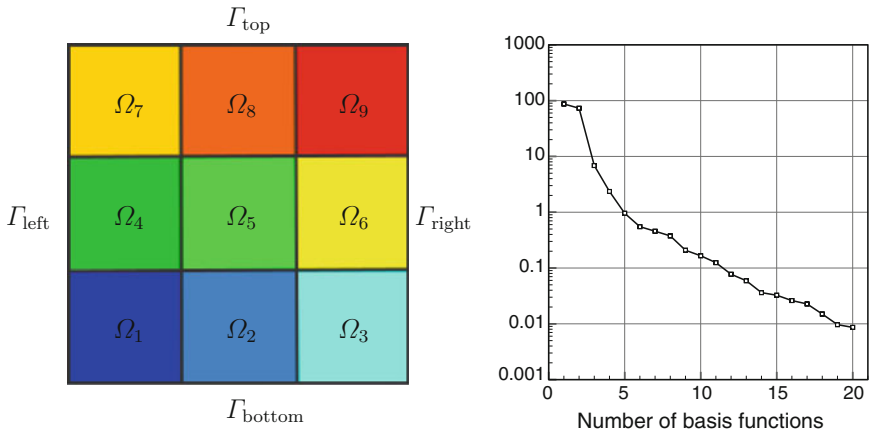

Fig. 6.1 Geometric set-up (*left*) and the convergence of the POD-greedy algorithm (*right*) for the time-dependent thermal block heat conductivity problem

Table 6.1 Error bounds and effectivity metrics for the field variable $u_{\mathtt{rb}}(\mu)$ with respect to the number of reduced basis functions N for the thermal block time-dependent problem

N	$\eta_{\mathtt{en,av}}$	$\mathtt{eff}_{\mathtt{en,max}}$	$\mathtt{eff}_{\mathtt{en,av}}$
5	0.18	24.67	7.51
10	0.07	26.27	7.69
15	0.03	25.82	6.79
20	0.02	31.63	9.53

Inhomogeneous parametrized Neumann boundary conditions, corresponding to heat fluxes, are imposed on the bottom boundary Γ_{bottom} and result in the following right-hand side:

$$f(v;\mu) = \mu_{[9]} \int_{\Gamma_{\text{bottom}}} v$$

where

$$\mu_{[9]} \in [-1, 1]\,.$$

Finally, homogeneous Neumann boundary conditions are applied on the remaining part of the boundary, i.e. on the left Γ_{left} and the right Γ_{right} of the square.

In Fig. 6.1 the convergence of the POD-greedy algorithm is illustrated confirming exponential convergence also in this time-dependent case ($T_{\mathtt{f}} = 3\,\mathrm{s}$, $\Delta t = 0.05\,\mathrm{s}$). The algorithm was performed on a training set $\mathbb{P}_h$ of cardinality 5,000. Table 6.1 shows the averaged a posteriori error estimate over the validation set $\mathbb{P}_h^{\mathrm{v}}$ of cardinality 3,000 as well as the maximum and average effectivities of the estimator confirming the quality of the estimator.

Finally, we show in Figs. 6.2, 6.3 and 6.4 the truth solution as well as the reduced basis approximation and the pointwise errors for three different parameter values.

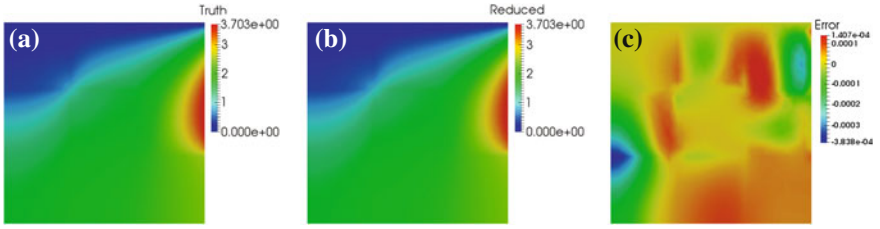

Fig. 6.2 Comparison between the truth model (**a**), the reduced basis approximation (**b**) for $\mu = (1.0, 10.0, 10.0, 1.0, 10.0, 1.0, 10.0, 1.0, 10.0, -1.0)$ at the final time $T_{\mathtt{f}} = 3\,\mathrm{s}$. The pointwise error between the two solutions is reported in (**c**)

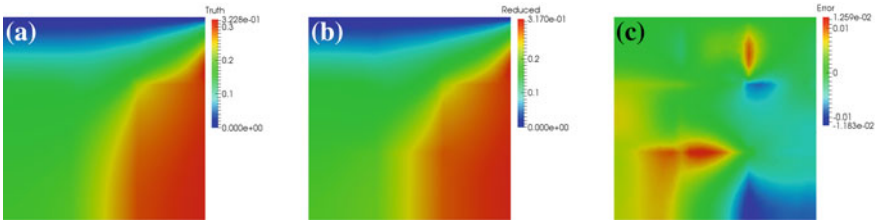

Fig. 6.3 Comparison between the truth model (**a**), the reduced basis approximation (**b**) for $\mu = (5.24, 1.34, 8.52, 5.25, 1.38, 7.98, 0.94, 2.54, 0.98)$ at the final time $T_{\mathrm{f}} = 3\,\mathrm{s}$. The pointwise error between the two solutions is reported in (**c**)

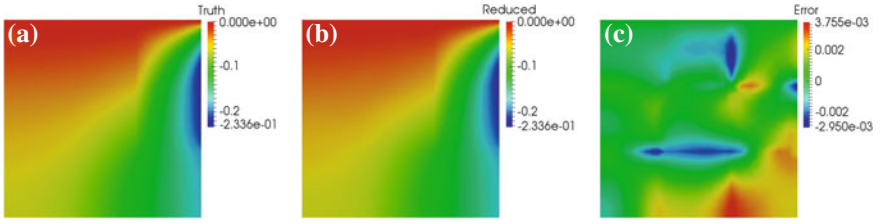

Fig. 6.4 Comparison between the truth model (**a**), the reduced basis approximation (**b**) for $\mu = (6.96, 4.79, 5.33, 9.42, 6.09, 1.87, 8.04, 9.22, -0.94)$ at the final time $T_{\mathrm{f}} = 3\,\mathrm{s}$. The pointwise error between the two solutions is reported in (**c**)

6.2 Geometric Parametrization

Reduced basis methods can be applied in many problems of industrial interest: material sciences and linear elasticity [7–10], heat and mass transfer [11–14], acoustics [15], potential flows [16], micro-fluid dynamics [17], electro-magnetism [18]. In many such problems, there are physical or engineering parameters which characterize the problem, but often also geometric parameters to consider. This combination is quite typical for many industrial devices, e.g., biomedical devices or complex aerodynamic shapes [19–22].

Let us consider a scalar field in d space dimension. We define an original problem, denoted by subscript $\circ$, posed over the parameter-dependent physical domain $\Omega_\circ = \Omega_\circ(\mu)$. We denote by $\mathbb{V}_\circ(\mu)$ a suitable Hilbert space defined on $\Omega_\circ(\mu)$ and consider an elliptic problem of the following form: Given $\mu \in \mathbb{P}$, evaluate

$$s_\circ(\mu) = \ell_\circ(u_\circ(\mu); \mu), \tag{6.10}$$

where $u_\circ(\mu) \in \mathbb{V}_\circ(\mu)$ satisfies

$$a_\circ(u_\circ(\mu), v; \mu) = f_\circ(v; \mu), \qquad \forall v \in \mathbb{V}_\circ(\mu). \tag{6.11}$$

The reduced basis framework requires a reference (μ-independent) domain Ω to compare and combine discrete solutions that otherwise are computed on different domains and grids. Hence, we map $\Omega_\circ(\mu)$ to a reference domain $\Omega = \Omega_\circ(\bar{\mu})$, $\bar{\mu} \in \mathbb{P}$ to recover a transformed problem of the form (2.1) to (2.2), which is the point of departure of the reduced basis approach. The reference domain Ω is related to the original domain $\Omega_\circ(\mu)$ through a parametric mapping $T(\cdot\,;\mu)$, such that $\Omega_\circ(\mu) = T(\Omega;\mu)$ and $T(\cdot\,;\bar{\mu})$ becomes the identity. It remains to place some restrictions on both the geometry (i.e. on $\Omega_\circ(\mu)$) and the operators (i.e. $a_\circ$, $f_\circ$, $\ell_\circ$) such that the transformed problem satisfies the basic hypotheses introduced above, in particular, the affine assumptions (3.11)–(3.13). For many problems, a domain decomposition [5] may be useful as we shall demonstrate shortly.

Let us first consider a simple class of admissible geometries. To build a parametric mapping related to geometrical properties, we introduce a conforming domain partition of $\Omega_o(\mu)$,

$$\Omega_\circ(\mu) = \bigcup_{l=1}^{L_\Omega} \Omega_\circ^l(\mu), \tag{6.12}$$

consisting of mutually nonoverlapping open subdomains $\Omega_\circ^l(\mu)$, such that $\Omega_\circ^l(\mu) \cap \Omega_\circ^{l'}(\mu) = \emptyset$, $1 \le l < l' \le L_\Omega$. The geometric input parameters, e.g. lengths, thicknesses, orientation, diameters or angles, allows for the definition of parametric mappings to be done in an intuitive fashion. The regions can represent different material properties, but they can also be used for algorithmic purposes to ensure well-behaved mappings. In the following we will identify $\Omega^l = \Omega_\circ^l(\bar{\mu})$, $1 \le l \le L_\Omega$, and denote (6.12) the reduced basis macro triangulation. It will play an important role in the generation of the affine representation (3.11)–(3.13). The original and the reference subdomains must be linked via a mapping $T(\cdot\,;\mu) : \Omega^l \to \Omega_\circ^l(\mu)$, $1 \le l \le L_\Omega$, as

$$\Omega_\circ^l(\mu) = T^l(\Omega^l;\mu), \qquad 1 \le l \le L_\Omega.$$

These maps must be bijective, collectively continuous, such that $T^l(x;\mu) = T^{l'}(x;\mu)$, $\forall\, x \in \overline{\Omega^l} \cap \overline{\Omega^{l'}}$, for $1 \le l < l' \le L_\Omega$.

If we consider the affine case, where the transformation is given, for any $\mu \in \mathbb{P}$ and $x \in \Omega^l$, as

$$T^l(x,\mu) = \mathbf{G}^l(\mu)\, x + \mathrm{c}^l(\mu),$$

for given translation vectors $\mathrm{c}^l : \mathbb{P} \to \mathbb{R}^d$ and linear transformation matrices $\mathbf{G}^l : \mathbb{P} \to \mathbb{R}^{d\times d}$. The linear transformation matrices can enable rotation, scaling and/or shear and must be invertible. The associated Jacobians are defined as $J^l(\mu) = |\det(\mathbf{G}^l(\mu))|$, $1 \le l \le L_\Omega$.

Let us now explain how to introduce the geometric parametrization in the operators. We consider the bilinear forms

$$a_\circ(w, v; \mu) = \sum_{l=1}^{L_\Omega} \int_{\Omega_\circ^l(\mu)} D(w)^T \, \mathbf{K}_\circ^l(\mu) \, D(v) \tag{6.13}$$

where $D(v) : \Omega \to \mathbb{R}^{d+1}$ is defined by $D(v) = \left[\frac{\partial v}{\partial x_1}, \ldots, \frac{\partial v}{\partial x_d}, v\right]^T$, and $\mathbf{K}_\circ^l : \mathbb{P} \to \mathbb{R}^{(d+1)\times(d+1)}$, $1 \le l \le L_\Omega$, are prescribed coefficients. Here, for $1 \le l \le L_\Omega$, $\mathbf{K}_\circ^l : \mathbb{P} \to \mathbb{R}^{(d+1)\times(d+1)}$ is a given symmetric positive definite matrix, ensuring coercivity of the bilinear form. The upper $d \times d$ principal submatrix of $\mathbf{K}_\circ^l$ is the usual tensor conductivity/diffusivity; the $(d+1, d+1)$ element of $\mathbf{K}_\circ^l$ represents the identity operator and the $(d+1, 1) - (d+1, d)$ (and $(1, d+1) - (d, d+1)$) elements of $\mathbf{K}_\circ^l$, which we can be zero if the operators are symmetric, represent first derivative terms to model convective terms.

Similarly, we require that $f_\circ(\cdot)$ and $\ell_\circ(\cdot)$ are expressed as

$$f_\circ(v; \mu) = \sum_{l=1}^{L_\Omega} \int_{\Omega_\circ^l(\mu)} F_\circ^l(\mu) \, v, \qquad \ell_\circ(v; \mu) = \sum_{l=1}^{L_\Omega} \int_{\Omega_\circ^l(\mu)} L_\circ^l(\mu) \, v,$$

where $F_\circ^l : \mathbb{P} \to \mathbb{R}$ and $L_\circ^l : \mathbb{P} \to \mathbb{R}$, for $1 \le l \le L_\Omega$, are prescribed coefficients. By identifying $u(\mu) = u_\circ(\mu) \circ T(\cdot\,; \mu)$ and tracing (6.13) back to the reference domain Ω by the mapping $T(\cdot\,; \mu)$, we can define a transformed bilinear form $a(\cdot\,, \cdot\,; \mu)$ as

$$a(w, v; \mu) = \sum_{l=1}^{L_\Omega} \int_{\Omega^l} D(w)^T \mathbf{K}^l(\mu) D(v), \tag{6.14}$$

where $\mathbf{K}^l : \mathbb{P} \to \mathbb{R}^{(d+1)\times(d+1)}$, $1 \le l \le L_\Omega$, is a parametrized tensor

$$\mathbf{K}^l(\mu) = J^l(\mu) \, \hat{\mathbf{G}}^l(\mu) \, \mathbf{K}_\circ^l(\mu) \, (\hat{\mathbf{G}}^l(\mu))^T$$

and $\hat{\mathbf{G}}^l : \mathbb{P} \to \mathbb{R}^{(d+1)\times(d+1)}$ is given by

$$\hat{\mathbf{G}}^l(\mu) = \begin{pmatrix} (\mathbf{G}^l(\mu))^{-1} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}, \quad 1 \le l \le L_\Omega.$$

The transformed linear forms can be expressed similarly as

$$f(v; \mu) = \sum_{l=1}^{L_\Omega} \int_{\Omega^l} F^l(\mu) \, v, \qquad \ell(v; \mu) = \sum_{l=1}^{L_\Omega} \int_{\Omega^l} L^l(\mu) \, v, \tag{6.15}$$

where $F^l : \mathbb{P} \to \mathbb{R}$ and $L^l : \mathbb{P} \to \mathbb{R}$ are given by $F^l = J^l(\mu)\, F^l_{\circ}(\mu)$, $L^l = J^l(\mu)\, L^l_{\circ}(\mu)$, for $1 \le l \le L_\Omega$. In this setting, the problem on the original domain has been recast on the reference configuration Ω, resulting in a parametrized problem where the effect of geometry variations is now expressed by its parametrized transformation tensors. With the above definitions and defining $\mathbb{V} = \mathbb{V}_{\circ}(\bar{\mu})$, (6.10)–(6.11) is equivalent to: given $\mu \in \mathbb{P}$, evaluate

$$s(\mu) = \ell(u(\mu); \mu),$$

where $u(\mu) \in \mathbb{V}$ satisfies

$$a(u(\mu), v; \mu) = f(v; \mu), \qquad \forall v \in \mathbb{V}.$$

The affine formulation (3.11) (resp. (3.12) and (3.13)) can be derived by simply expanding the expression (6.14) (and (6.15)) in terms of the subdomains Ω^l and the different entries of $\mathbf{K}^l$ (and F^l, L^l).

