# Reduced Models for Optimal Control, Shape Optimization and Inverse Problems in Haemodynamics

THÈSE N$^{O}$ 5402 (2012)

PRÉSENTÉE LE 5 JUILLET 2012

À LA FACULTÉ DES SCIENCES DE BASE
CHAIRE DE MODÉLISATION ET CALCUL SCIENTIFIQUE
PROGRAMME DOCTORAL EN MATHÉMATIQUES

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Andrea MANZONI

acceptée sur proposition du jury:

Prof. K. Hess Bellwald, présidente du jury
Prof. A. Quarteroni, Dr G. Rozza, directeurs de thèse
Prof. D. Kressner, rapporteur
Prof. Y. Maday, rapporteur
Prof. A. T. Patera, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2012

*Per chi viaggia in direzione*
*ostinata e contraria*

But there was something else he had had on the tip of his tongue,
something about mathematical problems that did not admit
of any general solution, though they did admit of particular solutions,
the combination of which could bring one closer to the general solution.
He might have added that he regarded the problem set by every human
life as one of these. What someone calls an age – without knowing whether
he should by that understand centuries, millennia, or the span of time
between schooldays and grandparenthood – this broad, unregulated
flux of conditions would then amount to much the same thing as a
chaotic succession of unsatisfactory and, when taken singly, false
attempts at a solution, attempts that might produce the correct
and total solution, but only after men had learned to combine them.
In the tram going home he remembered this...

*Robert Musil, Der Mann ohne Eigenschaften (1942), I. 358.*
In *Italo Calvino, Lezioni Americane (1985), Molteplicità.*

# Acknowledgements

First of all, I would like to thank Prof. Alfio Quarteroni and Dr. Gianluigi Rozza for being more than thesis directors during these years at EPFL. Alfio gave me wonderful opportunities to grow in my education, to learn new branches of numerical analysis and applied mathematics, to become a teaching assistant and to travel all around the world within these years. Spending three years with him next door is a unique experience: his point of view – as a teacher, a father, a man – is always stimulating, opening new and interesting perspectives.
Gianluigi was a precious trainer and a constant guide during this work and these years, thanks to his insights, passion and care. His human qualities and kindness have allowed a very fruitful collaboration, a constant exchange of ideas and a great friendship.
They are the best examples of people dealing with a *many-query* job, in a *real-time* fashion, with incredible passion – any reduction effort sounds completely useless in this special case!

I am very grateful to Prof. Anthony T. Patera (MIT) who has been the engine of the reduced basis community and is a constant reference. I thank the other members of the committee: Prof. Kathryn Hess Bellwald for presiding the jury, Prof. Daniel Kressner (EPFL) and Prof. Yvon Maday, the latter also for having organized several opportunities of exchange and discussion at Laboratoire Jacques-Louis Lions of Université Pierre et Marie Curie, Paris VI.

Several people contributed to the achievement of this work. A special thank goes to Dr. Toni Lassila, a precious colleague and coauthor, and a great friend. The possibility to share mathematical problems and the office with him, as well as his talent for *synthesis* and *irony*, have been wonderful gifts. This work has also benefited from the collaboration with Federico Negri and Francesco Ballarin, who carried out their master thesis under my co-supervision.

A special acknowledgement goes to all my colleagues of the Chair of Modelling and Scientific Computing (CMCS) of EPFL. In particular, many thanks to senior researchers Dr. Simone Deparis for his availability and kindness, Dr. Marco Discacciati for his passion in teaching and Dr. Luca Dedè for many stimulating discussions, about mathematical problems and our ill-conditioned country. Thanks to Matteo, my third office mate – the most sporting guy I have ever met – to Samuel, my personal French teacher, and to Jean, for his interesting feedbacks and medical insights. Thanks also to all other people with whom I interacted during my period at EPFL, especially all the members and colleagues of MATHICSE.

A particular recognition goes to MOX (Modelling and Scientific Computing Center) and to the Mathematics Department of Politecnico di Milano, where I spent five exciting years during my Bachelor and Master Degrees. Many thanks to all the people who have contributed to establish the Mathematical Engineering Degree and its popularity among students. In particular, thanks to Prof. Anna Maria Paganoni for her friendship, her incredible humanity and her immense passion. Thanks to Prof. Sandro Salsa, for his strong encouragement and his unique charism.

I acknowledge the Swiss National Science Foundation supporting my scientific activity during these years at EPFL (Project 122136, “Reduced basis methods for the optimization of complex systems” and Project 135444, “Numerical simulation of a sailing boat: dynamics and shape optimization”) and the European Research Council, which granted the research program “Mathcard, Mathematical Modelling and Simulation of the Cardiovascular System” (Project ERC-2008-AdG 227058). Several applications, as well as many collaborations within these years, stem from such a stimulating and fascinating framework.

I acknowledge the use of the `rbMIT` package for the reduced basis simulations and developed by Prof. Anthony T. Patera, Dr. Phuong Huynh, Dr. Cuong Nguyen (MIT) and Dr. Gianluigi Rozza. Phuong has always provided *real-time* support from MIT for my development activity within the library. I also acknowledge the use of `MLife` library for the finite element simulations, and developed by Prof. Fausto Saleri and colleagues at MOX - Politecnico di Milano. Many thanks also to Prof. Paola Gervasio, for her kindness and her effort in maintaining Fausto’s scientific heritage.

Thanks to all my friends – the list would be too long and incomplete – to be as they are. In particular, thanks to Chiara, Andrea, Gabriele, Valeria, Lorenzo, Marco, Niccolò, Pietro, Simone, Laura, Matteo, Alessandro, Elisa, Noemi, Francesco.
Last but not least, a very special thanks to my parents for their constant support and patience and to Francesca for enriching my life with her unconditional love.

*Lausanne, June 2012*

# Abstract

The objective of this thesis is to develop reduced models for the numerical solution of optimal control, shape optimization and inverse problems. In all these cases suitable functionals of state variables have to be minimized. State variables are solutions of a partial differential equation (PDE), representing a constraint for the minimization problem. The solution of these problems induce large computational costs due to the numerical discretization of PDEs and to iterative procedures usually required by numerical optimization (*many-query* context).
In order to reduce the computational complexity, we take advantage of the reduced basis (RB) approximation for parametrized PDEs, once the state problem has been reformulated in parametrized form. This method enables a rapid and reliable approximation of parametrized PDEs by constructing low-dimensional, problem-specific approximation spaces. In case of PDEs defined over domains of variable shapes (e.g. in shape optimization problems) we need to introduce suitable, low-dimensional shape parametrization techniques in order to tackle the geometrical complexity. Free-Form Deformations and Radial-Basis Functions techniques have been analyzed and successfully applied with this aim.
We analyze the reduced framework built by coupling these tools and apply it to the solution of optimal control and shape optimization problems. Robust optimization problems under uncertain conditions are also taken into consideration. Moreover, both deterministic and Bayesian frameworks are set in order to tackle inverse identification problems.
As state equations, we consider steady viscous flow problems described by Stokes or Navier-Stokes equations, for which we provide a detailed analysis and construction of RB approximation and *a posteriori* error estimation. Several numerical test cases are also illustrated to show efficacy and reliability of RB approximations.
We exploit this general reduced framework to solve some optimization and inverse problems arising in haemodynamics. More specifically, we focus on the optimal design of cardiovascular prostheses, such as bypass grafts, and on inverse identification of pathological conditions or flow/shape features in realistic parametrized geometries, such as carotid artery bifurcations.

*Keywords*: Reduced Order Modelling, Reduced Basis Methods, Shape Parametrization Techniques, Optimal Control, Shape Optimization, Inverse Problems, Fluid Dynamics.

# Version abrégée

Le but de cette thèse est le développement de modèles réduits pour l'approximation numérique de problèmes de contrôle, d'optimisation de forme et d'autres problèmes inverses. Dans tous ces situations, il faut minimiser une fonctionnelle coût, donnée par une fonction des variables d'état. Celles-ci sont obtenues par la solution d'une équation aux dérivées partielles (EDP), qui représente une contrainte pour le problème de minimisation. Une caractéristique commune aux problèmes susmentionnés est l'énorme coût computationnel, dû à la discrétisation numérique des EDP et à la résolution du problème d'optimisation par des méthodes itératives.
Pour réduire le coût computationnel, nous profitons de la méthode des bases réduites (*Reduced Basis*, RB) pour EDP paramétrisées, une fois que le problème a été reformulé sous forme paramétrique. Cette méthode permet de reconstruire des solutions approchées d'équations paramétrisées en fonction d'autres solutions déjà calculées et disponibles, ce qui réduit énormement les ressources de calcul nécessaires. Dans le cas où les EDP sont définies sur des domaines dont la forme est variable (par example, dans le cas des problèmes d'optimisation de forme) nous introduisons une paramétrisation des formes convenable, pour aborder la complexité géométrique. En particulier, on a considéré deux techniques possibles, les déformations de forme libre (plus connues sous le nom de *Free-Form Deformations*) et des fonctions à base radiale (aussi dites *Radial Basis Functions*). Ces techniques ont été analysées puis appliquées aux cas d'interêt pratique devéloppés dans cette thèse.
Le cadre de réduction donné par le couplage de ces outils est analysé et utilisé pour résoudre des problèmes de contrôle optimal et d'optimisation de forme. Des problèmes d'optimisation robuste sous conditions d'incertitude ont également été examinés. De plus, nous considérons un cadre déterministique et un cadre bayésien pour la solution de problèmes d'identification inverse.
Au niveau du problème d'état, on considère les équations de Stokes et de Navier-Stokes stationnaires. On fournit une analyse détaillée de l'approximation à bases réduites pour ces problèmes, ainsi que des techniques pour l'estimation à posteriori des erreurs. Nous présentons de nombreux tests numériques pour montrer l'efficacité de ces techniques.
Au niveau de l'application, on considère des problèmes d'optimisation et des problèmes inverses dans le cadre de l'hémodynamique. Deux exemples sont donnés par le design optimal des prothèses cardiovasculaires (ou pontages) et l'identification de conditions pathologiques qui dépendent de l'interaction entre l'écoulement sanguin et la géometrie des artères.

# Estratto

L'obiettivo di questa tesi è lo sviluppo di modelli ridotti per la soluzione numerica di problemi di controllo ottimo, ottimizzazione di forma e altri problemi inversi. In tutti questi casi occorre minimizzare un funzionale costo dipendente dalle variabili di stato del sistema, attraverso il controllo di una o più variabili che influenzano la soluzione del problema di stato, costituito da un sistema di Equazioni a Derivate Parziali (EDP). Questi problemi sono caratterizzati da una notevole complessità computazionale, a causa della discretizzazione numerica delle EDP e delle procedure iterative richieste per l'ottimizzazione numerica.
Per ridurre la complessità computazionale, riformuliamo il problema di stato in forma parametrizzata e consideriamo una approssimazione a basi ridotte (*Reduced Basis*, RB) per EDP parametrizzate. Questo metodo permette di ricostruire la soluzione approssimata di un problema parametrizzato combinando opportunamente un insieme di soluzioni – o *basi* – precedentemente calcolate (tramite una tecnica di discretizzazione tradizionale, come il metodo degli elementi finiti). In questo modo, la dimensione dello spazio di approssimazione risulta notevolmente ridotta.
Nel caso in cui l'equazione è definita su un dominio di forma variabile (come nel caso dell'ottimizzazione di forma) occorre introdurre una parametrizzazione del dominio, in modo da ridurre la complessità geometrica e riportare il calcolo delle basi su un dominio di riferimento, allo scopo di poterle combinare. In particolare, sono state considerate due differenti tecniche, basate su deformazioni a forma libera (*Free-Form Deformations*) o sull'impiego di funzioni a basi radiali (*Radial Basis Functions*).
Questo metodo di *riduzione*, basata sull'accoppiamento di approssimazioni a basi ridotte e di opportune tecniche per la parametrizzazione di forma, è stata ampiamente analizzata e applicata a problemi di controllo ottimo e ottimizzazione di forma. Successivamente, sono stati presi in considerazione anche problemi di ottimizzazione robusta in condizioni di incertezza e più generali problemi inversi, affrontati sia mediante tecniche deterministiche per l'ottimizzazione sia grazie a metodi della statistica bayesiana.
L'analisi presentata in questa tesi si concentra su problemi legati a flussi viscosi stazionari, descritti da equazioni di Stokes o di Navier-Stokes. Per questa classe di problemi sono state sviluppate (o ampliate, laddove già introdotte) sia la tecnica di approssimazione a basi ridotte che la metodologia per la stima a posteriori dell'errore. Numerosi test numerici sono stati forniti anche allo scopo di mostrare l'efficacia e l'attendibilità dell'approssimazione a basi ridotte in questo contesto.
Da ultimo, questo metodo di riduzione è stato sfruttato per risolvere numerosi problemi di ottimizzazione di forma e problemi inversi di interesse emodinamico, come il design ottimale di bypass aorto-coronarici e femorali o l'identificazione di potenziali condizioni di rischio patologico legate all'interazione tra il flusso di sangue e la forma di particolari segmenti di arteria, come nel caso di stenosi della biforcazione carotidea.

# Contents

# Introduction

The efficient solution of optimization problems or more general inverse problems involving Partial Differential Equations (PDEs) has become an important part of the computational science and engineering disciplines. These problems represent a challenging framework because they rely on both rigorous theoretical tools of functional analysis and efficient computational techniques, required to face the complexity arising from numerical approximation. In case of optimal control or shape optimization problems governed by PDEs, the goal is the minimization/maximization of some objective function depending on a system described by a PDE model, the control being represented by either suitable variables (such as sources, model coefficients or boundary values) or the shape of the domain itself. Instead, when dealing with inverse identification problems, the purpose is to recover some features of a system from output measurements. In this case, quantifying the effect of uncertainty driven by data on the identified features is a further issue to be considered. In this framework, computational costs arise from three distinct sources: *(i)* numerical approximation of the state system (usually a nonlinear system of PDEs); *(ii)* iterative optimization procedures, requiring recursive evaluations of the state solution and the objective function of interest; *(iii)* handling domains of arbitrary shapes in optimization/inverse problems related with shape-dependent quantities. In order to tackle the complexity implied by the interplay of these aspects, suitable reduced order models must be considered.

The goal of this thesis is to develop a reduction framework relying on the Reduced Basis (RB) method for the efficient solution of optimization and inverse problems. This method enables a rapid and reliable approximation of parametrized PDEs by constructing low-dimensional, problem-specific approximation spaces. Using these spaces for the discretization of the original problem, instead of high-dimensional, general approximation spaces like in the finite element method, we can construct a (reduced-order) model – depending only on a handful of degrees of freedom, with respect to the large amount required by traditional (full-order) approximations – which is a sufficiently accurate approximation to the original problem. Exploiting such a model within the optimization framework thus features a strong reduction of *computational complexity*. Further efforts are required when dealing with shape-dependent problems, in order to face the additional *geometrical complexity* arising from the treatment of shape deformation and representation.

The application of reduced strategies to speed up the solution of many-query problems such as optimal control and shape optimization problems is not completely new – and the RB method is not the unique available choice. As stated by Gunzburger ten years ago in his monograph *Perspective in Flow Control and Optimization* [128], *"many cost reduction schemes have been proposed, tested, and sometimes actually employed to solve specific yet useful flow control and optimization problems; however, very few of them are rooted in firm mathematical ground or have otherwise been shown to be of general use"*. Nonetheless, during the last decade, several efforts in theoretical foundations, numerical investigations and methodological improvements have made headway to develop general ideas in reduced order modelling.

This branch of scientific computing provides nowadays a solid mathematical background of general utility, which can be exploited to tackle several problems governed by PDEs, recast both in *real-time* and *many-query* contexts. In particular, RB methods provide not only an efficient approximation technique for parametrized PDEs (made possible by a suitable affine decomposition of PDE operators and a related computational splitting), but also a certified framework, thanks to *a posteriori* error estimates between reduced-order and full-order solutions/ quantities of interest. In this way, the parametrized framework provides a great opportunity, provided that suitable parametrization techniques are available, in order to describe our *input* of interest, such as geometrical configurations or control functions.

The most original contribution of this work is the development of a modular, reduced framework to solve optimal control, shape optimization and other inverse problems dealing with fluid flows in complex geometrical configurations. This framework is based on the coupling of two reduction ingredients, which have been set, analyzed and implemented in this work:

1. **Reduced Basis Methods** for Stokes and Navier-Stokes equations to account for the *computational (complexity) reduction*. We have extended the existing framework to global, nonaffine geometrical parametrizations (relying on the empirical interpolation method for recovering an affine structure). *A posteriori* error estimates for these problems have been developed, based on the Babuška stability theory for Stokes case and on the Brezzi-Rappaz-Raviart stability theory for Navier-Stokes case, in order to obtain joint velocity/pressure error estimates. Error bounds for general linear and quadratic outputs in Stokes case have been derived as well.
2. **Free-Form Deformation** (FFD) and **Radial Basis Function** (RBF) techniques to allow for *geometrical (complexity) reduction*. Both these strategies are low-dimensional techniques relying on a set of control points. They have been applied to geometrical configurations such as bifurcations or junctions, some of the most complex geometries managed by reduced basis methods. FFD parametrizations, which were already exploited within the RB context, have been extended to viscous flows, whereas RBF parametrizations have been employed within the RB context for the first time. New theoretical results related to these techniques and some automatic procedures for selecting control points have been proposed.

The parametrized reduced framework resulting from these techniques has been widely analyzed. Moreover, both a deterministic and a Bayesian framework to deal with inverse problems have been implemented. In the former, inverse identification problems have been reformulated as suitable optimization problems, whereas the latter has enabled to solve some statistical inverse problems and to address some issues related with uncertainty quantification.
At the end, the parametrized reduced framework has been exploited within several applications arising in fluid dynamics – and more specifically to face some (simplified versions of) problems dealing with blood flows, such as the optimal design of cardiovascular prostheses or the inverse identification of pathological conditions through measurements of risk indices related with flows. We have decided to test our framework – which is general, and can be used in a broad variety of applications – to some interesting applications arising in haemodynamics, in order to address a proof of its efficacy in a relevant applicative context. Moreover, we point out that:

- the results addressed by this work are the first examples of RB approximations of viscous flows dealing with complex geometries, parametrized through global nonaffine mappings;
- the developed framework provides a useful tool not only to address the *many-query* problems introduced above, but it is also relevant to sensitivity analysis (e.g. with respect to geometrical features) and to rapid (almost real-time) numerical simulations over parametrized settings (e.g. within a wide family of similar geometrical configurations).

The dissertation is organized in three Parts, each consisting of two Chapters, as follows.

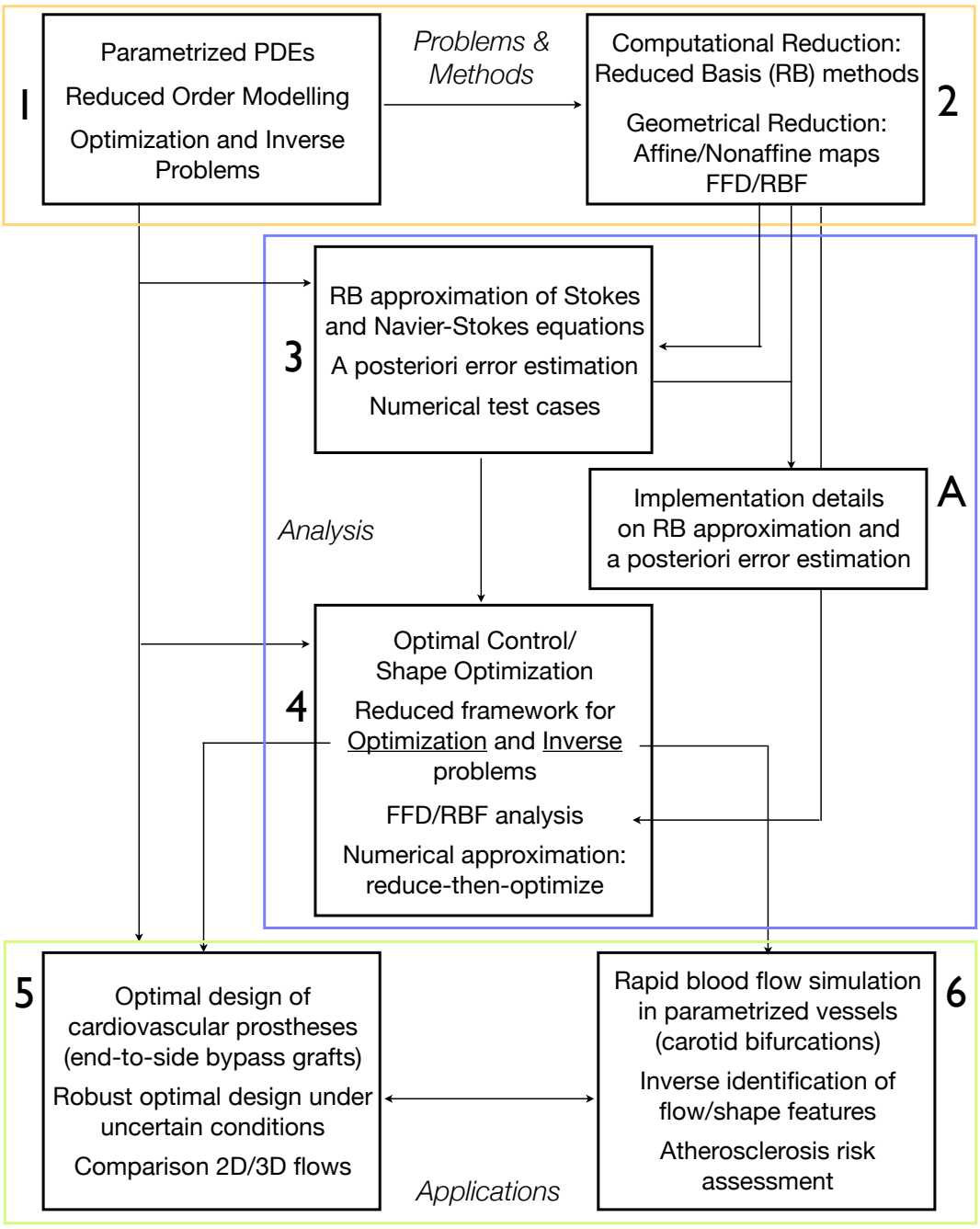
1. **Part I** is devoted to a general overview of **problems and methods** discussed and developed throughout this work.
   - In **Chapter 1** we set an abstract framework for parametrized PDEs and introduce optimal control, shape optimization and more general inverse problems governed by PDEs, providing some examples related to fluid dynamics. We highlight the main complexity points and sketch the guidelines of our reduced strategy. A more detailed and critical explanation of the contributions and scopes of this thesis is also provided.
   - In **Chapter 2** we outline the main features of computational and geometrical reduced order modelling. We detail the general construction of reduced basis methods for parametrized PDEs and *a posteriori* error estimations. Then, we recall basic affine and nonaffine shape parametrizations, focusing on two nonaffine paradigms, the Free-Form Deformations and Radial Basis Functions techniques. We show their construction and highlight some relevant properties.

2. **Part II** is devoted to the **analysis** of the reduced framework built over the Reduced Basis method for parametrized PDEs.
   - In **Chapter 3** we present the reduced basis approximation and *a posteriori* error estimation for steady Stokes and Navier-Stokes equations, as well as some numerical tests. We derive the abstract parametrized formulation and state the main assumptions ensuring its well-posedness. Then, we discuss the RB approximation, focusing on the concept of algebraic and approximation stability and on a suitable Offline/Online decomposition. We derive error bounds for velocity and pressure, both in Stokes and Navier-Stokes case. Moreover, we derive error bounds for both linear and quadratic outputs in Stokes case.
   - In **Chapter 4** we describe the reduced framework for optimization and inverse problems governed by parametrized PDEs, built over the RB approximation of the state problem. We also provide some results concerning optimal control and shape optimization problems, as well as FFD and RBF parametrization techniques. Moreover, we characterize a deterministic and a Bayesian framework for inverse problems, recasting them in a many-query context.

3. **Part III** is devoted to some **applications** of interest in haemodynamics.
   - In **Chapter 5** we provide several results concerning optimal design problems related with cardiovascular prostheses, such as bypass anastomoses. We introduce the main features related to this problem and show how the reduced framework discussed in Part II can be applied. We present some results concerning *(i)* a simplified version of an optimal flow control problem, *(ii)* two shape optimization problems dealing with Stokes and Navier-Stokes flows and *(iii)* a further robust shape optimization problem which can be of interest in this framework.
   - In **Chapter 6** we detail some inverse identification problems related to blood flows in parametrized geometrical configurations. We discuss some possible strategies to get low-dimensional representations of realistic vessel geometries such as carotid bifurcations, and some shape sensitivity problems. Then we consider three inverse identification problems related with pathological risk assessment and recovery of flow and shape features, solved within both a deterministic and a statistical inverse setting.

4. Some concluding remarks and perspectives on future developments of this work are briefly discussed in the final **Conclusion**. Further technical insights about the implementation of reduced basis methods and *a posteriori* error estimates are addressed in the **Appendix A**.

To help the reader following the logical flow of the work, we present in this page the general organization of the thesis, highlighting both methodological and conceptual links.



# Problems and Methods Part I

# 1 Optimization problems based on parametrized PDEs

In this chapter a general framework for forward and inverse problems based on parametrized partial differential equations is introduced. The abstract setting and the notation, as well as the main classes of problems faced throughout the thesis, are outlined. Optimal control, shape optimization and, more generally, inverse problems are described and recast in the context of *many-query* problems, featuring the key complexity issues arising in their numerical solution.

## 1.1 Forward and inverse problems in real-time and many query contexts

Scientific computing and numerical simulations in engineering have gained an ever increasing importance during the last decades. In several fields, from aerospace and mechanical engineering to life sciences, computer experiments (or experiments *in silico*) provide nowadays a virtual platform ancillary to material/mechanics testing or *in vitro* experiments, useful either for *(i)* inference about the underlying system (such as the prediction of input/output responses of interest) or *(ii)* its design, control and optimization [294]. A crucial factor contributing to a wide spread of computer simulations is the constant increase of available computational power, which has gone with the progressive improvement of algorithms for solving large linear systems. Indeed, numerical simulations of turbulent flows, multiscale and multiphysics phenomena, are nowadays possible by means of discretization techniques such as finite elements, finite volumes or spectral methods, but are very demanding, involving up to $\mathcal{O}(10^9)$ degrees of freedom and several hours (or even days) of CPU time also on powerful hardware parallel architectures.

From a mathematical point of view, problems arising from applied sciences are often modelled by partial differential equations (PDEs), which depend on a set of *input data*, such as material coefficients, boundary and initial conditions, source terms, as well as on a *geometrical configuration*, which can be considered as *input* itself.
Given a PDE model of a system, solving a *forward problem* consists in computing the solution of the PDE and some output of interest for specified combinations of the input data. On the other hand, whenever some parameters are unknown or uncertain, their values (and/or distributions) may be inferred from indirect observations or measures by solving an inverse problem: given an *observed* or *measured* output, the values of the input resulting in that observation can be found by driving the solution of the PDE and the corresponding *computed* output as close as possible to the observed output value, thus minimizing a distance functional in a proper sense. Very often, we may also be interested in minimizing or maximizing some physical indices expressing some desired properties or performances of the underlying system – and depending on the PDE

solution – by acting on some *control variables* (such as sources, boundary conditions, etc.) or on the shape of the domain itself. In the former case, we deal with *optimal control* problems, while we refer to the latter as *shape optimization* or *optimal design* problems.

Since the most common numerical strategies used to face previous examples are based on iterative optimization procedures, optimization and inverse problems under PDE constraints can be recast in the so-called *many-query context*, involving several input/output evaluations as well as many repeated PDE solutions. Other cases fitting the many-query context are for instance sensitivity analysis of PDE solutions with respect to input data, parametric studies and statistical analyses exploited in design of computer experiments. Thus, despite the computer resources nowadays available, it is still difficult – and often impossible – to deal with applications and scenarios involving the repeated solution of PDEs on different data settings (many-query context) or requiring a numerical solution within a *real time* context – or at least very rapidly. Both these contexts are crucial to computational engineering and their interest is growing also in view of more widespread application of numerical methods for PDEs in engineering practice and more specific industrial processes. They also feature a remarkable challenge to *classical* numerical techniques, such as – but not limited to – the finite element (FE) method; in fact, classical FE approximations may require big computational efforts (and also data/memory management) when the dimension of the discretization space becomes large. This makes both *real-time* and *many-query* simulations unaffordable: hence, looking also for computational efficiency in numerical methods becomes mandatory. As detailed in the following, suitable *computational reduction* techniques enable to solve these problems entailing an acceptable amount of CPU time and limited storage capacity.

## 1.2 Parametrized PDEs: abstract setting and notation

In this section we introduce an abstract setting for the formulation of parametrized partial differential equations, as well as the relevant notation. First of all, we introduce an *input-parameter* vector $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)^T$ and an input parameter domain – i.e. the set of all possible inputs – as a compact subset $\mathcal{D}$ of $\mathbb{R}^p$; the *input-parameter* vector typically characterizes physical properties and material, geometrical configuration, or even boundary conditions and force fields or sources.
In particular, we consider an output of interest $s(\boldsymbol{\mu}) \in \mathbb{R}$ expressed as a functional of a field variable $y(\boldsymbol{\mu})$ that is the solution of a PDE, *parametrized* with respect to the input parameter $p$-vector $\boldsymbol{\mu}$. The *outputs of interest* are physical quantities or indexes used to measure and assess the behavior of a system, i.e. related to fields variables or fluxes, as for example domain or boundary averages of the field variables, or other quantities such as energies, drag forces, flow rates, etc.

In abstract form, a parametrized PDE problem can be formulated as follows: given $\boldsymbol{\mu} \in \mathcal{D} \subset \mathbb{R}^p$, evaluate the output of interest

$$s(\boldsymbol{\mu}) = S(y(\boldsymbol{\mu}); \boldsymbol{\mu}), \tag{1.1}$$

where $y(\boldsymbol{\mu}) \in X(\Omega)$ satisfies

$$A(y(\boldsymbol{\mu}), w; \boldsymbol{\mu}) + \delta C(y(\boldsymbol{\mu}), y(\boldsymbol{\mu}), w; \boldsymbol{\mu}) = F(w; \boldsymbol{\mu}), \quad \forall\, w \in X(\Omega). \tag{1.2}$$

We denote $\Omega$ a suitably regular bounded spatial domain in $\mathbb{R}^d$ (for $d = 2$ or 3) and $X = X(\Omega)$ a suitable Hilbert space, defined on $\Omega$. We assume that $\Omega$ is a fixed, $\boldsymbol{\mu}$-independent, reference domain: when dealing with variable domains, we will distinguish between the reference domain $\Omega$ and the *original* domain, which will be denoted $\Omega_o$.
Correspondingly, $\Gamma = \partial\Omega$ (resp. $\Gamma_o = \partial\Omega_o$) will denote the boundary of the domain $\Omega$ (resp. $\Omega_o$),

$\mathbf{x} \in \mathbb{R}^d$ (resp. $\mathbf{x}_o \in \mathbb{R}^d$) a point in $\Omega$ (resp. in $\Omega_o$) and $\mathbf{n} \in \mathbb{R}^d$ (resp. $\mathbf{n}_o \in \mathbb{R}^d$) the normal unit vector on the boundary $\Gamma$ (resp. $\Gamma_o$). Furthermore, when input parameters affect the shape of the domain, i.e. when we deal with geometrical parametrizations, the original domain will be denoted $\Omega_o = \Omega_o(\boldsymbol{\mu})$, in order to underline the parametric dependence. For reasons that will be extensively discussed throughout the thesis, the parametrized PDEs setting needs to rely on a reference, parameter-independent domain; we will come back on this point several times.