The process by which we map this original problem to the reference problem can be automated [5]. There are many ways in which we can relax the given assumptions and treat an even broader class of problems. For example, we may consider curved triangular subdomains [5] or coefficient functions $\mathbf{K}$, $\mathbf{M}$ which are high order polynomials in the spatial coordinate or approximated by the Empirical Interpolation Method, see Chap. 5. In general, an increased complexity in geometry and operator will result in more terms in affine expansions with a corresponding increase in the online reduced basis computational costs.

6.2.1 *Illustrative Example 4: A 2D Geometric Parametrization for an Electronic Cooling Component*

We consider a first example to illustrate the geometrical parametrization in the context of reduced basis methods to apply the above theoretical considerations. The original domain $\Omega_{\circ}(\mu)$

$$\Omega_{\circ}(\mu) = (-2, 2) \times (-2, 2) \setminus (-\mu_{[1]}, \mu_{[1]}) \times (-\mu_{[2]}, \mu_{[2]}),$$

is provided as a square with a variable rectangular hole. The two geometric parameters correspond to the dimensions of the rectangular hole. The user-provided control points/edges are shown in Fig. 6.5a which yield the $L_\Omega = 8$ reduced basis macro triangulation of Ω shown in Fig. 6.5b. There are $Q_{\mathrm{a}} = 10$ different terms in our affine expansion (3.11) for the Laplacian. Due to the symmetry in the reduced basis triangulation, the number of terms in the affine expansion for the Laplacian reduces from the maximum possible of 24 to 10.

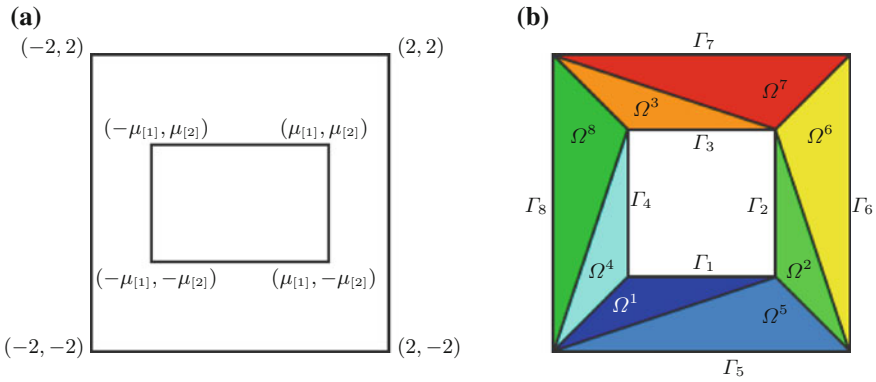

Fig. 6.5 The geometric set-up and control-points (**a**) and the reduced basis macro triangulation of the reference configuration Ω, used to define the piecewise affine transformation (**b**)

We consider this geometric configuration to study the performance of a device designed for heat exchange. The internal walls $\Gamma_1 - \Gamma_4$ are subject to a constant heat flux to be dissipated. The outer walls $\Gamma_5 - \Gamma_8$ of the square are characterized by a heat exchange rate with a fluid surrounding the device. We model the air flow by a simple convection heat transfer coefficient, i.e. the Biot number, used as third parameter $\mu_{[3]}$ for the model problem. The steady-state temperature distribution is governed by a Laplace equation. Our output of interest is the average conduction temperature distribution at the inner walls. We shall therefore consider the Laplace operator (thus with isotropic diffusivity) corresponding to $(\mathbf{K}_\circ^l)_{11} = (\mathbf{K}_\circ^l)_{22} = 1$ and all other entries of $\mathbf{K}_\circ^l$ zero for $1 \le l \le L_\Omega$. We are dealing with $P = 3$ parameters and the parameter domain is given by $\mathbb{P} = [0.5, 1.5] \times [0.5, 1.5] \times [0.01, 1]$.

From an engineering point of view, this problem illustrates the application of conduction analysis to an important class of cooling problems, for example for electronic components and systems.

We now turn the attention to the problem formulation on the "original" domain $\Omega_\circ(\mu)$. We consider (6.10)–(6.11) with $\mathbb{V}_\circ(\mu) = H^1(\Omega_\circ(\mu))$ and

$$a_\circ(w, v; \mu) = \int_{\Omega_\circ(\mu)} \nabla w \cdot \nabla v + \mu_{[3]} \left(\int_{\Gamma_{\circ,5}} w + \int_{\Gamma_{\circ,6}} w + \int_{\Gamma_{\circ,7}} w + \int_{\Gamma_{\circ,8}} w \right),$$

$$f_\circ(v) = \int_{\Gamma_{\circ,1}} v + \int_{\Gamma_{\circ,2}} v + \int_{\Gamma_{\circ,3}} v + \int_{\Gamma_{\circ,4}} v,$$

which represents the bilinear form associated with the Laplace operator, imposing the Robin-type boundary conditions on the outer wall, and the linear form, imposing the constant heat flux on the interior walls respectively. The problem is clearly coercive, symmetric, and compliant as $s_\circ(\mu) = f_\circ(u(\mu))$.

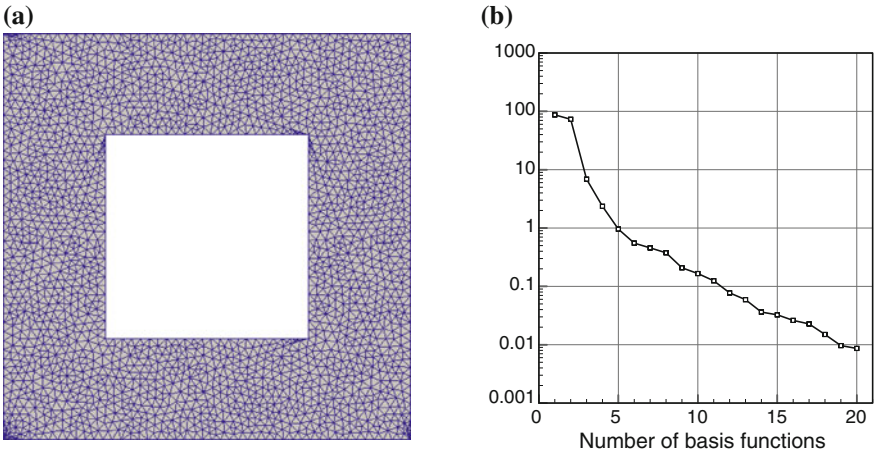

Fig. 6.6 The finite element mesh used for the truth solver (**a**) and the maximum error of the greedy-algorithm with respect to the number of basis functions employed (**b**) for the electronic cooling component problem

The Finite Element method has been computed employing first order elements. In Fig. 6.6a the mesh of the reference domain Ω featuring 4,136 elements is illustrated.

The basis functions are obtained by orthogonalization, through a standard Gram-Schimdt procedure, of snapshots selected by the greedy algorithm as discussed in Sect. 3.2.2. The training space $\mathbb{P}_h$ consists of 1,000 points in $\mathbb{P}$. In Fig. 6.6b, we plot the maximum error (over $\mathbb{P}_h$) with respect to the number of basis functions employed. The first four snapshots are depicted in Fig. 6.7. In Fig. 6.8, the outcomes provided by the truth solver and the reduced basis computations, for a randomly chosen combination $\mu = (1.176, 0.761, 0.530)$ and $N = 20$, are compared, and the pointwise difference between the two solutions is plotted, highlighting the overall accuracy.

In Table 6.2 the error bounds and effectivity metrics for the energy-norm estimates as well as for the estimates for the output functional with respect to the number of basis functions employed are shown. The values have been averaged over a validation set $\mathbb{P}_h^{\mathrm{v}}$ consisting of 3,000 sample points, i.e., using (4.23) for the energy-norm effectivities, and

$$\eta_{\mathrm{s,av}} = \frac{1}{|\mathbb{P}_h^{\mathrm{v}}|} \sum_{\mu \in \mathbb{P}_h^{\mathrm{v}}} \eta_{\mathrm{s}}(\mu)$$

where

$$\mathtt{eff}_{\mathrm{s,max}} = \max_{\mu \in \mathbb{P}_h^{\mathrm{v}}} \frac{\eta_{\mathrm{s}}(\mu)}{|s_\delta(\mu) - s_{\mathrm{rb}}(\mu)|} \quad \text{and} \quad \mathtt{eff}_{\mathrm{s,av}} = \frac{1}{|\mathbb{P}_h^{\mathrm{v}}|} \sum_{\mu \in \mathbb{P}_h^{\mathrm{v}}} \frac{\eta_{\mathrm{s}}(\mu)}{|s_\delta(\mu) - s_{\mathrm{rb}}(\mu)|}. \tag{6.16}$$

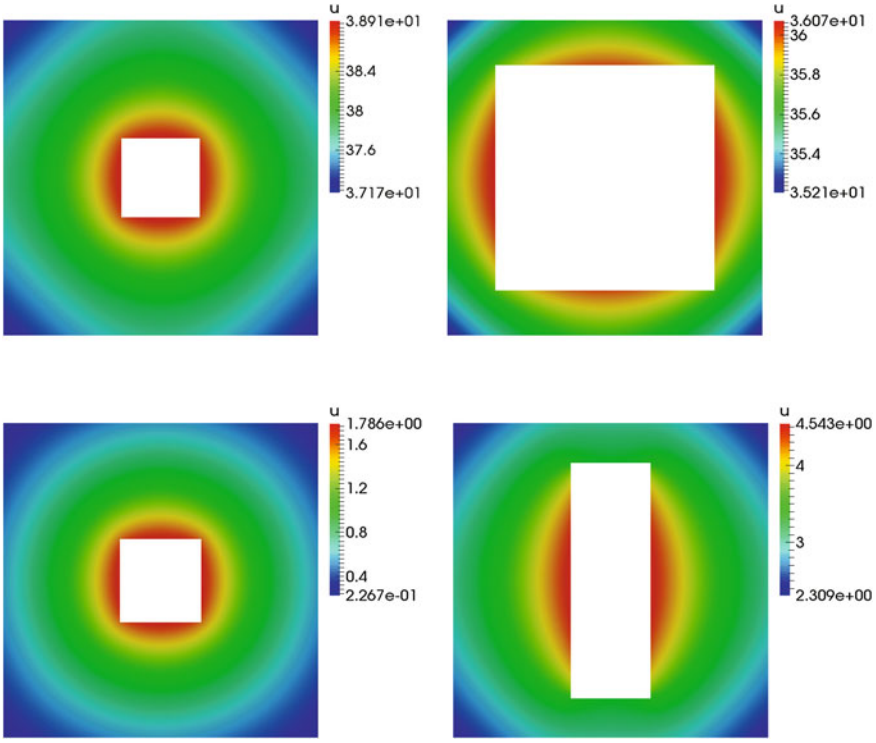

Fig. 6.7 The first four basis functions, representing the most important modes of the solution manifold, illustrated on the original domain $\Omega_{\circ}(\mu)$

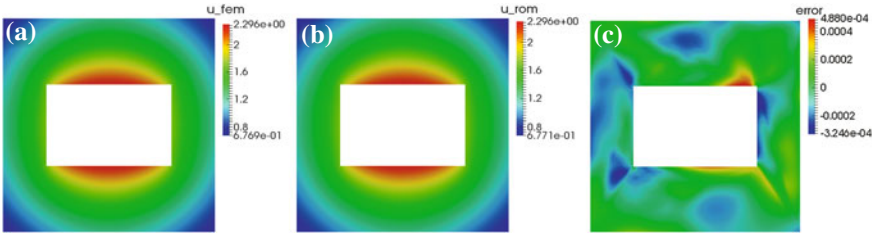

Fig. 6.8 Comparison between the truth solution (**a**) and the reduced basis approximation (**b**) for $\mu = (1.176, 0.761, 0.530)$. The pointwise difference between the two solutions is reported in (**c**)

Table 6.2 Error bounds and output error bounds with effectivity metrics as a function of N for the example with geometrical parametrization

N	$\eta_{\mathrm{en,av}}$	$\mathtt{eff}_{\mathrm{en,max}}$	$\mathtt{eff}_{\mathrm{en,av}}$
5	0.43	16.73	8.23
10	0.11	23.45	12.37
15	4.21e–2	33.65	14.78
20	5.67e–3	38.24	16.35
N	$\eta_{\mathrm{s,av}}$	$\mathtt{eff}_{\mathrm{s,max}}$	$\mathtt{eff}_{\mathrm{s,av}}$
5	0.13	12.20	4.53
10	1.17e–2	23.95	11.86
15	1.82e–3	34.03	18.48
20	3.25e–5	41.65	22.18

The error are shown in the solution $u_{\mathtt{rb}}(\mu)$ (*top*) and the output $s_{\mathtt{rb}}(\mu)$ (*bottom*)

6.2.2 *Illustrative Example 5: A 3D Geometric Parametrization for a Thermal Fin*

This problem addresses the performance of a heat sink for cooling electronic components. The heat sink is modelled as a spreader $\Omega_{\circ}^{1}$, see Fig. 6.9a, depicted in blue, which supports a plate fin $\Omega_{\circ}^{2}(\mu)$ exposed to flowing air, depicted in red, Fig. 6.9b. The heat transfer from the fin to the air is taken into account with the Biot number, which is the first parameter $\mu_{[1]}$. The second parameter is the relative length of the fin with respect to the spreader, and is labeled as $\mu_{[2]}$. The third parameter $\mu_{[3]}$ is the ratio between the thermal conductivity of the spreader and the fin. The parameters ranges are

$$\mu_{[1]} \in [0.1, 1.0], \qquad \mu_{[2]} \in [0.5, 10.0], \qquad \mu_{[3]} \in [1.0, 10.0].$$

A uniform heat flux is applied at the base of the spreader, denoted here by Γ_{bottom}, and Robin boundary conditions are imposed on the vertical faces of the fin, denoted here by Γ_{side}. Homogeneous Neumann conditions are imposed at all other surfaces. The bilinear and linear forms of this problem are given as

$$a_{\circ}(w, v; \mu) = \mu_{[3]} \int_{\Omega_{\circ}^{1}} \nabla w \cdot \nabla v + \int_{\Omega_{\circ}^{2}(\mu)} \nabla w \cdot \nabla v + \mu_{[1]} \int_{\Gamma_{\mathrm{side}}} w\, v,$$
$$f_{\circ}(v) = \int_{\Gamma_{\mathrm{bottom}}} v,$$

and the output of interest $s_{\circ}(\mu)$ is computed as

$$s_{\circ}(\mu) = f_{\circ}(u(\mu)), \tag{6.17}$$

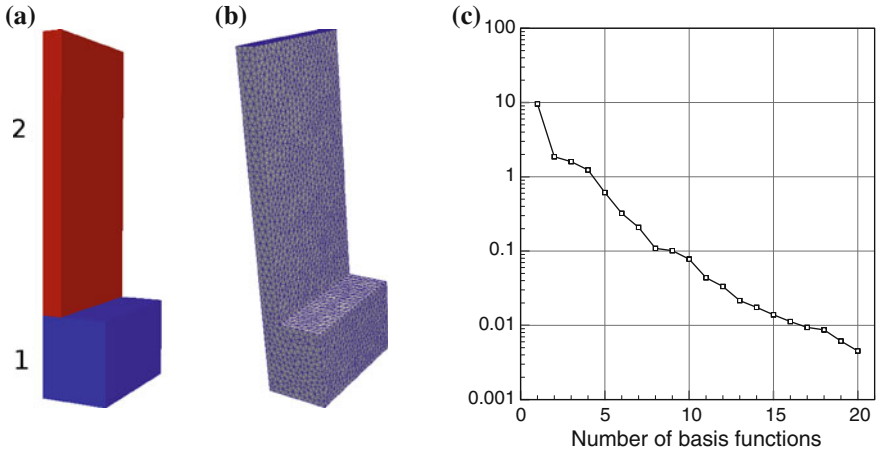

Fig. 6.9 The subdomain division (**a**), the finite element mesh used for the truth solver (**b**) and the maximum error over $\mathbb{P}_h$ of the greedy-algorithm with respect to the number of basis functions employed (**c**) for the thermal fin problem with a geometric parametrization

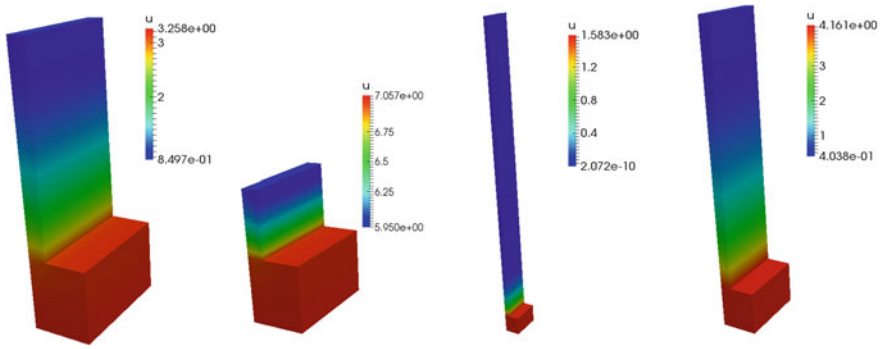

Fig. 6.10 First four basis functions for the thermal fin problem with a geometric parametrization

i.e., it is a compliant output. The finite element method, employing first order elements, is used as the truth model. In Fig. 6.9b the mesh of the reference domain is reported, featuring 22,979 elements.

The basis functions have been obtained by orthogonalization, through a Gram-Schimdt procedure, of snapshots computed by a greedy approach to select parameters $\mathbb{P}_h$ with cardinality of 3,000. In Fig. 6.9c the graph showing the maximum error with respect to the number of basis functions employed is reported.

In Fig. 6.10, the first four snapshots, corresponding to $u_\delta(\mu_1), \ldots, u_\delta(\mu_4)$, are depicted and in Fig. 6.11, the outcomes provided by the truth model and the reduced method, for a randomly chosen $\mu = (6.94, 1.36, 2.53)$ and $N = 20$, are compared. The pointwise difference between the two approximations is plotted as well, illustrating the excellent error behavior.

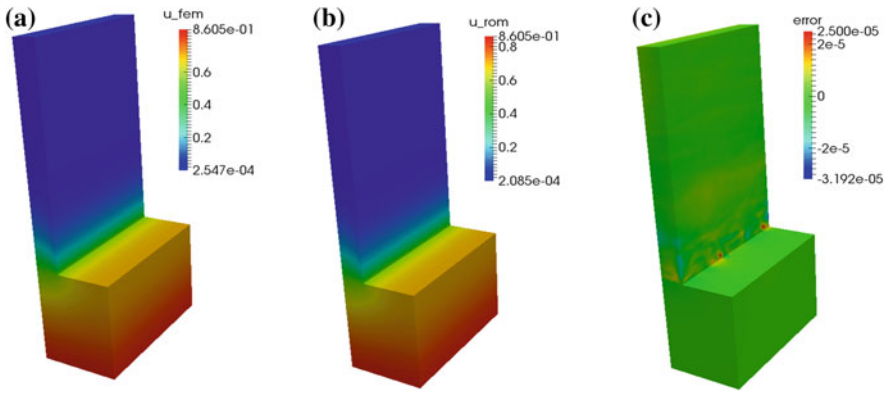

Fig. 6.11 Comparison between the truth approximation (**a**) and the reduced basis approximation (**b**) for $\mu = (1.176, 0.761, 0.530)$. The difference between the two solutions is reported in (**c**)

6.3 Non-compliant Output

For the sake of simplicity, we addressed in Chap. 3 the reduced basis approximation of affinely parametrized coercive problems in the compliant case. Let us now consider the more general non-compliant elliptic truth problem: given $\mu \in \mathbb{P}$, find

$$s_\delta(\mu) = \ell(u_\delta(\mu); \mu),$$

where $u_\delta(\mu) \in \mathbb{V}_\delta$ satisfies

$$a(u_\delta(\mu), v_\delta; \mu) = f(v_\delta; \mu), \quad \forall\, v_\delta \in \mathbb{V}_\delta.$$

We assume that $a(\cdot, \cdot; \mu)$ is coercive, continuous, and affine, but not necessarily symmetric. We further assume that both ℓ and f are bounded functionals but we no longer require that $\ell = f$.

Following the methodology outlined in Chap. 4, we can readily develop an a posteriori error bound for $s_{\mathrm{rb}}(\mu)$. By standard arguments [5, 23]

$$|s_\delta(\mu) - s_{\mathrm{rb}}(\mu)| \le \|\ell(\cdot; \mu)\|_{(\mathbb{V})'}\, \eta_{\mathrm{en}}(\mu)$$

where $\|u_\delta(\mu) - u_{\mathrm{rb}}(\mu)\|_\mu \le \eta_{\mathrm{en}}(\mu)$ and $\eta_{\mathrm{en}}(\mu)$ is given by (4.6a). We denote this approach, i.e., $\eta_{\mathrm{s}}(\mu) = \|\ell(\cdot; \mu)\|_{(\mathbb{V})'}\, \eta_{\mathrm{en}}(\mu)$, as primal-only. Although primal-only is perhaps the best approach in the case of many outputs in which each additional output, and the associated error bound, is an add-on, this approach has two drawbacks:

(i) We loose the quadratic convergence effect (see Proposition 4.1) for outputs, unless $\ell = f$ and $a(\cdot, \cdot; \mu)$ is symmetric, since the accuracy of the output is no longer the square of the accuracy of the field variable.

(ii) The effectivities $\eta_{\mathrm{s}}(\mu)/|s_\delta(\mu) - s_{\mathrm{rb}}(\mu)|$ may be unbounded. If $\ell = f$ we know from Proposition 4.1 that $|s_\delta(\mu) - s_{\mathrm{rb}}(\mu)| \sim \|\hat{r}_\delta(\mu)\|^2_{\mathbb{V}}$ and hence $\eta_{\mathrm{s}}(\mu)/|s_\delta(\mu) - s_{\mathrm{rb}}(\mu)| \sim 1/\|\hat{r}_\delta(\mu)\|_{\mathbb{V}} \to \infty$ as $N \to \infty$. Thus, the effectivity of the output error bound (4.6b) tends to infinity as ($N \to \infty$ and $u_{\mathrm{rb}}(\mu) \to u_\delta(\mu)$). We may expect similar behavior for any ℓ close to f. The problem is that (4.6b) does not reflect the contribution of the test space to the convergence of the output.

The introduction of reduced basis primal-dual approximations takes care of these issues and ensures a stable limit as $N \to \infty$. We introduce the dual problem associated with the functional $\ell(\cdot\,;\mu)$ as follows: find the adjoint or dual field $\psi(\mu) \in \mathbb{V}$ such that

$$a(v, \psi(\mu); \mu) = -\ell(v; \mu), \quad \forall\, v \in \mathbb{V}.$$

Let us define the reduced basis spaces for the primal and the dual problem, respectively, as

$$\mathbb{V}_{\mathrm{pr}} = \mathrm{span}\{u(\mu^n_{\mathrm{pr}}),\, 1 \le n \le N_{\mathrm{pr}}\},$$
$$\mathbb{V}_{\mathrm{du}} = \mathrm{span}\{\psi(\mu^n_{\mathrm{du}}),\, 1 \le n \le N_{\mathrm{du}}\}.$$

For our purpose, a single discrete truth space $\mathbb{V}_\delta$ suffices for both the primal and dual, although in many cases, the truth primal and dual spaces may be different. For a given $\mu \in \mathbb{P}$, the resulting reduced basis approximations $u_{\mathrm{rb}}(\mu) \in \mathbb{V}_{\mathrm{pr}}$ and $\psi_{\mathrm{rb}}(\mu) \in \mathbb{V}_{\mathrm{du}}$ solve

$$a(u_{\mathrm{rb}}(\mu), v_{\mathrm{rb}}; \mu) = f(v_{\mathrm{rb}}; \mu), \quad \forall v_{\mathrm{rb}} \in \mathbb{V}_{\mathrm{pr}},$$
$$a(v_{\mathrm{rb}}, \psi_{\mathrm{rb}}(\mu); \mu) = -\ell(v_{\mathrm{rb}}; \mu), \quad \forall v_{\mathrm{rb}} \in \mathbb{V}_{\mathrm{du}}.$$

Then, the reduced basis output can be evaluated as [6]

$$s_{\mathrm{rb}}(\mu) = \ell(u_{\mathrm{rb}}; \mu) - r_{\mathrm{pr}}(\psi_{\mathrm{rb}}; \mu),$$

where $r_{\mathrm{pr}}(\cdot; \mu), r_{\mathrm{du}}(\cdot; \mu) \in (\mathbb{V}_\delta)'$, are defined by

$$r_{\mathrm{pr}}(v_\delta; \mu) = f(v_\delta; \mu) - a(u_{\mathrm{rb}}, v_\delta; \mu), \qquad v_\delta \in \mathbb{V}_\delta,$$
$$r_{\mathrm{du}}(v_\delta; \mu) = -\ell(v_\delta; \mu) - a(v_\delta, \psi_{\mathrm{rb}}; \mu), \qquad v_\delta \in \mathbb{V}_\delta,$$

are the primal and the dual residuals. In the non-compliant case using a primal-dual strategy, the output error bound takes the form

$$\eta_{\mathrm{s}}(\mu) = \frac{\|r_{\mathrm{pr}}(\cdot\,;\mu)\|_{(\mathbb{V}_\delta)'}}{(\alpha_{\mathrm{LB}}(\mu))^{1/2}} \, \frac{\|r_{\mathrm{du}}(\cdot\,;\mu)\|_{(\mathbb{V}_\delta)'}}{(\alpha_{\mathrm{LB}}(\mu))^{1/2}}.$$

We thus recover the quadratic convergence of the output effect. Note that the offline-online procedure is very similar to the primal-only case, but now everything must be done for both the primal and the dual problem. Moreover, we need to evaluate both a primal and a dual residual for the a posteriori error bounds. Error bounds related to the gradient of computed quantities, such as velocity and pressure in potential flow problems, have been addressed in [16]. For parabolic problems, the treatment of non-compliant outputs follows the same strategy; we only note that the dual problem in this case must evolve backward in time [1].