We consider either linear ($\delta = 0$) or nonlinear ($\delta = 1$) PDEs, whence $A(\cdot, \cdot; \boldsymbol{\mu})$ and $C(\cdot, \cdot, \cdot; \boldsymbol{\mu})$ are the bilinear and the trilinear forms associated to the differential operator, respectively, while $F(\cdot; \boldsymbol{\mu})$ is a linear form associated to data. We shall focus on second-order PDEs, and hence $(H_0^1(\Omega))^\kappa \subset X(\Omega) \subset (H^1(\Omega))^\kappa$, where $\kappa = 1$ (respectively, $\kappa = d$) for a scalar (respectively, vector) field. Here $L^2(\Omega)$ is the space of square integrable functions over $\Omega$, while $H^1(\Omega)$ and $H_0^1(\Omega)$ denote the Sobolev spaces defined as $H^1(\Omega) = \{v : v \in L^2(\Omega), \nabla v \in (L^2(\Omega))^d\}$, $H_0^1(\Omega) = \{v \in H^1(\Omega) : v|_{\partial\Omega} = 0\}$, respectively. We denote by $(\cdot, \cdot)_X$ the inner product associated with the Hilbert space $X$, whose induced norm $\|\cdot\|_X = (\cdot, \cdot)_X^{1/2}$ is equivalent to the usual $(H^1(\Omega))^\kappa$ norm. Similarly, $(\cdot, \cdot)$ and $\|\cdot\|$ denote the $L^2(\Omega)$ inner product and induced norm, respectively. Moreover, given a Banach space $X$, $X'$ will denote its dual, i.e. the space of linear functionals on $X$, ${}_{X'}\langle \cdot, \cdot \rangle_X$ will denote the duality pairing of $X$ and $X'$, while $\mathcal{L}(X, X)$ will denote the space of linear operators from $X$ to itself.
We consider throughout the thesis both linear outputs of a field variable, i.e $s(\boldsymbol{\mu}) = L(y(\boldsymbol{\mu}); \boldsymbol{\mu})$ for a suitable linear form $L(\cdot; \boldsymbol{\mu})$, and quadratic outputs of a field variable, i.e $s(\boldsymbol{\mu}) = Q(y(\boldsymbol{\mu}), y(\boldsymbol{\mu}); \boldsymbol{\mu})$ for a suitable bilinear form $Q(\cdot, \cdot; \boldsymbol{\mu})$. Finally, the *field variables* $y(\boldsymbol{\mu})$ connecting the input parameters to the output depend on the selected PDE models and may represent velocity or pressure, temperature or concentration, displacements, potential functions, distribution functions.

We thus arrive at an *input-output* relationship $\boldsymbol{\mu} \to s(\boldsymbol{\mu})$, whose evaluation requires the solution of a parametrized PDE. In particular, we shall assume that the *bilinear* form $A(\cdot, \cdot; \boldsymbol{\mu}) : X \times X \to \mathbb{R}$ is *continuous* and *coercive* over $X$ for all $\boldsymbol{\mu}$ in $\mathcal{D}$, i.e.

$$\gamma_A(\boldsymbol{\mu}) := \sup_{w \in X} \sup_{v \in X} \frac{A(w, v; \boldsymbol{\mu})}{\|w\|_X \|v\|_X} < +\infty, \quad \forall\, \boldsymbol{\mu} \in \mathcal{D}, \tag{1.3}$$

$$\exists\ \bar{\alpha} > 0\ :\ \alpha(\boldsymbol{\mu}) := \inf_{w \in X} \frac{A(w, w; \boldsymbol{\mu})}{\|w\|_X^2} \geq \bar{\alpha}, \quad \forall\, \boldsymbol{\mu} \in \mathcal{D}. \tag{1.4}$$

If the coercivity condition is not satisfied, it will be replaced by a more general *inf-sup* condition; see e.g. Sect. 3.2.1. In the same way, we assume that the *trilinear* form $C(\cdot, \cdot, \cdot; \boldsymbol{\mu}) : X \times X \times X \to \mathbb{R}$ is *continuous*, i.e.

$$\gamma_C(\boldsymbol{\mu}) := \sup_{w \in X} \sup_{v \in X} \sup_{z \in X} \frac{C(w, v, z; \boldsymbol{\mu})}{\|w\|_X \|v\|_X \|z\|_X} < +\infty, \quad \forall\, \boldsymbol{\mu} \in \mathcal{D}, \tag{1.5}$$

that the linear forms $F(\cdot; \boldsymbol{\mu})$ and $L(\cdot; \boldsymbol{\mu})$ are continuous over $X$, i.e.

$$\gamma_F(\boldsymbol{\mu}) := \sup_{w \in X} \frac{|F(w; \boldsymbol{\mu})|}{\|w\|_X} < +\infty, \quad \gamma_L(\boldsymbol{\mu}) := \sup_{w \in X} \frac{|L(w; \boldsymbol{\mu})|}{\|w\|_X} < +\infty, \quad \forall\, \boldsymbol{\mu} \in \mathcal{D} \tag{1.6}$$

and that the bilinear form $Q(\cdot, \cdot; \boldsymbol{\mu})$ is continuous and coercive over $X$ for all $\boldsymbol{\mu}$ in $\mathcal{D}$. Under previous assumptions, in the linear case ($\delta = 0$) the parametrized PDE (1.2) admits a unique solution for each parameter value $\boldsymbol{\mu} \in \mathcal{D}$; in the nonlinear case, additional assumptions on operators and data are needed for ensuring the well-posedness of problem (1.2). This analysis is postponed to the cases of interest addressed in Part II.

We shall make one last assumption, crucial for the efficacy of computational reduction, by assuming that the parametric (bilinear/linear/trilinear) forms are *affine* in the parameter $\boldsymbol{\mu}$: for some finite $Q_A$, $Q_F$, $Q_C$, $A(\cdot,\cdot;\boldsymbol{\mu})$, $F(\cdot;\boldsymbol{\mu})$ and $C(\cdot,\cdot,\cdot;\boldsymbol{\mu})$ can be expressed as

$$A(w,v;\boldsymbol{\mu}) = \sum_{q=1}^{Q_A} \Theta_A^q(\boldsymbol{\mu})\, A^q(w,v), \qquad F(w;\boldsymbol{\mu}) = \sum_{q=1}^{Q_F} \Theta_F^q(\boldsymbol{\mu})\, F^q(w), \tag{1.7}$$

$$C(w,v,z;\boldsymbol{\mu}) = \sum_{q=1}^{Q_C} \Theta_C^q(\boldsymbol{\mu})\, C^q(w,v,z), \tag{1.8}$$

for given (typically very smooth) $\boldsymbol{\mu}$-dependent functions $\Theta_A^q : \mathcal{D} \to \mathbb{R}$, $1 \le q \le Q_A$, $\Theta_F^q : \mathcal{D} \to \mathbb{R}$, $1 \le q \le Q_F$, $\Theta_C^q : \mathcal{D} \to \mathbb{R}$, $1 \le q \le Q_C$ and continuous $\boldsymbol{\mu}$-independent (bilinear/linear/trilinear) forms $A^q$, $1 \le q \le Q_A$, $F^q$, $1 \le q \le Q_F$, $C^q$, $1 \le q \le Q_C$, respectively. As we shall see in the following, the assumption of affine parametric dependence – not so unusual in many applicative contexts, dealing for instance with different property materials or geometrical blocks – is *relevant* to many instances of both property *and* geometry parametric variation. Nevertheless, this assumption may be relaxed by introducing a suitable approximation of the operators based on the so-called *empirical interpolation method* [21], as we will see in Sect. 2.4.

## 1.3 Optimization and inverse problems governed by PDEs

In this section we introduce the main features of optimal control, shape optimization and other inverse problems governed by (general, non parametrized) PDEs, leaving some remarks on the case of parametrized PDEs to the following section.

### 1.3.1 Optimal control problems

In abstract terms, given a *state* or *controlled* system, the goal of an *optimal control* (OC) problem is to find the best *control input* such that the observation of some output of interest (depending on the solution of the state system) fulfills – in a suitable sense – an objective to be achieved. In our case, state systems are given by boundary value problems for PDEs, written[1] in abstract variational form as

$$A(y(u), w) = F(w; u), \qquad \forall w \in X; \tag{1.9}$$

here $A(\cdot,\cdot) : X \times X \to \mathbb{R}$ and $F(\cdot;u) : X \to \mathbb{R}$ are the bilinear form associated to the differential operator and the linear form associated to the data and the control input, respectively. The PDE is defined over a fixed spatial domain $\Omega$, which for the sake of clarity has been omitted.
The state variable $y = y(u) \in X$ depends on a control function $u \in \mathcal{U}_{ad}$ acting on the system, where $\mathcal{U}_{\rm ad} \subseteq \mathcal{U}$; $X$ and $\mathcal{U}$ denote general reflexive Banach spaces, $X_{\rm ad} \subseteq X$ and $\mathcal{U}_{\rm ad} \subseteq \mathcal{U}$ the space of admissible controls. In the case $\mathcal{U}_{\rm ad} \subsetneq \mathcal{U}$, the problem is said to be control-constrained; similarly, when $y \in X_{\rm ad} \subsetneq X$, it is said to be state-constrained. The control is operated through a source term (when it is called *distributed control*) or through the boundary data (when it is called instead *boundary control*). The objective is usually given by a cost functional $\mathcal{J} : \mathcal{U}_{\rm ad} \to \mathbb{R}$ expressed as a function of the state variable as

$$\mathcal{J}(u) = \tilde{\mathcal{J}}(y(u), u),$$

[1] We consider, for the sake of simplicity, the case of a scalar state variable and of a linear differential operator; the formulation of optimization and inverse problems related to systems modelled by nonlinear PDEs, as well as systems involving vector state variables, follows the same setting.

where $\tilde{\mathcal{J}} : X \times \mathcal{U}_{\text{ad}} \to \mathbb{R}$. Then the OC problem can be expressed as a constrained optimization problem as follows:

$$\hat{u} = \arg\min_{u \in \mathcal{U}_{\text{ad}}} \mathcal{J}(u) \quad \text{s.t.} \quad A(y(u), w) = F(w; u), \quad \forall w \in X, \tag{1.10}$$

where the constraint is given by the PDE modelling the state system. In the simplest case, we deal with linear/quadratic elliptic control problems, in which the weak form of a second order, linear elliptic operator (usually referred to as a general *advection-diffusion-reaction* problem) appearing in (1.9) is given by

$$A(y, w) = \int_\Omega \left( \frac{\partial y}{\partial x_i} a_{ij}(\mathbf{x}) \frac{\partial w}{\partial x_j} + b_i(\mathbf{x}) \frac{\partial y}{\partial x_i} w + c(\mathbf{x}) y w \right),$$

where $\Omega \subset \mathbb{R}^d$ is a bounded, Lipschitz domain; $a_{ij}(\cdot) \in L^\infty(\Omega)$, $b_i(\cdot) \in L^\infty(\Omega)$, for $1 \leq i, j \leq d$, $c(\cdot) \in L^\infty(\Omega)$, $c \geq 0$ some bounded functions, representing the diffusivity (or conductivity) tensor, the advection field and the reaction coefficient, respectively. Moreover, $\mathbb{A} = (a_{ij}(\mathbf{x}))$ is uniformly elliptic, i.e.

$$(\mathbb{A}(\mathbf{x})\boldsymbol{\xi}, \boldsymbol{\xi})_{\mathbb{R}^d} \geq \alpha |\boldsymbol{\xi}|^2, \qquad \forall \mathbf{x} \in \Omega, \quad \forall \boldsymbol{\xi} \in \mathbb{R}^d.$$

The control can be defined as a distributed control or a boundary control; in the former case, the right-hand side of (1.9) is given by

$$F(w; u) = \langle f + B_\Omega u, w \rangle,$$

where $B_\Omega : L^2(\partial\Omega) \to H^{-1}(\Omega)$ is an injection or restriction operator and the control $u$ might act only in a subdomain $\Omega_c \subset \Omega$, so that $B_\Omega u = u \mathbf{1}_{\Omega_c}$, where $\mathbf{1}_{\Omega_c} = \mathbf{1}_{\Omega_c}(\mathbf{x})$ is the characteristic function of $\Omega_c$. In the latter case the control might affect Dirichlet, Neumann or Robin data and in general the right-hand side of (1.9) can be expressed as

$$F(w; u) = \langle f + B_\Gamma u, w \rangle,$$

where (e.g. in the case of a Neumann control) $B_\Gamma : L^2(\partial\Omega) \to H^{-1}(\Omega)$. On the other hand, the general expression of a cost functional is given by

$$\mathcal{J}(u) = \tilde{\mathcal{J}}(y(u), u) = \frac{1}{2} \mathcal{J}_1(Cy(u)) + \frac{\varepsilon}{2} \mathcal{J}_2(u),$$

where $\mathcal{J}_1 : Z \to \mathbb{R}$ is usually a quadratic functional of the observation of the system $z(u) := Cy(u)$. This is typically given by a linear function of the state solution, i.e. $z(y) = Cy$. Here we denote $C : X \to Z$ a suitable observation operator and $Z$ a Banach space (of observations); $z(y)$ might be the state variable itself, its restriction to a subdomain $\Omega_{\text{obs}} \subset \Omega$ or to a portion of the boundary $\Gamma_{\text{obs}} \subset \Gamma$, or (in the cases of our interest) some physical indices depending on the gradient of the state variable. In the more common cases,

$$\mathcal{J}_1(Cy(u)) = \int_\Omega (y(u) - z_d)^2 d\mathbf{x}, \qquad \text{or} \qquad \mathcal{J}_1(Cy(u)) = \int_\Omega |\nabla(y(u) - z_d)|^2 d\mathbf{x},$$

or a portion of its boundary $\partial\Omega$, so that

$$\mathcal{J}_1(Cy(u)) = \int_{\partial\Omega} (y(u) - z_d)^2 d\sigma, \qquad \text{or} \qquad \mathcal{J}_1(Cy(u)) = \int_{\partial\Omega} (\partial_{\mathbf{n}} y(u) - z_d)^2 d\sigma.$$

Here, $z_d \in Z$ is the target state or in general a condition to which we aim to drive the system. In the same way, and without loss of generality, we can assume that the cost functionals we deal with can be written under the following form:

$$\mathcal{J}_1 : X \to \mathbb{R}, \qquad \mathcal{J}_1(y) = \int_\Omega \Phi_\Omega(\mathbf{x}, y, \nabla y)\, d\Omega + \int_{\partial\Omega} \Phi_{\partial\Omega}(\sigma, y, \nabla y)\, d\sigma,$$

where $\Phi_\Omega : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and $\Phi_{\partial\Omega} : \mathbb{R}^{d-1} \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ are two $\mathcal{C}^1$ functions. Instead, $\mathcal{J}_2 : \mathcal{U} \to \mathbb{R}$ is a regularization term given by

$$\mathcal{J}_2(u) = (Nu, u)_\mathcal{U},$$

where $N \in \mathcal{L}(\mathcal{U},\mathcal{U})$ is a symmetric and coercive operator. A simple example is provided by a spatial region or a body occupying the domain $\Omega$ that is to be heated (or cooled); in this case, $y$ represents its temperature, and the state system is modelled by the heat equation, fitting the general expression (1.9). To heat or cool the system, we act either on its boundary or in the whole volume by means of a heat source $u$, which thus plays the role of a boundary control or a distributed control. Our goal is to choose the control in such a way that the corresponding temperature distribution $y = y(u)$ in $\Omega$ approaches as much as possible the desired temperature distribution $z_d$, and the effort in controlling the system is as little as possible. Further examples arising from fluid dynamics will be discussed in Sect. 1.5.

The classical approach for analysis and solution of OC problems governed by PDEs is based on the theory developed by J.L. Lions [192]; an alternative and more general approach is based on the Lagrangian formalism (see for instance [311, 143, 39]). In the latter case, the OC problem is seen as a constrained minimization problem, for which a Lagrangian functional can be defined by introducing a Lagrange multiplier (denoted also ad *adjoint state* or *co-state*); the optimum, if it does exist, is a stationary "point" of the Lagrangian functional. A third possible formulation is inherited from *saddle-point* problems (see e.g. [129]). The solution of an OC problem can be characterized by the so-called Karush-Kuhn-Tucker system; in the case of problems governed by PDEs, it takes the form of a system of PDEs (also known as system of Euler-Lagrange equations) and it is made by the state equation, the adjoint equation which has to be fulfilled by the adjoint state and a *minimum principle* expressing the optimality condition. In Sect. 4.1.2 the case of the optimal control of stationary Navier-Stokes flows will be briefly discussed, recalling the main existence results. We will not address the case of OC problems for parabolic systems; some general insights can be found e.g. in the book by J.L. Lions [192] and in many other more recent monographs (see e.g. [311]).

### 1.3.2 Shape optimization problems

Another class of optimization problems under PDE constraints is given by *optimal design* or *shape optimization* (SO) problems, where the goal is to achieve some objective which involves the solution of a state system, by controlling the shape of the spatial domain itself. As in the OC case, the state system is modelled through a boundary value problem for PDEs, under the abstract variational form

$$A(y(\Omega_o), w; \Omega_o) = F(w; \Omega_o), \qquad \forall w \in X(\Omega_o),$$

while the objective is a shape functional

$$\mathcal{J}(\Omega_o) = \tilde{\mathcal{J}}(y(\Omega_o), \Omega_o),$$

depending on the state solution $y(\Omega_o) \in X(\Omega_o)$; for the sake of clarity, the dependence on the original, variable shape $\Omega_o$ has been highlighted also in the definition of linear/bilinear forms and functional spaces. Here we denote $A(\cdot,\cdot;\Omega_o) : X \times X \to \mathbb{R}$ and $F(\cdot;\Omega_o) : X \to \mathbb{R}$ the bilinear form and the linear form associated to the differential operator and to the data, respectively, $\mathcal{O}_{\text{ad}} \subseteq \mathcal{O}$ the set of all admissible domains, where $\mathcal{O}$ is a larger set of domains, $X(\Omega_o)$ a reflexive Banach space defined on $\Omega_o$; as before, the cost functional can be expressed as a function of the state variable, with $\tilde{\mathcal{J}} : X \times \mathcal{O} \to \mathbb{R}$. Thus, the SO problem can be expressed as the following constrained optimization problem:

$$\hat{\Omega}_o = \arg \min_{\Omega_o \in \mathcal{O}_{\text{ad}}} \mathcal{J}(\Omega_o) \quad \text{s.t.} \quad A(y(\Omega_o), w; \Omega_o) = F(w; \Omega_o), \quad \forall w \in X(\Omega_o), \tag{1.11}$$

where the constraint is given, as in the OC case, by the PDE modelling the state system. In this case, we assume to deal with *shape functionals* under the form

$$\mathcal{J}(\Omega) = \tilde{\mathcal{J}}(y(\Omega), \Omega) = \frac{1}{2}\mathcal{J}_1(y(\Omega)) + \frac{\varepsilon}{2}\mathcal{J}_2(\Omega),$$

where, as in the OC case and without loss of generality,

$$\mathcal{J}_1 : X(\Omega) \to \mathbb{R}, \qquad \mathcal{J}_1(y) = \int_\Omega \Phi_\Omega(\mathbf{x}, y, \nabla y)\, d\Omega + \int_{\partial\Omega} \Phi_{\partial\Omega}(\sigma, y, \nabla y)\, d\sigma.$$

The functional $\mathcal{J}_2 : \mathcal{O} \to \mathbb{R}$ is a regularization term which usually enforces some constraints dealing with the shape (e.g. volume constraints, perimeter constraints, etc.) and may help in ensuring the well-posedness of the shape optimization problem (1.11).
For instance, regarding the case of a region (or a body) to be heated (or cooled) within a shape optimization context, we act on the shape of the domain by considering fixed the heat sources at the boundary and/or in the volume. Thus, we seek the *best* shape – under some constraints, e.g. regarding its volume or its perimeter – which allows to approach as much as possible the desired temperature distribution $z_d$, i.e. the shape providing the best thermal response provided some heating/cooling sources. Further examples of shape optimization problems dealing with fluid dynamics will be discussed in Sect. 1.5.

In general, this class of problems features more difficulties than OC problems, concerning both the theoretical analysis and the numerical solution [248, 137, 139]. For instance, verifying the well-posedness of shape optimization problems involves additional assumptions of regularity on admissible shapes and continuity of the state solution with respect to shape deformations. Like in the OC case, also the solution of a SO problem can be characterized through a Lagrangian approach (see e.g. [7]). Nevertheless, besides numerical approximation of PDEs and optimization, it requires also a suitable approach for representing and deforming efficiently the shape of the underlying geometry, as well as for computing – if needed – shape derivatives. A crucial aspect is thus the geometrical treatment of the shapes during the optimization process and the definition of a set of admissible shapes; following an increasing difficulty, three approaches can be envisaged [7]:

- *parametric shape optimization*: shapes are naturally described by means of a set of $p$ *shape* parameters [7]. In this way, the shape optimization problem is a simple constrained optimization problem in $\mathbb{R}^p$.

- *geometrical shape optimization*: we seek for the solution on a set of shapes with fixed topological properties, being the position of the boundary of the shape under control changed during the optimization process [75, 303, 139].

- *topological shape optimization*: in the most difficult case, shape can be optimized by acting, during the optimization process, both on the position of its free boundary and on its topology [7, 60], by allowing for example the inclusion of holes.

We will focus on some problems which traditionally arise in the context of geometrical shape optimization – some tools for the theoretical analysis will be provided in Sect. 4.2. Nevertheless, as it will be illustrated in the next section and detailed in Sect. 2.5, an effective strategy to reduce computational costs of SO problems relies on suitable *shape parametrizations*, enabling to represent a set of admissible shapes by means of a set of shape parameters, which are iteratively update during the optimization process until convergence to the optimal shape. In order to take advantage of computational reduction techniques for parametrized PDEs, we will set a fully parametrized framework for the numerical solution of this kind of problems – thus recasting some instances of geometrical SO problems into a suitable parametric SO framework – as detailed in the next section and in Sect. 4.3.
In particular, the case of optimal shape design problems related with Stokes/Navier-Stokes flows will be briefly discussed in Sect. 4.2.2, recalling the main existence results, a general way to define shape deformations and a possible characterization of the shape parametrization we deal with. As for the optimal control problems, we will not consider the case of SO problems governed by parabolic PDEs; some details can be found e.g. in the book by Haslinger and Mäkinen [137].

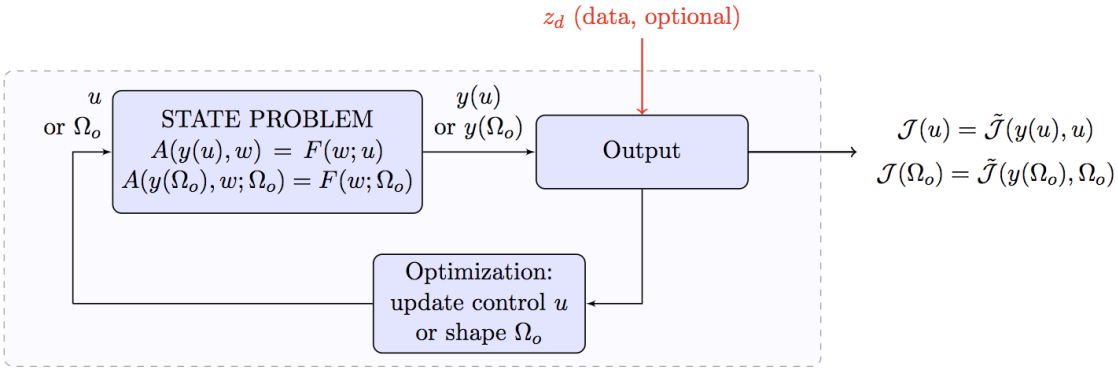



Figure 1.1: Optimal control and shape optimization problems.

### 1.3.3 More general inverse problems

Besides optimal control and shape optimization problems, there are many other examples of inverse problems that can be recast into an optimization framework based on PDEs or, more generally, into a *many-query* context. Far from being exhaustive, we focus in this thesis on three different cases that can fit our general parametrized framework.

A first case is given by the extension of the previous problems to a more general *robust optimization* setting – i.e., robust optimal control or robust shape optimization; these problems occur whenever the state system depends also on some factors which are either uncertain or cannot be controlled. Denoting $q \in \mathcal{Q}$ this factor, we look for an optimal control or an optimal shape which is *robust* with respect to possible variations of $q \in \mathcal{Q}$, i.e. we seek the *worst case scenario* being able to cover a range of possible realizations of this factor. This leads to a *min-max optimization* problem, which in the case of *robust optimal control* can be written as follows:

$$\hat{u} = \arg \min_{u \in \mathcal{U}_{\text{ad}}} \mathcal{H}(u) := \max_{q \in \mathcal{Q}} \mathcal{J}(u, q) \quad \text{s.t.} \quad A(y(u, q), w) = F(w; u, q), \quad \forall w \in X; \tag{1.12}$$

in the same way, a *robust shape optimization* problem can be expressed under the following form:

$$\hat{\Omega}_o = \arg \min_{\Omega_o \in \mathcal{O}_{\text{ad}}} \mathcal{H}(\Omega_o) := \max_{q \in \mathcal{Q}} \mathcal{J}(\Omega_o, q) \quad \text{s.t.} \quad A(y(\Omega_o, q), w; \Omega_o) = F(w; \Omega_o, q), \quad \forall w \in X(\Omega_o).$$

In both cases, cost functional and state problem follow the setting presented in Sect. 1.3.1-1.3.2; in addition, we only highlight their dependence on the uncertain/uncontrollable factor $s$.

A second scenario deals with the solution of some *inverse identification problems*, where we aim at recovering some uncertain features $q \in \mathcal{Q}_{\text{ad}} \subseteq \mathcal{Q}$ of a system described through PDEs, by matching the values of the system observation $z(y) \in Z$ (in our case, computed through numerical simulation) with experimental measurements $z^*$ referred to the same quantity. As in the previous cases, denoting $y = y(q)$ the state of the system, this problem can be written as follows:

$$\hat{q} = \arg \min_{q \in \mathcal{Q}_{\text{ad}}} \mathcal{J}(q) := \tilde{\mathcal{J}}(y(q), q) := \|z(y(q)) - z^*\|_Z^2 \quad \text{s.t.} \quad A(y(q), w; q) = F(w; q), \quad \forall w \in X. \tag{1.13}$$

In particular, in order to exploit our parametrized framework, we aim at describing the quantity $q$ to be identified – which may represent a boundary condition, some physical coefficients, etc. – by means of a parameter (or a set of parameters), thus yielding an *inverse parameter identification* problem. In this way, by minimizing the distance between the experimental measurements $z^*$ – which are usually affected by experimental noise – and the simulated output, we shall identify *possibility regions* that characterize parameter uncertainty, thus yielding an *uncertainty quantification* problem[2]. Some basic premises of uncertainty quantification are, for instance, to *(i)* understand which parts of the system (boundary conditions, material parameters, constitutive laws, etc.) being modelled are not known exactly and which kind of variation we can expect to observe in these quantities; *(ii)* quantify how these variations propagate into the solution of the equations and how they affect the predicted quantities of the simulations, and *(iii)* minimize the negative effects of the uncertainty in the desired outcomes. To deal with inverse identification problems, after recasting them into our parametrized context, we have developed a framework (see Sect. 4.5) for both *deterministic inversion* and *statistical inversion*, in order to address some insights in uncertainty quantification.

A third case is about the possibility to rewrite some simplified instances of *geometrically coupled multiphysics problems* as suitable *inverse problems*. We refer to the former group of problems whenever the coupling conditions are expressed through suitable interface or boundary conditions, such as in the case of *fluid-structure interaction*. By introducing a suitable *shape parametrization*, the coupling conditions between fluid and structure can be formulated in the low-dimensional space of geometric parameters, rather than directly on the larger space of the geometrical displacement fields. In this case, the coupling process is typically iterative – thus falling within a *many-query* context – and the reconstruction of the geometrical displacement from the parameter update requires the solution of an inverse identification problem (see Sect. 4.6.2).

## 1.4 Complexity issues and reduced order modelling

Optimization problems governed by PDEs, as well as other inverse problems, entail very large computational efforts, since they involve iterative optimization procedures that require several PDE solves and cost functional evaluations. Standard techniques rely on iterative optimization schemes based on the gradient of the cost functional, such as the steepest descent method: in this case, for each iteration, the control variable is updated in order to step along the opposite direction of the gradient of the cost functional. This entails the repeated solution of the PDEs system obtained as first order necessary optimality conditions.

[2]Uncertainty quantification has become a popular topic in the mathematical modelling community during the last decade. Moving on the surface of this huge research field, our goal is to exploit the parametrized framework – and computational reduction techniques – in a possible investigation of uncertainty quantification problems, which are naturally *many-query* problems.

Not only, the forward state problem is typically a nonlinear PDE, possibly incorporating coupled multiphysics phenomena. State-of-the-art discretization methods and parallel codes are therefore required to solve them up to a reasonable accuracy – this is exacerbated by the fact that solving the inverse problem requires multiple solutions of the forward problem. Hence, if the forward problem can be replaced with an inexpensive (but reliable) surrogate, solving the inverse problem is much more feasible.
Incorporating geometrical configurations into the framework, e.g. when dealing with shape optimization problems under the form (1.11), makes the inverse problem even less affordable. In fact, even though a SO problem can be solved through the same approaches used for OC problems, discretization procedures result expensive when problem geometry keeps changing; this fact is rather evident also within geometrically coupled multiphysics problems, like fluid-structure interaction (FSI) problems.

For instance, both FSI and SO problems deal traditionally with *(i)* discretization techniques over domains of variable shape; *(ii)* mesh motion and possibly remeshing during iterative procedures for coupling or optimization; and *(iii)* the need of exchanging information over the free boundary (either by transferring loads and displacements between fluid and structure, or moving the fluid boundary according to the minimization of a cost functional). In particular:

- In FSI problems, stresses or pressure loads are transmitted from the fluid domain to the structural nodes on the fluid-structure interface. Once the new structure motion is determined, the displacement of the fluid mesh points (on both the interface and the whole fluid domain) has to be imposed in order to take into account the structural deformation;

- In shape optimization problems, the nodes lying on the boundary to be optimized are usually displaced according to the shape gradient of the chosen cost functional to be minimized [249]. Thus, a classical shape optimization algorithm based on explicit boundary displacement can provide accurate solutions at the expense of being very inefficient, since at each step PDEs have to be solved for a new geometrical configuration.