6.3.1 *Illustrative Example 6: A 2D Graetz Problem with Non-compliant Output*

This example, known as the two-dimensional Graetz problem, aims to illustrate the computational details of a problem with a non-compliant output of interest. In particular, we consider forced heat convection in a geometrically parametrized channel divided into two parts such that $\Omega_o(\mu) = \Omega_o^1 \cup \Omega_o^2(\mu)$ as illustrated in Fig. 6.12. Within the first part Ω_o^1 (in blue) the temperature of the flow at the wall is kept constant and the flow has a prescribed convective field. The length along the axis x of the subdomain $\Omega_o^2(\mu)$, relative to the length of the subdomain Ω_o^1, is given by the parameter $\mu_{[1]}$. The heat transfer between the domains can be taken into account by means of the Péclet number, measuring the ratio between the convective and the conduction heat transfer, which will be labeled as the parameter $\mu_{[2]}$. The ranges of the two parameters are

$$\mu_{[1]} \in [0.1, 10], \qquad \mu_{[2]} \in [0.01, 10],$$

so that $\mathbb{P} = [0.1, 10] \times [0.01, 10]$.

This problem represents an example of a non-symmetric operator, associated with a primal-dual formulation for the non-compliant output, as well as a geometric parametrization treated by a transformation to a reference domain. This is a simplified version of the larger problem, discussed in detail in Sect. 6.5. We first need to reformulate the problem to obtain a Galerkin formulation, based on an identical trial and test space. We introduce the two parameter-dependent spaces

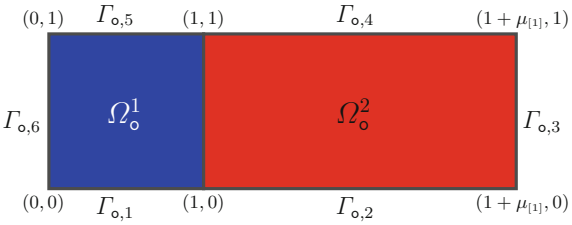

Fig. 6.12 The geometric set-up for the non-compliant and geometrically parametrized Graetz problem

$$\hat{\mathbb{V}}(\mu) = \left\{ v \in H^1\big(\Omega_\circ(\mu)\big) \,\middle|\, v|_{\Gamma_{\circ,1,5,6}} = 0,\, v|_{\Gamma_{\circ,2,4}} = 1 \right\},$$
$$\mathbb{V}(\mu) = \left\{ v \in H^1\big(\Omega_\circ(\mu)\big) \,\middle|\, v|_{\Gamma_{\circ,1,2,4,5,6}} = 0 \right\},$$

where we use the notation $\Gamma_{\circ,1,5,6} = \Gamma_{\circ,1} \cup \Gamma_{\circ,5} \cup \Gamma_{\circ,6}$ and similarly for $\Gamma_{\circ,2,4}$ and $\Gamma_{\circ,1,2,4,5,6}$. The original problem can be stated as follows: for any $\mu = (\mu_{[1]}, \mu_{[2]})$, find $\hat{u}_\circ(\mu) \in \hat{\mathbb{V}}(\mu)$ such that

$$\frac{1}{\mu_{[2]}} \int_{\Omega_\circ(\mu)} \nabla \hat{u}_\circ(\mu) \cdot \nabla v + \int_{\Omega_\circ(\mu)} x_2(1-x_2)\partial_{x_1} \hat{u}_\circ(\mu)\, v = 0, \qquad \forall v \in \mathbb{V}(\mu).$$

Further, set $\bar{\mu} \in \mathbb{P}$ such that $\bar{\mu}_{[1]} = 1$ and define $\Omega = \Omega_\circ(\bar{\mu})$, $\hat{\mathbb{V}} = \hat{\mathbb{V}}(\bar{\mu})$, $\mathbb{V} = \mathbb{V}(\bar{\mu})$. After a mapping onto the reference domain Ω, the problem can be reformulated on Ω as follows: find $\hat{u}(\mu) \in \hat{\mathbb{V}}$ such that

$$\frac{1}{\mu_{[2]}} \int_{\Omega^1} \nabla \hat{u}(\mu) \cdot \nabla v + \frac{1}{\mu_{[1]}\mu_{[2]}} \int_{\Omega^2} \partial_{x_1} \hat{u}(\mu)\, \partial_{x_1} v + \mu_{[1]}\mu_{[2]} \int_{\Omega^2} \partial_{x_2} \hat{u}(\mu)\, \partial_{x_2} v$$
$$+ \int_{\Omega} x_2(1-x_2)\partial_{x_1} \hat{u}(\mu)\, v = 0, \quad \forall v \in \mathbb{V},$$

with $\Omega^1 = \Omega^1_\circ$, $\Omega^2 = \Omega^2_\circ(\bar{\mu})$. Finally, using the (continuous) finite element interpolation R_δ of the lifting function $R = \mathbf{1}_{\Omega^2}$ (the characteristics functions of Ω^2), the discrete problem is written as: find $u_\delta(\mu) \in \mathbb{V}_\delta \subset \mathbb{V}$ such that

$$a(u_\delta(\mu), v_\delta; \mu) = f(v_\delta; \mu), \qquad \forall v_\delta \in \mathbb{V}_\delta,$$

where

$$a(w, v; \mu) = \mu_{[2]} \int_{\Omega^1} \nabla w \cdot \nabla v + \frac{\mu_{[2]}}{\mu_{[1]}} \int_{\Omega^2} \partial_x w\, \partial_{x_1} v + \mu_{[2]}\mu_{[1]} \int_{\Omega^2} \partial_{x_2} w\, \partial_{x_2} v$$
$$+ \int_{\Omega} x_2(1-x_2)\, \partial_{x_1} w v,$$
$$f(v; \mu) = -a(R_\delta, v; \mu).$$

Finally, the non-compliant output of interest $s(\mu)$ is given by

$$s_\delta(\mu) = \int_{\Gamma_{\circ,3}} u_\delta(\mu).$$

In Fig. 6.13 we first illustrate the exponential convergence of the successive constraint method (SCM) with respect to the number of eigenvalues for this more complicated problem. As expected, one observes a rapidly converging model. We also show the convergence of both the primal and the dual reduced basis approximation with

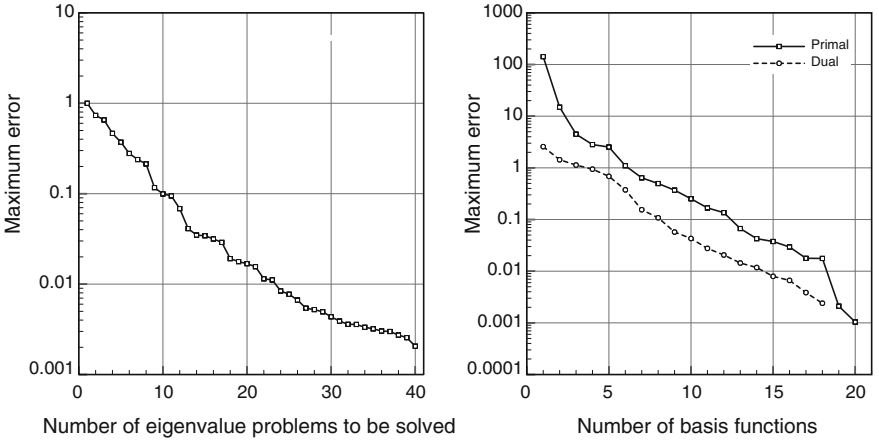

Fig. 6.13 Convergence of the SCM-algorithm with respect to the number of eigenvalue problems to be solved (*left*) and convergence of the maximum absolute error bound with respect to the number of selected basis functions for the primal and the dual reduced basis (*right*) for the non-compliant numerical example

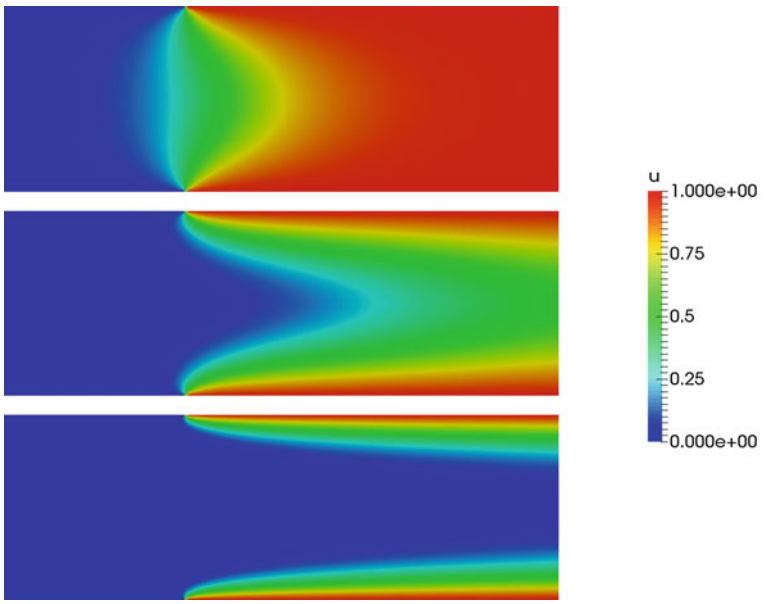

Fig. 6.14 Three representative solutions for a fixed geometry ($\mu_{[1]} = 2$) and different Péclet numbers $\mu_{[2]} = 0.1, 1, 10$ (from *top* to *bottom*) for the 2D Graetz problem

Table 6.3 Error bounds and output error bounds with effectivity metrics as function of N for the non-compliant example for the error in the solution $u_{\mathrm{rb}}(\mu)$ (*top*) and the output $s_{\mathrm{rb}}(\mu)$ (*bottom*)

N	$\eta_{\mathrm{en,av}}$	$\mathtt{eff}_{\mathrm{en,max}}$	$\mathtt{eff}_{\mathrm{en,av}}$
5	1.41e−1	13.31	3.37
10	7.90e−2	8.52	3.24
15	7.76e−3	10.47	2.97
20	1.86e−3	13.01	3.60
N	$\eta_{\mathrm{s,av}}$	$\mathtt{eff}_{\mathrm{s,max}}$	$\mathtt{eff}_{\mathrm{s,av}}$
5	1.89e−2	12.01	4.32
10	6.20e−3	18.73	6.78
15	5.74e−5	27.26	9.67
20	3.46e−6	48.15	15.43

increasing number of basis elements. Again we observe exponential convergence to the truth approximation. In Fig. 6.14 we illustrate the reduced basis approximation for different values of the Péclet number for this two-dimensional Graetz problem, illustrating the substantial dynamic variation of the solution. Finally, we show in Table 6.3 the effectivities of the error estimator for both the solution and the output of interest. We observe in particular the quadratic convergence of the error estimate of the output as expected from the primal-dual approach. Reduced basis approximation of convection-conduction and diffusion-transport(advection)-reaction problems is still an open research field, above all concerning stabilization techniques required also in the online part of the problem, taking into account parametrized propagating fronts and boundary layers at higher Péclet numbers [24–27] and in perspective higher Reynolds numbers for viscous flows.

6.4 Non-coercive Problems

Let us now discuss wider classes of problems to treat with the reduced basis method, including non-coercive problems, such as Helmholtz, Maxwell's equations, saddle-point problems like Stokes problem.

The reduced basis framework can be effectively applied to problems involving operators which do not satisfy the coercivity assumption [4]. Examples include the (Navier)-Stokes problem, where stability is fulfilled in the general sense of the inf-sup condition [28]. For the sake of simplicity, we restrict our discussion to the elliptic scalar case (2.1)–(2.2) and assume that the (parametrized) bilinear form $a(\cdot\,,\,\cdot\,;\mu) : \mathbb{V} \times \mathbb{W} \to \mathbb{R}$ is continuous and satisfies the inf-sup condition

$$\exists \beta > 0 \;:\; \beta(\mu) = \inf_{v \in \mathbb{V}\setminus\{0\}} \sup_{w \in \mathbb{W}\setminus\{0\}} \frac{a(v,w;\mu)}{\|v\|_{\mathbb{V}}\|w\|_{\mathbb{W}}} \geq \beta \qquad \forall\, \mu \in \mathbb{P}.$$

We refer to the Appendix for a brief overview view on the numerical analysis of this class of problems in a unparametrized setting. The discrete, and thus the subsequent reduced basis, approximation is based on a more general Petrov-Galerkin approach. Given two truth spaces $\mathbb{V}_\delta \subset \mathbb{V}$, $\mathbb{W}_\delta \subset \mathbb{W}$, the truth approximation $u_\delta(\mu) \in \mathbb{V}_\delta$ satisfies

$$a(u_\delta(\mu), w_\delta, \mu) = f(w_\delta; \mu), \qquad \forall w_\delta \in \mathbb{W}_\delta,$$

and the output can be evaluated as[1]

$$s_\delta(\mu) = \ell(u_\delta(\mu); \mu).$$

To have a stable truth approximation, we require that there exists $\beta_\delta > 0$ such that

$$\beta_\delta(\mu) = \inf_{v_\delta \in \mathbb{V}_\delta \setminus \{0\}} \sup_{w_\delta \in \mathbb{W}_\delta \setminus \{0\}} \frac{a(v_\delta, w_\delta; \mu)}{\|v_\delta\|_{\mathbb{V}} \|w_\delta\|_{\mathbb{W}}} \geq \beta_\delta \qquad \forall\, \mu \in \mathbb{P}. \tag{6.18}$$

The reduced basis approximation inherits the same Petrov-Galerkin structure: Given some $\mu \in \mathbb{P}$, find $u_{\mathtt{rb}}(\mu) \in \mathbb{V}_{\mathtt{rb}}$ such that

$$a(u_{\mathtt{rb}}(\mu), w_{\mathtt{rb}}; \mu) = f(w_{\mathtt{rb}}; \mu), \quad \forall w_{\mathtt{rb}} \in \mathbb{W}^\mu_{\mathtt{rb}},$$

and evaluate

$$s_{\mathtt{rb}}(\mu) = \ell(u_{\mathtt{rb}}(\mu); \mu),$$

where the test and trial spaces are taken to be of the form

$$\mathbb{V}_{\mathtt{rb}} = \mathrm{span}\{u_\delta(\mu_n) \mid 1 \leq n \leq N\}, \qquad \mathbb{W}^\mu_{\mathtt{rb}} = \mathrm{span}\{\mathtt{A}^\mu_\delta u_\delta(\mu_n) \mid 1 \leq n \leq N\},$$

for a common set of parameter points $\{\mu_n\}_{n=1}^N$. The so-called inner supremizer operator $\mathtt{A}^\mu_\delta : \mathbb{V}_\delta \to \mathbb{W}_\delta$ is defined by

$$(\mathtt{A}^\mu_\delta v_\delta, w_\delta)_{\mathbb{W}} = a(v_\delta, w_\delta; \mu), \qquad \forall\, v_\delta \in \mathbb{V}_\delta,\ \forall\, w_\delta \in \mathbb{W}_\delta,$$

which can be shown to realize the supremum:

$$\beta_\delta(\mu) = \inf_{v_\delta \in \mathbb{V}_\delta \setminus \{0\}} \sup_{w_\delta \in \mathbb{W}_\delta \setminus \{0\}} \frac{a(v_\delta, w_\delta; \mu)}{\|v_\delta\|_{\mathbb{V}} \|w_\delta\|_{\mathbb{W}}} = \inf_{v_\delta \in \mathbb{V}_\delta \setminus \{0\}} \frac{a(v_\delta, \mathtt{A}^\mu_\delta v_\delta; \mu)}{\|v_\delta\|_{\mathbb{V}} \|\mathtt{A}^\mu_\delta v_\delta\|_{\mathbb{W}}}.$$

Applying the inner supremizer and the Cauchy-Schwarz inequality implies

[1] We consider here a primal approximation. However, we can readily extend the approach to a primal-dual formulation as described for coercive problems in Sect. 6.3. See also [18].

$$\sup_{w_\delta \in \mathbb{W}_\delta \setminus \{0\}} \frac{a(v_\delta, w_\delta; \mu)}{\|v_\delta\|_{\mathbb{V}} \|w_\delta\|_{\mathbb{W}}} = \sup_{w_\delta \in \mathbb{W}_\delta \setminus \{0\}} \frac{(\mathtt{A}_\delta^\mu v_\delta, w_\delta)_{\mathbb{W}}}{\|v_\delta\|_{\mathbb{V}} \|w_\delta\|_{\mathbb{W}}}$$
$$\leq \sup_{w_\delta \in \mathbb{W}_\delta \setminus \{0\}} \frac{\|\mathtt{A}_\delta^\mu v_\delta\|_{\mathbb{W}} \|w_\delta\|_{\mathbb{W}}}{\|v_\delta\|_{\mathbb{V}} \|w_\delta\|_{\mathbb{W}}} = \frac{\|\mathtt{A}_\delta^\mu v_\delta\|_{\mathbb{W}}^2}{\|v_\delta\|_{\mathbb{V}} \|\mathtt{A}_\delta^\mu v_\delta\|_{\mathbb{W}}} = \frac{a(v_\delta, \mathtt{A}_\delta^\mu v_\delta; \mu)}{\|v_\delta\|_{\mathbb{V}} \|\mathtt{A}_\delta^\mu v_\delta\|_{\mathbb{W}}},$$

for all $\forall v_\delta \in \mathbb{V}_\delta$ such that

$$\beta_\delta(\mu) \leq \inf_{v_\delta \in \mathbb{V}_\delta \setminus \{0\}} \frac{a(v_\delta, \mathtt{A}_\delta^\mu v_\delta; \mu)}{\|v_\delta\|_{\mathbb{V}} \|\mathtt{A}_\delta^\mu v_\delta\|_{\mathbb{W}}} \leq \beta_\delta(\mu).$$

The particular ansatz of the test and trial spaces ensures stability of the reduced basis approximation as seen by the following development

$$\inf_{v_{\mathtt{rb}} \in \mathbb{V}_{\mathtt{rb}} \setminus \{0\}} \sup_{w_{\mathtt{rb}} \in \mathbb{W}_{\mathtt{rb}}^\mu \setminus \{0\}} \frac{a(v_{\mathtt{rb}}, w_{\mathtt{rb}}; \mu)}{\|v_{\mathtt{rb}}\|_{\mathbb{V}} \|w_{\mathtt{rb}}\|_{\mathbb{W}}} = \inf_{v_{\mathtt{rb}} \in \mathbb{V}_{\mathtt{rb}} \setminus \{0\}} \frac{a(v_{\mathtt{rb}}, \mathtt{A}_\delta^\mu v_{\mathtt{rb}})_{\mathbb{W}}}{\|v_{\mathtt{rb}}\|_{\mathbb{V}} \|\mathtt{A}_\delta^\mu v_{\mathtt{rb}}\|_{\mathbb{W}}}$$
$$\geq \inf_{v_\delta \in \mathbb{V}_\delta \setminus \{0\}} \frac{a(v_\delta, \mathtt{A}_\delta^\mu v_\delta)_{\mathbb{W}}}{\|v_\delta\|_{\mathbb{V}} \|\mathtt{A}_\delta^\mu v_\delta\|_{\mathbb{W}}} = \beta_\delta(\mu),$$

where we used again that $\mathtt{A}_\delta^\mu v_{\mathtt{rb}}$ is the supremizer but this time of the supremum over $\mathbb{W}_{\mathtt{rb}}^\mu \setminus \{0\}$ rather than $\mathbb{W}_\delta \setminus \{0\}$. The arguments are identical. Therefore, this ultimately guarantees that the reduced basis inf-sup constant

$$\beta_{\mathtt{rb}}(\mu) = \inf_{v_{\mathtt{rb}} \in \mathbb{V}_{\mathtt{rb}} \setminus \{0\}} \sup_{w_{\mathtt{rb}} \in \mathbb{W}_{\mathtt{rb}}^\mu \setminus \{0\}} \frac{a(v_{\mathtt{rb}}, w_{\mathtt{rb}}; \mu)}{\|v_{\mathtt{rb}}\|_{\mathbb{V}} \|w_{\mathtt{rb}}\|_{\mathbb{W}}} \geq \beta_\delta \qquad \forall \mu \in \mathbb{P}, \tag{6.19}$$

is bounded from below by the truth uniform inf-sup constant β_δ, for any $\mu \in \mathbb{P}$ and it holds

$$\|u_\delta(\mu) - u_{\mathtt{rb}}(\mu)\|_{\mathbb{V}} \leq \left(1 + \frac{\gamma(\mu)}{\beta_\delta(\mu)}\right) \inf_{v_{\mathtt{rb}} \in \mathbb{V}_{\mathtt{rb}}} \|u_\delta(\mu) - v_{\mathtt{rb}}\|_{\mathbb{V}},$$

which is the analogue of (3.7) for non-coercive problems. Observe that the approximation is provided by $\mathbb{V}_{\mathtt{rb}}$ and stability (through $\beta_{\mathtt{rb}}$) by $\mathbb{W}_{\mathtt{rb}}^\mu$.

The offline-online computational strategies, as well as the a posteriori error estimation, are based on the same arguments as in Chap. 4 for the coercive case. We note that the inner supremizer operator can be written in the affine form under the affinity assumption (3.11) on $a(\cdot, \cdot; \mu)$. In particular, from (6.19), we can obtain that

$$\eta_{\mathbb{V}} = \frac{\|\hat{r}_\delta(\mu)\|_{\mathbb{V}}}{\beta_{\mathtt{LB}}(\mu)},$$

where $\beta_{\text{LB}}(\mu)$ is a lower bound of the inf-sup constant $\beta_\delta(\mu)$ defined in (6.18) and can be computed by means of the successive constrain method (SCM) procedure used for the lower bound of coercivity constants [17, 29].

An interesting case of non-coercive problems is given by Stokes problems where approximation stability is guaranteed by the fullfillment of an equivalent inf-sup stability condition on the pressure term with reduced basis approximation spaces properly enriched [30, 31]. Error bounds can be developed in the general non-coercive framework [17, 32] or within a penalty setting [33].

6.5 Illustrative Example 7: A 3D Parametrized Graetz Channel

In this final numerical example, we return to a classic problem, dealing with forced steady heat convection combined with heat conduction in a duct with walls at different temperature and of different lengths. The first segment of the duct has cold walls, while the second segment has hot walls. The flow has an imposed temperature at the inlet and a known convection field, i.e., a given parabolic Poiseuille-type velocity profile. From an engineering point of view, this example illustrates the application of convection and conduction analysis to an important class of heat transfer problems in fluidic devices. From a physical point of view, the problem illustrates many aspects of steady convection-diffusion phenomena such as heat transfer into a channel, forced convection with an imposed velocity profile and heat conduction through walls and insulation [34]. The Péclet number, providing a measure of relative axial transport velocity field, and the length of the hot portion of the duct are only some of the interesting parameters used to extract average temperatures.

The forced heat convection flows into a channel that is divided into two parts, illustrated in Fig. 6.15 (left), such that $\Omega_\circ(\mu) = \Omega_\circ^1 \cup \Omega_\circ^2(\mu)$. Within the first part $\Omega_\circ^1$ (in blue) the temperature is kept constant and the flow has a given convective field. The length of the axis x of $\Omega_\circ^2(\mu)$, relative to the length of $\Omega_\circ^1$, is given by the

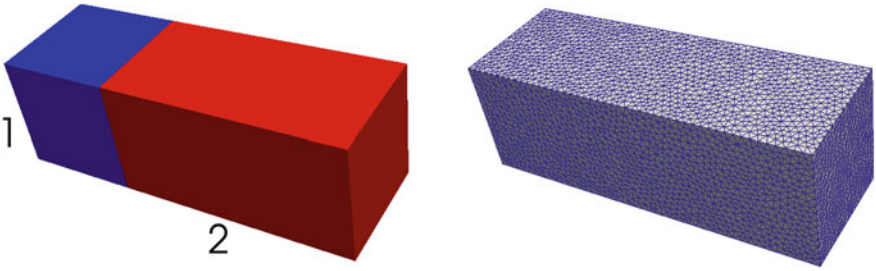

Fig. 6.15 Subdomain division (*left*) and the finite element mesh (*right*) for the three-dimensional Graetz convective problem

geometrical parameter $\mu_{[1]}$. To maintain a common notation as in Sect. 6.2 we denote the domains $\Omega_\circ(\mu)$ and $\Omega^2_\circ(\mu)$ despite the fact that the only geometrical parameter is $\mu_{[1]}$. The heat transfer between the domains can be taken into account by means of the Péclet number, labeled as the physical parameter $\mu_{[2]}$. A constant heat flux is imposed on the walls of the right domain $\Omega^2_\circ(\mu)$, which is the third parameter $\mu_{[3]}$. The ranges of the three parameters are the following:

$$\mu = (\mu_{[1]}, \mu_{[2]}, \mu_{[3]}) \in \mathbb{P} = [1.0, 10.0] \times [0.1, 100.0] \times [-1.0, 1.0].$$

At the inlet of $\Omega^1_\circ$ and on its walls, homogeneous boundary conditions are imposed. At the interface between $\Omega^1_\circ$ and $\Omega^2_\circ(\mu)$, the continuity of the temperature and the heat flux are imposed. At the outlet of $\Omega^2_\circ(\mu)$, denoted by $\Gamma_{2,\text{outlet}}(\mu)$, a homogeneous Neumann boundary condition is imposed. On the lateral walls of $\Omega^2_\circ(\mu)$, denoted by $\Gamma_{2,\text{side}}(\mu)$, the heat flux is imposed trough a Neumann boundary condition given by the parameter $\mu_{[3]}$. The non-symmetric bilinear and linear forms of the problem are given by:

$$a_\circ(w, v; \mu) = \frac{1}{\mu_{[2]}} \int_{\Omega_\circ(\mu)} \nabla w \cdot \nabla v + \int_{\Omega_\circ(\mu)} 10 \left(x_2(1 - x_2) + x_3(1 - x_3)\right) \partial_{x_1} w \, v,$$

$$f_\circ(v) = \mu_{[3]} \int_{\Gamma_{2,\text{side}}(\mu)} v.$$

The output of interest $s_\circ(\mu)$ is

$$s_\circ(\mu) = \int_{\Gamma_{2,\text{outlet}}(\mu)} u_\circ(\mu).$$

It is worth mentioning that the aforementioned output is non-compliant and the state equation is a non-symmetric operator. The finite element method, employing first order elements, is used as the truth model. In Fig. 6.15 (right) the mesh of the reference domain is reported, and features 51,498 elements. The basis functions have been obtained by orthogonalization, through a Gram-Schimdt procedure, of snapshots selected by a greedy approach across $\mathbb{P}_h$ with cardinality 10,000. In Fig. 6.16 (left) the graph shows the maximum error (over $\mathbb{P}_h$) with respect to the number of basis functions employed and Fig. 6.16 (right) depicts the average (over $\mathbb{P}_h$) error and estimator for the field variable and the output functional with respect to the number of basis functions employed. One can clearly observe the quadratic convergence of the output functional due to the primal-dual reduced basis approximation.

In Fig. 6.17, the outcomes provided by the RBM, for two different values of the Péclet number $\mu_{[2]}$ are reported. In particular, $\mu = (2.0, 0.1, 1.0)$ and $\mu = (2.0, 100.0, 1.0)$ have been considered. These two cases show the very different physics induced by the Péclet number: a conduction phenomenon without a thermal boundary layer and a well developed thermal boundary layer when the convective field is dominating, respectively.