In both cases, once the boundary displacement has been computed, we need to extend this displacement also to the interior nodes to maintain a good mesh quality. Several methods can be used in order to handle the shape deformation. Usual techniques are based on the so-called *continuous approach*, in which the mesh nodes are displaced according to e.g. the harmonic extension of the boundary displacement to the entire domain; although this strategy is computationally cheap, it can run into trouble when large domain deformations occur. Alternatives include solving the elasticity equations, the biharmonic equation, the use of spring models etc. A second possibility is based on the so-called *point-by-point schemes*, in which the deformation of the mesh nodes is explicitly prescribed (e.g. making the deformation for a node proportional to its distance to the boundary, or using interpolation schemes such as radial basis functions). For a review of coupling methods for non-matching meshes we refer the reader to [70].
Hence, it is rather evident that dealing with a *fixed domain* approach and accounting for the shape deformations directly onto the equations modelling the state problem would entail a substantial computational saving concerning the optimization procedure. Incidentally, we point out that the reduction techniques we have in mind to speedup the solution of the state problem need to rely on a fixed, reference domain, which has to be suitably parametrized, so that the introduction of a shape parametrization and the reformulation on a *fixed domain* can be seen as a very good chance in view of a combined attempt of reduction – both geometrical and computational at the same time.

From a more general standpoint, optimization problems presented in the previous section can be recast in the *many-query* context for at least two reasons:

- In several applications, the state system is modelled by a parametrized PDE, so that the computation of optimal controls or optimal shapes are required for many different values of other input parameters, representing for instance physical or geometrical properties of the underlying system which do not act on the state system as controllers. In this case, the OC problem (in the same way for SO or other inverse problems) takes the following form: given $\boldsymbol{\mu} \in \mathcal{D}$, find

$$\hat{u}(\boldsymbol{\mu}) = \arg \min_{u(\boldsymbol{\mu}) \in \mathcal{U}_{\text{ad}}} \mathcal{J}(u(\boldsymbol{\mu})) \quad \text{s.t.} \quad A(y(u(\boldsymbol{\mu})), w; \boldsymbol{\mu}) = F(w; u(\boldsymbol{\mu}), \boldsymbol{\mu}), \quad \forall w \in X.$$

  The numerical solution of this problem is even less affordable, since an optimization problem is required to be solved for each new input instance. To deal with this additional degree of complexity, we can rely for instance on fast solvers for OC problems based on the so-called *one-shot* approach [268], rather than on more common iterative optimization procedures. Nevertheless, computational reduction techniques for parametrized PDEs are necessary to tackle the need of repeated solutions. However, we will not consider this case in the present work; many details on a recent application in this respect can be found e.g. in [223].

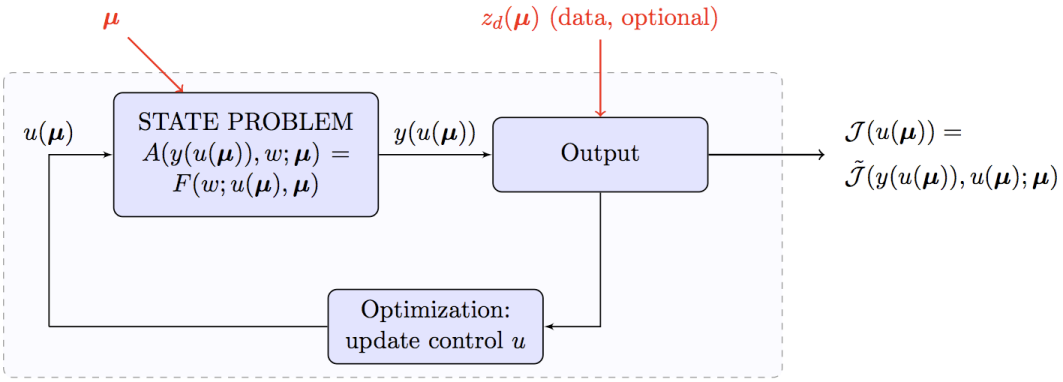



Figure 1.2: Parametrized optimal control problems.

- For the reasons discussed in the previous sections, the solution of OC or SO problems can be very time-consuming also if the underlying system is not parametrized. In this case, the parametric context represents a computational opportunity, rather than a computational burden: whenever it is possible to represent either control inputs, admissible shapes or uncertain elements as parametrized quantities, the reduction framework developed for parametrized PDEs guarantees a significative improvement in computational performances.

The following subsection is devoted to the illustration of the reduced framework analysed and developed in the thesis for solving optimization problems based on parametrized PDEs; care has been devoted to shape optimization problems and other shape related inverse problems, for which geometrical reduction through shape parametrization cannot be renounced.

### 1.4.1 A reduced framework for optimization and inverse problems

The main focus of this thesis is the development, analysis and application of a reduced integrated setting for efficient solution of optimization problems based on parametrized PDEs. This framework exploits the Reduced Basis (RB) method [253, 237, 280] for a rapid and reliable approximation of parametrized PDEs, built upon a full-order discretization technique such as the

FE method. Of course, this is not the only available option, since many other techniques – such as, for instance, Proper Orthogonal Decomposition (POD) [15, 33, 144] – can be exploited to achieve computational reduction. In practice, for inverse problems such as optimal control and shape optimization, very often the standard adjoint-based approach turns to be too computationally expensive; the presence of uncertainty e.g. in measurements or in model parameters make the solution of inverse problems even less feasible.
Substantial computational saving becomes possible thanks to a *reduced order model* (ROM) which relies on two reduction steps: *(i)* parametrization of the control inputs and *(ii)* substitution of the full-order solution of the state problem with a reduced-order solution, which in our case is obtained through the reduced basis method. In particular, the parametrization of control variables in OC problems, as well as the parametrization of (simple) uncertain quantities in a context where shape variation are not involved, is usually straightforward, and is made *a priori* when the model is set. In this case, by expressing for instance the set of admissible controls as a family of functions depending on the parameter vector $\boldsymbol{\mu} \in \mathcal{D}$, i.e.

$$\mathcal{U}_{\text{ad}} = \{u = u(\boldsymbol{\mu}) \in \mathcal{U} \,:\, \boldsymbol{\mu} \in \mathcal{D}_{\text{ad}} \subseteq \mathcal{D}\},$$

the optimal control problem (1.10) can be rewritten in a parametric framework as follows:

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu} \in \mathcal{D}_{\text{ad}}} \mathcal{J}(\boldsymbol{\mu}) \quad \text{s.t.} \quad A(y(\boldsymbol{\mu}), w) = F(w; \boldsymbol{\mu}), \quad \forall w \in X, \tag{1.14}$$

where $y = y(\boldsymbol{\mu})$ is the state of the system – now expressed directly as a function of the *control* parameter vector $\boldsymbol{\mu}$, $A(\cdot, \cdot)$ the (in case, parameter dependent) operator corresponding to a parametrized PDE (see equation (1.2)) and $\mathcal{J}(\boldsymbol{\mu}) = \tilde{\mathcal{J}}(y(\boldsymbol{\mu}), \boldsymbol{\mu})$ the parametrized version of the cost functional to be minimized. In the same way, by expressing the set of uncertain features affecting the state system as

$$\mathcal{Q}_{\text{ad}} = \{q = q(\boldsymbol{\mu}) \in \mathcal{Q} \,:\, \boldsymbol{\mu} \in \mathcal{D}_{\text{ad}} \subseteq \mathcal{D}\},$$

the inverse identification problem (1.13) can be rewritten as follows:

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu} \in \mathcal{Q}_{\text{ad}}} \mathcal{J}(\boldsymbol{\mu}) = \|z(y(\boldsymbol{\mu})) - z^*\|_Z^2 \quad \text{s.t.} \quad A(y(\boldsymbol{\mu}), w; \boldsymbol{\mu}) = F(w; \boldsymbol{\mu}), \quad \forall w \in X. \tag{1.15}$$

Thus, once the control function or the uncertain features have been expressed with respect to some parameters, the problems at hand are rewritten as parametric optimization problems, where each input/output evaluation requires the solution of a parametrized PDE fitting the framework presented in Sect. 1.2. These problems are defined on a fixed geometry, which does not affect the problem and does not require any treatment. On the other hand, parametrization of shapes requires particular care (and additional work) in case of shape optimization and shape-related inverse problems – our ultimate goal. Here we address the main blocks and guidelines of this framework; specific ingredients will be detailed within the next chapters.

### Shape parametrization techniques for geometrical complexity reduction

A first key to simplification consists of making use of a *fixed-domain approach* in shape-related problems. To do this, we introduce a reference configuration on which every problem is brought back and solved at each iteration of the optimization process, whereas geometry variations are accounted for the equation coefficients. In order to represent a set of admissible shapes through a set of geometrical parameters, we rely on suitable (low-dimensional) shape parametrizations: in this way, domains $\Omega_o(\boldsymbol{\mu})$ corresponding to different shape configurations are obtained from the fixed, reference configuration $\Omega$ through a parametric map $T(\,\cdot\,; \boldsymbol{\mu}) : \Omega \to \Omega_o(\boldsymbol{\mu})$. Several options and computational requirements for shape parametrization will be detailed in Sect. 2.5.

Hence, by mapping the problem (1.11) back to the reference domain, and considering the set of admissible shapes $\mathcal{O}_{\text{ad}}$ as

$$\mathcal{O}_{\text{ad}} = \{\Omega_o(\boldsymbol{\mu}) = T(\Omega; \boldsymbol{\mu}) \in \mathcal{O} \,:\, \boldsymbol{\mu} \in \mathcal{D}_{\text{ad}} \subseteq \mathcal{D}\},$$

we obtain the following formulation[3]:

$$\hat{\boldsymbol{\mu}} = \arg \min_{\boldsymbol{\mu} \in \mathcal{D}_{\text{ad}}} \mathcal{J}(\boldsymbol{\mu}) \quad \text{s.t.} \quad A(y(\boldsymbol{\mu}), w; \boldsymbol{\mu}) = F(w; \boldsymbol{\mu}), \quad \forall w \in X, \tag{1.16}$$

where $\mathcal{D}_{\text{ad}} \subseteq \mathcal{D}$ is the subset of admissible parameters. We underline that at this stage the PDE problem – as well as the functional space – has been defined on the reference domain $\Omega$, where the new linear/bilinear forms can be expressed from the original ones through a suitable change of variables. Clearly, if the parametric map $T(\,\cdot\,; \boldsymbol{\mu})$ is sufficiently flexible and low-dimensional, shape deformations can be easily handled by acting on a small number of geometrical parameters – thus yielding to an effective *geometric reduction*.
The introduction of such a map is also instrumental to the use of reduced basis methods, since basis solutions corresponding to different shape configurations $\Omega_o(\boldsymbol{\mu})$ can only be compared and combined on a reference, parameter-independent domain. Once the original problem has been traced back to a reference configuration (and the admissible shapes have been parametrized), it results in a parametrized problem where the effect of geometry variations is traced back onto its parametrized transformation tensors – thus fitting the parametrized framework of Sect. 1.2.

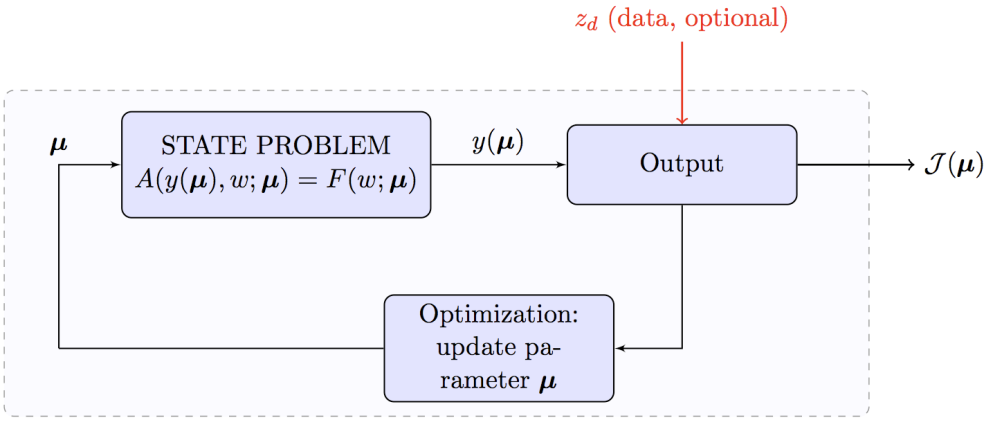



Figure 1.3: Parametric optimization problems (optimal control or shape optimization).

### Reduced basis method for computational reduction

To solve the parametrized PDEs appearing in the formulations (1.14)–(1.16) in a very efficient way, we rely on a suitable reduced-order method, which in our case is the *reduced basis method*. Thanks to a suitable computational *Offline-Online* decomposition (see Sect. 2.2.4), these methods provide rapid results at a greatly reduced cost. At the outer level, the solution of parametrized PDEs and the rapid input/output evaluation is exploited within a suitable iterative procedure for the (now, parametric) optimization, provided e.g. by some nonlinear programming techniques such as sequential quadratic programming and various quasi-Newton methods [23, 143].

[3] As before, we consider for the sake of simplicity in the exposition a linear state problem. The analysis introduced in the next chapters will deal with a more general nonlinear state problem.

Roughly, reduced-order methods for parametrized PDE problems are low-dimensional approximations to the solution of a parametrized PDE, built by choosing a basis $\zeta_i,\ i = 1, \ldots, N$ of full-order discretized solutions computed for some selected parameter values, i.e. $\zeta_i = y(\boldsymbol{\mu}^i)$, where $N$ is – hopefully – much smaller than the dimension (say $\mathcal{N}$) of the full-order discretization. Next, the reduced approximation to the state solution for a new parameter value $\boldsymbol{\mu}$ is defined as

$$y_N(\boldsymbol{\mu}) = \sum_{j=1}^{N} y_{N\,j}(\boldsymbol{\mu})\zeta_j$$

so that $y_N(\boldsymbol{\mu}) \in X_N = \text{span}\{\zeta_1 \ldots, \zeta_N\}$, where the coefficients $\{y_{N\,j}\}$ are determined through a Galerkin projection of the state equation onto the reduced space $X_N$. The cost of such a computation is thus very small – compared to a full-order approximation – if $N$ is small. Hence, if the state solution is approximated in the reduced space $X_N$, then the cost of each optimization step will be very small compared to that involving the full-order state approximation. Hence, throughout the optimization, only the small RB-based discrete solutions are used; this entails relevant savings in CPU costs when compared to the use of a full-order discrete system within the optimization process. We emphasize that a reduced basis method requires the solution of some full-order and therefore very expensive discrete equations. The *key-idea* is that these demanding calculations can be done *Offline*, before the optimization with respect to the control parameters is attempted. Hence, a RB-based strategy will be convenient whenever the optimization procedure requires as many steps as the number of basis functions which have to be computed Offline. However, very often the same reduced basis for the approximation of the state system can be exploited for the optimization with respect to many different cost functionals – this greatly enhances the computational saving offered by RB-based strategies. The principal features of RB methods will be introduced in Chapter 2; the RB approximation of fluid dynamics problems modelled by Stokes and Navier-Stokes equations will be presented in Chapter 3.

Computational reduction strategies such as RB methods or POD have already been employed to speedup the solution of optimal control, simpler shape optimization and other inverse problems dealing with parametrized PDEs. First examples of optimal control problems solved by exploiting computational reduction techniques have been addressed by Ito and Ravindran, in the context either of (a preliminary version of) the reduced basis method [152, 153, 154] or of the proper orthogonal decomposition method [264, 265, 267]. A more systematic approach has been firstly provided by Maday and Patera [198, 253]. More recent contributions dealing with RB methods have been presented in both the elliptic case by Quarteroni, Rozza and Quaini [257], by Tonn and Urban [309], Grepl and Kärcher [121], and the parabolic case by Dedè [73, 74]. Other recent works dealing with optimal control problems through POD techniques have been addressed for instance by Kunisch and Volkwein [176], Tonn, Urban and Volkwein [310] dealing with POD techniques.
Shape optimization problems solved by means of RB techniques have been addressed for the first time by Rozza [273] in the case of simple shape parametrizations and afterwards by Lassila and Rozza [185] in a more complex case; other inverse design problems solved through POD techniques have been presented by Willcox and Ghattas [51, 52], whereas a more recent treatment of uncertainties in aerodynamic design dealing with POD-based reduced order models to inverse has been addressed by Schulz and Schillings [295, 296]. Other reduced models, such as balanced truncation [14], have been applied in the last two years to tackle shape optimization problems by Heinkenschloss, Hoppe, Sorensen, Antil [13, 12].
Furthermore, reduced models have been recently applied to real-time parameter estimation by Nguyen and many others [225, 123], as well as to statistical inverse problems related with uncertainty quantification by Willcox, Ghattas et al. [106, 191], Patera, Huynh, Knezevic, [147], Nguyen and Rozza [227], Grepl and Veroy [124]. An exhaustive review of reduced order models in this context can be found in [101].

## 1.5 Optimization problems in fluid dynamics

The control of fluid flows for the purpose of achieving some desired objective is crucial to many applications. The reduction of drag forces acting on airfoils [159] or the vorticity minimization in cardiovascular prostheses [254, 273] are just two examples from very different fields, which can be recast in the same control framework. In particular, *design problems* are of particular interest whenever the performances of a system can be greatly enhanced by acting on its shape, such as in aircraft, ship or automotive design, as well as in biomedical engineering.

In fluid mechanics, cost functionals are expressed in terms of flow variables (such as velocity, pressure, temperature, etc.), while constraints are represented by the PDE (advection-diffusion, potential, Euler, Stokes, Navier-Stokes or Boussinesq equations, etc.) describing the flow, and by topological constraints on the shape of the domain, whenever necessary. In the applications of our interest we will focus on steady, incompressible viscous fluid flows, governed by either Stokes or Navier-Stokes equations for laminar Newtonian flows. These equations – extensively analyzed in Sect. 3.1 – are the steady version of the classical mathematical description of Newton's law of motion and conservation of mass for an incompressible viscous flow, and read as follows:

$$\begin{cases} -\nu\Delta\mathbf{v} + \delta(\mathbf{v}\cdot\nabla)\mathbf{v} + \nabla p = \mathbf{f} + \mathbf{u}_\Omega & \text{in } \Omega \\ \nabla\cdot\mathbf{v} = 0 & \text{in } \Omega \\ \mathbf{v} = \mathbf{u}_{\mathrm{c}} & \text{on } \Gamma_{\mathrm{c}} \\ \mathbf{v} = \mathbf{v}_{\mathrm{in}} & \text{on } \Gamma_{\mathrm{in}} \\ \mathbf{v} = \mathbf{0} & \text{on } \Gamma_{\mathrm{w}} \\ -p\mathbf{n} + \nu\dfrac{\partial\mathbf{v}}{\partial\mathbf{n}} = \mathbf{g} & \text{on } \Gamma_{\mathrm{out}}. \end{cases} \tag{1.17}$$

Here $\delta = 1$ in the case of Navier-Stokes (resp. $\delta = 0$ in the case of Stokes) equations, $\mathbf{v}$ is the velocity vector, $p$ is the pressure, $\nu > 0$ is the kinematic viscosity – defined as the ratio between the dynamic viscosity $\mu > 0$ and the density $\rho > 0$ of the fluid – and $\Omega \subset \mathbb{R}^d$ for $d = 2, 3$ is the domain occupied by the fluid. The Dirichlet portion $\Gamma_D$ of the boundary is further divided into an inlet $\Gamma_{\mathrm{in}}$ where we prescribe an inflow condition, a wall $\Gamma_{\mathrm{w}}$ where we prescribe a no-slip boundary condition and a portion $\Gamma_{\mathrm{c}}$ where we prescribe a *boundary control* $\mathbf{u}_{\Gamma_{\mathrm{c}}}$. Furthermore, we prescribe the value of the normal stress $\mathbb{T}(\mathbf{v}, p)\mathbf{n}$ at the outlet $\Gamma_{\mathrm{out}} \neq \emptyset$, where $\mathbb{T}(\mathbf{v}, p)$ is the stress tensor defined by

$$\mathbb{T}(\mathbf{v}, p) = -p\mathbb{I} + 2\nu\sigma(\mathbf{v}), \qquad \sigma(\mathbf{v}) = \frac{1}{2}(\nabla\mathbf{v} + \nabla^T\mathbf{v});$$

$\sigma(\mathbf{v})$ is called *strain rate tensor*. The control $\mathbf{u}$ can be applied at the boundary, i.e. $\mathbf{u} = \mathbf{u}_{\mathrm{c}} \in \mathcal{U}_c \subseteq (H^{1/2}(\Gamma_c))^d$, or can be distributed in the domain, i.e. $\mathbf{u} = \mathbf{u}_\Omega \in \mathcal{U}_\Omega \subset (L^2(\Omega))^d$, as a source term. Typical problems are related with the reduction of the *vorticity* $\nabla \times \mathbf{v}$, e.g. in turbulent flows, for which a model problem can be stated as follows:

$$\hat{\mathbf{u}} = \arg\min_{u\in\mathcal{U}_{\mathrm{ad}}} \mathcal{J}(\mathbf{u}) \;\text{ s.t. }\; (\mathbf{v}, p) = (\mathbf{v}, p)(\mathbf{u}) \text{ is the solution of (1.17) associated with } \mathbf{u}, \tag{1.18}$$

where

$$\mathcal{J}(\mathbf{u}) = \frac{\nu}{2}\int_\Omega |\nabla\times\mathbf{v}(\mathbf{u})|^2\, d\Omega + \frac{\varepsilon}{2}\int_\Omega |\mathbf{u}|^2 d\Omega$$

is the cost functional related to a distributed observation of the vorticity and a distributed control $\mathbf{u} = \mathbf{u}_\Omega$. The second integral appearing in the cost functional plays the role of a regularization term, where $\varepsilon > 0$ is a suitable (sometimes called *Tikhonov*) regularization factor.

In the case of a boundary (Dirichlet) control $\mathbf{u} = \mathbf{u}_c$, $\mathcal{J}(\mathbf{u})$ takes the following form:

$$\mathcal{J}(\mathbf{u}) = \frac{\nu}{2}\int_\Omega |\nabla \times \mathbf{v}(\mathbf{u})|^2\, d\Omega + \frac{\varepsilon}{2}\int_{\Gamma_c} |\mathbf{u}|^2 d\Omega.$$

Another cost functional which can be used is the following *energy* functional

$$\mathcal{J}(\mathbf{u}) = \frac{\nu}{2}\int_\Omega |\nabla \mathbf{v}(\mathbf{u})|^2\, d\Omega + \frac{\varepsilon}{2}\int_{\Omega} |\mathbf{u}|^2 d\Omega,$$

if we are interested in the minimization of the viscous energy dissipation $\nu|\nabla \mathbf{v}|^2$ (referred also as strain rate) for flow regularization. Other examples will be discussed for instance in Sect. 5.2.4.

In the case of optimal design problems, we act instead on a portion $\Gamma_c \subseteq \partial\Omega_o$ of the boundary (also denoted as *free-boundary*), in order to minimize a shape functional depending on velocity $\mathbf{v}$ and/or pressure $p$, which – following the same notation as before – solve:

$$\begin{cases} -\nu\Delta\mathbf{v} + \delta(\mathbf{v}\cdot\nabla)\mathbf{v} + \nabla p = \mathbf{f} & \text{in } \Omega_o \\ \nabla\cdot\mathbf{v} = 0 & \text{in } \Omega_o \\ \mathbf{v} = \mathbf{v}_{in} & \text{on } \Gamma_{\text{in}} \\ \mathbf{v} = \mathbf{0} & \text{on } \Gamma_c \cup \Gamma_{\text{w}} \\ -p\mathbf{n} + \nu\dfrac{\partial \mathbf{v}}{\partial \mathbf{n}} = \mathbf{g} & \text{on } \Gamma_{\text{out}}. \end{cases} \tag{1.19}$$

Thus, a model problem can be stated as follows:

$$\hat{\Omega}_o = \arg\min_{\Omega_o \in \mathcal{O}_{\text{ad}}} \mathcal{J}(\Omega_o) \quad \text{s.t.} \quad (\mathbf{v}, p) = (\mathbf{v}, p)(\Omega_o) \text{ is the solution of (1.19) on } \Omega_o, \tag{1.20}$$

where the set of admissible shapes can be defined as

$$\mathcal{O}_{\text{ad}} = \{\Omega \in \mathcal{O} \,:\, |\Omega| \le V,\ \Gamma_{\text{in}} \cup \Gamma_{\text{w}} \cup \Gamma_{\text{out}} \text{ is given } \},$$

where $\Omega_o$ is the fluid domain and in this case $\mathcal{O}$ is the set of domains $\Omega_o \subset \mathbb{R}^d$ piecewise $C^2$ with convex corners. When dealing with internal flows, $\partial\Omega_o$ represents the external wall of the branch $\Omega_o$ whereas in the case of external flows the target is the shape of a body $B$ embedded into a fictitious fluid volume $D$, so that $\Omega_o = D \setminus B$ and $\Gamma_c \subseteq \partial B$.
Depending on the objective, several shape functionals can be defined; omitting for the sake of simplicity the regularization term, in case of energy or vorticity minimization, we can introduce the following functionals:

$$\mathcal{J}(\Omega) = \frac{\nu}{2}\int_\Omega |\nabla \times \mathbf{v}(\Omega)|^2\, d\Omega, \qquad \mathcal{J}(\Omega) = \frac{\nu}{2}\int_\Omega |\nabla \mathbf{v}(\Omega)|^2\, d\Omega,$$

respectively. Another very well studied problem deals with the minimization of drag forces (or resistances) on a body $B$ in relative motion within a fluid, for which the usual functional to be minimized is given by the *drag* acting on the body:

$$\mathcal{J}(\Omega) = -\int_{\partial B} (\mathbb{T}(\mathbf{v}(\Omega), p(\Omega))\mathbf{n}) \cdot \hat{\mathbf{v}}_\infty d\Gamma,$$

where $\mathbf{v}_\infty = U\hat{\mathbf{v}}_\infty$ is the horizontal component of the fluid flow acting on the body and $\hat{\mathbf{v}}_\infty$ a unit vector in the horizontal direction. An equivalent choice is the adimensional drag coefficient:

$$\mathcal{J}(\Omega) = -\frac{1}{q_\infty d}\int_{\partial B} (\mathbb{T}(\mathbf{v}(\Omega), p(\Omega)))\mathbf{n}) \cdot \hat{\mathbf{v}}_\infty d\Gamma,$$

where $d$ is a characteristic length of the body and $q_\infty = \frac{1}{2}\rho U^2$. Also concerning the case of shape optimization problems, several other examples will be introduced in Sect. 5.2.4.
Indeed, in the context of optimal flow control problems, the challenges stem from the nature of the state system: the nonlinearity of Navier-Stokes equations leads to nonconvex problems which cause major difficulties concerning existence and uniqueness analysis. The case of optimal design problems is even more difficult, due to further critical aspects related with shape regularity and deformations, and to continuity of the nonlinear state solution with respect to shape variations. Furthermore, on the computational side, applications within these contexts lead to large-scale nonlinear optimization problems, which are among the most challenging optimization problems in computational science and engineering.

Early contributions on optimal control problems associated with Navier-Stokes equations (both in the stationary and in the time-dependent case) date back to 90's and are addressed by Abergel and Temam [2], Casas [1], a series of articles by Gunzburger, Hou, Svobodny ( [130,131], e.g.) and Kim [132, 167], Ghattas [112], Heinkenschloss [138], Berggren [30, 31, 32], Ito and Ravindran [81, 266]. These authors are mainly concerned with important questions such as the formulation of feasible problems, existence of optimal controls, first-order necessary conditions for optimality, and discretization issues. A complete review of challenges and features related to optimal flow control problems can be found in the monograph by Gunzburger [128]. More recent contributions – mainly concerning the numerical approximation, and among a very long list including many authors – have been presented by Hintermuller, Kunisch, Volkwein [142] and Vexler [174], Agoshkov, Quarteroni and Rozza [5, 6] and Dedè [71].
Concerning optimal design problems related with fluid flows, the first theoretical contributions date back to the 70's and are due to Pironneau [246, 247] and Glowinski [116], and subsequently to many other authors such as Simon [301, 27], Zolésio and coauthors [37, 89], Gunzburger [133] and Kim [166], and more recently by Gao, Ma and Zhuang [108, 107]. These works are mainly concerned with the formulation of feasible problems, existence of optimal shapes, regularity and differentiability of state solution and cost functionals with respect to shape and optimality conditions. Starting from a pioneering work by Bourot [42], a long list of authors have proposed many tools for tackling the numerical difficulties arising in shape optimization for fluid dynamics problems. We just mention Jameson [158, 159, 160], whose role has been fundamental for the application of shape optimization techniques to aerodynamics problems such has the optimal design of airfoils. An exhaustive survey on both theoretical and numerical aspects is provided e.g. in the book by Pironneau and Mohammadi [214] and other reviews by these authors [249, 215].

On the other hand, the first attempts to solve flow control problems by reduced order models date back to last decade and are given to Ito [153, 154] and Ravindran [267, 265]; subsequent developments in flow control problems are given by Quarteroni, Rozza and Quaini [257], Tonn and Urban [309], Dedè [73, 74]. Concerning reduced order models applied to shape optimization problems – at the best of our knowledge – the only contributions are due to Antil, Heinkenschloss and Hoppe [13, 12], beyond the works by Rozza [276, 273] and Lassila [185], which constitute the seeds from which applications and methods presented in this thesis actually stem.
Some general remarks about the use of reduced models in flow control problems have been addressed by Gunzburger in his monograph [128]. Nevertheless, whether several strategies for geometrical reduction have been fully exploited in optimal design problems, the introduction of reduced-order models is somehow very recent and largely still to be explored. In our opinion, this is due basically to *(i)* the lack of adequate methods for the certification of reduced-order solutions up to the last decade, and *(ii)* to the lack of advanced parametrization techniques within a parametrized PDEs context. We underline that efficient and rigorous *a posteriori* error bounds are still missing for a large class of optimal control problems: the main difficulty stands in the construction of rigorous error bounds not just for reduced state variables, but also for the reduced cost functional and the reduced control – an aspect which is not taken into account in this thesis, but represents an interesting topic for forthcoming research.

Indeed, our ultimate goal is the investigation of a reduced framework based on the reduced basis method and suitable shape parametrization techniques for optimal design problems arising in fluid dynamics. In particular, our driving applications deal with haemodynamics, such as design of cardiovascular prostheses and assessment of pathological risks through inverse identification of flow and shape features. To do this, we will face both optimal control and shape optimization, as well as with more general inverse identification problems. Clinical motivations, modelling features and detailed results will be addressed in Part III.
Nevertheless, techniques and methods developed and discussed in this thesis prove to be useful also in facing other optimization problems arising in computational fluid dynamics and, more generally, in science and engineering contexts modelled by PDEs.

## 1.6 Thesis contributions, scopes and limitations

In this thesis we focus on the development of suitable reduced models for the numerical solution of optimal control, shape optimization and other inverse problems. We rely on a certified reduced basis methodology for parametrized PDEs for the sake of computational reduction, whereas we exploit suitable shape parametrization techniques in problems dealing with variable shapes for the sake of geometrical reduction. The coupling of these techniques has pushed us to build a modular framework, where several shape parametrizations can be taken into account – depending on the applicative context – and to develop further methodologies to face additional approximation and reduction issues.