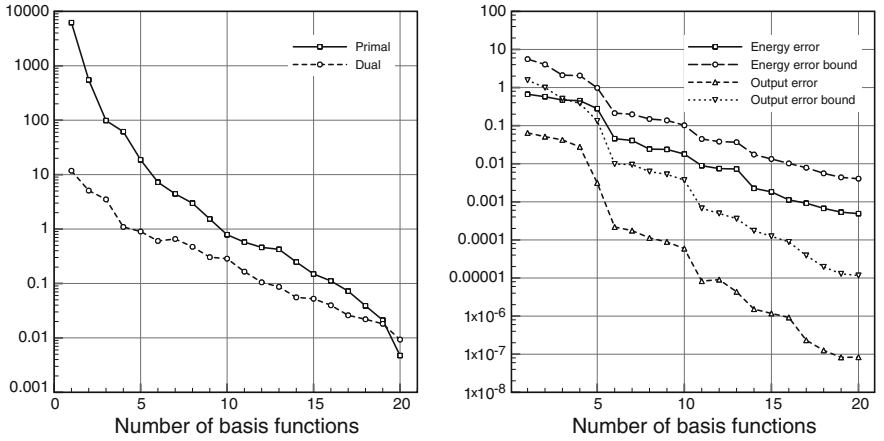

Fig. 6.16 Maximum (over $\mathbb{P}_h$) error for the primal and dual problem (*left*) and the average (over $\mathbb{P}_h$) error and estimator for the field variable and output functional (*right*) with respect to the number of basis functions employed for the three-dimensional Graetz channel

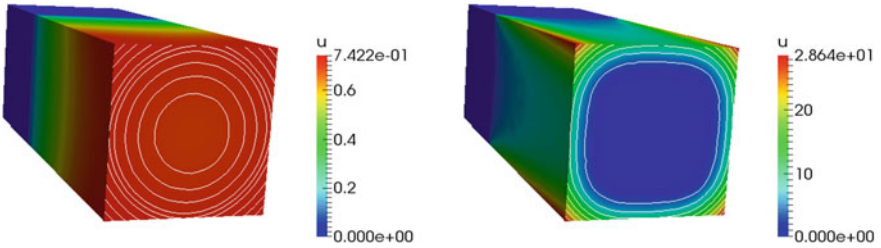

Fig. 6.17 Temperature field, provided by the reduced basis solution when (*left*) $\mu = (2.0, 0.1, 1.0)$, without a thermal boundary layer and (*right*) $\mu = (2.0, 100.0, 1.0)$, showing a well developed thermal boundary layer

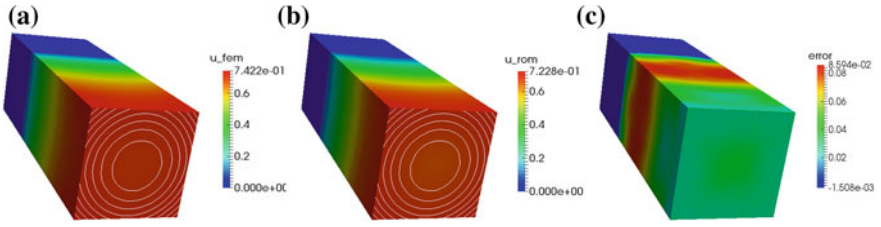

Fig. 6.18 Comparison between the truth model (**a**), the reduced basis approximation (**b**) for $\mu = (2.0, 0.1, 1.0)$. The pointwise error between the two solutions is reported in (**c**)

In the following, the representative solutions provided by the truth (finite element) model and the RBM for three different values of the Péclet number $\mu_{[2]}$ are reported in Figs. 6.18, 6.19, 6.20. In particular, $\mu = (2.0, 0.1, 1.0)$, $\mu = (2.0, 35.0, 1.0)$ and $\mu = (2.0, 100.0, 1.0)$ have been chosen. The pointwise error is reported too.

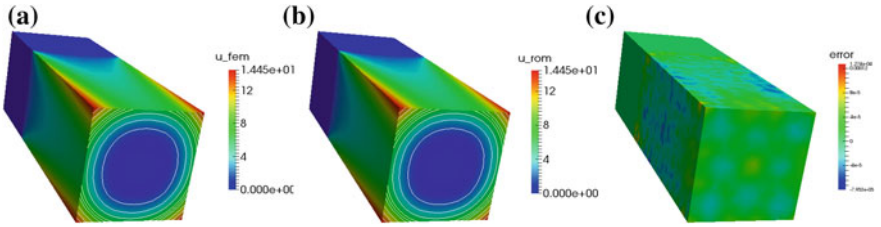

Fig. 6.19 Comparison between the truth model (**a**), the reduced basis approximation (**b**) for $\mu =$ (2.0, 35.0, 1.0). The pointwise error between the two solutions is reported in (**c**)

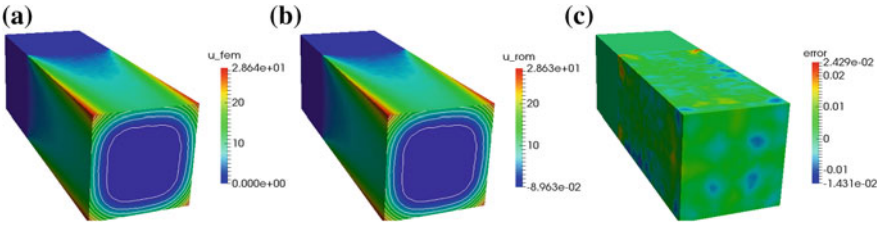

Fig. 6.20 Comparison between the truth model (**a**), the reduced basis approximation (**b**) for $\mu =$ (2.0, 100.0, 1.0). The pointwise error between the two solutions is reported in (**c**)

In Table 6.4 the output error bounds and solution error bounds, respectively, as well as effectivity metrics as function of basis functions employed are presented. The values have been averaged over 2,000 samples and we observe again the expected quadratic convergence of the output and very reasonable values of the effectivities.

Table 6.4 Error bounds and output error bounds with effectivity metrics as function of N for the three-dimensional Graetz problem for the error in the solution $u_{\mathrm{rb}}(\mu)$ (*top*) and the output $s_{\mathrm{rb}}(\mu)$ (*bottom*)

N	$\eta_{\mathrm{en,av}}$	$\mathtt{eff}_{\mathrm{en,max}}$	$\mathtt{eff}_{\mathrm{en,av}}$
5	0.98	21.54	3.51
10	0.10	26.05	5.69
15	0.013	20.78	7.29
20	4.0e–3	35.33	8.27
N	$\eta_{\mathrm{s,av}}$	$\mathtt{eff}_{\mathrm{s,max}}$	$\mathtt{eff}_{\mathrm{s,av}}$
5	0.96	18.45	5.78
10	0.012	24.82	7.26
15	0.0017	19.64	9.22
20	0.16e–5	26.86	11.38

References

1. M.A. Grepl, A.T. Patera, A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations. ESAIM Math. Model. Numer. Anal. **39**, 157–181 (2005)
2. N. Nguyen, G. Rozza, P. Huynh, A.T. Patera, in *Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Parabolic Pdes; Application to Real-Time Bayesian Parameter Estimation*, ed. by L. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, L. Tenorio, B. van Bloemen Waanders, K. Willcox. Computational Methods for Large Scale Inverse Problems and Uncertainty Quantification (Wiley, UK, 2009)
3. B. Haasdonk, M. Ohlberger, Reduced basis method for finite volume approximations of parametrized linear evolution equations. ESAIM Math. Model. Numer. Anal. **42**, 277–302 (2008)
4. K. Veroy, C. Prud'homme, D. Rovas, A. Patera, A posteriori error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations, in *Proceedings of the 16th AIAA Computational Fluid Dynamics Conference*, vol. 3847 (2003)
5. G. Rozza, P. Huynh, A.T. Patera, Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: Application to transport and continuum mechanics. Arch. Comput. Methods Eng. **15**, 229–275 (2008)
6. N.A. Pierce, M.B. Giles, Adjoint recovery of superconvergent functionals from pde approximations. SIAM Rev. **42**, 247–264 (2000)
7. R. Milani, A. Quarteroni, G. Rozza, Reduced basis method for linear elasticity problems with many parameters. Comput. Methods Appl. Mech. Eng. **197**, 4812–4829 (2008)
8. N.N. Cuong, K. Veroy, A.T. Patera, in *Certified Real-Time Solution of Parametrized Partial Differential Equations*. Handbook of Materials Modeling (Springer, 2005), pp. 1529–1564
9. P. Huynh, A. Patera, Reduced basis approximation and a posteriori error estimation for stress intensity factors. Int. J. Numer. Methods Eng. **72**, 1219–1259 (2007)
10. P. Huynh, G. Rozza, *Reduced Basis Method and A Posteriori Error Estimation: Application to Linear Elasticity Problems* (2014)
11. S. Deparis, G. Rozza, Reduced basis method for multi-parameter-dependent steady Navier-Stokes equations: applications to natural convection in a cavity. J. Comput. Phys. **228**, 4359–4378 (2009)
12. G. Rozza, N.C. Nguyen, A.T. Patera, S. Deparis, Reduced Basis Methods and A Posteriori Error Estimators for Heat Transfer Problems, in *ASME, Heat Transfer Summer Conference collocated with the InterPACK09 and 3rd Energy Sustainability Conferences, American Society of Mechanical Engineers* (2009), pp. 753–762
13. G. Rozza, P. Huynh, N.C. Nguyen, A.T. Patera, Real-Time Reliable Simulation of Heat Transfer Phenomena, in *ASME, Heat Transfer Summer Conference collocated with the InterPACK09 and 3rd Energy Sustainability Conferences, American Society of Mechanical Engineers* (2009), pp. 851–860
14. F. Gelsomino, G. Rozza, Comparison and combination of reduced-order modelling techniques in 3D parametrized heat transfer problems. Math. Comput. Model. Dyn. Syst. **17**, 371–394 (2011)
15. S. Sen, K. Veroy, D. Huynh, S. Deparis, N.C. Nguyen, A.T. Patera, Natural norm a posteriori error estimators for reduced basis approximations. J. Comput. Phys. **217**, 37–62 (2006)
16. G. Rozza, Reduced asis approximation and error bounds for potential flows in parametrized geometries. Commun. Comput. Phys. **9**, 1–48 (2011)
17. G. Rozza, P. Huynh, A. Manzoni, Reduced basis approximation and a posteriori error estimation for Stokes flows in parametrized geometries: roles of the inf-sup stability constants. Numer. Math. **125**, 115–152 (2013)
18. Y. Chen, J.S. Hesthaven, Y. Maday, J. Rodríguez, Certified reduced basis methods and output bounds for the harmonic Maxwell's equations. SIAM J. Sci. Comput. **32**, 970–996 (2010)
19. A. Manzoni, A. Quarteroni, G. Rozza, Model reduction techniques for fast blood flow simulation in parametrized geometries. Int. J. Numer. Methods Biomed. Eng. **28**, 604–625 (2012)

20. A. Manzoni, T. Lassila, A. Quarteroni, G. Rozza, in *A Reduced-Order Strategy for Solving Inverse Bayesian Shape Identification Problems in Physiological Flows*, ed. by H.G. Bock, X.P. Hoang, R. Rannacher, J.P. Schloder, Modeling, Simulation and Optimization of Complex Processes—HPSC 2012 (Springer International Publishing, 2014), pp. 145–155
21. T. Lassila, A. Manzoni, G. Rozza, in *Reduction Strategies for Shape Dependent Inverse Problems in Haemodynamics*, ed. by D. Homberg, F. Troltzsch, System Modeling and Optimization. IFIP Advances in Information and Communication Technology, vol. 391 (Springer, Berlin, 2013), pp. 397–406
22. G. Rozza, T. Lassila, A. Manzoni, in *Reduced Basis Approximation for Shape Optimization in Thermal Flows with a Parametrized Polynomial Geometric Map*. Spectral and High Order Methods for Partial Differential Equations (Springer, 2011), pp. 307–315
23. A.T. Patera, G. Rozza, *Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations*, Copyright MIT 2007, MIT Pappalardo Graduate Monographs in Mechanical Engineering (2007). http://augustine.mit.edu
24. P. Pacciarini, G. Rozza, Stabilized reduced basis method for parametrized advection-diffusion PDEs. Comput. Methods Appl. Mech. Eng. **274**, 1–18 (2014)
25. P. Pacciarini, G. Rozza, in *Reduced Basis Approximation of Parametrized Advection-Diffusion PDEs with High Peclet Number*, ed. by A. Abdulle, S. Deparis, D. Kressner, F. Nobile, M. Picasso, Numerical Mathematics and Advanced Applications—ENUMATH 2013. Lecture Notes in Computational Science and Engineering, vol. 103 (Springer International Publishing, 2015), pp. 419–426
26. W. Dahmen, C. Plesken, G. Welper, Double greedy algorithms: reduced basis methods for transport dominated problems. ESAIM: M2AN **48**, 623–663 (2014)
27. P. Pacciarini, G. Rozza, Stabilized Reduced Basis Method for Parametrized Scalar Advection-Diffusion Problems at Higher Péclet Number: Roles of the Boundary Layers and Inner Fronts, in *Proceedings of the Jointly Organized 11th World Congress on Computational Mechanics—WCCM XI, 5th European Congress on Computational Mechanics—ECCM V, 6th European Congress on Computational Fluid Dynamics—ECFD VI* (2014), pp. 5614–5624
28. A. Quarteroni, A. Valli, in *Numerical approximation of partial differential equations*. Springer Series in Computational Mathematics, vol. 23 (Springer, Berlin, 1994)
29. P. Huynh, D. Knezevic, Y. Chen, J.S. Hesthaven, A. Patera, A natural-norm successive constraint method for inf-sup lower bounds. Comput. Methods Appl. Mech. Eng. **199**, 1963–1975 (2010)
30. G. Rozza, K. Veroy, On the stability of the reduced basis method for Stokes equations in parametrized domains. Comput. Methods Appl. Mech. Eng. **196**, 1244–1260 (2007)
31. G. Rozza, Reduced basis methods for Stokes equations in domains with non-affine parameter dependence. Comput. Vis. Sci. **12**, 23–35 (2009)
32. A.-L. Gerner, K. Veroy, Certified reduced basis methods for parametrized saddle point problems. SIAM J. Sci. Comput. **34**, A2812–A2836 (2012)
33. A.-L. Gerner, K. Veroy, Reduced basis a posteriori error bounds for the Stokes equations in parametrized domains: a penalty approach. Math. Models Methods Appl. Sci. **21**, 2103–2134 (2011)
34. V. Arpaci, in *Conduction Heat Transfer*. Addison-Wesley Series in Mechanics and Thermodynamics (Addison-Wesley Pub. Co., 1966)

Appendix A
Mathematical Preliminaries

A.1 Banach and Hilbert Spaces

A.1.1 Basic Definitions

Let us start with a couple of definitions: Let $\mathbb{V}$ be a vector space on $\mathbb{R}$ (all the following definitions and results can also be extended to the field of complex numbers $\mathbb{C}$).