Concerning the reduced basis methodology, we introduce some advances to the current framework for two-dimensional steady fluid dynamics equations, for steady Stokes and Navier-Stokes equations. First, we extend the certified framework to the treatment of general parametrized *nonaffine* Stokes and Navier-Stokes problems, exploiting the empirical interpolation method. Second, we provide both a stability study based on the role of the Brezzi's inf-sup constant in the RB context and an error analysis and certification for velocity and pressure fields by jointly exploiting either the Babuška's or the Brezzi-Rappaz-Raviart theory in the Stokes and Navier-Stokes cases, respectively. Moreover, we also develop *a posteriori* error bounds for quadratic functionals in the Stokes case. Third, we improve the Offline/Online computational procedure (in particular for the nonlinear case) and apply for the first time the *natural-norm* Successive Constraint Method for estimating lower bounds of parametrized stability factors in the framework of general noncoercive problems, providing a preliminary extension to the nonlinear case.
Regarding shape parametrization techniques, we manage complex geometrical configurations hosting internal flows in problems dealing with variable shapes by exploiting the Free-Form Deformation (FFD) technique and applying, for the first time, the Radial Basis Function (RBF) paradigm in the context of parametrized shape deformations. Furthermore, we provide some theoretical results about these two techniques which can be useful also in a more classic shape optimization framework.

The problems and the goals presented in Sect. 1.3-1.4 are quite challenging, and the variety of methods (for discretization, optimization) that could be addressed is extensive. For instance, we do not consider adjoint-based optimization methods for the solution of optimal control or shape optimization problems: instead, we treat the optimization procedure relying on fast and certified input/output evaluations within a *many-query* context. This work can be seen as a part of a much larger effort in developing RB methods and error estimation procedures for optimization problems based on parametrized PDEs. Nevertheless, we consider also new applications in the context of inverse problems, such as robust optimal control problems, robust shape optimization problems, inverse identification problems. In particular, we develop a Bayesian framework relying on RB input/output evaluations, for the analysis of statistical inversion and uncertainty

quantification problems within a more general *many-query* context. The efforts in setting this coupled computational/geometrical reduction framework, driven by the solution of optimization and inverse problems, have led to develop suitable methods for fast numerical simulations and related output evaluations which can be of interest in many applicative fields in science and engineering.

Examples and applications provided to justify and support methodological developments deal with problems of great interest in haemodynamics, such as the optimal design of better prosthetic devices (e.g. bypass graft prostheses) and the characterization of physical and geometrical flow properties which may be related to pathological risks. Under some (not negligible) restrictions and modelling assumptions, we are able to provide some preliminary results which may be of applicative interest after extending the current framework to more complex fluid dynamics models. Results have already provided at this step some interesting information on fine physical details and reduction of computational efforts. Of course, methods and tools developed within this work are general and can be applied to several other contexts and problems of interest in science and engineering.
However, RB methods for fluid dynamics equations have also been applied to other physical parametrization of great interest, to three-dimensional problems [79, 148] and to time-dependent problems [168, 228]. These problems are not addressed in any detail in this work, with the exception of a short analysis in Chapter 6 and a discussion of future work in the final Conclusions.

Concerning the reduced basis methodology, this thesis builds on earlier works on Stokes [278, 282] and Navier-Stokes [79, 255, 312] equations. On the other hand, concerning mathematical problems and applications addressed within, this thesis extends the earlier, precursor work [276, 273] by G. Rozza, sharing many focuses but addressing more general features and results. Throughout this work, results refer to computations executed on a personal computer with $2 \times 2$GHz Dual Core AMD Opteron (tm) processors 2214 HE and 16 GB of RAM. Reduced basis computations have been performed within the `rbMIT` [149] library, created by A.T. Patera and his group at MIT. Practicing and extending this library has taken a non-negligible part of the activity carried out within this work and represent a further achievement of this thesis.

# 2 Computational/geometrical reduced order modelling

In this chapter we describe the computational and geometrical reduction techniques involved in our general framework for the solution of optimal control, shape optimization and other inverse problems. After a brief insight on the main features shared by several computational reduction techniques, we focus on reduced basis methods for parametrized PDEs, introducing the setting and the notation useful for the formulation and the analysis of the parametrized problems we will deal with. The second part of the chapter is devoted to the illustration of several available approaches for shape parametrization, which will be exploited afterwards in the numerical solution of optimization and inverse problems related with shape variations.

## 2.1 Computational reduction: main features and goals

As pointed out in Chapter 1, problems arising in several *many-query* contexts and/or dealing with PDE-constrained optimization entail very large computational costs and require a strong computational reduction to be tackled and solved.
The goal of computational reduction techniques is to capture the essential features of the input/output behavior of a system in a rapid and reliable way, i.e. *(i)* by improving the computational performances and *(ii)* by keeping under control the approximation error between the reduced-order solution and the full-order one. In particular, we aim at approximating the solution of a PDE system using a handful of degrees of freedom instead of the millions of degrees of freedom needed for a full-order approximation. In this way, we need to solve the full-order problem only for few instances of the input (through a demanding *Offline* stage), in order to be able to perform many low-cost reduced-order simulations (very inexpensive *Online* stage) for new instances of the input. We may distinguish between two general *paradigms* in computational reduction – *projection* vs. *interpolation* – yielding to the following groups of techniques:

1. *Computational Reduction Techniques* (CRT) are *problem-dependent* methods which aim at reducing the dimension of the algebraic system arising from the discretization of a PDE problem. The reduced solution is thus obtained through a *projection* of the original problem onto a small subspace made by *global* basis functions, constructed for the specific problem, rather than onto a large space of generic, *local* basis functions.
2. *Surrogate Models*, also known as *metamodels* or *emulators*, are instead *problem-transparent* methods, which provide an approximation of the input/output map by fitting a set of input/output data obtained by numerical simulation. These methods are widely employed in analysis and post-processing of computer experiments, to extract and synthetize information from large sets of simulations. In this case, PDEs are still solved through full-order discretization techniques – thus reduction acts just at a subsequent stage.

Our focus is on computational reduction techniques for parametrized PDEs, aiming at computing, in a cheap way, a low-dimensional approximation of the PDE solution that would be obtained through a higher-fidelity, computationally expensive discretization scheme. The most common choices, like reduced basis (RB) or proper orthogonal decomposition (POD) methods, seek for a reduced solution through a *projection* onto suitable low-dimensional subspaces[1]. Of course, several other reduction strategies for computational reduction are available; we cite low-rank tensor methods [165, 170], Krylov subspace methods, balanced truncation [188, 126], modal truncation and other balancing-related truncation techniques employing the idea of spectral projection [28]. On the other hand, surrogate models have been initially exploited within numerical experiments in substitution of expensive simulations whenever interested to parametric studies and sensitivity analyses. Nowadays, they provide an essential tool in planning and post-processing of large computer experiments and optimization problems. A general survey of these techniques can be found for instance in [284, 316], while some applications to analysis of computer experiments are addressed e.g. in [289, 314]. Some comparisons between surrogate and reduced order models are provided in [101] and in a recent work developed with Quarteroni and Rozza [207].

RB and POD methods are not bound to Galerkin approximations – such as the ones obtained through finite elements methods – in fact they can also be built over other discretization techniques, we present the key reduction issues starting from a differential problem under strong form. For the sake of simplicity, we focus here on a linear, time-independent problem: evaluate $s(\boldsymbol{\mu}) = S(y(\boldsymbol{\mu}); \boldsymbol{\mu})$ where the solution $y(\boldsymbol{\mu}) \in X = X(\Omega)$ satisfies

$$\mathcal{L}(\boldsymbol{\mu}) y(\boldsymbol{\mu}) = \mathcal{F}(\boldsymbol{\mu}), \tag{2.1}$$

where $\mathcal{L}(\boldsymbol{\mu}) : X \to X'$ is a second-order parametrized differential operator and $\mathcal{F}(\boldsymbol{\mu}) \in X'$, for any $\boldsymbol{\mu} \in \mathcal{D}$. The weak formulation of (2.1) reads (recalling (1.2)): find $y(\boldsymbol{\mu}) \in X = X(\Omega)$ such that

$$A(y(\boldsymbol{\mu}), w; \boldsymbol{\mu}) = F(w; \boldsymbol{\mu}), \quad \forall\, w \in X, \tag{2.2}$$

where $A(\cdot, \cdot; \boldsymbol{\mu}) : X \times X \to \mathbb{R}$ and $F(\cdot; \boldsymbol{\mu}) : X \to \mathbb{R}$ are given[2], respectively, by

$$A(y, w; \boldsymbol{\mu}) := {}_{X'}\langle \mathcal{L}(\boldsymbol{\mu}) y, w \rangle_X, \qquad F(w; \boldsymbol{\mu}) = {}_{X'}\langle \mathcal{F}(\boldsymbol{\mu}), w \rangle_X, \qquad \forall\, y, w \in X. \tag{2.3}$$

Here the attention is thus restricted to a typically *smooth* and rather *low-dimensional* parametrically induced manifold $\mathcal{M} = \{y(\boldsymbol{\mu}) \in X : \ \boldsymbol{\mu} \in \mathcal{D}\}$, spanned by the set of fields engendered as the input $\boldsymbol{\mu}$ varies over the parameter domain $\mathcal{D}$ – clearly, generic approximation spaces are unnecessarily rich (and expensive) within the parametric framework. The essential components of a CRT can be thus summarized as follows:

- **Full-order discretization technique**

  In any case, a CRT is premised upon, and does not replace completely, a full-order or high-fidelity (sometimes also denoted as *truth*) discretization method, such as the finite element method[3] (FEM), which in the linear, time-independent case (2.1) reads as: given $\boldsymbol{\mu} \in \mathcal{D}$, evaluate $s^{\mathcal{N}}(\boldsymbol{\mu}) = S(y^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu})$, where $y^{\mathcal{N}}(\boldsymbol{\mu}) \in X^{\mathcal{N}}$ is such that

  $$\mathcal{L}^{\mathcal{N}}(\boldsymbol{\mu}) y^{\mathcal{N}}(\boldsymbol{\mu}) = \mathcal{F}^{\mathcal{N}}(\boldsymbol{\mu}), \tag{2.4}$$

[1] Indeed, we remark that several CRTs, like POD, have been originally introduced and developed in order to speed-up the solution of very complex time-dependent and nonlinear problems, like for turbulent flows, without being addressed to parametrized problems (i.e. time was considered somehow as the only parameter).

[2] In a more rigorous way, we should introduce the Riesz identification operator $R : X' \to X$ by which we identify $X$ and its dual, so that, given a third Hilbert space $H$ such that $X \hookrightarrow H'$ and $H' \hookrightarrow X'$, ${}_{X^*}\langle \mathcal{L}(\boldsymbol{\mu}) u, v \rangle_X = (R\, \mathcal{L}(\boldsymbol{\mu}) u, v)_H$. However, the Riesz operator will be omitted for the sake of simplicity.

[3] Historically [97] RB methods have been built upon *finite element* discretizations. However, spectral elements [196, 194], finite volumes [135], finite differences and other discretization methods may be considered too.

and $X^{\mathcal{N}} \subset X$ is a finite dimensional (*truth* approximation) space of typically very large dimension $\mathcal{N} = \mathcal{N}_h$, $\mathcal{L}^{\mathcal{N}}(\boldsymbol{\mu})$, $\mathcal{F}^{\mathcal{N}}(\boldsymbol{\mu})$ are discrete operators, and $h$ denotes the spacing of the computational FE grid. The RB method consists thus in a low-order approximation of the "truth" manifold $\mathcal{M}^{\mathcal{N}} = \{y^{\mathcal{N}}(\boldsymbol{\mu}) \in X^{\mathcal{N}} : \ \boldsymbol{\mu} \in \mathcal{D}\}$.

We assume that

$$\|y(\boldsymbol{\mu}) - y^{\mathcal{N}}(\boldsymbol{\mu})\|_X \leq \mathcal{E}(h), \qquad \forall\, \boldsymbol{\mu} \in \mathcal{D},$$

where $\mathcal{E}(h)$ is an estimation of the discretization error, which can be made as small as desired by choosing a suitable spacing size $h$. From an abstract standpoint, introducing the projection operators $\Pi_{\mathcal{N}} : X \to X^{\mathcal{N}}$ and $\Pi^*_{\mathcal{N}} : X' \to X^{\mathcal{N}\,\prime}$ onto $X^{\mathcal{N}}$ and $X^{\mathcal{N}\,\prime}$, respectively, and denoting $y^{\mathcal{N}}(\boldsymbol{\mu}) = \Pi_{\mathcal{N}}\, y(\boldsymbol{\mu})$, we have that

$$\Pi^*_{\mathcal{N}}(L(\boldsymbol{\mu})\Pi^{-1}_{\mathcal{N}} y^{\mathcal{N}}(\boldsymbol{\mu}) - \mathcal{F}(\boldsymbol{\mu})) = 0, \tag{2.5}$$

so that we can identify $\mathcal{L}^{\mathcal{N}}(\boldsymbol{\mu}) = \Pi^*_{\mathcal{N}} L(\boldsymbol{\mu})\Pi^{-1}_{\mathcal{N}}$ and $\mathcal{F}^{\mathcal{N}}(\boldsymbol{\mu}) = \Pi^*_{\mathcal{N}}\mathcal{F}(\boldsymbol{\mu})$. Equivalently, thanks to (2.3) the weak formulation of the problem (2.4) reads:

$$A(y^{\mathcal{N}}(\boldsymbol{\mu}), w^{\mathcal{N}}; \boldsymbol{\mu}) = F(w^{\mathcal{N}}; \boldsymbol{\mu}), \quad \forall\, w^{\mathcal{N}} \in X^{\mathcal{N}}. \tag{2.6}$$

- **Space construction and (Galerkin) projection**

  Any CRT usually consists of selecting a (reduced) basis of few full-order PDE solutions $\{y^{\mathcal{N}}(\boldsymbol{\mu}^i)\}_{i=1}^N$ and seeking for a reduced approximation $y_N(\boldsymbol{\mu})$ expressed as a linear combination of the basis functions [237]. Given a positive integer $N_{\max}$, we then introduce an associated sequence of approximation spaces: for $N = 1, \ldots, N_{\max}$, $X_N^{\mathcal{N}}$ is a $N$-dimensional subspace of $X^{\mathcal{N}}$; we further suppose that they are nested (or hierarchical), i.e. $X_1^{\mathcal{N}} \subset X_2^{\mathcal{N}} \subset \cdots X_{N_{\max}}^{\mathcal{N}} \subset X^{\mathcal{N}}$; this condition is fundamental in ensuring (memory) efficiency of the resulting reduced approximation. In order to define a (hierarchical) sequence of Lagrange[4] spaces $X_N^{\mathcal{N}}$, $1 \leq N \leq N_{\max}$, we first introduce a "master set" of properly selected parameter points $\boldsymbol{\mu}^n \in \mathcal{D}$, $1 \leq n \leq N_{\max}$. We then define, for given $N \in \{1, \ldots, N_{\max}\}$, the Lagrange parameter samples

  $$S_N = \{\boldsymbol{\mu}^1, \ldots, \boldsymbol{\mu}^N\}\ , \tag{2.7}$$

  and associated Lagrange reduced spaces

  $$X_N^{\mathcal{N}} = \text{span}\{y^{\mathcal{N}}(\boldsymbol{\mu}^n),\ 1 \leq n \leq N\}\ ; \tag{2.8}$$

  the $y^{\mathcal{N}}(\boldsymbol{\mu}^n)$, $1 \leq n \leq N_{\max}$, are often referred to as "(retained) snapshots" of the parametric manifold $\mathcal{M}^{\mathcal{N}}$ and are obtained by solving the FE problem (2.6) for $\boldsymbol{\mu}^n$, $1 \leq n \leq N_{\max}$; the sampling strategy used to build the set $S_N$ is based on a *greedy* algorithm (see Sect. 2.2.2). From now on, we omit the superscript $^{\mathcal{N}}$ when denoting the space $X_N^{\mathcal{N}} \equiv X_N$.

  Then, we compute the reduced approximation $y_N(\boldsymbol{\mu})$ through a (Galerkin-like) projection of the PDE onto $X_N$: given $\boldsymbol{\mu} \in \mathcal{D}$, evaluate $s_N(\boldsymbol{\mu}) = S(y_N(\boldsymbol{\mu}); \boldsymbol{\mu})$, where $y_N(\boldsymbol{\mu}) \in X_N$ solves

  $$\mathcal{L}_N(\boldsymbol{\mu}) y_N(\boldsymbol{\mu}) - \mathcal{F}_N(\boldsymbol{\mu}) = 0; \tag{2.9}$$

  clearly, the smaller is the dimension $N$ of the reduced space compared to the dimension $\mathcal{N}$ of the full-order approximation space, the cheaper will be the solution of the reduced problem. In the same way as before, introducing the projectors $\Pi_N : X^{\mathcal{N}} \to X_N$ onto the reduced space $X_N$ and $\Pi^*_N : X^{\mathcal{N}\,\prime} \to X'_N$ onto its dual $X'_N$, we have

[4] However, this formulation, as well as Offline-Online procedures and *a posteriori* error estimation, are relevant also to other reduced approximations, such as the ones based on Taylor and Hermite spaces.

$$\Pi_N^*(\mathcal{L}^{\mathcal{N}}(\boldsymbol{\mu})\Pi_N^{-1} y_N(\boldsymbol{\mu}) - \mathcal{F}^{\mathcal{N}}(\boldsymbol{\mu})) = 0, \tag{2.10}$$

so that $\mathcal{L}_N(\boldsymbol{\mu}) = \Pi_N^* \mathcal{L}^{\mathcal{N}}(\boldsymbol{\mu})\Pi_N^{-1}$ and $\mathcal{F}_N(\boldsymbol{\mu}) = \Pi_N^* \mathcal{F}^{\mathcal{N}}(\boldsymbol{\mu})$. Equivalently, the weak formulation of the reduced PDE problem (2.9) reads as

$$A(y_N(\boldsymbol{\mu}), w_N; \boldsymbol{\mu}) = F(w_N; \boldsymbol{\mu}), \quad \forall\, w_N \in X_N. \tag{2.11}$$

The computational procedure used to find the reduced solution $y_N$ – ultimately, the coefficients in the linear combination of the basis functions – will be discussed in Sect. 2.2.4. We point out that the classical *optimality result* of the Galerkin method holds also in the reduced approximation, since in the energy norm[5] we have:

$$|||y^{\mathcal{N}}(\boldsymbol{\mu}) - y_N^{\mathcal{N}}(\boldsymbol{\mu})|||_{\boldsymbol{\mu}} \leq \inf_{w \in X_N^{\mathcal{N}}} |||y^{\mathcal{N}}(\boldsymbol{\mu}) - w|||_{\boldsymbol{\mu}}\ , \tag{2.12}$$

i.e. the Galerkin procedure automatically selects the *best* combination of snapshots ("best fit" approximation).

- **Offline/Online procedure**

  Under the *affinity assumptions* (1.7), the parametric setting enables to decouple the computational effort in two stages: a very extensive (parameter independent) pre-processing performed *Offline* once, consisting in the generation of the snapshots database, that prepares the way for subsequent very inexpensive calculations performed *Online* for each new input-output evaluation required. In the *real-time* or *many-query* contexts, where the goal is to achieve a very low *marginal cost* per input-output evaluation, we can accept an increased "Offline" cost – not tolerable for a single or few evaluations – in exchange for greatly decreased "Online" cost for each new/additional input-output evaluation. As pointed out in Sect. 2.4, whenever the affinity assumption is not naturally induced by the problem, it can be recovered through a further interpolation stage.
  Nevertheless, the chance to decouple *completely* the Online computational cost from the dimension $\mathcal{N}$ of the full-order problem is unavoidable whenever interested in *real-time* output or PDE evaluations or problems fitting a *many-query* context. In particular, we will detail the Offline/Online procedure in Sect. 2.2.4.

- **Error estimation procedure**

  Effective, sharp, inexpensive *a posteriori* error bounds for field variables

  $$\|y^{\mathcal{N}}(\boldsymbol{\mu}) - y_N(\boldsymbol{\mu})\|_X \leq \Delta_N(\boldsymbol{\mu}), \qquad \forall\, \boldsymbol{\mu} \in \mathcal{D}, \quad N = 1, \ldots, N_{\max},$$

  and outputs of interest

  $$|s^{\mathcal{N}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu})| \leq \Delta_N^s(\boldsymbol{\mu}), \qquad \forall\, \boldsymbol{\mu} \in \mathcal{D}, \quad N = 1, \ldots, N_{\max},$$

  are crucial for both the efficiency and the reliability of RB approximations. As regards *efficiency*, *a posteriori* error estimation permits us to control the error, as well as to minimize the computational effort by controlling the dimension of the RB space. Moreover, in the greedy algorithm (see Sect. 2.2.2) the application of error bounds – as surrogates for the actual error – allows a better parameter space exploration at greatly reduced cost.
  Concerning *reliability*, *a posteriori* error bounds allow a confident exploitation of the rapid predictive power of the RB approximation, making up for an error quantification for each new parameter value $\boldsymbol{\mu}$ in the Online stage and assuring the feasibility of the solution.

[5] Under the coercivity and the symmetry assumptions, the bilinear form $A(\cdot, \cdot; \boldsymbol{\mu})$ defines an *(energy)* scalar product given by $((w, v))_{\boldsymbol{\mu}} := A(w, v; \boldsymbol{\mu})\ \forall w, v \in X$; the induced *energy* norm is given by $|||w|||_{\boldsymbol{\mu}} = ((w, w))_{\boldsymbol{\mu}}^{1/2}$.

We remark that CRTs do not replace, but rather are built upon – and measured (as regards accuracy) relative to – the high-fidelity discretization technique, so that an algorithmic collaboration is pursued, expressed simply by means of a triangular inequality as follows:

$$\|y(\boldsymbol{\mu}) - y_N(\boldsymbol{\mu})\|_X \leq \|y(\boldsymbol{\mu}) - y^{\mathcal{N}}(\boldsymbol{\mu})\|_X + \|y^{\mathcal{N}}(\boldsymbol{\mu}) - y_N(\boldsymbol{\mu})\|_X \leq \mathcal{E}(h) + \Delta_N(\boldsymbol{\mu}), \quad \forall \boldsymbol{\mu} \in \mathcal{D}.$$

The ingredients presented in this section form a common paradigm shared by several CRTs; we will focus on certified reduced basis methods for parametrized PDEs, which exploit effective greedy algorithms for reduced spaces construction. Next sections are meant to give all the details concerning RB approximations for linear elliptic parametrized PDEs. The reader interested to the treatment of fluid dynamics equations can refer directly to Chapter 3.

## 2.2 Reduced Basis Methods for parametrized PDEs

Reduced Basis discretization is, in brief, a Galerkin projection on an $N$-dimensional approximation space that focuses on the parametrically induced manifold $\mathcal{M}^{\mathcal{N}}$. We restrict the attention to the Lagrange RB spaces, which are based on the use of "snapshot" FE solutions of the PDEs, corresponding to certain (properly selected) parameter values, as global approximation basis functions previously computed and stored; other possible approaches, such as Taylor [250] or Hermite spaces [153], take into account also partial derivatives of these basis solutions.

Initial ideas grew out of two related research topics dealing with linear/nonlinear structural analysis in the late 70's: the need for more effective many-query design evaluation and more efficient parameter continuation methods [8, 230, 231]. These first works were soon extended to (*i*) general finite-dimensional systems as well as certain classes of ODEs/PDEs [97,251,22,270], and (*ii*) a variety of different RB approximation spaces, such as Taylor and Lagrange expansions. The next decade saw further expansion into different applications and classes of equations, such as fluid dynamics and, more specifically, the incompressible Navier-Stokes equations [127, 152, 154, 242]. In these early methods the approximation spaces tended to be rather local and typically low-dimensional in parameter (often $p = 1$ physical parameters), due also to the absence of *a posteriori* error estimators and effective sampling procedures. Much current effort in the last ten years has thus been devoted in the RB framework to the development of these tools, which are crucial to guarantee reliability, accuracy and efficiency. The *a posteriori* error bounds are of course mandatory for rigorous certification of any particular RB Online output prediction. However, they also play an important role in effective (greedy) sampling procedures [313, 280]: they allow us to explore efficiently the parameter domain in search of most representative "snapshots," and to determine when we have *just enough* basis functions.

Concerning instead computational reduction and decoupling stratagems, early work on the RB method exploited (but not fully) the Offline-Online procedure. Consequently, the computational savings (relative to classical FE evaluation) were typically rather modest [230, 250, 251] *despite* the very small size of the RB linear systems. Much work has thus been devoted to *full* decoupling of the FE and RB spaces through Offline-Online procedures, above all concerning the efficient *a posteriori* error estimation; in this way, the complexity of the Offline stage depends on $\mathcal{N}$, while the complexity of the Online stage depends only on $N$ and $Q_\bullet$, so that we can reach the accuracy of a high-fidelity FE model but at the very low cost of a reduced-order model.
In the context of *affine parameter dependence*, the Offline-Online idea is quite self-apparent and has been naturally exploited [19, 242] and extended more recently in order to obtain efficient *a posteriori* error estimation. In the case of *nonaffine parameter dependence* the development of Offline-Online strategies is even more challenging and only in the last few years effective procedures have been studied and applied [21] to allow more complex parametrizations; clearly, Offline-Online procedures are unavoidable both in the real-time and the many-query contexts.

In the following we present the main features of the approximation of a simple, scalar ($\kappa = 1$) elliptic problem in two spatial dimensions ($d = 2$), which turn to be useful for the subsequent RB approximation of Stokes and Navier-Stokes equations, extensively discussed in Chapter 3. A survey of RB methods can be found in a review paper by Rozza, Huynh and Patera [280], as well as in the book by Patera and Rozza [237] and in the collection of material provided by the `augustine.mit.edu` portal [149] of Massachusetts Institute of Technology (MIT), held by the group of A.T. Patera.

### 2.2.1 Parametrized formulation

We briefly recall how the parametrized formulation (1.1)-(1.2) can be obtained, in the case of a simple linear ($\delta = 0$) elliptic problem; we shall first define an "original" problem (subscript $o$), posed over the *parameter-dependent* domain $\Omega_o = \Omega_o(\boldsymbol{\mu})$: given $\boldsymbol{\mu} \in \mathcal{D}$, evaluate

$$s_o(\boldsymbol{\mu}) = L_o(y_o(\boldsymbol{\mu}); \boldsymbol{\mu}) \,,$$

where $y_o(\boldsymbol{\mu}) \in X_o(\boldsymbol{\mu})$ satisfies

$$A_o(y_o(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = F_o(v; \boldsymbol{\mu}), \quad \forall v \in X_o(\boldsymbol{\mu}) \,, \tag{2.13}$$

where $X_o(\boldsymbol{\mu})$ is a suitable Hilbert space defined on $\Omega_o(\boldsymbol{\mu})$; definition of bilinear and linear forms follows the same notation of Sect. 1.2. Since the RB framework requires a reference ($\boldsymbol{\mu}$-independent) domain $\Omega$ in order to compare, and combine, FE solutions that would be otherwise computed on different domains and grids, the original $\Omega_o(\boldsymbol{\mu})$ has to be mapped to a reference domain[6], in order to get the parametrized "transformed" problem (2.2), which is the point of departure of RB approach. For the sake of the applications addressed, the construction of a parametric mapping related to geometrical properties can be managed following two different strategies (both detailed in Sect. 2.5), based on *(i)* patches of local *affine mappings* combined within a domain decomposition approach or *(ii)* global nonaffine mappings. In general, we can introduce a conforming domain decomposition of $\Omega_o(\boldsymbol{\mu})$,

$$\Omega_o(\boldsymbol{\mu}) = \bigcup_{k=1}^{K_{\text{dom}}} \Omega_o^k(\boldsymbol{\mu}), \tag{2.14}$$

consisting of mutually nonoverlapping open subdomains $\Omega_o^k(\boldsymbol{\mu})$, s.t. $\Omega_o^k(\boldsymbol{\mu}) \cap \Omega_o^{k'}(\boldsymbol{\mu}) = \emptyset$, $1 \le k < k' \le K_{\text{dom}}$. If related to geometrical properties used as input parameters (e.g. lengths, thicknesses, diameters or angles) the definition of parametric mappings can be done in a quite intuitive fashion[7]. In the following we will identify $\Omega^k = \Omega_o^k(\boldsymbol{\mu}_{\text{ref}})$, $1 \le k \le K_{\text{dom}}$, and denote (2.14) the "RB triangulation"; it will play an important role in the generation of the affine representation (1.7). Hence, original and reference subdomains must be linked via a mapping $T^k(\cdot; \boldsymbol{\mu}) : \Omega^k \to \Omega_o^k(\boldsymbol{\mu})$, $1 \le k \le K_{\text{dom}}$, such that

$$\Omega_o^k(\boldsymbol{\mu}) = T^k(\Omega^k; \boldsymbol{\mu}), \quad 1 \le k \le K_{\text{dom}}; \tag{2.15}$$

conditions ensuring the well posedness of the map and some recipes for its construction, will be detailed in Appendix A. In the case of a single global mapping ($K_{\text{dom}} = 1$) the superscript $k$ will be omitted. The more difficult case of nonaffine mappings will be discussed, presenting several options particularly suitable for the applications we are interested in.

[6] The reference domain can be seen as a particular instance of the parametrized ones, $\Omega = \Omega_o(\boldsymbol{\mu}_{ref})$, being $\boldsymbol{\mu}_{ref} \in \mathcal{D}$ a selected parameter value.

[7] These regions can represent different material properties, but they can also be used for algorithmic purposes to ensure well-behaved mappings.

It remains to introduce the class of *admissible operators* such that (upon mapping) the transformed problem satisfies the hypotheses introduced above – in particular, the affinity assumption (1.7)-(1.8). Nevertheless, when this assumption is not fulfilled by the resulting transformed formulation, it can be recovered through a suitable approximation – see Sect. 2.4 for details. We may now consider the associated bilinear forms $A_o : X_o(\boldsymbol{\mu}) \times X_o(\boldsymbol{\mu}) \to \mathbb{R}$

$$A_o(w, v; \boldsymbol{\mu}) = \sum_{k=1}^{K_{\text{dom}}} \int_{\Omega_o^k(\boldsymbol{\mu})} \left( \frac{\partial w}{\partial x_{o,i}} \nu_{ij}^{o,k}(\boldsymbol{\mu}) \frac{\partial v}{\partial x_{o,j}} + v \chi_i^{o,k}(\boldsymbol{\mu}) \frac{\partial w}{\partial x_{o,i}} + w \eta_i^{o,k}(\boldsymbol{\mu}) \frac{\partial v}{\partial x_{o,i}} + wv \right) d\Omega_o,$$

where $\boldsymbol{\nu}^{o,k} : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}^{2\times 2}$, $1 \le k \le K_{\text{dom}}$, are parametrized (symmetric positive definite) conductivity/diffusivity tensors, while $\boldsymbol{\chi}^{o,k} : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}^2$, $\boldsymbol{\eta}^{o,k} : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}^2$, $1 \le k \le K_{\text{dom}}$, are two parametrized vectors, representing transport/convective terms.
Similarly, we require that $F_o(\cdot; \boldsymbol{\mu}) : X_o(\boldsymbol{\mu}) \to \mathbb{R}$ and $L_o(\cdot; \boldsymbol{\mu}) : X_o(\boldsymbol{\mu}) \to \mathbb{R}$ are given by

$$F_o(v; \boldsymbol{\mu}) = \sum_{k=1}^{K_{\text{dom}}} \int_{\Omega_o^k(\boldsymbol{\mu})} f^{o,k}(\boldsymbol{\mu}) v \, d\Omega_o, \qquad L_o(v; \boldsymbol{\mu}) = \sum_{k=1}^{K_{\text{dom}}} \int_{\Omega_o^k(\boldsymbol{\mu})} l^{o,k}(\boldsymbol{\mu}) v \, d\Omega_o \tag{2.16}$$

where $f^{o,k} : \mathcal{D} \to \mathbb{R}$ and $l^{o,k} : \mathcal{D} \to \mathbb{R}$, for $1 \le l \le K_{\text{dom}}$, are prescribed coefficients [275]. By identifying $y(\boldsymbol{\mu}) = y_o(\boldsymbol{\mu}) \circ T(\cdot; \boldsymbol{\mu})$ and tracing $A_o(w, v; \boldsymbol{\mu})$ back on the reference domain $\Omega$, it follows that the transformed bilinear form $A(\cdot, \cdot; \boldsymbol{\mu}) : X \times X \to \mathbb{R}$ can be expressed as

$$A(w, v; \boldsymbol{\mu}) = \sum_{k=1}^{K_{\text{dom}}} \int_{\Omega^k} \left( \frac{\partial w}{\partial x_i} \nu_{ij}^k(\boldsymbol{\mu}) \frac{\partial v}{\partial x_j} + v \chi_i^k(\boldsymbol{\mu}) \frac{\partial w}{\partial x_i} + w \eta_i^k(\boldsymbol{\mu}) \frac{\partial v}{\partial x_i} + wv |J_T^k(\boldsymbol{\mu})| \right) d\Omega, \tag{2.17}$$

where $\boldsymbol{\nu}^k(\mathbf{x}; \cdot) : \mathcal{D} \to \mathbb{R}^{2\times 2}$, $1 \le k \le K_{\text{dom}}$, are parametrized tensors given by

$$\boldsymbol{\nu}^k(\mathbf{x}; \boldsymbol{\mu}) = (J_T^k(\mathbf{x}; \boldsymbol{\mu}))^{-T} \boldsymbol{\nu}^{o,k}(\boldsymbol{\mu}) (J_T^k(\mathbf{x}; \boldsymbol{\mu}))^{-1} |J_T^k(\mathbf{x}; \boldsymbol{\mu})| \tag{2.18}$$

where $J_T^k(\mathbf{x}; \boldsymbol{\mu}) : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}^{2\times 2}$ is the Jacobian matrix of the map $T^k(\cdot; \boldsymbol{\mu})$, defined as

$$(J_T^k(\mathbf{x}; \boldsymbol{\mu}))_{ij} = \frac{\partial (T^k)_i}{\partial x_j}(\mathbf{x}; \boldsymbol{\mu}) \tag{2.19}$$

and $|J_T^k(\mathbf{x}; \boldsymbol{\mu})| : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}$ its determinant. In the same way, $\boldsymbol{\chi}^k(\mathbf{x}; \boldsymbol{\mu}) : \mathcal{D} \to \mathbb{R}^2$, $\boldsymbol{\eta}^k(\mathbf{x}; \boldsymbol{\mu}) : \mathcal{D} \to \mathbb{R}^2$, $1 \le k \le K_{\text{dom}}$, are parametrized vectors, given by

$$\begin{aligned} \boldsymbol{\chi}^k(\mathbf{x}; \boldsymbol{\mu}) &= (J_T^k(\mathbf{x}; \boldsymbol{\mu}))^{-T} \boldsymbol{\chi}^{o,k}(\boldsymbol{\mu}) |J_T^k(\mathbf{x}; \boldsymbol{\mu})|, \\ \boldsymbol{\eta}^k(\mathbf{x}; \boldsymbol{\mu}) &= (J_T^k(\mathbf{x}; \boldsymbol{\mu}))^{-T} \boldsymbol{\eta}^{o,k}(\boldsymbol{\mu}) |J_T^k(\mathbf{x}; \boldsymbol{\mu})|. \end{aligned} \tag{2.20}$$

The transformed linear forms $F(\cdot; \boldsymbol{\mu}) : X \to \mathbb{R}$ and $L(\cdot; \boldsymbol{\mu}) : X \to \mathbb{R}$ can be expressed similarly as

$$F(v; \boldsymbol{\mu}) = \sum_{k=1}^{K_{\text{dom}}} \int_{\Omega^k} f^k(\mathbf{x}; \boldsymbol{\mu}) v \, d\Omega, \qquad L(v; \boldsymbol{\mu}) = \sum_{k=1}^{K_{\text{dom}}} \int_{\Omega^k} l^k(\mathbf{x}; \boldsymbol{\mu}) v \, d\Omega,$$

where $f^k(\mathbf{x}; \cdot) : \mathcal{D} \to \mathbb{R}$ and $l^k(\mathbf{x}; \cdot) : \mathcal{D} \to \mathbb{R}$ are given, for $1 \le k \le K_{\text{dom}}$, by

$$f^k(\mathbf{x}; \boldsymbol{\mu}) = f^{o,k}(\boldsymbol{\mu}) |J_T^k(\mathbf{x}; \boldsymbol{\mu})|, \qquad f^k(\mathbf{x}; \boldsymbol{\mu}) = f^{o,k}(\boldsymbol{\mu}) |J_T^k(\mathbf{x}; \boldsymbol{\mu})|.$$

Hence, the original problem has been reformulated on a reference configuration, resulting in a parametrized problem where the effect of geometry variations is traced back onto its parametrized transformation tensors. *A priori*, parametrized tensors depend both on the parameter $\boldsymbol{\mu}$ and the spatial variables $\mathbf{x} \in \Omega$; however, in the affine case they depend just on the parameter components,

allowing to write the affine formulation (1.7) by simply expanding the expression (2.17) in terms of the subdomains $\Omega^l$ and the different entries of the tensors. This results, for example, in

$$A(w,v;\boldsymbol{\mu}) = \nu_{11}^1(\boldsymbol{\mu}) \int_{\Omega^1} \frac{\partial w}{\partial x_1}\frac{\partial v}{\partial x_1} + \nu_{12}^1(\boldsymbol{\mu}) \int_{\Omega^1} \frac{\partial w}{\partial x_1}\frac{\partial v}{\partial x_2} + \cdots \tag{2.21}$$

The affine representation is now clear: for each term in (2.21), the ($\boldsymbol{\mu}$-independent) integral represents $A^q(w,v)$, while the ($\boldsymbol{\mu}$-dependent) prefactor represents $\Theta^q(\boldsymbol{\mu})$; the linear forms $F(\cdot;\boldsymbol{\mu})$ and $L(\cdot;\boldsymbol{\mu})$ admit a similar treatment. The same transformation can be operated if the problem accounts for nonlinear terms and/or the output is quadratic; see for instance Chapter 3 for the case of a quadratic output related to a Stokes problem or for the case of Navier-Stokes operator.

### 2.2.2 Reduced spaces construction: Greedy algorithms

In order to sample the parameter space and construct the master set (2.7) – and ultimately the reduced basis space – several strategies are available; we rely on the so-called *greedy* algorithms, based on the idea that a proper selection is made by choosing at each step the locally optimal element. Together with the Offline/Online computational strategy, greedy procedures are nowadays the kernel of RB methods for parametrized PDEs [280, 237]. In particular, the current procedure for constructing reduced subspaces in parametrized PDEs like (2.2) is based on the following greedy procedure. Other available strategies will be briefly recalled in Sect. 2.2.3.

In an abstract setting, given a compact subspace $\mathcal{K}$ of a Hilbert space $\mathcal{X}$, the goal is to seek for a set of functions $\{\varphi_0, \varphi_1, \ldots, \varphi_{N-1}\}$ such that each element $\varphi \in \mathcal{X}$ can be well approximated by the elements of the subspace $\mathcal{K}_N = \text{span}\{\varphi_0, \ldots, \varphi_{N-1}\}$. Starting from a first element $\varphi_0$ such that $\|\varphi_0\|_{\mathcal{X}} = \max_{\varphi\in\mathcal{K}} \|\varphi\|_{\mathcal{X}}$, at the $N$-th step a greedy algorithm selects

$$\varphi_N = \arg\max_{\varphi\in\mathcal{K}} \|\varphi - \Pi_N\varphi\|_{\mathcal{X}},$$

where $\Pi_N$ is the projector onto $\mathcal{K}_N$, i.e. $\varphi_N$ is the worst case element, which maximizes the error in approximating the subspace $\mathcal{K}$ using the elements of $\mathcal{K}_N$. A more feasible variant of this algorithm – called *weak greedy algorithm* in [36] – replaces the true error $\|\varphi - \Pi_N\varphi\|_{\mathcal{X}}$ by a surrogate $\eta_N(\varphi)$ – in our case, the *a posteriori* error bound $\Delta_N$ – satisfying

$$c_\eta \eta_N(\varphi) \le \|\varphi - \Pi_N^{RB}\varphi\|_{\mathcal{X}} \le C_\eta \eta_N(\varphi), \qquad f \in \mathcal{X}\,;$$

in this way, $\varphi_N = \arg\max_{\varphi\in\mathcal{K}} \eta_N(\varphi)$ can be more easily computed, provided that the surrogate error $\eta_N(\varphi)$ is cheap to evaluate. Following this setting, we now detail the *greedy* algorithm exploited for the construction of our (Lagrangian) RB spaces. First of all, let us denote by $\Xi$ a finite sample of points in $\mathcal{D}$, which shall serve as surrogates for $\mathcal{D}$ in the calculation of errors (and error bounds) over the parameter domain. For a generic element $z\colon \mathcal{D} \to X^{\mathcal{N}}$, we denote

$$\|z\|_{L^\infty(\Xi;X)} \equiv \operatorname*{ess\,sup}_{\boldsymbol{\mu}\in\Xi} \|z(\boldsymbol{\mu})\|_X \ .$$

Moreover, we denote the particular samples which shall serve to select the RB space – or "train" the RB approximation – by $\Xi_{\text{train}}$. The cardinality of $\Xi_{\text{train}}$ will be denoted $|\Xi_{\text{train}}| = n_{\text{train}}$. We note that although the "test" samples $\Xi$ serve primarily to understand and assess the quality of the RB approximation and *a posteriori* error estimators, the "train" samples $\Xi_{\text{train}}$ serve to *generate* the RB approximation. The choice of $n_{\text{train}}$ and $\Xi_{\text{train}}$ thus have important Offline and Online computational implications. Roughly, at iteration $N$ of the greedy algorithm, we would append to the master set (2.7) that parameter value $\boldsymbol{\mu}^{N+1}$ which maximizes the true error $\|y(\boldsymbol{\mu}) - y^{\mathcal{N}}(\boldsymbol{\mu})\|_{L^\infty(\Xi;X)}$. In actual practice, we never compute this error, rather we exploit

our *a posteriori error bound* to select that parameter value $\boldsymbol{\mu}^{N+1}$ which maximizes the error bound $\Delta_N(\boldsymbol{\mu})$. Thus, denoting by $\epsilon^*_{\text{tol}}$ a chosen tolerance for the stopping criterium, the greedy algorithm can be implemented as follows:

$S_1 = \{\boldsymbol{\mu}^1\}$; `compute` $y^{\mathcal{N}}(\boldsymbol{\mu}^1)$;
$X_1^{\text{greedy}} = \text{span}\{y^{\mathcal{N}}(\boldsymbol{\mu}^1)\}$;
`for` $N = 2 : N_{\max}$
  $\boldsymbol{\mu}^N = \arg\max_{\boldsymbol{\mu}\in\Xi_{\text{train}}} \Delta_{N-1}(\boldsymbol{\mu})$;
  $\varepsilon_{N-1} = \Delta_{N-1}(\boldsymbol{\mu}^N)$;
  `if` $\varepsilon_{N-1} \leq \epsilon^*_{\text{tol}}$
    $N_{\max} = N - 1$;
  `end`;
  `compute` $y^{\mathcal{N}}(\boldsymbol{\mu}^N)$;
  $S_N = S_{N-1} \cup \{\boldsymbol{\mu}^N\}$;
  $X_N^{\text{greedy}} = X_{N-1}^{\text{greedy}} \cup \text{span}\{y^{\mathcal{N}}(\boldsymbol{\mu}^N)\}$;
`end`.

By means of this procedure, at iteration $N$ our greedy algorithm appends to the *retained* snapshots space $X_N^{\text{greedy}}$ that particular candidate snapshot – over all candidate snapshots $y^{\mathcal{N}}(\boldsymbol{\mu})$, $\boldsymbol{\mu} \in \Xi_{\text{train}}$ – which is (predicted by the *a posteriori* error bound to be the) least well approximated by (the RB prediction associated to) $X_{N-1}^{\text{greedy}}$. In this way, the greedy algorithm provides hierarchical spaces and in the strong $L^\infty(\Xi_{\text{train}}; X)$ norm in parameter. Not only, the cost of this formulation enables a very exhaustive search – large $n_{\text{train}}$ – with corresponding high quality approximation spaces. Clearly, the accuracy and cost of the *a posteriori* error estimator $\Delta_N(\boldsymbol{\mu})$ are crucial to the success of the algorithm. We will provide in the following chapters a computational evidence that this greedy selection process generates spaces which are, if not optimal, at least very good – a theoretical proof of this result, based on the exponential convergence of the Kolmogorov $N$-width, can be found in [36]. Recent enhanced greedy algorithms using either a saturation assumption (so that only error estimators are computed for those points in $\Xi_{\text{train}}$ with a large enough predicted error) or an adaptive enrichment of $\Xi_{\text{train}}$ have been proposed in [141]. From now on we refer to RB subspaces built through a greedy sampling if the superscript *greedy* is omitted.

### 2.2.3 Reduced spaces construction: other options

An alternative technique used for the construction of reduced spaces in dynamical or parametrized systems is the proper orthogonal decomposition (POD). This method is very popular also in other fields, such as multivariate statistical analysis (where it is called *principal component analysis*) or theory of stochastic processes (*Karhunen-Loève decomposition*). The first applications of POD were concerned with the analysis of turbulent flows and date back to the early '90s [15,33]; more recently, applications in computational fluid dynamics can be found, for instance, in [144,175,177,197,161], as well as in [63,54,134] for parametrized flows.
POD techniques reduce the dimensionality of a system by transforming the original variables onto a new set of uncorrelated variables (which are called POD modes, or principal components) such that the first few modes retain most of the *energy* present in all of the original variables. This allows to obtain a reduced, modal representation through a spectral decomposition which requires basic matrix computations (a *singular value decomposition*) also for complex nonlinear problems. Nevertheless, *(i) a posteriori* error bounds are in general not available, making the choice of the reduction size and the quality assessment of the reduced solution sometimes critical. Moreover, *(ii)* space reduction through spectral decomposition entails larger computational costs than RB's.

We shortly review the main features of the POD in the context of parametrized PDEs. Given a finite sample $\Xi$ of points in $\mathcal{D}$, a train sample $\Xi_{\text{train}}$ (which shall serve to select the POD space), for a generic element $z\colon \mathcal{D} \to X^{\mathcal{N}}$, we denote

$$\|z\|_{L^2(\Xi;X)} \equiv \left(|\Xi|^{-1} \sum_{\boldsymbol{\mu}\in\Xi} \|z(\boldsymbol{\mu})\|_X^2\right)^{1/2}.$$

The POD method seeks an $N$-dimensional subspace $X_N^{POD} \subset X^{\mathcal{N}}$ approximating the data in an optimal *least-squares* sense; thus, we seek an orthogonal projector $\Pi_N^{POD} : X^{\mathcal{N}} \to X_N^{POD}$, of prescribed rank $N$ minimizing the *least-square* distance as follows:

$$X_N^{\text{POD}} = \underset{X_N \subset \text{span}\{y^{\mathcal{N}}(\boldsymbol{\mu}) \,|\, \boldsymbol{\mu}\in\Xi_{\text{train}}\}}{\arg\inf} \quad \|y^{\mathcal{N}}(\boldsymbol{\mu}) - \Pi_N^{POD}\, y^{\mathcal{N}}(\boldsymbol{\mu})\|_{L^2(\Xi_{\text{train}};X)}\,. \tag{2.22}$$

Following the so-called *method of snapshots*, introduced by Sirovich [302], we compute the $n_{\text{train}}$ full-order approximations $\{y(\boldsymbol{\mu}^m)\}_{m=1}^{n_{\text{train}}}$ corresponding to $\boldsymbol{\mu}^1, \dots, \boldsymbol{\mu}^{n_{\text{train}}}$, the mean

$$\bar{y} = \frac{1}{n_{\text{train}}} \sum_{j=1}^{n_{\text{train}}} y(\boldsymbol{\mu}^j)$$

and the *correlation* matrix $\mathbb{C} \in \mathbb{R}^{n_{\text{train}}\times n_{\text{train}}}$ whose components are

$$\mathbb{C}_{ij} = \frac{1}{n_{\text{train}}} \sum_{m=1}^{n_{\text{train}}} \big(y(\boldsymbol{\mu}^i) - \bar{y}, y(\boldsymbol{\mu}^j) - \bar{y}\big)_X\,, \qquad 1 \le i,j \le n_{\text{train}}.$$

Then, we compute the eigenvalues $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_{n_{\text{train}}} \ge 0$ and the eigenvectors of the correlation matrix, which solve: $\mathbb{C}\boldsymbol{\psi}_k = \lambda_k \boldsymbol{\psi}_k$, for $k = 1, \dots, n_{\text{train}}$. The central result of POD states that the *optimal* subspace $X_N^{POD}$ of dimension $N$ minimizing (2.22) is given by

$$X_N^{POD} = \text{span}\{\zeta_n,\ 1 \le n \le N\}, \quad 1 \le N \le N_{\max},$$

where the POD basis functions are defined as

$$\zeta_k = \frac{\tilde{\zeta}_k}{\|\tilde{\zeta}_k\|_X}, \qquad \tilde{\zeta}_k = \sum_{m=1}^{n_{\text{train}}} \psi_{k,m}(y(\boldsymbol{\mu}^m) - \bar{y}), \qquad 1 \le k \le n_{\text{train}},$$

where $\psi_{k,m} = (\boldsymbol{\psi}_k)_m$ is the $m$-th component of the $k$-th eigenvector. In particular, the basis functions $\{\zeta_k\}_{k=1}^{n_{\text{train}}}$ are orthonormal, i.e. they are such that $(\zeta_n, \zeta_m)_X = \delta_{n\,m}$, for $1 \le n,m \le n_{\text{train}}$. In particular, $N_{\max}$ is chosen as the smallest $N$ such that

$$\varepsilon_N^{POD} = \left(\sum_{k=N+1}^{n_{\text{train}}} \lambda_k\right)^{1/2} \le \epsilon_{\text{tol}}^*,$$

i.e. the energy $\varepsilon_N^{POD}$ retained by the last $(n_{\text{train}} - N_{\max})$ modes is negligible. Concerning the convergence properties of greedy and POD spaces, the two algorithms perform in a similar way if measured in the *native* norm over $\Xi_{\text{train}}$ which defines the respective objective function – $L^2(\Xi_{\text{train}})$ for POD, $L^\infty(\Xi_{\text{train}})$ for greedy. From the viewpoint of Offline computational cost, a POD approach is much more expensive than the Greedy approach: in the latter, we only need to compute the $N$ – typically very few – FE *retained* snapshots; in the POD approach, we must compute all $n_{\text{train}}$ – typically/desirably very many – FE *candidate* snapshots, as well as the solution of an eigenproblem for the correlation matrix $\mathbb{C} \in \mathbb{R}^{n_{\text{train}}\times n_{\text{train}}}$.
Not surprisingly, the POD has found most application in the reduction of dynamical systems – the first (and most used) application of this strategy, see e.g. [245,315] – described in the time-domain (which has single dimension if seen as a parameter space) in which $n_{\text{train}}$ typically remains quite

small. Furthermore, the correlations between the solution at different times can be effectively captured by the global nature of the POD optimization; in this context, the greedy approach is not as successful. For this reason, in the RB approximation of parabolic problems, an efficient combined POD (in time) - greedy (in parameter space) approach has been proposed [135, 228] and is currently used [227, 120].
Among many other methods for generating reduced spaces, we mention the Centroidal Voronoi Tessellation [53, 54] and the Proper Generalized Decomposition [62, 232, 86], which are indeed quite close to POD.

### 2.2.4 Offline/Online computational strategy

We now consider the discrete equations associated with the Galerkin approximation (2.11) under the compliance assumption (i.e. if $A(\cdot,\cdot;\boldsymbol{\mu})$ is a symmetric bilinear form and $L(\cdot;\boldsymbol{\mu}) \equiv F(\cdot;\boldsymbol{\mu})$): given $\boldsymbol{\mu} \in \mathcal{D}$, evaluate

$$s_N(\boldsymbol{\mu}) = F(y_N(\boldsymbol{\mu});\boldsymbol{\mu}), \tag{2.23}$$

where $y_N(\boldsymbol{\mu}) \in X_N \subset X^{\mathcal{N}}$ satisfies

$$A(y_N(\boldsymbol{\mu}), w_N;\boldsymbol{\mu}) = F(w_N;\boldsymbol{\mu}), \quad \forall\, w_N \in X_N. \tag{2.24}$$

First of all, we apply the Gram-Schmidt orthogonalization procedure with respect to the $(\cdot,\cdot)_X$ inner product to snapshots $y^{\mathcal{N}}(\boldsymbol{\mu}^n)$, $1 \le n \le N_{\max}$, to obtain mutually $(\cdot,\cdot)_X$–orthonormal basis functions $\zeta_n$, $1 \le n \le N_{\max}$. Then, the RB solution can be expressed as a linear combination of the basis functions:

$$y_N(\boldsymbol{\mu}) = \sum_{m=1}^{N} y_{N\,m}(\boldsymbol{\mu})\zeta_m; \tag{2.25}$$

by taking $w = \zeta_n$, $1 \le n \le N$, into (2.24) and using (2.25), we obtain the RB "stiffness" equations

$$\sum_{m=1}^{N} A(\zeta_m, \zeta_n;\boldsymbol{\mu})\, y_{N\,m}(\boldsymbol{\mu}) = F(\zeta_n;\boldsymbol{\mu}), \quad 1 \le m, n \le N\ , \tag{2.26}$$

for the RB coefficients $y_{N\,m}(\boldsymbol{\mu})$; we can subsequently evaluate the linear RB output as

$$s_N(\boldsymbol{\mu}) = \sum_{m=1}^{N} y_{N\,m}(\boldsymbol{\mu}) F(\zeta_m;\boldsymbol{\mu})\ . \tag{2.27}$$

The system (2.26) is nominally of small size: a set of $N$ linear algebraic equations in $N$ unknowns. However, the formation of the stiffness matrix, and indeed the load vector, involves entities $\zeta_n$, $1 \le n \le N$, associated with our $\mathcal{N}$-dimensional FE approximation space. Fortunately, we can appeal to affine parameter dependence to construct very efficient Offline-Online procedures.
In particular, system 2.26 can be expressed, thanks to the affinity assumption (1.7), as

$$\sum_{m=1}^{N} \left( \sum_{q=1}^{Q_A} \Theta_A^q(\boldsymbol{\mu})\, A^q(\zeta_m, \zeta_n) \right) y_{N\,m}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_F} \Theta_F^q(\boldsymbol{\mu}) F^q(\zeta_n),$$

for $1 \le n \le N$. The equivalent matrix form is

$$\left( \sum_{q=1}^{Q_a} \Theta_A^q(\boldsymbol{\mu}) \mathbb{A}_N^q \right) \mathbf{y}_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_F} \Theta_F^q(\boldsymbol{\mu}) \mathbf{f}_N^q, \tag{2.28}$$

where $(\mathbf{y}_N(\boldsymbol{\mu}))_m = y_{N\,m}(\boldsymbol{\mu})$ and

$$(\mathbb{A}_N^q)_{mn} = A^q(\zeta_m, \zeta_n), \qquad (\mathbf{f}_N^q)_n = F^q(\zeta_n), \qquad 1 \le m, n \le N_{\max}.$$

Since each basis function $\zeta_n$ belongs to the FE space $X^{\mathcal{N}}$, it can be written as a linear combination of the FE basis functions $\{\phi_i^{\mathcal{N}}\}_{i=1}^{\mathcal{N}}$:

$$\zeta_n = \sum_{i=1}^{\mathcal{N}} \zeta_n^i \phi_i^{\mathcal{N}}, \quad 1 \le n \le N_{\max};$$

therefore, the RB "stiffness" matrix can be assembled once the corresponding FE "stiffness" matrix has been computed. In fact, we have:

$$A^q(\zeta_m, \zeta_n) = \sum_{i=1}^{\mathcal{N}} \sum_{j=1}^{\mathcal{N}} \zeta_n^i A^q(\phi_i^{\mathcal{N}}, \phi_j^{\mathcal{N}}) \zeta_n^j, \qquad 1 \le m, n \le N, \qquad 1 \le q \le Q_A, \tag{2.29}$$

$$F^q(\zeta_n) = \sum_{i=1}^{\mathcal{N}} \zeta_n^i F^q(\phi_i^{\mathcal{N}}), \qquad 1 \le n \le N, \qquad 1 \le q \le Q_F; \tag{2.30}$$

then, we can introduce *basis* matrices $\mathbb{Z}_N \in \mathbb{R}^{\mathcal{N} \times N}$, $1 \le N \le N_{\max}$:

$$(\mathbb{Z}_N)_{jn} = \zeta_j^n, \qquad \text{i.e.} \qquad \mathbb{Z}_N = [\ \boldsymbol{\zeta}_1 \mid \ldots \mid \boldsymbol{\zeta}_N\ ] \in \mathbb{R}^{\mathcal{N} \times N}, \quad 1 \le N \le N_{\max},$$

so that the $n^{\text{th}}$ column of $\mathbb{Z}_{N_{\max}}$ contains the vector $\boldsymbol{\zeta}_n \in \mathbb{R}^{\mathcal{N}}$ of FE basis coefficients associated with the $n^{\text{th}}$ RB basis function. In this way, (2.29)-(2.30) can be rewritten as follows:

$$\mathbb{A}_N^q = \mathbb{Z}^T \mathbb{A}_{\mathcal{N}}^q \mathbb{Z}, \qquad \mathbf{f}_N^q = \mathbb{Z}^T \mathbf{f}_{\mathcal{N}}^q, \tag{2.31}$$

where

$$(\mathbb{A}_{\mathcal{N}}^q)_{ij} = A^q(\phi_j, \phi_i), \qquad (\mathbf{f}_{\mathcal{N}}^q)_i = F^q(\phi_i) \tag{2.32}$$

is the corresponding FE "stiffness" matrices and right-hand-sides. Reduced basis structures can thus be computed, starting from the full-order discretization structures, through a pre- and post-multiplication by rectangular basis matrices; see Fig. 2.1 from a schematic representation.

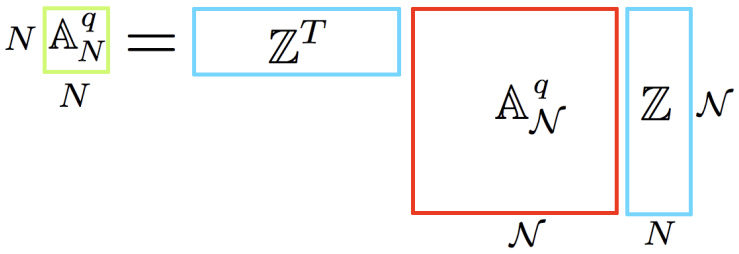



Figure 2.1: Schematic representation of the reduced basis projection.

Although being dense (rather than sparse as in the FE case), since the RB basis elements have global support, the system matrix is very small because its size $N$ is presumably small – and, of course, independent of the FE space dimension $\mathcal{N}$ (see e.g. chapter 3 of [237]). However, we point out that the computational savings provided by this projection operation would be rather modest – despite the very small size of the RB system –if system (2.28) were generated

by direct appeal to the full-order truth approximation structures (2.32). In this way, thanks to the *Offline/Online* strategy, based on the affinity assumption, computation entails an expensive $\boldsymbol{\mu}$-independent Offline stage performed only once and an Online stage for any chosen parameter value $\boldsymbol{\mu} \in \mathcal{D}$, whose complexity depends only on $N$ and no more on the dimension of the full-order system $\mathcal{N}$.

We close this section by pointing out the connection between the RB approximation and the FE approximation from an algebraic standpoint. Let us denote by $\mathbf{y}^{\mathcal{N}} \in \mathbb{R}^{\mathcal{N}}$ and $\mathbf{y}_N \in \mathbb{R}^N$ the vectors of degrees of freedom of the FEM and of the RB approximation, associated to the functions $y^{\mathcal{N}} \in X^{\mathcal{N}}$ and $y_N \in X_N$, respectively. Moreover, let $\mathbb{A}_{\mathcal{N}}(\boldsymbol{\mu})$ and $\mathbb{A}_N(\boldsymbol{\mu})$ be the matrices corresponding to the FEM and to the RB discretization, respectively, for any given parameter value $\boldsymbol{\mu} \in \mathcal{D}$. In this way, the reduced linear system (2.28) can be rewritten as

$$\mathbb{A}_N(\boldsymbol{\mu})\mathbf{y}_N(\boldsymbol{\mu}) = \mathbf{f}_N(\boldsymbol{\mu}), \tag{2.33}$$

whereas the full-order FEM linear system would read

$$\mathbb{A}_{\mathcal{N}}(\boldsymbol{\mu})\mathbf{y}^{\mathcal{N}}(\boldsymbol{\mu}) = \mathbf{f}_{\mathcal{N}}(\boldsymbol{\mu}). \tag{2.34}$$

In order to make a connection between the RB and the FE linear systems, we can express (for the sake of simplicity, without considering a basis orthonormalization)

$$\mathbf{y}^{\mathcal{N}}(\boldsymbol{\mu}) = \mathbb{Z}_N(\mathbf{y}_N(\boldsymbol{\mu}) + \boldsymbol{\delta}_N(\boldsymbol{\mu})), \tag{2.35}$$

where the error term $\boldsymbol{\delta}_N \in \mathbb{R}^N$ accounts for the fact that $\mathbb{Z}_N\mathbf{y}_N$ is not the exact solution of the full-order system and *a priori* is not vanishing. By plugging the expression (2.35) into (2.34) and multiplying the system by $\mathbb{Z}_N^T$, we obtain $\mathbb{Z}_N^T\mathbb{A}_{\mathcal{N}}(\boldsymbol{\mu})\mathbb{Z}_N(\mathbf{y}_N(\boldsymbol{\mu}) + \boldsymbol{\delta}_N(\boldsymbol{\mu})) = \mathbb{Z}_N^T\mathbf{f}_{\mathcal{N}}(\boldsymbol{\mu})$. Thanks to (2.31) and to (2.33), we thus find that

$$\mathbb{A}_N(\boldsymbol{\mu})\boldsymbol{\delta}_N(\boldsymbol{\mu}) = \mathbf{0},$$

i.e. the algebraic counterpart of the *Galerkin orthogonality* property, fulfilled by the RB approximation, that is: $a(y^{\mathcal{N}}(\boldsymbol{\mu}) - y_N(\boldsymbol{\mu}), w_N) = 0$, for any $w_N \in X_N$. On the other hand, setting $\boldsymbol{\delta}^{\mathcal{N}}(\boldsymbol{\mu}) = \mathbb{Z}_N\boldsymbol{\delta}_N(\boldsymbol{\mu})$, we have

$$\mathbf{y}^{\mathcal{N}}(\boldsymbol{\mu}) = \mathbb{Z}_N\mathbf{y}_N(\boldsymbol{\mu}) + \boldsymbol{\delta}^{\mathcal{N}}(\boldsymbol{\mu}), \tag{2.36}$$

where now the error term $\boldsymbol{\delta}^{\mathcal{N}} \in \mathbb{R}^{\mathcal{N}}$ is represented in the reduced vector space $\mathbb{R}^{\mathcal{N}}$. Plugging (2.36) into (2.34), we end up with

$$\mathbb{A}_{\mathcal{N}}(\boldsymbol{\mu})\boldsymbol{\delta}^{\mathcal{N}}(\boldsymbol{\mu}) = \mathbf{f}_{\mathcal{N}}(\boldsymbol{\mu}) - \mathbb{A}_{\mathcal{N}}(\boldsymbol{\mu})\mathbb{Z}_N\mathbf{y}_N(\boldsymbol{\mu}),$$

i.e. the algebraic counterpart of the error residual relationship (2.37), which constitutes the basis of *a posteriori* error estimation procedures, discussed in the forthcoming section.