- For a set $\{w_1, \ldots, w_N\} \subset \mathbb{V}$ of elements of $\mathbb{V}$ we denote by

$$\operatorname{span}\{w_1, \ldots, w_N\} = \left\{ v \in \mathbb{V} \;\middle|\; v = \sum_{n=1}^{N} \alpha_n w_n,\ \alpha_n \in \mathbb{R} \right\}$$

  the *linear subspace spanned by the elements* $w_1, \ldots, w_N$.
- The space $\mathbb{V}$ is of *finite dimension* if there exists a maximal set of linearly independent elements $v_1, \ldots, v_N$, otherwise $\mathbb{V}$ is of *infinite dimension*.
- A *norm* $\|\cdot\|_{\mathbb{V}}$ on $\mathbb{V}$ is an application $\|\cdot\|_{\mathbb{V}} : \mathbb{V} \to \mathbb{R}$ such that

  (i) $\|v\|_{\mathbb{V}} \geq 0, \quad \forall v \in \mathbb{V} \qquad$ and $\qquad \|v\|_{\mathbb{V}} = 0$ if and only if $v = 0$.
  (ii) $\|\alpha v\|_{\mathbb{V}} = |\alpha| \, \|v\|_{\mathbb{V}}, \qquad \forall \alpha \in \mathbb{R}, v \in \mathbb{V}$.
  (iii) $\|u + v\|_{\mathbb{V}} \leq \|u\|_{\mathbb{V}} + \|v\|_{\mathbb{V}}, \qquad \forall u, v \in \mathbb{V}$.

- The pair $(\mathbb{V}, \|\cdot\|_{\mathbb{V}})$ is a *normed space* and we can define a distance function $d(u, v) = \|u - v\|_{\mathbb{V}}$ to measure the distance between two elements $u, v \in \mathbb{V}$.
- A *semi-norm* on $\mathbb{V}$ is an application $|\cdot|_{\mathbb{V}} : \mathbb{V} \to \mathbb{R}$ such that $|v|_{\mathbb{V}} \geq 0$ for all $v \in \mathbb{V}$ and (*ii*) and (*iii*) above are satisfied. In consequence, a semi-norm is a norm if and only if $|v|_{\mathbb{V}} = 0$ implies $v = 0$.
- Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if there exists two constants $C_1, C_2 > 0$ such that

$$C_1 \|\cdot\|_1 \leq \|\cdot\|_2 \leq C_2 \|\cdot\|_1, \qquad \forall v \in V. \tag{A.1}$$

J.S. Hesthaven et al., *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, SpringerBriefs in Mathematics,
DOI 10.1007/978-3-319-22470-1

A.1.2 Linear Forms

Let $(\mathbb{V}, \|\cdot\|_{\mathbb{V}})$ be a normed space. Then, we define the following notions.

- An application $F : \mathbb{V} \to \mathbb{R}$ is said to be *linear*, a *functional* or a *linear form* if

$$\begin{aligned} F(u+v) &= F(u) + F(v), \qquad && \forall u, v \in \mathbb{V}, \\ F(\alpha u) &= \alpha F(u), && \forall \alpha \in \mathbb{R}, u \in \mathbb{V}. \end{aligned}$$

- F is *bounded* if there exists a constant $C > 0$ such that

$$|F(v)| \leq C \|v\|_{\mathbb{V}}, \qquad \forall v \in \mathbb{V}.$$

- F is *continuous* if for all $\varepsilon > 0$ there exists a $\delta_\varepsilon > 0$ such that

$$\|u - v\|_{\mathbb{V}} \leq \delta_\varepsilon \qquad \Rightarrow \qquad |F(u) - F(v)| < \varepsilon.$$

As a consequence of these definitions, one can show that the notion of continuity and boundedness is equivalent for linear forms.

- The *dual* space of the normed space $(\mathbb{V}, \|\cdot\|_{\mathbb{V}})$ denoted by $(\mathbb{V}', \|\cdot\|_{\mathbb{V}'})$ is defined by

$$\mathbb{V}' = \left\{ F : \mathbb{V} \to \mathbb{R} \;\middle|\; F \text{ is linear and continuous} \right\},$$

  endowed with the norm

$$\|F\|_{\mathbb{V}'} = \sup_{v \in \mathbb{V}, v \neq 0} \frac{|F(v)|}{\|v\|_{\mathbb{V}}}, \qquad \forall F \in \mathbb{V}'.$$

A.1.3 Bilinear Forms

- A *bilinear form* $a(\cdot, \cdot)$ acting on the vector spaces $\mathbb{V}$ and $\mathbb{W}$ is given as

$$\begin{aligned} a : \mathbb{V} \times \mathbb{W} &\to \mathbb{R}, \\ (u, v) &\mapsto a(u, v), \end{aligned}$$

  and is linear with respect to each of its arguments.
- Let $\mathbb{V}$ and $\mathbb{W}$ be endowed with the norms $\|\cdot\|_{\mathbb{V}}$ and $\|\cdot\|_{\mathbb{W}}$. A bilinear form $a(\cdot, \cdot)$ is *continuous* if there exists a constant $\gamma > 0$ such that

$$|a(u, v)| \leq \gamma \|u\|_{\mathbb{V}} \|v\|_{\mathbb{W}}, \qquad \forall u, v \in \mathbb{V}.$$

- If $\mathbb{V} = \mathbb{W}$, a bilinear form a is *coercive* if there exists a constant $\alpha > 0$ such that

$$a(v, v) \geq \alpha \|v\|_{\mathbb{V}}^2, \qquad \forall v \in \mathbb{V}.$$

A.1.3 Banach and Hilbert Spaces

We first introduce the notation of a Banach space.

- Let $\{v_n\}_{n\in\mathbb{N}}$ by a sequence in a normed space $(\mathbb{V}, \|\cdot\|_{\mathbb{V}})$. This sequence is a *Cauchy sequence* if

$$\lim_{n,m\to\infty} \|v_n - v_m\|_{\mathbb{V}} = 0.$$

- A normed space $(\mathbb{V}, \|\cdot\|_{\mathbb{V}})$ is a *Banach space* if any Cauchy sequence in $\mathbb{V}$ converges to an element in $\mathbb{V}$ (where the convergence is measured with respect to the norm $\|\cdot\|_{\mathbb{V}}$).

Hilbert spaces are particular Banach spaces. Let us begin with the following definitions.

- A inner product in a vector space $\mathbb{V}$ is an application $(\cdot, \cdot)_{\mathbb{V}} : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ with the properties:

(i) $(u, v)_{\mathbb{V}} \geq 0, \quad \forall v \in \mathbb{V} \qquad$ and $\qquad (u, u)_{\mathbb{V}} = 0$ if and only if $u = 0$.
(ii) $(u, v)_{\mathbb{V}} = (v, u)_{\mathbb{V}}, \qquad \forall u, v \in \mathbb{V}$.
(iii) $(\alpha u + \beta v, w)_{\mathbb{V}} = \alpha(u, w)_{\mathbb{V}} + \beta(v, w)_{\mathbb{V}}, \qquad \forall \alpha, \beta \in \mathbb{R}, u, v, w \in \mathbb{V}$.

The *Cauchy-Schwarz inequality*

$$|(u, v)_{\mathbb{V}}| \leq \sqrt{(u, u)_{\mathbb{V}}}\sqrt{(v, v)_{\mathbb{V}}}, \qquad \forall u, v \in \mathbb{V},$$

is a simple consequence of the definition of the inner product. Let $\mathbb{V}$ be a vector space endowed with an inner product. Then, define

$$\|v\|_{\mathbb{V}} = \sqrt{(v, v)_{\mathbb{V}}}, \qquad \forall v \in \mathbb{V},$$

and as an immediate consequence of the Cauchy-Schwarz inequality, there holds

$$|(u, v)_{\mathbb{V}}| \leq \|u\|_{\mathbb{V}} \|v\|_{\mathbb{V}}, \qquad \forall u, v \in \mathbb{V}.$$

We therefore recognize that a inner product on $(\mathbb{V}, \|\cdot\|_{\mathbb{V}})$ is a continuous and coercive bilinear form.

In addition, $(\mathbb{V}, \|\cdot\|_{\mathbb{V}})$ is a normed space, since (relying on the definition of the inner product)

(i) $\|v\|_{\mathbb{V}} = \sqrt{(v,v)_{\mathbb{V}}} \geq 0, \quad \forall v \in \mathbb{V} \quad \text{and} \quad \|v\|_{\mathbb{V}} = 0 \quad \Leftrightarrow \quad (v,v)_{\mathbb{V}} = 0 \quad \Leftrightarrow \quad v = 0.$

(ii) $\|\alpha v\|_{\mathbb{V}} = \sqrt{(\alpha v, \alpha v)_{\mathbb{V}}} = \sqrt{\alpha^2 (v,v)_{\mathbb{V}}} = |\alpha| \sqrt{(v,v)_{\mathbb{V}}}, \qquad \forall \alpha \in \mathbb{R}, v \in \mathbb{V}.$

(iii) As a consequence of the Cauchy-Schwarz inequality we obtain

$$\begin{aligned} \|u+v\|_{\mathbb{V}}^2 &= (u+v, u+v)_{\mathbb{V}} = (u,u)_{\mathbb{V}} + (v,v)_{\mathbb{V}} + 2(u,v)_{\mathbb{V}} \\ &\leq (u,u)_{\mathbb{V}} + (v,v)_{\mathbb{V}} + 2\sqrt{(u,u)_{\mathbb{V}}}\sqrt{(v,v)_{\mathbb{V}}} = \left(\|u\|_{\mathbb{V}} + \|v\|_{\mathbb{V}}\right)^2, \qquad \forall u, v \in \mathbb{V}, \end{aligned}$$

so that $\|u+v\|_{\mathbb{V}} \leq \|u\|_{\mathbb{V}} + \|v\|_{\mathbb{V}}$ for all $u, v \in \mathbb{V}$.

- A *Hilbert space* is a Banach space whose norm is induced by an inner product.

Let us now recall some elementary results from functional analysis.

Theorem A.1 (Riesz representation) *Let $\mathbb{V}$ be a Hilbert space. For any $f \in \mathbb{V}'$, there exists an element $v \in \mathbb{V}$ such that*

$$(v,w)_{\mathbb{V}} = f(w), \qquad \forall w \in \mathbb{V},$$

and

$$\|f\|_{\mathbb{V}'} = \|v\|_{\mathbb{V}}.$$

Theorem A.2 (Hilbert Projection Theorem) *Let $\mathbb{V}$ be a Hilbert space and $M \subset \mathbb{V}$ a closed subspace of $\mathbb{V}$. Then, for any $v \in \mathbb{V}$, there exists a unique $P_M v \in M$ such that*

$$\|P_M v - v\|_{\mathbb{V}} = \inf_{w \in M} \|w - v\|_{\mathbb{V}}.$$

In addition, the infimum is characterized by $P_M \in M$ such that $(v - P_M v, w)_{\mathbb{V}} = 0$ for any $w \in M$ (thus $v - P_M v \in M^{\perp}$) and $P_M v = v$ for all $v \in M$. The operator $P_M : \mathbb{V} \to M$ is linear and called the orthogonal projection onto M.

A.2 Lax-Milgram and Banach-Nečas-Babuška Theorem

Many problems in science and engineering can be formulated as: find $u \in \mathbb{V}$ such that

$$a(u,v) = f(v), \qquad \forall v \in \mathbb{V}, \tag{A.2}$$

where $\mathbb{V}$ is a Hilbert space, $a : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ a bilinear form and $f \in \mathbb{V}'$ a linear form. Then, the following theorem provides conditions to ensure a solution to such a problem exists, is unique and stable with respect to the data.

Theorem A.3 (Lax-Milgram Theorem) *Let $\mathbb{V}$ be a Hilbert space, $a : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ a continuous and coercive bilinear form and $f \in \mathbb{V}'$ a continuous linear form. Then, problem* (A.2) *admits a unique solution $u \in \mathbb{V}$. Further, there holds that*

$$\|u\|_{\mathbb{V}} \leq \frac{1}{\alpha}\|f\|_{\mathbb{V}'},$$

where $\alpha > 0$ is the coercivity constant of $a(\cdot\,,\,\cdot)$.

A second class of problems can be defined as: find $u \in \mathbb{V}$ such that

$$a(u, v) = f(v), \qquad \forall v \in \mathbb{W}, \tag{A.3}$$

where $\mathbb{V}$ and $\mathbb{W}$ are Hilbert spaces, $a : \mathbb{V} \times \mathbb{W} \to \mathbb{R}$ is a bilinear form and $f \in \mathbb{W}'$ a linear form. Then, the following theorem provides conditions so that a solution to such a problem exists, is unique and stable with respect to the data. Note that we assume here that $\mathbb{V}$ and $\mathbb{W}$ are Hilbert spaces, which is not necessary as $\mathbb{V}$ can be taken as a Banach space and $\mathbb{W}$ a reflexive Banach space.