## 2.3 A posteriori error estimation

The motivations for error estimation – *efficiency* and *reliability* – introduced in Sect. 2.1 lead to some necessary requirements on the construction and evaluation of the error bounds. First, the error bounds must be *rigorous*, i.e. valid for all $N$ and for all parameter values in the parameter domain $\mathcal{D}$ (non-rigorous error "indicators" may suffice for adaptivity during basis assembling, but not for reliability of the Online evaluation); reasonably *sharp*, and very *efficient*, relying – such as the solution and output approximations – on a suitable Offline/Online procedure, so that the Online evaluation of the RB error bounds results independent of $\mathcal{N}$.

Let us now consider *a posteriori* error bounds for the field variable $y_N(\boldsymbol{\mu})$ and the output $s_N(\boldsymbol{\mu})$ in the elliptic coercive (compliant) case (2.23)-(2.24); extensions to noncompliant, noncoercive and nonlinear problems will be detailed in the following. Here we introduce two basic ingredients of our error bounds – the error residual relationship and lower bounds of stability (here, coercivity) factors – which will be exploited also in the error estimation for more general classes of problems.

The central equation in *a posteriori* theory is the error residual relationship. In particular, it follows from the problem statements for $y^{\mathcal{N}}(\boldsymbol{\mu})$, (2.6), and $y_N(\boldsymbol{\mu})$, (2.11), that the error $e(\boldsymbol{\mu}) := y^{\mathcal{N}}(\boldsymbol{\mu}) - y_N(\boldsymbol{\mu}) \in X^{\mathcal{N}}$ satisfies

$$A(e(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = r(v; \boldsymbol{\mu}), \quad \forall\, v \in X^{\mathcal{N}} \ . \tag{2.37}$$

Here $r(v; \boldsymbol{\mu}) \in (X^{\mathcal{N}})'$ (the dual space to $X^{\mathcal{N}}$) is the residual,

$$r(v; \boldsymbol{\mu}) := F(v; \boldsymbol{\mu}) - A(y_N(\boldsymbol{\mu}), v; \boldsymbol{\mu}), \quad \forall\, v \in X^{\mathcal{N}}. \tag{2.38}$$

Indeed, (2.37) directly follows from the definition (2.38), $F(v; \boldsymbol{\mu}) = A(y^{\mathcal{N}}(\boldsymbol{\mu}), v; \boldsymbol{\mu})$, $\forall\, v \in X^{\mathcal{N}}$, the bilinearity of $A(\cdot, \cdot; \boldsymbol{\mu})$, and the definition of $e(\boldsymbol{\mu})$. It shall prove convenient [237] to introduce the Riesz representation of $r(v; \boldsymbol{\mu})$, $\hat{e}(\boldsymbol{\mu}) \in X^{\mathcal{N}}$, satisfying

$$(\hat{e}(\boldsymbol{\mu}), v)_X = r(v; \boldsymbol{\mu}), \quad \forall\, v \in X^{\mathcal{N}} \ . \tag{2.39}$$

This allows us to write the error residual equation (2.37) as

$$A(e(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = (\hat{e}(\boldsymbol{\mu}), v)_X, \quad \forall\, v \in X^{\mathcal{N}} \tag{2.40}$$

and it follows that the dual norm of the residual can be evaluated through the Riesz representation:

$$\|r(\,\cdot\,; \boldsymbol{\mu})\|_{(X^{\mathcal{N}})'} := \sup_{v \in X^{\mathcal{N}}} \frac{r(v; \boldsymbol{\mu})}{\|v\|_X} = \|\hat{e}(\boldsymbol{\mu})\|_X; \tag{2.41}$$

this expression is important for the Offline-Online strategy developed below.

As a second ingredient, we need a positive, parametric lower bound function

$$0 < \alpha_{\mathrm{LB}}^{\mathcal{N}}(\boldsymbol{\mu}) \leq \alpha^{\mathcal{N}}(\boldsymbol{\mu}) \quad \forall \boldsymbol{\mu} \in \mathcal{D}, \tag{2.42}$$

for the FE coercivity constant[8] defined as

$$\alpha^{\mathcal{N}}(\boldsymbol{\mu}) = \inf_{w \in X^{\mathcal{N}}} \frac{A(w, w; \boldsymbol{\mu})}{\|w\|_X^2}; \tag{2.43}$$

Instead of computing $\alpha^{\mathcal{N}}(\boldsymbol{\mu})$ by solving an eigenproblem for each parameter value $\boldsymbol{\mu}$, we use the lower bound $\alpha_{\mathrm{LB}}^{\mathcal{N}}(\boldsymbol{\mu})$ since its Online evaluation, for any value of $\boldsymbol{\mu} \in \mathcal{D}$ is independent of $\mathcal{N}$ and thus fulfills the efficiency requirements on the error bounds articulated before.

An efficient algorithm for the computation of $\alpha_{LB}^{\mathcal{N}}(\boldsymbol{\mu})$ is given by the so-called Successive Constraint Method (SCM), widely analyzed in [150, 280, 146]. We present the *natural norm* version for noncoercive problems in the analysis of the Stokes problem (Sect. A.3 of Appendix A), as well as a possible extension to nonlinear problems for the Navier-Stokes case (Sect. A.3.2). Here, we only remark that the SCM algorithm – which is based on the successive solution of suitable linear optimization problems – has been developed for the special requirements of the RB method, and features an efficient Offline-Online strategy.

[8] As we assumed that the bilinear form is coercive and the FE approximation spaces are conforming, it follows that $\alpha^{\mathcal{N}}(\boldsymbol{\mu}) \geq \alpha(\boldsymbol{\mu}) \geq \bar{\alpha} > 0$, $\forall \boldsymbol{\mu} \in \mathcal{D}$.

We define error estimators for the solution (in the energy norm) as

$$\Delta_N(\boldsymbol{\mu}) := \frac{\|\hat{e}(\boldsymbol{\mu})\|_X}{(\alpha_{\mathrm{LB}}^{\mathcal{N}}(\boldsymbol{\mu}))^{1/2}} \ . \tag{2.44}$$

Moreover, under the compliance assumption (i.e. if $A(\cdot,\cdot;\boldsymbol{\mu})$ is a symmetric bilinear form and $L(\cdot;\boldsymbol{\mu}) \equiv F(\cdot;\boldsymbol{\mu})$), the output converges as the "square" of the energy error, i.e. it verifies the following property:

$$s^{\mathcal{N}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu}) = |||y^{\mathcal{N}}(\boldsymbol{\mu}) - y_N(\boldsymbol{\mu})|||^2_{\boldsymbol{\mu}}; \tag{2.45}$$

in this case, the error estimator for the output are given by

$$\Delta_N^s(\boldsymbol{\mu}) := \Delta_N^2(\boldsymbol{\mu}) \equiv \frac{\|\hat{e}(\boldsymbol{\mu})\|_X^2}{\alpha_{\mathrm{LB}}^{\mathcal{N}}(\boldsymbol{\mu})} \ . \tag{2.46}$$

Although this latter result depends critically on the compliance assumption, extensions *via* adjoint approximations to the noncompliant case are also possible (see e.g. [237, 256] and the application to the Stokes case in Sect. 3.4.2-3.4.3). Nevertheless, the error bounds (2.44)–(2.46) are of no utility without an accompanying Offline-Online computational approach. The computationally crucial component of all the error bounds introduced is the dual norm $\|\hat{e}(\boldsymbol{\mu})\|_X$ of the residual, which can be computed through an Offline-Online procedure; details about these expressions are reported in Appendix A for the Stokes (Sect. A.2.1) and the Navier-Stokes (Sect. A.2.2) case.

## 2.4 Extension to nonaffine problems

We briefly discuss in this section the extension of the RB methodology to nonaffine problems, which are particularly relevant to our geometrical parametrizations and related applications. Also other (more general) classes of problems, such as noncompliant and noncoercive problems, can be tackled by the RB methodology. In the next chapter the case of Stokes and Navier-Stokes problems – representing one of the most remarkable cases of noncoercive problems – will be extensively treated, In the same chapter, an example of noncompliant problem is presented as well, when dealing with linear (noncompliant) or quadratic outputs of Stokes variables. For a general introduction to noncompliant and noncoercive problems the interested reader can refer to [237, 256].

As already remarked in Sect. 2.2.4, the assumption of affine parametric dependence (1.7) is fundamental in order to exploit the Offline-Online strategy and then minimize the marginal cost associated with each input-output evaluation. However, also nonaffine problems, i.e. problems in which condition (1.7) is not still valid, can be efficiently treated in the RB framework. In this case, we rely on the Empirical Interpolation Method (EIM) [21, 122, 90], which is based on adaptively chosen interpolation points and global shape functions, and allows to recover the assumption of affine parametric dependence in nonaffinely parametrized operators (e.g. linear, bilinear forms, etc.). In the case of a nonaffinely parametrized bilinear form $A(v, w; \boldsymbol{\mu})$, the latter is replaced by an affinely parametrized approximation of the form

$$A(v, w; \boldsymbol{\mu}) = \sum_{q=1}^{Q} \Theta^q(\boldsymbol{\mu}) A^q_{\mathrm{EIM}}(v, w) + \varepsilon_{\mathrm{EIM}}(v, w; \boldsymbol{\mu}), \tag{2.47}$$

where the error term $\varepsilon_{\mathrm{EIM}}$ needs to be controlled to an acceptable tolerance. We provide a short presentation of the EIM procedure based on [21]. Let us denote by $g(\mathbf{x}, \boldsymbol{\mu}) \in C^0(\mathcal{D}; L^\infty(\Omega))$ a scalar function depending on both the spatial coordinates $\mathbf{x}$ and the parameters vector $\boldsymbol{\mu}$ in a

nonaffine way; the extension to tensors through an element-wise procedure is straightforward. The goal is to find an approximate expansion of the form

$$g_M(\mathbf{x},\boldsymbol{\mu}) = \sum_{j=1}^{M} \Theta^j(\boldsymbol{\mu})\zeta_j(\mathbf{x}), \tag{2.48}$$

where $\Theta^j(\boldsymbol{\mu})$, $j=1,\dots,M$, are parameter-dependent functions and $\zeta_j(\mathbf{x})$, $j=1,\dots,M$, are parameter-independent functions, denoted also *shape functions*. Being an interpolation procedure, the EIM procedure seeks a sequence of (nested) sets of interpolation points $T_M = \{\mathbf{p}_1,\dots,\mathbf{p}_M\}$ (*magic points*), with $\mathbf{p}_j \in \Omega$ for each $j=1,\dots,M$, and a set of shape functions $\zeta_j(\mathbf{x})$, in order to compute the expansion (2.48) by solving the following Lagrange interpolation problem:

$$\sum_{j=1}^{M} \mathbb{B}^M_{i,j}\Theta^j(\boldsymbol{\mu}) = g(\mathbf{p}_i,\boldsymbol{\mu}), \quad \forall\, i=1,\dots,M,$$

where $\mathbb{B}^M \in \mathbb{R}^{M\times M}$ is defined as $(\mathbb{B}^M)_{ij} := \zeta_j(\mathbf{p}_i)$, $i,j=1,\dots,M$. Let us denote by $\Xi^{EIM}_{train} \subset \mathcal{D}$ a large training set, $M_{max}$ the maximum number of terms, $\varepsilon^*_{EIM}$ a fixed tolerance, and select an initial parameter value $\boldsymbol{\mu}^1$. The EIM procedure [21] is as follows:

$\zeta_1(\mathbf{x}) := g(\mathbf{x},\boldsymbol{\mu}^1)$; `compute` $\mathbf{p}_1 := \arg\operatorname{ess\,sup}_{\mathbf{x}\in\Omega} |\zeta_1(\mathbf{x})|$;
$q_1 = \zeta_1(\mathbf{x})/\zeta_1(\mathbf{p}_1)$; $G_1 := \operatorname{span}(\zeta_1)$, `set` $\mathbb{B}^1_{11} = 1$;
`for` $M = 2 : M_{\max}$
  `solve` $\boldsymbol{\mu}^M := \arg\max_{\boldsymbol{\mu}\in\Xi^{EIM}_{train}} \inf_{v\in G_{M-1}} ||g(\cdot,\boldsymbol{\mu}) - v||_{L^\infty(\Omega)}$ *(linear programming)*
  `set` $\zeta_M(\mathbf{x}) := g(\mathbf{x},\boldsymbol{\mu}^M)$, $G_M := \operatorname{span}(\zeta_1,\dots,\zeta_M)$
  `solve` $\sum_{j=1}^{M-1}\sigma_j^{M-1} q_j(\mathbf{p}_i) = \zeta_M(\mathbf{p}_i)$, $i=1,\dots,M-1$;
  `compute` *(residual)* $r_M(\mathbf{x}) := \zeta_M(\mathbf{x}) - \sum_{j=1}^{M-1}\sigma_j^{M-1}\zeta_j(\mathbf{x})$;
  `compute` $\mathbf{p}_M := \arg\operatorname{ess\,sup}_{\mathbf{x}\in\Omega} |r_M(\mathbf{x})|$;
  `set` $q_M(\mathbf{x}) = r_M(\mathbf{x})/r_M(\mathbf{p}_M)$, $\mathbb{B}^M_{ij} = q_j(\mathbf{p}_i)$, $i,j=1,\dots,M$;
  `if` $\max_{\boldsymbol{\mu}\in\Xi^{EIM}_{train}} \inf_{v\in G_M} ||g(\cdot,\boldsymbol{\mu}) - v||_{L^\infty(\Omega)} < \varepsilon^*_{EIM}$
    $M_{\max} = M-1$;
  `end`;
`end`.

In the end, given an approximation $g_M(\mathbf{x},\boldsymbol{\mu})$, $M < M_{max}$, we denote the *one point* error estimator the following quantity (very inexpensive to compute):

$$\hat{\epsilon}_M(\boldsymbol{\mu}) = |g(\mathbf{p}_{M+1};\boldsymbol{\mu}) - g_M(\mathbf{p}_{M+1};\boldsymbol{\mu})|, \tag{2.49}$$

corresponding to the difference between the function and the interpolant at the point $\mathbf{p}_{M+1}$, which gives the largest residual $r_M(\mathbf{x})$. While not rigorous as *a posteriori* error bound, this quantity proves to be an intuitive measure of the error committed by the EIM procedure [21]. Advances in error bounds developments have been presented in [90, 122, 200, 226].

In practice, if the problem[9] is not affinely parametrized (e.g. when the geometrical transformation (2.15) is not affine, see Sect. 2.5.2, or the physical coefficients appearing in the tensors $\boldsymbol{\nu}^{o,k}$, $\boldsymbol{\chi}^{o,k}$, $\boldsymbol{\eta}^{o,k}$ are nonaffine functions of $\mathbf{x}$ and $\boldsymbol{\mu}$), the parametrized tensors in (2.17) depend both on the parameter $\boldsymbol{\mu}$ and the spatial coordinate $\mathbf{x}$. In this case, the operators can not be expressed as in

[9] For the sake of simplicity, we consider here a parametrized bilinear form $A(\cdot,\cdot;\boldsymbol{\mu})$ corresponding to a pure diffusion operator; for more general operators, the same procedure is applied to each parametrized tensor.

(1.7) and we thus need an additional pre-processing, before the FE assembling stage, in order to recover the affinity assumption. According to EIM – considering for instance the tensor $\boldsymbol{\nu}^k$ – each component $\boldsymbol{\nu}^k_{ij}(\mathbf{x}, \boldsymbol{\mu})$ is approximated by an affine expression given by

$$\boldsymbol{\nu}^k_{ij}(\mathbf{x}, \boldsymbol{\mu}) = \sum_{l=1}^{K^k_{ijl}} \beta_l^{ijk}(\boldsymbol{\mu}) \eta_l^{ijk}(\mathbf{x}) + \varepsilon^k_{ij}(\mathbf{x}, \boldsymbol{\mu}); \tag{2.50}$$

all the functions $\beta_l^{ijk}$'s and $\eta_l^{ijk}$'s are efficiently computable scalar functions and the error terms are guaranteed to be under some tolerance,

$$\|\varepsilon^k_{ij}(\cdot; \boldsymbol{\mu})\|_\infty \leq \varepsilon^{EIM}_{tol} \qquad \forall \boldsymbol{\mu} \in \mathcal{D}.$$

In this way, we can identify the $\boldsymbol{\mu}$-dependent functions $\beta_l^{ijk}(\boldsymbol{\mu})$ in (2.50) as the functions $\Theta^q_A(\boldsymbol{\mu})$ in (1.7), where $q$ is a condensed index for $(i, j, k, l)$, while the $\boldsymbol{\mu}$-independent functions will be treated as pre-factors in the integrals which give the $\boldsymbol{\mu}$-independent bilinear forms $A^q(w, v)$.
The nonaffine treatment is really important since many problems involving more complex geometrical parametrizations and/or more complex physical properties are hold by nonaffine parametric dependence. Not only, EIM has been exploited also to deal with nonlinear operators involving polynomial functions of the variables, for instance in [122].

## 2.5 Geometrical reduction: main features and goals

The complexity of geometrical models is nowadays increasing in optimal design together with engineering applications, implying the importance of automation in computer-aided design (CAD) and the need of a strong *geometrical reduction*. Since shape representation is highly specific problem-dependent, various methods have been proposed in several engineering contexts; for a review of shape parametrization techniques we refer the reader to [286, 104, 214].
Two different approaches in *geometrical shape optimization* are the *local boundary variation* [249] and explicit *boundary shape parametrizations* [59, 26, 286]. The former uses the discrete nodes of the triangulation lying on the boundary as design variables and their displacement for the shape deformation. Since very fine meshes are needed for complex shapes or flows, the number of design variables may become very large, leading to very high computational costs, and sometimes it is difficult to preserve the regularity of the geometry. The latter uses polynomial or *spline* representations (e.g. Bézier curves, B-spline curves, etc.) depending on some weights and a set of control points, which are treated as parameters. The drawbacks of this approach are that *(i)* the number of parameters used to describe the shapes may become *high-dimensional* for arbitrarily complex shapes and that *(ii)* after each shape deformation, remeshing is still required.

Neither the *boundary variation* nor the *boundary parametrization* represent the ultimate choice to define a shape parametrization fitting a reduced basis framework for parametrized PDEs. Rather, we turn to *global domain parametrizations* in order to solve efficiently shape optimization problems and other shape-related inverse problems by means of a *fixed domain* approach. Not only, the introduction of a shape parametrization has to be seen as a first *geometrical reduction* step, provided that the parametric map is sufficiently flexible and low-dimensional. In particular, we can evaluate shape parametrization methods according to five criteria:

1. *computational efficiency*: the parametric shapes must be computable efficiently, by means of a compact and effective set of design variables, with little loss of numerical precision, and preferably allows a parallel implementation;

2. *generality*: the parametrization method must be able to describe smooth, regular shapes with arbitrary precision as the number of parameters used to describe the shape is increased;

3. *flexibility*: the number of parameters and their effect on the shapes must be tunable by the user, e.g. to add control points near areas of the shape where the design is more sensitive to small variations;
4. *transferability*: it should be possible to transfer the parametric shapes between the CAD program used by designers, and the finite element software used by the engineers, without manual pre-processing steps in between. This also includes being able to transfer the optimized shape back to the CAD program since many industrial manufacturing processes nowadays can be directly driven by the CAD;
5. *quality of shapes*: of course, we typically want to avoid self-intersection of shapes, singular points, cusps, etc. We might also demand some minimum level of regularity from the shapes (e.g. $\mathcal{C}^k$ for $k \geq 1$), depending on the problem at hand, as well as lower continuity requirements for some parts of the geometry (e.g. NURBS patches can have repeated knots with only $\mathcal{C}^0$ continuity locally).

In view of exploiting the RB method for the sake of computational reduction, a suitable shape parametrization should not only realize a good compromise between these criteria – in particular, *flexibility* and *computational efficiency* – but also allow to fit some unavoidable requirements of the parametrized framework, such as the *affinity* assumption – or at least not to cause an overwhelming extracost if this assumption is not automatically verified and has to be properly recovered. Based on these features, different options for the construction of the parametric map $T(\,\cdot\,;\boldsymbol{\pi})$ have been considered in the framework of reduced basis applications. For example, affine or nonaffine maps built by hands and based on a domain decomposition approach have been introduced in order to deal with simple cartesian geometries; several extensions to geometries with curved boundaries have been also implemented, but leading to large parametric complexities (in terms of affine expansions of the operators) for more involved geometries.
An alternative way, exploited in our applications for constructing flexible but low-dimensional maps, is based on the so-called *free shape representations*. The most popular technique within this group is the Free-Form Deformation (FFD) technique (see Sect. 2.6) which is based on tensor products of splines and gives a global nonaffine transformation map. Another technique which we have successfully coupled with the RB framework and allows to construct such a map is given by the Radial Basis Function (RBF), which is a general paradigm for interpolation of scattered data in high dimensions. In the following sections we present these approaches, which will be exploited in the examples discussed in Part III for the description of complex geometries arising from some haemodynamics applications.

### 2.5.1 Automatic piecewise affine maps

Affine maps represent the simplest example of parametric maps to transform the original problem (2.13) into a parametrized problem defined on the fixed, reference domain. Their definition is elementary on a simple, Cartesian geometry where the geometrical parameters are shape properties such as lengths, thicknesses, diameters or angles. Moreover, it allows the parametrized PDE to straightforward inherit the *affinity* assumption, crucial for the efficacy of the Offline/Online strategy. Not only, the generation of this kind of maps can be made automatic thanks to the introduction of a suitable *domain decomposition* paradigm also on curved and more general geometries; clearly, increased complexity in geometry and operator will result in more terms in affine expansions, with a corresponding increase in the RB (Online) computational costs. As already mentioned in Sect. 2.2.1, by introducing the RB triangulation (2.14), original and reference subdomains are linked via a mapping $T(\cdot;\boldsymbol{\mu}) : \Omega^k \to \Omega_o^k(\boldsymbol{\mu})$, $1 \leq k \leq K_{\text{dom}}$. These maps must be individually bijective, collectively continuous, which means they have to fulfill the following interface condition:

$$T^k(\mathbf{x};\boldsymbol{\mu}) = T^{k'}(\mathbf{x};\boldsymbol{\mu}), \quad \forall \mathbf{x} \in \bar{\Omega}^k \cap \bar{\Omega}^{k'}, \quad 1 \leq k < k' \leq K_{\text{dom}}. \tag{2.51}$$

In the affine case, for the $k^{th}$ subdomain ($1 \leq k \leq K_{\text{dom}}$) the precise affine transformation is then given, for any $\boldsymbol{\mu} \in \mathcal{D}$ and $\mathbf{x} \in \Omega^k$, by

$$T_i^k(\mathbf{x}, \boldsymbol{\mu}) = C_i^k(\boldsymbol{\mu}) + \sum_{j=1}^{d} G_{ij}^k(\boldsymbol{\mu}) x_j, \ 1 \leq i \leq d \tag{2.52}$$

for given translation vectors $\mathbf{C}^k : \mathcal{D} \to \mathbb{R}^d$ and linear transformation matrices $\mathbf{G}^k : \mathcal{D} \to \mathbb{R}^{d \times d}$. The linear transformation matrices can effect rotation, scaling and/or shear and have to be invertible. The associated Jacobians are defined as

$$J_T^k(\boldsymbol{\mu}) = |\det(\mathbf{G}^k(\boldsymbol{\mu}))|, \quad 1 \leq k \leq K_{\text{dom}};$$

for invertible mappings they are strictly positive. In particular, the software `rbMIT` [149] used for the RB computations allows to build efficient affine mappings in an automatic fashion based on a domain decomposition made up by some "building blocks" introduced in this section; for clarity, we concentrate on a single subdomain in the two-dimensional case ($d = 2$). In this case, straight lines are mapped into straight lines and parallelism is preserved, a parallelogram is mapped into a parallelogram and hence a triangle into a triangle; moreover, affine transformations map ellipses into ellipses. These features are exploited for the development of an automatic domain decomposition technique that is suitable for the RB context; the interested reader can refer to Sect. A.5 of Appendix A for more details related to the construction of basic RB triangulations built on (standard) triangles, elliptical triangles and general "curvy" triangles.

### 2.5.2 Automatic nonaffine maps: volume-based parametrizations

For problems with simple domains and/or simple sizing operations, *piecewise affine maps* defined over a suitable domain decomposition are workable. For more complex problems, although the construction can be performed automatically, the computational cost might sensibly increase, because of the splitting generated by the domain decomposition. In this case, it is preferable to turn to global *nonaffine maps*, defined on the whole domain $\Omega$, even if this causes the parametrized formulation of the PDE problem not to satisfy automatically the affinity assumption, and thus imply an additional step to recover this crucial feature.
We denote as *nonaffine* any transformation which cannot be written under the form (2.52). In this case, for the $k^{th}$ subdomain ($1 \leq k \leq K_{\text{dom}}$), for any $\boldsymbol{\mu} \in \mathcal{D}$ and for any $\mathbf{x} \in \Omega^k$, $\mathbf{x}_o \in \Omega_o^k(\boldsymbol{\mu})$, the nonaffine transformation is given by a generic expression

$$\mathbf{x}_{o\,i} = T_i^{\text{naff},k}(\mathbf{x}, \boldsymbol{\mu}), \quad 1 \leq i \leq d, \tag{2.53}$$

and the Jacobians $J_T^k(\mathbf{x}, \boldsymbol{\mu}) = |\det(\mathbf{G}^{\text{naff},k}(\mathbf{x}, \boldsymbol{\mu}))|$, $1 \leq k \leq K_{\text{dom}}$, are strictly positive, where $\mathbf{G}^{\text{naff},k} : \Omega^k \times \mathcal{D} \to \mathbb{R}^{d \times d}$ are the Jacobian matrices of the mappings $T^{\text{naff},k}(\cdot, \boldsymbol{\mu})$ defined in (2.19). As in the affine case, the interface condition (2.51) allows us to interpret the set of local mappings as a global bijective piecewise affine transformation $T(\cdot; \boldsymbol{\mu}) : \Omega \to \Omega_o(\boldsymbol{\mu})$, given for any $\boldsymbol{\mu} \in \mathcal{D}$ by

$$T(\mathbf{x}, \boldsymbol{\mu}) = T^{\text{aff} \vee \text{naff},k}(\mathbf{x}; \boldsymbol{\mu}), \quad k = \min_{k' \in \{1,\ldots,K_{\text{dom}}\} | \mathbf{x} \in \Omega^{k'}} k'. \tag{2.54}$$

Several techniques are available in literature for constructing nonaffine parametric maps which are smooth, sufficiently flexible but low-dimensional; in particular, computer graphics techniques (e.g. in the so-called *soft object animation* field) provide powerful algorithms for modifying shapes. In the following sections, we focus on a widespread paradigm developed in this field – the so-called *volume-based representations* [104] – which allow to perform deformations whatever the complexity of the shape. In particular, by means of maps relying on such a representation, geometry can be parametrized independently from PDE models, discrete formulations and computational meshes.

Volume based representations operate on a control volume, regardless of the object to be deformed, embedded in the volume. Parametric maps are thus defined by introducing a set of control points over the control volume and considering their displacements (which actually induce a shape deformation) as geometrical parameters, rather than geometrical properties directly related with the shape itself. In this way, shifting a control point causes a deformation of the embedding space – the parametric map is defined inside the whole control volume – and thereby induces a global modification of each shape located inside this volume. Two possible techniques leading to *free shape representations* are detailed in Sect. 2.6 and 2.7.

## 2.6 Free-Form Deformation techniques

A first case of *free shape* representation is given by the Free-Form Deformation (FFD) technique, introduced first on computer graphics by Sederberg and Parry in the late 80s [297]. This technique consists of embedding the shape to be deformed inside a control volume and then of modifying – by acting on a lattice of control points – the metrics of this space and thus the shape embedded into it, rather than modifying the shape directly. A modification of the control points position thus results in a deformation inside the control volume and, automatically, of the computational FE mesh. A simple physical analogy for FFD is to consider a parallelepiped of clear matter, flexible (at the limit, like a *jelly*) material in which we embed an object we wish to deform. The object is also flexible, so that it can be deformed along with the external parallelepiped.

Based on tensor product of splines, FFD inherits from boundary parametrization techniques the possibility to handle with global deformations by acting on a set of control points [10, 40], but provides an easier tool – since any explicit parametrization is required – which can be applied to virtually any geometrical model.
A simple version of the FFD construction is defined as follows, as a mapping from $\mathbb{R}^d$ to $\mathbb{R}^d$ through a $d$-variate tensor product Bernstein polynomial. For the sake of simplicity, we restrict ourselves to the two-dimensional case ($d = 2$), extension to three dimensions is straightforward.

Given a fixed rectangular domain $D$ containing the reference domain $\Omega \subset D$ we wish to deform, we introduce an affine map

$$\Psi : D \to \hat{D} \equiv (0,1)^2, \qquad \hat{\mathbf{x}} = \Psi(\mathbf{x}), \quad \mathbf{x} \in D,$$

with $\Psi(\cdot)$ a monotonic function, in order to define FFD – in a simpler way – in the coordinates $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2)$ of the spline parameter space $(0,1)^2$. We thus select an ordered lattice of control points $\mathbf{P}_{l,m} \in (0,1)^2$ (the unperturbed configuration), so that

$$\mathbf{P}_{l,m} = \begin{bmatrix} l/L \\ m/M \end{bmatrix}, \qquad l = 0, \dots, L, \quad m = 0, \dots, M.$$

A perturbation of the control points positions is specified by a set of $(L+1)(M+1)$ parameter vectors $\boldsymbol{\mu}_{l,m} \in \mathbb{R}^2$ so that the perturbed configuration of the control points results in

$$\mathbf{P}^o_{l,m}(\boldsymbol{\mu}_{l,m}) = \mathbf{P}_{l,m} + \boldsymbol{\mu}_{l,m}, \tag{2.55}$$

giving in total $2(L+1)(M+1)$ possible degrees of freedom. Very often, only small subsets of these are selected as design variables if we want to perform a sensible geometrical model order reduction; moreover, several rows or columns of control points can be fixed to obtain desired levels of continuity or to "anchor" certain parts of the domain. In general, we indicate the effectively free scalar-valued parameters chosen as design variables (or actual degrees of freedom) as $\mu_1, \dots, \mu_p$, each corresponding to the displacement of a control point in either the $\hat{x}_1$ or the $\hat{x}_2$ direction, i.e. to one of the components of a vector $\boldsymbol{\mu}_{l,m}$, and call $p$ the number of degrees of freedom.

From now on we assume that the shape parametrization involves only those $\boldsymbol{\mu}_{l,m}$ corresponding to the actual degrees of freedom, omitting the parameters that have been fixed. in this way, we denote by $\boldsymbol{\mu} = (\mu_1 \dots, \mu_p)$, even if also other control points – which do not correspond to effective design variables – obviously go under displacement.
We thus construct a parametric domain map $\hat{T}(\cdot;\boldsymbol{\mu}) : \hat{D} \to \hat{D}_o(\boldsymbol{\mu})$ by which the updated geometry is computed as follows:

$$\hat{T}(\hat{\mathbf{x}};\boldsymbol{\mu}) = \sum_{l=0}^{L}\sum_{m=0}^{M} b_{l,m}^{L,M}(\hat{\mathbf{x}})\mathbf{P}_{l,m}^{o}(\boldsymbol{\mu}_{l,m}) = \sum_{l=0}^{L}\sum_{m=0}^{M} b_{l,m}^{L,M}(\hat{\mathbf{x}})(\mathbf{P}_{l,m} + \boldsymbol{\mu}_{l,m}), \tag{2.56}$$

where

$$b_{l,m}^{L,M}(\hat{\mathbf{x}}) = b_l^L(\hat{x}_1)b_m^M(\hat{x}_2) = \binom{L}{l}\binom{M}{m}(1-\hat{x}_1)^{L-l}\hat{x}_1^l(1-\hat{x}_2)^{M-m}\hat{x}_2^m \tag{2.57}$$

are (*Bézier*) tensor products of the univariate *Bernstein basis polynomials*

$$b_l^L(\hat{x}_1) = \binom{L}{l}\hat{x}_1^l(1-\hat{x}_1)^{L-l}, \qquad b_m^M(\hat{x}_2) = \binom{M}{m}\hat{x}_2^m(1-\hat{x}_2)^{M-m}.$$

defined on the unit square $\hat{D}$ with local variables $(\hat{x}_1, \hat{x}_2) \in (0,1)^2$. The shape and continuity of a deformation within the volume is related to the degree $L$, $M$ of the Bernstein polynomials. Finally, the FFD mapping $T(\cdot,\boldsymbol{\mu})$ is obtained as the composition

$$T(\cdot;\boldsymbol{\mu}) : D \to D_o(\boldsymbol{\mu}), \qquad T(\mathbf{x};\boldsymbol{\mu}) = \Psi^{-1} \circ \hat{T} \circ \Psi(\mathbf{x};\boldsymbol{\mu}); \tag{2.58}$$

in particular, the parametrized domain $\Omega_o(\boldsymbol{\mu})$ is obtained as $\Omega_o(\boldsymbol{\mu}) = \Psi^{-1} \circ \hat{T} \circ \Psi(\Omega;\boldsymbol{\mu})$; see Fig. 2.2 for a representation of the mapping construction. An example of FFD mapping to represent deformations of a two-dimensional airfoil of the NACA family is shown in Fig. 2.3. In this case, the design parameters are given by the $p = 8$ vertical displacements of the control points located in the interior part of the domain.

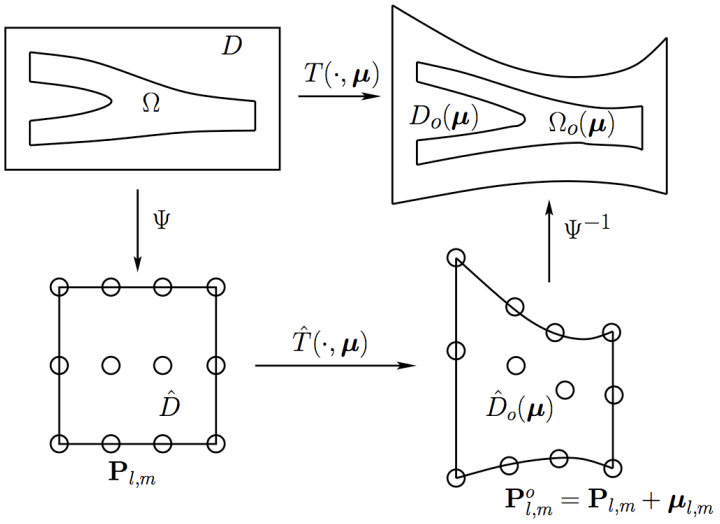



Figure 2.2: Schematic diagram of FFD technique: unperturbed control points $\mathbf{P}_{l,m}$, perturbed control points $\mathbf{P}_{l,m}^{o}(\boldsymbol{\mu}_{l,m})$, map $\hat{\mathbf{x}} = \Psi(\mathbf{x})$, $\hat{T}(\hat{\mathbf{x}};\boldsymbol{\mu})$ and resulting FFD map $T(\mathbf{x};\boldsymbol{\mu}) = (\Psi^{-1} \circ \hat{T} \circ \Psi)(\mathbf{x};\boldsymbol{\mu})$.

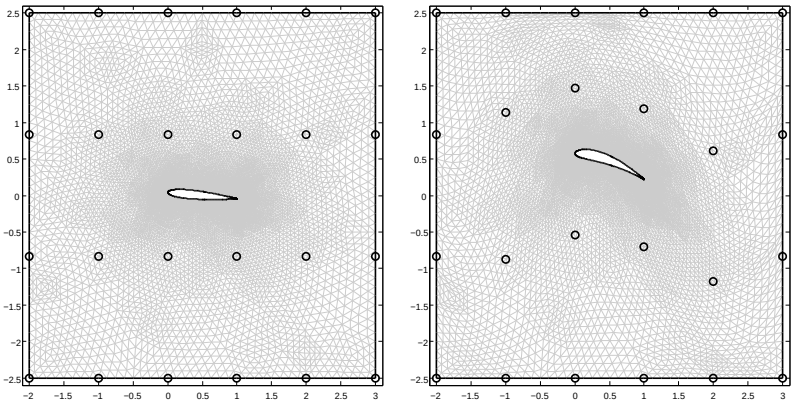


Figure 2.3: Example of shape deformation of a NACA0012 airfoil obtained through FFD mapping: reference configuration and control points (left), deformed configuration and control points (right).

Bernstein basis polynomials are used mainly because they fulfill the *partition of unity*

$$\sum_{l=1}^{L} b_l^L(\hat{x}_1) = \sum_{m=1}^{M} b_m^M(\hat{x}_2) = 1 \tag{2.59}$$

and positivity

$$b_{l,m}^{L,M}(\hat{\mathbf{x}}) \geq 0 \quad \forall \hat{x} \in (0,1)^2$$

properties. Using the first property, it can be shown that FFD are a particular case of the transformations defined as *perturbation of the identity*, i.e. $T(\mathbf{x};\mathbf{0}) = \mathbf{x}$. A more detailed analysis of this class of mappings will be addressed in sect. 4.4. One benefit of the simple FFD defined above is that the parametric transformations of the reference domain are simple polynomial functions in spatial coordinates and parameter, and they can be evaluated numerically through the stable De Casteljau algorithm [95], while their derivatives can be expressed in terms of tensor products of lower order Bernstein polynomials by means of the following formula [95]:

$$\nabla b_{l,m}^{L,M}(\hat{\mathbf{x}}) = \begin{bmatrix} L(b_{l-1}^{L-1}(\hat{x}_1) - b_l^{L-1}(\hat{x}_1) b_m^M(\hat{x}_2)) \\ M(b_{m-1}^{M-1}(\hat{x}_2) - b_m^{M-1}(\hat{x}_2) b_l^L(\hat{x}_1)) \end{bmatrix}^T. \tag{2.60}$$

This gradient formula is particularly useful for evaluating the parametrized tensors (2.18)–(2.20) appearing in the parametrized formulation (1.1)-(1.2) of our PDE problems, which depend on the Jacobian (2.19) of the map. Exploiting the partition of unity property (2.59), thanks to the formula (2.60), in the FFD case we obtain

$$J_T(\mathbf{x};\boldsymbol{\mu}) = J_\Psi^{-1}(\mathbf{x}) \left[ I + \sum_{l=0}^{L} \sum_{m=0}^{M} \nabla b_{l,m}^{L,M}(\Psi(\mathbf{x})) \boldsymbol{\mu}_{l,m} \right] J_\Psi(\mathbf{x}). \tag{2.61}$$

Thus, not only the FFD map $T(\cdot;\boldsymbol{\mu})$ can be stably evaluated, but also its Jacobian $J_T(\cdot;\boldsymbol{\mu})$ as well as the parametrized tensors (2.18)–(2.20). Moreover, the determinant of the Jacobian provides us with control over the volume change that a body experiences under FFD: by computing $|\det(J_T)|$, it can be shown that the largest and smallest polynomial coefficients provide an upper and a lower bound, respectively, on the volume change FFD. A condition on the well-posedness of the FFD map, ensuring also that the connectivity of the lattice of control points is preserved, will be briefly discussed in Sect. 4.4.

Concerning *generality* and *quality of shapes*, FFD techniques lead to low-dimensional parametrizations without loss of accuracy, since perturbations on parameters yield smooth shape deformations even if control points are not related to the shape boundary. In particular, parametric curves and surfaces remain parametric under FFD, geometrical singularities can be taken into account (since the initial shape including its singularities is deformed) and the smoothness of the deformation is controlled thanks to the Bernstein polynomials appearing in the map.
Regarding *efficiency*, the number of design variables depends on the user's choice: we can keep fixed a subset of control points or only allow them to move in one direction; this allows the user to keep the number of FFD parameters to a desired low level (in our case $P < 10$ is typical). In particular, the number and position of control points chosen have a deep impact on FFD flexibility: it is crucial to maximize the influence of the control points by placing them close to the more sensitive regions of the configuration. Adaptive procedures for the selection of the control points based on sensitivity analysis and correlations are also available (see e.g Sect. 4.8).

FFD have been largely employed for parametrization and optimal design of aerodynamic surfaces such as wings, for instance by the group of Désidéri [88, 82, 11] and by Samareh [287]. A nice study of shape optimization of fluid domains by means of FFD techniques for some examples of engineering interest can be found in a paper by Lehnhäuser and Schäfer [189].
In the RB context, FFD has previously been proposed as a parametrization technique for inverse airfoils design in potential flows by Lassila and Rozza [185] and for thermal flows control by the author with Lassila and Rozza [281]. In this Thesis, we aim at enhancing the computational performance in the shape optimization process for viscous flows, coupling FFD and RB methods for Stokes/Navier-Stokes flows, developing what started in these two previous works.

## 2.7 Radial Basis Function techniques

Whereas Free-Form Deformation show great flexibility and easiness of handling, they suffer from some limitations, since *(i)* the control points cannot be chosen freely, being the nodes of a lattice *(ii)* it is not possible to perform a boundary control and *(iii)* the process is not interpolatory. In particular, using a rectangular lattice to describe deformations of irregular or complex shapes makes the choice of control points crucially important; for example, control points located far from the boundary to be optimized have less influence. Choosing the subset of active points is therefore a critical issue, highly problem-dependent.

In order to overcome these limitations, an alternative method can be used for shape parametrization – still based on a set of control points, but which is in addition interpolatory: the Radial Basis Function (RBF) technique. Originally used in neural networks and later applied to the solution of PDEs in what is known as *meshless methods*, RBFs are now a widely used method for scattered interpolation and reconstruction of surfaces and volumes. In particular, they provide a general and flexible way of interpolating data in multi-dimensional spaces, even for unstructured data where it is often impossible to apply polynomial or spline interpolation [155]. For a general introduction on RBF method see for instance the monographs by Buhmann [50] or Wendland [320].

We first describe the RBF interpolation procedure (which is employed in the following chapters also for different goals than shape parametrization, see e.g. Sect. A.3.2), focusing afterwards on the construction of parametric maps based on this strategy. Without loss of generality, we restrict ourselves to the two-dimensional case, where we denote by $\mathbf{X} = \{\mathbf{x}_1, \dots \mathbf{x}_k\} \subseteq \mathbb{R}^2$ a set of $k > 2$ non-collinear points (usually called *centers*, with $\mathbf{x}_i = (x_{i1}, x_{i2})$) and by $Y = \{y_1, \dots y_k\} \subseteq \mathbb{R}$ the values at the centers which have to be interpolated.

| Radial Basis Function | $\Phi(r)$ |
|---|---|
| Linear | $r$ |
| Spline type ($R_\lambda$) | $\lvert r\rvert^\lambda, \quad \lambda$ odd |
| Thin-Plate Spline ($TPS_\lambda$) | $\lvert r\rvert^\lambda \log \lvert r\rvert, \quad \lambda$ even |
| Multiquadric (MQ) | $(1+r^2)^{1/2}$ |
| Inverse Multiquadric (IMQ) | $(1+r^2)^{-1/2}$ |
| Inverse Quadratic (IQ) | $(1+r^2)^{-2}$ |
| Gaussian (GS) | $\exp(-r^2)$ |

Table 2.1: Several options for radial basis functions

Then, the RBF interpolant $\tau : \mathbb{R}^2 \to \mathbb{R}$ has the form

$$\tau(\mathbf{x}) = \pi(\mathbf{x}) + \sum_{i=1}^{k} w_i \ \Phi(\|\mathbf{x} - \mathbf{x}_i\|). \tag{2.62}$$

In (2.62), the function $\Phi$ is a fixed basis function, which is radial with respect to the Euclidean distance $\|\mathbf{x}\|$, $\pi(\cdot) : \mathbb{R}^2 \to \mathbb{R}$ is a bivariate polynomial function of (low) degree $p$, $\{w_i\}_{i=1}^k,\ w_i \in \mathbb{R}$ is a set of coefficients corresponding to the centers $\mathbf{x}_i$. Common choices for RBFs are listed in Tab. 2.1. Moreover, in many cases it is appropriate to scale the basis function with a so-called shape parameter $\epsilon$, so that the basis function is replaced by $\Phi_\epsilon(r) = \Phi(\epsilon r)$.

The coefficients $\{w_i\}_{i=1}^k$ and the polynomial $\pi(\cdot)$ are determined so that $\tau$ interpolates the data $y_1, \dots, y_k$ (*interpolation constraints*)

$$\tau(\mathbf{x}_i) = y_i, \qquad 1 \le i \le k \tag{2.63}$$

and satisfies the additional requirements (*side constraints*)

$$\sum_{i=1}^{k} w_i \varphi(\mathbf{x}_i) = 0 \qquad \forall \varphi \in \Pi_2^q,\ q \le p, \tag{2.64}$$

where $\Pi_2^q$ is the space of all polynomials of degree up to $q \ge 1$ in 2 unknowns. By imposing the side constraints, we require that each polynomial up to degree $p$ is interpolated exactly; we discuss in Sect. 4.4 the geometrical meaning of this requirement, providing a general analysis of RBF schemes and some conditions on the well-posedness of RBF parametric mappings. For the sake of simplicity, we consider $\pi(\mathbf{x})$ to be a polynomial function of degree $p = 1$, so that

$$\pi(\mathbf{x}) = \gamma_0 + \gamma_1 x_1 + \gamma_2 x_2;$$

moreover, we denote by $\mathbf{w} = (w_1, \dots, w_k)^T \in \mathbb{R}^k$ and $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \gamma_2)^T \in \mathbb{R}^3$ the vectors whose components are the coefficients $w_1, \dots, w_k$ for the RBFs and for the polynomial, respectively. Constraints (2.63)-(2.64) thus lead to the following linear system for the coefficients vector $[\mathbf{w}\ \boldsymbol{\gamma}]^T \in \mathbb{R}^{k+3}$:

$$\begin{bmatrix} \mathbb{M} & \mathbb{P} \\ \mathbb{P}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\gamma} \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix} \tag{2.65}$$

where $\mathbf{y} = (y_1, \dots, y_k)^T \in \mathbb{R}^k$, $\mathbb{M} \in \mathbb{R}^{k \times k}$ is the interpolation matrix of components

$$(\mathbb{M})_{ij} = \Phi(\|\mathbf{x}_i - \mathbf{x}_j\|), \qquad 1 \le i, j \le k$$

and $\mathbb{P} \in \mathbb{R}^{k \times 3}$ the matrix defined by imposing the side constraints:

$$\mathbb{P} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{k1} & x_{k2} \end{bmatrix}.$$

Provided that the interpolation matrix $\mathbb{M}$ is nonsingular, we can invert (2.65) and obtain

$$\mathbf{w} = (\mathbb{I} - \mathbb{M}^{-1}\mathbb{P}\mathbb{M}_{\mathbb{P}}\mathbb{P}^T)\mathbb{M}^{-1}\mathbf{y}, \qquad \boldsymbol{\gamma} = \mathbb{M}_{\mathbb{P}}\mathbb{P}^T\mathbb{M}^{-1}\mathbf{y},$$

where $\mathbb{M}_{\mathbb{P}} = (\mathbb{P}^T\mathbb{M}\mathbb{P})^{-1}$, so that the RBF interpolant (2.62) is uniquely identified.

For shape parametrization in $d$ dimensions – as well as for mesh deformation – we need $d$ interpolating functions, one for each coordinate direction. In the two-dimensional case, we shall consider the centers $\{\mathbf{P}_1, \dots \mathbf{P}_k\} \subseteq \mathbb{R}^2$ as the unperturbed configuration of the *control points* in the reference domain $\Omega$, being $\mathbf{P}_i = (P_{i1}, P_{i2})$ for any $i = 1, \dots, k$, and $\{\mathbf{P}_1^o, \dots \mathbf{P}_k^o\} \subseteq \mathbb{R}^2$ as the perturbed configuration of the control points in the original domain $\Omega_o$.
By introducing a set of (now vector-valued) coefficients $\{\mathbf{w}_i\}_{i=1}^k$, $\mathbf{w}_i = (w_{i1}, w_{i2})^T \in \mathbb{R}^2$, and denoting by $\pi(\mathbf{x}) = \mathbf{c} + \mathbb{A}\mathbf{x}$ a polynomial function of degree 1, with $\mathbf{c} \in \mathbb{R}^2$ and $\mathbb{A} \in \mathbb{R}^{2\times 2}$, the RBF map (2.62) can be rewritten in a compact form as

$$\tau(\mathbf{x}) = \mathbf{c} + \mathbb{A}\mathbf{x} + \mathbb{W}^T s(\mathbf{x}), \tag{2.66}$$

where $s(\mathbf{x}) = (\Phi(\|\mathbf{x} - \mathbf{P}_1\|), \dots, \Phi(\|\mathbf{x} - \mathbf{P}_k\|))^T \in \mathbb{R}^k$ and $\mathbb{W} = [\mathbf{w}_1, \dots, \mathbf{w}_k]^T \in \mathbb{R}^{k\times 2}$. Regarding shape parametrizations, the polynomial of degree 1 represents the affine part of a deformation (rotation and/or scaling), while the term $\mathbb{W}^T s(\mathbf{x})$ depending on the control points adds a nonaffine contribution. The RBF map (2.66) is thus a function of $2k + 6$ coefficients in the two-dimensional case, which are determined by looking for a transformation such that: *(i)* each control point $\mathbf{P}_i$ in the unperturbed configuration is mapped onto the corresponding control point $\mathbf{P}_i^o$ in the perturbed configuration, and that *(ii)* each affine transformation is recovered exactly (i.e. the RBF interpolation is invariant with respect to rigid motions). This is equivalent to impose both the interpolation and the side constraints (2.63)-(2.64), which now read:

$$\tau(\mathbf{P}_i) = \mathbf{P}_i^o, \qquad i = 1, \dots, k, \tag{2.67}$$

$$\sum_{i=1}^k \mathbf{w}_i = 0, \qquad \sum_{i=1}^k P_{i1}\mathbf{w}_i = \sum_{i=1}^k P_{i2}\mathbf{w}_i = \mathbf{0}. \tag{2.68}$$

We point out that when dealing with mesh deformation or coupling between two different phases, interpolation and side constraints (2.67)-(2.68) guarantee the conservation of virtual work, total load and total momentum when the displacements are transferred from one phase to another [155, 217].

In order to fit the RBF technique in our parametrized framework, let us express the deformed positions $\mathbf{P}_i^o$ of the control points as

$$\mathbf{P}_i^o(\boldsymbol{\mu}_i) = \mathbf{P}_i + \boldsymbol{\mu}_i, \quad i = 1, \dots, k,$$

where $\boldsymbol{\mu}_i = (\mu_{i1}, \mu_{i2})$ is the displacement of the $i$-th control points. As in the FFD case, only small subsets of $p \leq 2k$ displacements are selected as design variables – accordingly to some problem-dependent criteria – if we want to perform a sensible geometrical reduction. We indicate the actual degrees of freedom as $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)$, each corresponding to the displacement of a control point in either the $x_1$ or the $x_2$ direction, i.e. to one of the components of a vector $\boldsymbol{\mu}_i$. The parametric map $T(\cdot; \boldsymbol{\mu}) : \Omega \to \Omega_o(\boldsymbol{\mu})$ is thus given by

$$T(\mathbf{x}; \boldsymbol{\mu}) = \mathbf{c}(\boldsymbol{\mu}) + \mathbb{A}(\boldsymbol{\mu})\mathbf{x} + \mathbb{W}(\boldsymbol{\mu})^T s(\mathbf{x}), \tag{2.69}$$

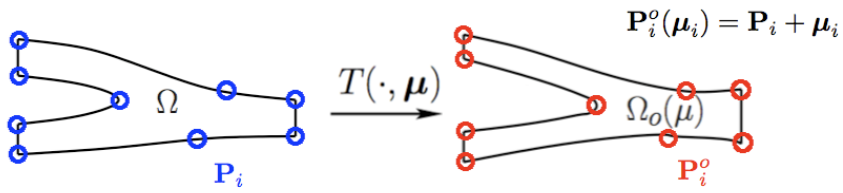



Figure 2.4: Schematic diagram of the RBF technique: on the left, the reference (or initial) configuration $\Omega$ and the unperturbed control points $\mathbf{X}_i$, $i = 1, \ldots, k$; on the right, the deformed (or target) configuration $\Omega_o(\boldsymbol{\mu})$ and the displaced control points $\mathbf{Y}_i$, $i = 1, \ldots, k$.

where the coefficients $\mathbf{c}(\boldsymbol{\mu})$, $\mathbb{A}(\boldsymbol{\mu})$, $\mathbb{W}(\boldsymbol{\mu})$ satisfy the constraints (2.67)-(2.68) and now depend on $\boldsymbol{\mu} \in \mathbb{R}^p$. As before, by denoting $\mathbb{M} \in \mathbb{R}^{k \times k}$ the interpolation matrix whose components $(\mathbb{M})_{ij} = M_{ij}$ are given by

$$M_{ij} = \Phi(\|\mathbf{P}_i - \mathbf{P}_j\|), \qquad 1 \leq i, j \leq k \tag{2.70}$$

and $\mathbb{P} \in \mathbb{R}^{k \times 3}$ the matrix defined by imposing the side constraints:

$$\mathbb{P} = \begin{bmatrix} 1 & P_{11} & P_{12} \\ \vdots & \vdots & \vdots \\ 1 & P_{k1} & P_{k2} \end{bmatrix}.$$

the coefficients $(\mathbf{c}(\boldsymbol{\mu}))_m = c_m(\boldsymbol{\mu})$, $(\mathbb{A}(\boldsymbol{\mu}))_{mn} = A_{mn}(\boldsymbol{\mu})$ and $(\mathbb{W}(\boldsymbol{\mu}))_{im} = w_{im}(\boldsymbol{\mu})$, for $1 \leq m, n \leq 2$, $1 \leq i \leq k$ appearing in the map (2.69) are obtained by solving the following linear systems:

$$\begin{bmatrix} \mathbb{M} & \mathbb{P} \\ \mathbb{P}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbb{W}_1(\boldsymbol{\mu}) \\ \boldsymbol{\gamma}_1(\boldsymbol{\mu}) \end{bmatrix} = \begin{bmatrix} \mathbf{P}_1^o(\boldsymbol{\mu}) \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} \mathbb{M} & \mathbb{P} \\ \mathbb{P}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbb{W}_2(\boldsymbol{\mu}) \\ \boldsymbol{\gamma}_2(\boldsymbol{\mu}) \end{bmatrix} = \begin{bmatrix} \mathbf{P}_2^o(\boldsymbol{\mu}) \\ 0 \end{bmatrix} \tag{2.71}$$

being

$$\mathbb{W}_m(\boldsymbol{\mu}) = (w_{1m}, (\boldsymbol{\mu}) \ldots, w_{km}(\boldsymbol{\mu}))^T, \qquad \mathbf{P}_m^o(\boldsymbol{\mu}) = (P_{1m}(\boldsymbol{\mu}), \ldots, P_{km}(\boldsymbol{\mu}))^T, \qquad m = 1, 2$$

and

$$\boldsymbol{\gamma}_1(\boldsymbol{\mu}) = (c_1(\boldsymbol{\mu}), A_{11}(\boldsymbol{\mu}), A_{21}(\boldsymbol{\mu}))^T, \qquad \boldsymbol{\gamma}_2(\boldsymbol{\mu}) = (c_2(\boldsymbol{\mu}), A_{12}(\boldsymbol{\mu}), A_{22}(\boldsymbol{\mu}))^T.$$

A representation of the mapping construction is shown in Fig. 2.4. An example of RBF mapping used to model deformations of a two-dimensional airfoil of the NACA family is shown in Fig. 2.5. In this case, the design parameters are given by the $p = 6$ vertical displacements of the control points located along the boundary sides.

Control points are usually chosen close to the boundary of the shape to be deformed (see Fig. 2.4), even if in our case the parametrization is constructed independently from the computational mesh. For a small number of control points, as in our approach, linear systems (2.71) can be efficiently solved by a suitable direct method (we remark that matrix factorization is not depending on the parameters). When using a large number of control points – as for example in fluid-structure interaction coupled problems or, more generally, when dealing with mesh motion through RBF – the matrix appearing in (2.71) may be badly conditioned and some difficulties may arise. In these cases, suitable scaling or preconditioning strategies may help, as discussed for example in [50].

In general, RBF maps provide more local control of the shape due to the ability to fine tune the shape by adding locally more control points and by choosing the type and support radius of the shape functions. The asymptotic behavior of $\Phi(h)$ is different among the selected basis functions, and the choice on the various possibilities is performed according to shape regularity and to convergence properties of the numerical method used to compute the coefficients appearing in (2.62). For example, the linear RBFs exhibit better convergence properties, while the cubic RBFs guarantee an enhanced shape smoothness.

RBFs can be divided into two groups, functions with compact support and functions with global support. Whenever $\Phi(\cdot)$ is an increasing function with global support, the influence of a center on an evaluation node increases with the distance of the two nodes. Thus, the global character of these functions tend to smooth out local effects, so that a large support radius yields a good approximation order. On the other hand, a full matrix system has to be solved[10]. Introduced by Wendland [319] and Wu [322], compactly supported RBFs attempt to reduce the bandwidth of the otherwise full interpolation matrix (2.70). However, in this case the choice of the support radius of the RBF functions, which is an additional parameter to be tuned, may become a critical issue. Increasing the support radius tends to produce ill-conditioned interpolation systems and lead to the Runge phenomenon [44]. On the other side, using a support radius that is too small yields a large interpolation error, even though the system can be more easily solved. To conclude, a drawback of the many different types of RBFs proposed in literature is that often the final shape is quite sensitive to the choice of the shape function and its support radius parameter.

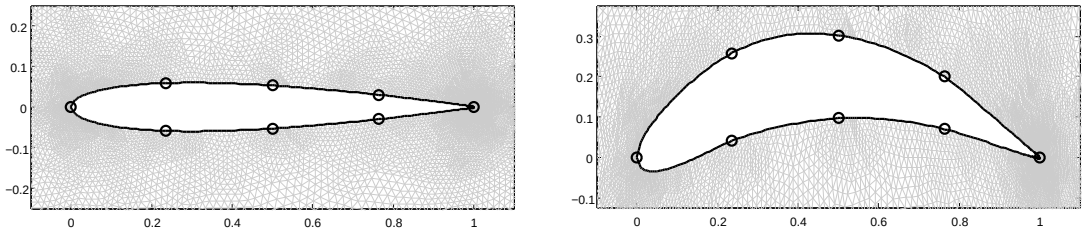



Figure 2.5: Example of shape deformation of a NACA0012 airfoil obtained through RBF mapping: reference configuration and control points (left), deformed configuration and control points (right).

Depending on the application at hand, we will discuss some possible choices of RBFs functions. In any case, the versatility of RBFs comes at a price of complexity in guaranteeing the quality of the remeshing interpolant.
Because of their excellent approximation properties, RBFs have been successfully applied to many different areas, such as computer graphics, mesh deformation [269, 69] and interpolation between nonmatching meshes in FSI computations [25, 70]. Another application of RBFs is in simulation based optimization, where they are used to build a surrogate model for less expensive black-box optimization; see e.g. [156] and references therein. In optimal design problems, RBFs have been used as shape parametrization for airfoil shape optimization in [217, 155], as well as for topological optimization of structure in [318, 317].

[10] Actually, this is not a big issue in our context, since $k$ is rather small – $k = \mathcal{O}(10)$ if we desire a low-dimensional shape parametrization – but may become somehow undesirable in applications of larger dimensions, such as mesh deformations or interpolation between non-matching meshes.

## 2.8 Other available options

In the previous sections we have illustrated just two paradigms: several extensions and modifications are possible. First of all, different FFD or RBF mappings can be recast in a domain decomposition framework, so that the parametric mapping is defined on a subregion close to the shape which has to be deformed.
In this respect, several tools to adaptively refine or coarsen the FFD lattice in some parts of the parametrized shape have been developed, as well as self-adaptive multilevel algorithms based on a geometrical hierarchy of nested parametrizations of Bézier type, for instance, in order to deal with *multi-level representations*, whenever interested in *multi-grid* optimization strategies. Different FFD and/or RBF maps can be defined on different subdomains and coupled in order to define piecewise a global map, even if this kind of procedures require careful handling and implies large efforts for the computation of the parametrized tensors involved in the parametrized formulation of the PDE problems.

On the other hand, FFD parametrizations can be built by considering more complex basis functions instead of Bernstein polynomials (also denoted as *Bézier parametrization* case): extensions of FFD to Non-Uniform and Rational B-Splines (NURBS) basis functions are also available [10, 290, 179, 104]. In this case, the FFD lattice is modelled by using B-splines functions instead of a $d$-variate tensor product of Bernstein polynomials; further, the distribution of the control points is not uniform, providing better local control of the object – this allows an exact representation of shapes given by conic sections, for instance – and local deformations can be implemented by locating the lattice on a specific part of the object.