Theorem A.4 (Banach-Nečas-Babuška Theorem) *Let $\mathbb{V}$ and $\mathbb{W}$ be Hilbert spaces. Let $a : \mathbb{V} \times \mathbb{W} \to \mathbb{R}$ be bilinear and continuous and $f \in \mathbb{W}'$. Then,* (A.3) *admits a unique solution if and only if:*

(i) There exists a constant $\beta > 0$ such that for all $v \in \mathbb{V}$ there holds

$$\beta\,\|v\|_{\mathbb{V}} \leq \sup_{w \in \mathbb{W}\setminus\{0\}} \frac{a(v, w)}{\|w\|_{\mathbb{W}}}. \tag{A.4}$$

(ii) For all $w \in \mathbb{W}$,

$$\big\{a(v, w) = 0, \quad \forall v \in \mathbb{V}\big\} \qquad \textit{implies that} \qquad w = 0.$$

Condition (A.4) is referred to as *inf-sup* condition as it is equivalent to the following statement:

$$\beta \leq \inf_{v \in \mathbb{V}\setminus\{0\}} \sup_{w \in \mathbb{W}\setminus\{0\}} \frac{a(v, w)}{\|v\|_{\mathbb{V}}\|w\|_{\mathbb{W}}}.$$

A.3 Sobolev Spaces

Let Ω be an open subset of $\mathbb{R}^d$ and k a positive integer. Let $L^2(\Omega)$ denote the space of square integrable functions on Ω.

- The *Sobolev space of order k* on Ω is defined by

$$H^k(\Omega) = \big\{f \in L^2(\Omega) \;\big|\; D^\alpha f \in L^2(\Omega),\ |\alpha| \leq k\big\},$$

where D^α is the partial derivative

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_d^{\alpha_1} \cdots \partial x_d^{\alpha_d}},$$

in the sense of distributions for the multi-index $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$ using the notation $|\alpha| = \alpha_1 + \cdots + \alpha_d$.

It holds by construction that $H^{k+1}(\Omega) \subset H^k(\Omega)$ and that $H^0(\Omega) = L^2(\Omega)$. $H^k(\Omega)$ is a Hilbert space with the inner product

$$(f, g)_{H^k(\Omega)} = \sum_{\alpha \in \mathbb{N}^d, |\alpha| \leq k} \int_\Omega (D^\alpha f)(D^\alpha g),$$

and the induced norm

$$\|f\|_{H^k(\Omega)} = \sqrt{(f, f)_{H^k(\Omega)}} = \sqrt{\sum_{\alpha \in \mathbb{N}^d, |\alpha| \leq k} \int_\Omega |D^\alpha f|^2}.$$

We also can endow $H^k(\Omega)$ by a semi-norm given by

$$\|f\|_{H^k(\Omega)} = \sqrt{\sum_{\alpha \in \mathbb{N}^d, |\alpha| = k} \int_\Omega |D^\alpha f|^2}.$$

Finally, the following (simplified) theorem provides more intuitive information on the regularity of such spaces.

Theorem A.5 (Sobolev embedding Theorem) *Let Ω be an open subset of $\mathbb{R}^d$. Then,*

$$H^k(\Omega) \subset C^m(\overline{\Omega}),$$

for any $k > m + \frac{d}{2}$.

A.4 Galerkin Approximation and Cea's Lemma

One way to find an approximation to the solution of (A.2) (resp. (A.3) with some modifications) is to introduce a finite dimensional subspace $\mathbb{V}_\delta \subset \mathbb{V}$ and consider the solution of the following problem: find $u_\delta \in \mathbb{V}_\delta$ such that

$$a(u_\delta, v_\delta) = f(v_\delta), \qquad \forall v_\delta \in \mathbb{V}_\delta. \tag{A.5}$$

We denote the dimension of the discrete space $\mathbb{V}_\delta$ by $N_\delta = \dim(\mathbb{V}_\delta)$. Given any basis function $\{\varphi_1, \ldots, \varphi_{N_\delta}\}$ of $\mathbb{V}_\delta$, we can represent the approximation $u_\delta \in \mathbb{V}_\delta$ by the vector $\mathrm{u}_\delta \in \mathbb{R}^{N_\delta}$ through the relation

$$u_\delta = \sum_{i=1}^{N_\delta} (\mathrm{u}_\delta)_i \, \varphi_i .$$

Then, we can write (A.5) as: find $\mathrm{u}_\delta \in \mathbb{R}^{N_\delta}$ such that

$$\sum_{i=1}^{N_\delta} (\mathrm{u}_\delta)_i \, a(\varphi_i, v_\delta) = f(v_\delta), \qquad \forall v_\delta \in \mathbb{V}_\delta .$$

By linearity of $a(\cdot, \cdot)$ in the second variable, we observe that testing the equation for any $v_\delta \in \mathbb{V}_\delta$ is equivalent to testing for all basis functions only, i.e., find $\mathrm{u}_\delta \in \mathbb{R}^{N_\delta}$ such that

$$\sum_{i=1}^{N_\delta} (\mathrm{u}_\delta)_i \, a(\varphi_i, \varphi_j) = f(\varphi_j), \qquad \forall j = 1, \ldots, N_\delta .$$

We now recognize that this is a system of N_δ linear algebraic relations that can be written in matrix form as

$$\mathbf{A}_\delta \, \mathrm{u}_\delta = \mathrm{f}_\delta ,$$

where $(\mathbf{A}_\delta)_{ij} = a(\varphi_j, \varphi_i)$ and $(\mathrm{f}_\delta)_i = f(\varphi_i)$ for all $i, j = 1, \ldots, N_\delta$.

We can now start to think about quantifying the quality of the approximation u_δ of u. For this, we introduce a sequence of finite dimensional approximation spaces $\{\mathbb{V}_\delta\}_{\delta > 0}$ such that:

- $\mathbb{V}_\delta \subset \mathbb{V}, \quad \forall \delta > 0.$
- $\dim(\mathbb{V}_\delta) < \infty, \quad \forall \delta > 0.$
- $\lim_{\delta \to 0} \inf_{v_\delta \in \mathbb{V}_\delta} \|v - v_\delta\|_{\mathbb{V}} = 0, \quad \forall v \in \mathbb{V}.$

Note that in the spirit of reduced basis methods, the approximation space $\mathbb{V}_\delta$ is supposed to be sufficiently rich to have an acceptable error of the Galerkin approximation $\|u_\delta - u\|_{\mathbb{V}}$. The considerations of numerical analysis that are outlined within this appendix are slightly different as here one is interested in proving that $\lim_{\delta \to 0} u_\delta = u$ and quantify convergence rates with respect to the discretization parameter δ.

We endow the discrete space with the (inherited) norm $\| \cdot \|_{\mathbb{V}}$ and observe that the Lax-Milgram Theorem also applies to (A.5) as $\mathbb{V}_\delta \subset \mathbb{V}$. We have a guarantee of existence and uniqueness of the approximation u_δ as well as the stability result

$$\|u_\delta\|_{\mathbb{V}} \leq \frac{1}{\alpha} \|f\|_{\mathbb{V}'} .$$

As a consequence, the $\|u_\delta\|_{\mathbb{V}}$ is uniformly bounded with respect to δ and stable with respect to perturbations in the data f. Indeed, let $\hat{u}_\delta \in \mathbb{V}_\delta$ be the solution to the perturbed system

$$a(\hat{u}_\delta, v_\delta) = \hat{f}(v_\delta), \qquad \forall v_\delta \in \mathbb{V}_\delta.$$

Then it holds that

$$a(u_\delta - \hat{u}_\delta, v_\delta) = f(v_\delta) - \hat{f}(v_\delta), \qquad \forall v_\delta \in \mathbb{V}_\delta,$$

and the stability result implies

$$\|u_\delta - \hat{u}_\delta\|_{\mathbb{V}} \leq \frac{1}{\alpha}\|f - \hat{f}\|_{\mathbb{V}'}.$$

This means that any perturbation of the data results in a controllable error in the approximation. Finally, we recall the following result.

Lemma A.6 (Cea's Lemma) *Let $\mathbb{V}$ be a Hilbert space, $a : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ a continuous and coercive bilinear form and $f \in \mathbb{V}'$ a continuous linear form. Let further $\mathbb{V}_\delta$ be a conforming approximation space $\mathbb{V}_\delta \subset \mathbb{V}$. Then, it holds that*

$$\|u - u_\delta\|_{\mathbb{V}} \leq \frac{\gamma}{\alpha} \inf_{v_\delta \in \mathbb{V}_\delta} \|u - v_\delta\|_{\mathbb{V}},$$

where γ, α denote the continuity and coercivity constants respectively.

Corollary A.7 *As we assumed that $\lim_{\delta \to 0} \inf_{v_\delta \in \mathbb{V}_\delta} \|v - v_\delta\|_{\mathbb{V}} = 0$, for all $v \in \mathbb{V}$ we immediately conclude that*

$$\lim_{\delta \to 0} u_\delta = u.$$

Cea's Lemma is actually not difficult to prove. Recall indeed that

$$\begin{aligned} a(u, v) &= f(v), \qquad \forall v \in \mathbb{V}, \\ a(u_\delta, v_\delta) &= f(v_\delta), \qquad \forall v_\delta \in \mathbb{V}_\delta, \end{aligned}$$

so that

$$a(u - u_\delta, v_\delta) = 0, \qquad \forall v_\delta \in \mathbb{V}_\delta,$$

which is known as *Galerkin orthogonality*. Note that the conditions on the bilinear form $a(\cdot, \cdot)$, i.e., coercivity, symmetry and bilinearity, imply that it is a inner product on $\mathbb{V}$. Therefore, u_δ is the othogonal projector, see Theorem A.2, of u onto $\mathbb{V}_\delta$ using the inner product $a(\cdot, \cdot)$ and its induced norm which, however, differs from $\|\cdot\|_{\mathbb{V}}$ in the general case. Cea's lemma quantifies the relation between the best approximation for the norm $\|\cdot\|_{\mathbb{V}}$ and the one induced by $a(\cdot, \cdot)$ through the norm equivalence, see (A.1), established by the coercivity and continuity constants α and γ.

Due to the coercivity, the Galerkin orthogonality and the continuity we can develop

$$\alpha\|u-u_\delta\|_{\mathbb{V}}^2 \le a(u-u_\delta,u-u_\delta) = a(u-u_\delta,u-v_\delta) + a(u-u_\delta,\underbrace{v_\delta-u_\delta}_{\in\mathbb{V}_\delta})$$
$$= a(u-u_\delta,u-v_\delta) \le \gamma\|u-u_\delta\|_{\mathbb{V}}\|u-v_\delta\|_{\mathbb{V}},$$

which yields Cea's lemma by taking the infimum over all $v_\delta \in \mathbb{V}_\delta$.

The approximation to problem of the second class (A.3) is similar with the difference that we also need to define a basis for a discrete test space $\mathbb{W}_\delta \subset \mathbb{W}$. We then define the approximation by seeking $u_\delta \in \mathbb{V}_\delta$ such that

$$a(u_\delta, w_\delta) = f(w_\delta), \qquad \forall w_\delta \in \mathbb{W}_\delta. \tag{A.6}$$

If the dimensions of the trial and test spaces $\mathbb{V}_\delta$ and $\mathbb{W}_\delta$ are equal, then this also results in a square linear system. Solvability and stability is in this case not inherited from the continuous formulation in contrast to coercive problems. However, the Banach-Nečas-Babuška Theorem can be applied to the discrete formulation so that the two conditions:

(i) There exists a constant $\beta_\delta > 0$ such that for all $v_\delta \in \mathbb{V}_\delta$ it holds

$$\beta_\delta \|v_\delta\|_{\mathbb{V}} \le \sup_{w_\delta\in\mathbb{W}_\delta\setminus\{0\}} \frac{a(v_\delta,w_\delta)}{\|w_\delta\|_{\mathbb{W}}}.$$

(ii) For all $w_\delta \in \mathbb{W}_\delta$,

$$\big\{a(v_\delta,w_\delta)=0, \quad \forall v_\delta\in\mathbb{V}_\delta\big\} \qquad \text{implies that} \qquad w_\delta = 0;$$

imply existence of a unique solution. Again, the former condition is equivalent to

$$\beta_\delta \le \inf_{v_\delta\in\mathbb{V}_\delta\setminus\{0\}} \sup_{w_\delta\in\mathbb{W}_\delta\setminus\{0\}} \frac{a(v_\delta,w_\delta)}{\|v_\delta\|_{\mathbb{V}}\|w_\delta\|_{\mathbb{W}}}.$$

Assume that these conditions are satisfied, then for any $v_\delta \in \mathbb{V}_\delta$ it holds

$$\|u-u_\delta\|_{\mathbb{V}} \le \|u-v_\delta\|_{\mathbb{V}} + \|v_\delta-u_\delta\|_{\mathbb{V}}.$$

Combining the discrete inf-sup stability, the Galerkin orthogonality and the continuity yields

$$\beta_\delta\|v_\delta-u_\delta\|_{\mathbb{V}} \le \sup_{w_\delta\in\mathbb{W}_\delta\setminus\{0\}} \frac{a(v_\delta-u_\delta,w_\delta)}{\|w_\delta\|_{\mathbb{W}}} = \sup_{w_\delta\in\mathbb{W}_\delta\setminus\{0\}} \frac{a(v_\delta-u,w_\delta)}{\|w_\delta\|_{\mathbb{W}}}$$
$$\le \gamma \sup_{w_\delta\in\mathbb{W}_\delta\setminus\{0\}} \frac{\|v_\delta-u\|_{\mathbb{V}}\|w_\delta\|_{\mathbb{W}}}{\|w_\delta\|_{\mathbb{W}}} = \gamma\,\|v_\delta-u\|_{\mathbb{V}}$$

so that

$$\|u - u_\delta\|_{\mathbb{V}} \leq \left(1 + \frac{\gamma}{\beta_\delta}\right) \inf_{v_\delta \in \mathbb{V}_\delta} \|u - v_\delta\|_{\mathbb{V}},$$

which is the analogue of Cea's lemma for non-coercive but inf-sup stable approximations. Note that optimally convergent approximations can thus be obtained if $\beta_\delta \geq \hat{\beta} > 0$ as $\delta \to 0$ for some $\hat{\beta}$ that is independent of the discretization parameter δ.

Index

J.S. Hesthaven et al., *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, SpringerBriefs in Mathematics,
DOI 10.1007/978-3-319-22470-1