Several other options for constructing different shape parametrizations are available, however we will focus on the volume-based representations discussed above. We just mention the *transfinite interpolation method*, used to build parametric mapping – the so-called *transfinite mapping* – within the reduced basis element method. The reduced basis element method is a domain decomposition method where the computational domain is decomposed into smaller blocks that are topologically similar to a few reference shapes, and the reduced solution is expressed as a gluing of basis functions defined on each local block (see e.g. [201, 202]).
The idea of the transfinite mapping is to define the interior points of the original domain as linear combinations of points on the boundaries [117]. In particular, each edge of the original domain $\Omega_o(\boldsymbol{\mu})$ is obtained as a one-to-one mapping of the corresponding edge on the reference domain $\Omega$, through a vector of geometrical parameters $\boldsymbol{\mu}$. In this way, we extend the possibility to act on the shape by moving its boundary already provided by *boundary shape parametrizations*, but defining a global map between domains. By introducing suitable weights and projection operators for each curvilinear side, the idea can be extended to general reference domains. More recent applications to flow problems in parametrized networks exploiting transfinite mappings can be found e.g. in [196, 195, 151].

# Analysis Part II

# 3 Reduced basis methods in computational fluid dynamics

In this chapter we present the reduced basis approximation and *a posteriori* error estimation methods for the steady Stokes and Navier-Stokes equations. After deriving the abstract parametrized formulation, we briefly recall the FE approximation and the main assumptions to ensure well-posedness at both continuous and discrete level. Then, we introduce the RB approximation, focusing on algebraic and approximation stability conditions, the Offline/Online decomposition and the treatment of nonaffine problems.
We provide new *a posteriori* error estimates for Stokes and Navier-Stokes solutions, based on the Babuška stability theory and on the Brezzi-Rappaz-Raviart theory, respectively. Moreover, we prove some error estimates for both (general) linear and (velocity) quadratic outputs in the Stokes case. We extend the Successive Constraint Method for the estimation of stability factors to the Stokes and Navier-Stokes cases, for which more insights are provided in Appendix A. Further, we address fluid dynamics problems in domains that are more general than those considered so far in the RB literature. Finally, we present several numerical test cases, in order to show some approximation features and computational performances.

## 3.1 Formulation of Stokes and Navier-Stokes problems

In this section we recall the necessary ingredients for the RB approximation of Stokes and Navier-Stokes problems, and related *a posteriori* error estimation. The Navier-Stokes equations provide a model of the flow motion for a viscous Newtonian incompressible fluid. In the steady case they can be stated as follows:

$$\begin{cases} -\nu\Delta\mathbf{v}_o + \delta(\mathbf{v}_o\cdot\nabla)\mathbf{v}_o + \nabla p_o = \mathbf{f}^o & \text{in } \Omega_o \\ \nabla\cdot\mathbf{v}_o = 0 & \text{in } \Omega_o \\ \mathbf{v}_o = \mathbf{0} & \text{on } \Gamma^o_{D_0} \\ \mathbf{v}_o = \mathbf{g}^o_D & \text{on } \Gamma^o_{D_g} \\ -p_o\mathbf{n}_o + \nu\dfrac{\partial\mathbf{v}_o}{\partial\mathbf{n}_o} = \mathbf{g}^o_N & \text{on } \Gamma^o_N, \end{cases} \tag{3.1}$$

where $(\mathbf{v}_o, p_o)$ are the velocity and the pressure fields defined on the original domain $\Omega_o$, for some given $\mathbf{f}^o$, $\mathbf{g}^o_D$, $\mathbf{g}^o_N$. Here $\Omega_o \subset \mathbb{R}^d$ for $d = 2, 3$ is assumed to be piecewise $C^2$ with convex corners, whereas the subscript/superscript "$o$" stands for "original". The first equation expresses the linear momentum conservation, the second one the mass conservation, which is also called the continuity equation.

Here $\mathbf{f}^o = (f_1^o, f_2^o)$ denotes a forcing term per unit mass, $\nu = \mu/\rho$ denotes the kinematic viscosity, where $\rho$ is the (constant) density and $\mu$ the dynamic viscosity of the fluid. Moreover, we define the Reynolds number as the non-dimensional ratio of convection to diffusion $\mathrm{Re} = L|\bar{\mathbf{v}}|/\nu$, where $L$ is a characteristic length of the domain $\Omega_o$ and $\bar{\mathbf{v}}$ a typical velocity of the flow. Navier-Stokes equations correspond to the case $\delta = 1$; we concentrate on *laminar* flows, with Reynolds number in the range $[1, 10^3]$, in two-dimensional domains ($d = 2$). If $\delta = 0$, the (quadratic) convective term is neglected, obtaining the steady Stokes equations, which provide a model in the case of slow motion of fluids with very high viscosity (low Reynolds number).

In what follows, we consider a partition $\partial\Omega_o = \Gamma_{D_0}^o \cup \Gamma_{D_g}^o \cup \Gamma_N^o$, then we prescribe homogeneous Dirichlet conditions on $\Gamma_{D_0}^o$, non-homogeneous Dirichlet conditions on $\Gamma_{D_g}^o$ and Neumann conditions on $\Gamma_N^o$, such that the Dirichlet portion is $\Gamma_D^o = \Gamma_{D_0}^o \cup \Gamma_{D_g}^o$; $\mathbf{n}_o$ is the normal unit vector to the boundary $\partial\Omega_o$. We denote the functional spaces for velocity and pressure fields by $V_o = (H^1_{0,\Gamma_D^o}(\Omega_o))^2$, $Q_o = L^2(\Omega_o)$, respectively, where $H^1_{0,\Gamma_D^o}(\Omega_o) = \{v \in H^1(\Omega_o) : v|_{\Gamma_D^o} = 0\}$. Moreover, we introduce the space $V_o^{\mathbf{g}} = \{\mathbf{v} \in (H^1(\Omega_o))^2 : \mathbf{v}|_{\Gamma_{D_0}^o} = \mathbf{0}, \ \ \mathbf{v}|_{\Gamma_{D_g}^o} = \mathbf{g}_D^o\}$. The corresponding abstract (or weak) form of Navier-Stokes equations (3.1) reads: find $(\mathbf{v}_o, p_o) \in V_o^{\mathbf{g}} \times Q_o$ such that

$$\begin{aligned}
\nu \int_{\Omega_o} \nabla\mathbf{v}_o : \nabla\mathbf{w} \, d\Omega_o \quad - \int_{\Omega_o} p_o \, \nabla\cdot\mathbf{w} \, d\Omega_o + \delta \int_{\Omega_o} (\mathbf{v}_o\cdot\nabla)\mathbf{v}_o\cdot\mathbf{w} \, d\Omega_o = \\
\int_{\Omega_o} \mathbf{f}^o\cdot\mathbf{w} \, d\Omega_o + \int_{\Gamma_N^o} \mathbf{g}_N^o\cdot\mathbf{w} \, d\Gamma_o, \qquad\qquad \forall\, \mathbf{w} \in V_o \\
\int_{\Omega_o} q\,\nabla\cdot\mathbf{v}_o \, d\Omega_o = 0, \qquad\qquad \forall\, q \in Q_o.
\end{aligned}$$

In order to incorporate the boundary conditions directly into the equations, we introduce a lift function $L_{o,\mathbf{g}} \in (H^1(\Omega_o))^2$ such that $L_{o,\mathbf{g}}|_{\Gamma_{D_g}^o} = \mathbf{g}_D^o$, $L_{o,\mathbf{g}}|_{\Gamma_{D_0}^o} = \mathbf{0}$ (extension of non-homogeneous boundary conditions to the interior of the domain) and denote $\hat{\mathbf{v}} = \mathbf{v} - L_{o,\mathbf{g}}$, so that $\hat{\mathbf{v}}|_{\Gamma_D} = \mathbf{0}$; for the sake of simplicity, we still denote $\hat{\mathbf{v}}$ with $\mathbf{v}$, as no ambiguity occurs. Hence, the abstract formulation of (3.1) reads: find $(\mathbf{v}_o, p_o) \in V_o \times Q_o$ such that, for all $\mathbf{w} \in V_o$ and $q \in Q_o$,

$$\begin{aligned}
\int_{\Omega_o} [\nu\nabla\mathbf{v}_o : \nabla\mathbf{w} - p_o\nabla\cdot\mathbf{w} + \delta(\mathbf{v}_o\cdot\nabla)\mathbf{v}_o\cdot\mathbf{w}] \, d\Omega_o + \delta\int_{\Omega_o} [(L_{o,\mathbf{g}}\cdot\nabla)\mathbf{v}_o\cdot\mathbf{w} + (\mathbf{v}_o\cdot\nabla)L_{o,\mathbf{g}}\cdot\mathbf{w}] \, d\Omega_o \\
= \int_{\Omega_o} \mathbf{f}^o\cdot\mathbf{w} \, d\Omega_o + \int_{\Gamma_N^o} \mathbf{g}_N^o\cdot\mathbf{w} \, d\Gamma_o - \int_{\Omega_o} \nu\nabla L_{o,\mathbf{g}} : \nabla\mathbf{w} \, d\Omega_o - \delta\int_{\Omega_o} (L_{o,\mathbf{g}}\cdot\nabla)L_{o,\mathbf{g}}\cdot\mathbf{w} \, d\Omega_o, \\
- \int_{\Omega_o} q\nabla\cdot\mathbf{v}_o \, d\Omega_o = \int_{\Omega_o} q\nabla\cdot L_{o,\mathbf{g}} \, d\Omega_o.
\end{aligned} \tag{3.2}$$

We assume to deal with two-dimensional, original domains $\Omega_o \subset \mathbb{R}^2$ made up of $K_{\mathrm{dom}}$ mutually nonoverlapping open subdomains $\{\Omega_o^k\}_{k=1}^{K_{\mathrm{dom}}}$, so that the bilinear and linear forms of the weak formulation, for $1 \le i, j \le 2$, read as follows:

$$a_o(\mathbf{v},\mathbf{w}) = \sum_{k=1}^{K_{\mathrm{dom}}} \int_{\Omega_o^k} \nu_{ij}^o \frac{\partial\mathbf{v}}{\partial x_i^o}\cdot\frac{\partial\mathbf{w}}{\partial x_j^o} \, d\Omega_o, \qquad b_o(q,\mathbf{w}) = -\sum_{k=1}^{K_{\mathrm{dom}}} \int_{\Omega_o^k} q\nabla\cdot\mathbf{w} \, d\Omega_o,$$

$$c_o(\mathbf{v},\mathbf{w},\mathbf{z}) = \sum_{k=1}^{K_{\mathrm{dom}}} \int_{\Omega_o^k} (\mathbf{v}\cdot\nabla)\mathbf{w}\cdot\mathbf{z} \, d\Omega_o, \qquad \langle F_s^o, \mathbf{w}\rangle = \sum_{k=1}^{K_{\mathrm{dom}}} \int_{\Omega_o^k} \mathbf{f}^o\cdot\mathbf{w} \, d\Omega_o + \sum_{k=1}^{\tilde{K}_{\mathrm{dom}}} \int_{\Gamma_N^{o,k}} \mathbf{g}^N\cdot\mathbf{w} \, d\Gamma_o,$$

denoting with $\nu_{ij}^o = \nu\delta_{ij}$ (where $\delta_{ij}$ indicates the Kronecker symbol) and with $\tilde{K}_{\mathrm{dom}}$ the number

of subportions in which the Neumann portion $\Gamma_N$ is divided through the domain decomposition, where $\Gamma_N^{o,k} = \partial\Omega_o^k \cap \Gamma_N^o$. Until stated otherwise, summation over repeated indices $i, j$ is understood. The terms due to non-homogeneous Dirichlet boundary conditions can thus be expressed as:

$$d_o(\mathbf{v}, \mathbf{w}) = c_o(L_{o,\mathbf{g}}, \mathbf{v}, \mathbf{w}) + c_o(\mathbf{v}, L_{o,\mathbf{g}}, \mathbf{w})$$

$$\langle F_d^o, \mathbf{w}\rangle = -a_o(L_{o,\mathbf{g}}, \mathbf{w}) - \delta c_o(L_{o,\mathbf{g}}, L_{o,\mathbf{g}}, \mathbf{w}), \qquad \langle G^o, q\rangle = -b_o(q, L_{o,\mathbf{g}}\rangle,$$

so that, denoting by $\langle F^o, \mathbf{w}\rangle = \langle F_s^o, \mathbf{w}\rangle + \langle F_d^o, \mathbf{w}\rangle$, the abstract formulation (3.2) can be rewritten, in a more compact form, as follows: find $(\mathbf{v}_o, p_o) \in V_o \times Q_o$ such that

$$\begin{cases} a_o(\mathbf{v}_o, \mathbf{w}) + \delta d_o(\mathbf{v}_o, \mathbf{w}) + b_o(p_o, \mathbf{w}) + \delta c_o(\mathbf{v}_o, \mathbf{v}_o, \mathbf{w}) = \langle F^o, \mathbf{w}\rangle\,, & \forall \mathbf{w} \in V_o \\ b_o(q, \mathbf{v}_o) = \langle G^o, q\rangle\,, & \forall q \in Q_o. \end{cases} \tag{3.3}$$

Starting from the abstract formulation (3.3), we derive in the next section the parametrized abstract formulation underlying our reduced basis approximation.

## 3.2 Parametrized abstract formulation and analysis

We now introduce an *input-parameter* vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)^T \in \mathcal{D}$ of $\mathbb{R}^p$, which may characterize either the geometrical configuration or physical properties, boundary data and sources[1].
Hence, we assume that the original domain can be expressed as $\Omega_o = \Omega_o(\boldsymbol{\mu})$ and it is obtained as the image of a reference domain $\Omega = \Omega_o(\boldsymbol{\mu}_{\text{ref}})$ through a parametrized mapping $T(\cdot; \boldsymbol{\mu}) : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}^2$. In the most general case, original and reference subdomains can be linked via either an affine or a nonaffine mapping $T^k(\cdot; \boldsymbol{\mu}) : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}^2$, such that $\Omega_o^k(\boldsymbol{\mu}) = T^k(\Omega^k; \boldsymbol{\mu})$, $1 \le k \le K_{\text{dom}}$ (see Sect. 2.5.2). For the sake of notation, we do not explicit the dependence on $\boldsymbol{\mu}$ in all the other *virtually* $\boldsymbol{\mu}$-dependent quantities as in Sect. 2.2.1.
We now derive the parametrized formulation by tracing (3.3) back on the reference domain $\Omega$. We point out that this operation is necessary when dealing with geometrical parameters. Otherwise, if parameters characterize just physical properties, we do not need to map the problem back to a reference configuration (i.e., in this case $\Omega \equiv \Omega_o$).

In what follows, we consider a partition $\partial\Omega = \Gamma_{D_0} \cup \Gamma_{D_g} \cup \Gamma_N$, homogeneous Dirichlet conditions on $\Gamma_{D_0}$, non-homogeneous Dirichlet conditions (with data $\mathbf{g}_D \in (H^{1/2}(\Gamma_D))^2$) on $\Gamma_{D_g}$ and Neumann conditions (with data $\mathbf{g}_N \in (H^{1/2}(\Gamma_N))^2$) on $\Gamma_N$, such that the Dirichlet portion is $\Gamma_D = \Gamma_{D_0} \cup \Gamma_{D_g}$; $\mathbf{n}$ is the normal unit vector to the boundary $\partial\Omega$. We introduce a lift function $L_{\mathbf{g}} \in (H^1(\Omega))^2$ such that $L_{\mathbf{g}}|_{\Gamma_{D_g}} = \mathbf{g}_D$, $L_{\mathbf{g}}|_{\Gamma_{D_0}} = \mathbf{0}$ and we denote the functional spaces for velocity and pressure fields by $V = (H^1_{0,\Gamma_D}(\Omega))^2$, $Q = L^2(\Omega)$, respectively, where $H^1_{0,\Gamma_D}(\Omega) = \{v \in H^1(\Omega) : v|_{\Gamma_D} = 0\}$.
We equip $V$ and $Q$ with the following notions of scalar products and corresponding norms:

$$\|\cdot\|_V = (\cdot,\cdot)_V^{1/2}, \qquad (\mathbf{v}, \mathbf{w})_V = (\nabla\mathbf{v}, \nabla\mathbf{w})_{(L^2(\Omega))^2}, \qquad \forall \mathbf{v}, \mathbf{w} \in X,$$

$$\|\cdot\|_Q = (\cdot,\cdot)_Q^{1/2}, \qquad (p, q)_Q = (p, q)_{L^2(\Omega)}, \qquad \forall p, q \in Q,$$

respectively. Then, by identifying $(\mathbf{v}(\boldsymbol{\mu}), p(\boldsymbol{\mu})) = (\mathbf{v}_o, p_o) \circ T(\cdot; \boldsymbol{\mu})$, we can express the problem as a system of parametrized PDEs by tracing (3.3) back on the reference domain $\Omega$, under the following abstract form:

[1] Here we do not make any distinction between geometrical and physical parameters – it will be made in Sect. 4.5, for the sake of the general analysis of the framework for our forward and inverse problems.

find $(\mathbf{v}, p) \in V \times Q$ such that

$$\begin{cases} a(\mathbf{v}(\boldsymbol{\mu}), \mathbf{w}; \boldsymbol{\mu}) + \delta d(\mathbf{v}(\boldsymbol{\mu}), \mathbf{w}; \boldsymbol{\mu}) + b(p(\boldsymbol{\mu}), \mathbf{w}; \boldsymbol{\mu}) + \delta c(\mathbf{v}(\boldsymbol{\mu}), \mathbf{v}(\boldsymbol{\mu}), \mathbf{w}; \boldsymbol{\mu}) = F(\mathbf{w}; \boldsymbol{\mu}), & \forall \mathbf{w} \in V \\ b(q, \mathbf{v}(\boldsymbol{\mu}); \boldsymbol{\mu}) = G(q; \boldsymbol{\mu}), & \forall q \in Q. \end{cases} \tag{3.4}$$

Here the parametrized bilinear and trilinear forms are defined as follows [255, 278, 282]:

$$a(\mathbf{v}, \mathbf{w}; \boldsymbol{\mu}) = \sum_{k=1}^{K_{\text{dom}}} \int_{\Omega^k} \frac{\partial \mathbf{v}}{\partial x_i} \nu_{ij}^k(\cdot; \boldsymbol{\mu}) \frac{\partial \mathbf{w}}{\partial x_j} d\Omega, \quad b(q, \mathbf{w}; \boldsymbol{\mu}) = -\sum_{k=1}^{K_{\text{dom}}} \int_{\Omega^k} q \chi_{ij}^k(\cdot; \boldsymbol{\mu}) \frac{\partial w_j}{\partial x_i} d\Omega, \tag{3.5}$$

$$c(\mathbf{v}, \mathbf{w}, \mathbf{z}; \boldsymbol{\mu}) = \sum_{k=1}^{K_{\text{dom}}} \int_{\Omega^k} v_i \, \chi_{ij}^k(\cdot; \boldsymbol{\mu}) \, \frac{\partial w_m}{\partial x_j} z_m \, d\Omega. \tag{3.6}$$

The transformation tensors for the bilinear viscous terms $\boldsymbol{\nu}^k : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}^{2\times 2}$, $1 \le k \le K_{\text{dom}}$, are defined as follows:

$$\boldsymbol{\nu}^k(\mathbf{x}; \boldsymbol{\mu}) = (J_T^k(\cdot; \boldsymbol{\mu}))^{-T} \boldsymbol{\nu}^o(\boldsymbol{\mu}) (J_T^k(\mathbf{x}; \boldsymbol{\mu}))^{-1} |J_T^k(\mathbf{x}; \boldsymbol{\mu})| \tag{3.7}$$

where $J_T^k : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}^{2\times 2}$ is the Jacobian matrix of the map $T^k(\cdot; \boldsymbol{\mu})$, and $|J_T^k(\mathbf{x}; \boldsymbol{\mu})| : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}$ its determinant. In the same way, the transformation tensors for the pressure/divergence terms and the nonlinear convective terms $\boldsymbol{\chi}^k : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}^2$, $1 \le k \le K_{\text{dom}}$, are given by

$$\boldsymbol{\chi}^k(\mathbf{x}; \boldsymbol{\mu}) = (J_T^k(\mathbf{x}; \boldsymbol{\mu}))^{-T} |J_T^k(\mathbf{x}; \boldsymbol{\mu})|. \tag{3.8}$$

The linear form corresponding to source terms is given by

$$F_s(\mathbf{w}; \boldsymbol{\mu}) = \sum_{k=1}^{K_{\text{dom}}} \int_{\Omega^k} \mathbf{f} \cdot \mathbf{w} \; |J_T^k(\cdot; \boldsymbol{\mu})| d\Omega + \sum_{k=1}^{\tilde{K}_{\text{dom}}} \int_{\Gamma_N^k} \mathbf{g}^N \cdot \mathbf{w} |J_T^k(\cdot; \boldsymbol{\mu}) \mathbf{t}| d\Gamma,$$

where $\mathbf{f} \in (L^2(\Omega))^2$ and $\mathbf{t}$ is the tangential unit vector to the boundary and $\Gamma_N^k = \partial\Omega^k \cap \Gamma_N$, while the terms due to non-homogeneous Dirichlet conditions on $\Gamma_D$ can be expressed as:

$$d(\mathbf{v}, \mathbf{w}; \boldsymbol{\mu}) = c(L_{\mathbf{g}}, \mathbf{v}, \mathbf{w}; \boldsymbol{\mu}) + c(\mathbf{v}, L_{\mathbf{g}}, \mathbf{w}; \boldsymbol{\mu}) \tag{3.9}$$

$$F_d(\mathbf{w}; \boldsymbol{\mu}) = -a(L_{\mathbf{g}}, \mathbf{w}) - \delta c(L_{\mathbf{g}}, L_{\mathbf{g}}, \mathbf{w}; \boldsymbol{\mu}), \qquad G(q; \boldsymbol{\mu}) = -b(q, L_{\mathbf{g}}; \boldsymbol{\mu}), \tag{3.10}$$

so that $F(\mathbf{w}; \boldsymbol{\mu}) = F_s(\mathbf{w}; \boldsymbol{\mu}) + F_d(\mathbf{w}; \boldsymbol{\mu})$.

Then, denoting by $X = X(\Omega)$ the product space given by $X = V \times Q$, by $Y(\boldsymbol{\mu}) = (\mathbf{v}(\boldsymbol{\mu}), p(\boldsymbol{\mu})) \in X$ and $W = (\mathbf{w}, q)$, the parametrized abstract formulation (3.4) can be rewritten in the following form, fitting the general parametrized formulation (1.2): find $Y(\boldsymbol{\mu}) = (\mathbf{v}(\boldsymbol{\mu}), p(\boldsymbol{\mu})) \in X(\Omega)$ s.t.

$$A(Y(\boldsymbol{\mu}), W; \boldsymbol{\mu}) + \delta C(Y(\boldsymbol{\mu}), Y(\boldsymbol{\mu}), W; \boldsymbol{\mu}) = \tilde{F}(W; \boldsymbol{\mu}), \quad \forall W \in X \tag{3.11}$$

where

$$A(Y, W; \boldsymbol{\mu}) = a(\mathbf{v}, \mathbf{w}; \boldsymbol{\mu}) + \delta d(\mathbf{v}, \mathbf{w}; \boldsymbol{\mu}) + b(p, \mathbf{w}; \boldsymbol{\mu}) + b(q, \mathbf{v}; \boldsymbol{\mu}), \tag{3.12}$$

$$C(Y, Y, W; \boldsymbol{\mu}) = c(\mathbf{v}, \mathbf{v}, \mathbf{w}; \boldsymbol{\mu}), \tag{3.13}$$

$$\tilde{F}(W; \boldsymbol{\mu}) = F(\mathbf{w}; \boldsymbol{\mu}) + G(q; \boldsymbol{\mu}). \tag{3.14}$$

Finally, we introduce two linear functionals $L_{\mathbf{v}}(\cdot;\boldsymbol{\mu}) \in V'$ and $L_p(\cdot;\boldsymbol{\mu}) \in Q'$ and a bilinear form $Q(\cdot,\cdot;\boldsymbol{\mu}) : V \times V \to \mathbb{R}$. We may then define the following output evaluation problem: given $\boldsymbol{\mu} \in \mathcal{D}$, evaluate the scalar output of interest

$$s(\boldsymbol{\mu}) = S(Y(\boldsymbol{\mu});\boldsymbol{\mu}) = L(Y(\boldsymbol{\mu});\boldsymbol{\mu}) + M(Y(\boldsymbol{\mu});\boldsymbol{\mu}), \tag{3.15}$$

where $L(Y(\boldsymbol{\mu});\boldsymbol{\mu}) = L_{\mathbf{v}}(\mathbf{v}(\boldsymbol{\mu});\boldsymbol{\mu}) + L_p(p(\boldsymbol{\mu});\boldsymbol{\mu})$ and $M(Y(\boldsymbol{\mu});\boldsymbol{\mu}) = Q(\mathbf{v}(\boldsymbol{\mu}),\mathbf{v}(\boldsymbol{\mu});\boldsymbol{\mu})$ are a linear functional in both velocity and pressure fields and a quadratic functional of the velocity fields, respectively, and $Y(\boldsymbol{\mu}) = (\mathbf{v}(\boldsymbol{\mu}), p(\boldsymbol{\mu})) \in X \times Q$ is the solution of (3.11).

Before addressing the analysis of the parametrized abstract formulation (3.11)-(3.15) we require that the linear, bilinear and trilinear forms introduced fulfill the assumption (crucial for our Offline/Online computational stratagem) of affine parametric dependence. Under the assumption that the transformation mappings are affine in the sense of (2.52), parametrized tensors depend just on the parameter $\boldsymbol{\mu}$ and in the Stokes/Navier-Stokes cases the affine decompositions (1.7)-(1.8) read as follows. For the bilinear forms we have, for some integers $Q_a$ and $Q_b$:

$$a(\mathbf{v},\mathbf{w};\boldsymbol{\mu}) = \sum_{q=1}^{Q_a} \Theta_a^q(\boldsymbol{\mu}) a^q(\mathbf{v},\mathbf{w}), \qquad b(q,\mathbf{w};\boldsymbol{\mu}) = \sum_{q=1}^{Q_b} \Theta_b^q(\boldsymbol{\mu}) b^q(q,\mathbf{w}); \tag{3.16}$$

where $q$ is a condensed index of $i,j,k$ quantities and, for $1 \le k \le K_{\text{dom}}$, $1 \le i,j \le 2$,

$$\Theta_a^{q(i,j,k)}(\boldsymbol{\mu}) = \nu_{ij}^k(\boldsymbol{\mu}), \qquad a^{q(i,j,r)}(\mathbf{v},\mathbf{w}) = \int_{\Omega^k} \frac{\partial \mathbf{v}}{\partial x_i}\frac{\partial \mathbf{w}}{\partial x_j} d\Omega, \tag{3.17}$$

$$\Theta_b^{q(i,j,k)}(\boldsymbol{\mu}) = \chi_{ij}^k(\boldsymbol{\mu}), \qquad b^{q(i,j,r)}(q,\mathbf{w}) = -\int_{\Omega^k} q\frac{\partial w_i}{\partial x_j} d\Omega. \tag{3.18}$$

For the trilinear form and the source term (e.g. in the case $\mathbf{g}^N = \mathbf{0}$), we have instead:

$$c(\mathbf{v},\mathbf{w},\mathbf{z};\boldsymbol{\mu}) = \sum_{q=1}^{Q_c} \Theta_c^q(\boldsymbol{\mu}) c^q(\mathbf{v},\mathbf{w},\mathbf{z}), \qquad F_s(\mathbf{w};\boldsymbol{\mu}) = \sum_{q'=1}^{Q_s} \Theta_s^{q'}(\boldsymbol{\mu}) F_s^{q'}(\mathbf{w}), \tag{3.19}$$

for some integers $Q_c$ and $Q_s$, where $q$ and $q'$ are condensed indexes of $i,j,k$ quantities and $k$ quantities, respectively, and

$$\Theta_c^{q(i,j,k)}(\boldsymbol{\mu}) = \chi_{ij}^k(\boldsymbol{\mu}), \qquad c^{q(i,j,k)}(\mathbf{v},\mathbf{w},\mathbf{z}) = \int_{\Omega^k} v_i \frac{\partial w_k}{\partial x_j} z_k \; d\Omega,$$

$$\Theta_s^{q'(k)}(\boldsymbol{\mu}) = |J_T^k(\cdot;\boldsymbol{\mu})|, \qquad F_s^{q'(k)}(\mathbf{w}) = \int_{\Omega^k} \mathbf{f}\cdot\mathbf{w} d\Omega.$$

On the other hand, if the parametric map $T(\cdot;\boldsymbol{\mu})$ is nonaffine, parametrized tensors are function of both spatial coordinates $\mathbf{x}$ and parameter $\boldsymbol{\mu}$. Hence, we rely on EIM (see Sect. 2.4) in order to recover the affinity assumption: by assuming for the sake of simplicity to deal with a global map ($K_{\text{dom}} = 1$), each component of $\boldsymbol{\nu} : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}^{2\times 2}$ and $\boldsymbol{\chi} : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}^{2\times 2}$ has to be approximated by an affine expansion of the form

$$\nu_{ij}(\mathbf{x},\boldsymbol{\mu}) = \sum_{k=1}^{K_{ij}^a} \beta_k^{i,j}(\boldsymbol{\mu})\xi_k^{i,j}(\mathbf{x}) + \varepsilon_{i,j}^a(\mathbf{x};\boldsymbol{\mu}), \qquad \chi_{ij}(\mathbf{x},\boldsymbol{\mu}) = \sum_{k=1}^{K_{ij}^b} \gamma_k^{i,j}(\boldsymbol{\mu})\eta_k^{i,j}(\mathbf{x}) + \varepsilon_{i,j}^b(\mathbf{x};\boldsymbol{\mu}), \tag{3.20}$$

for $1 \le i,j \le 2$, where $K_{ij}^a$ and $K_{ij}^b$ represent the total number of terms obtained for each tensorial component after the empirical interpolation procedure. The same expansion (made now by $K^s$ terms) is set up for the tensor appearing at the right-hand-side of velocity equation:

$$|\det(J_T(\mathbf{x};\boldsymbol{\mu}))| = \sum_{k=1}^{K^s} \delta_k(\boldsymbol{\mu})\psi_k(\mathbf{x}) + \varepsilon^s(\mathbf{x};\boldsymbol{\mu}). \tag{3.21}$$

All the coefficients $\beta_k^{i,j}$'s, $\gamma_k^{i,j}$'s, $\delta_k$'s, $\xi_k^{i,j}$'s, $\eta_k^{i,j}$'s and $\psi_k$'s are efficiently computable scalar functions and the error terms are guaranteed to be under some prescribed tolerance:

$$\|\varepsilon_{i,j}^{(a,b)}(\cdot;\boldsymbol{\mu})\|_\infty \leq \varepsilon_{tol}^{EIM}, \quad \|\varepsilon_i(\cdot;\boldsymbol{\mu})\|_\infty \leq \varepsilon_{tol}^{EIM}, \quad \forall \boldsymbol{\mu} \in \mathcal{D}.$$

In this way, we can recover the affine expansions (3.16)-(3.19) by setting

$$a^{q(i,j,k)}(\mathbf{v},\mathbf{w}) = \int_\Omega \xi_k^{i,j}(\mathbf{x}) \frac{\partial \mathbf{v}}{\partial x_i}\frac{\partial \mathbf{w}}{\partial x_j} d\Omega, \qquad b^{q(i,j,k)}(p,\mathbf{w}) = -\int_\Omega \eta_k^{i,j}(\mathbf{x}) p \frac{\partial \mathbf{w}_i}{\partial x_j} d\Omega, \tag{3.22}$$

$$c^{q(i,j,k)}(\mathbf{v},\mathbf{w},\mathbf{z}) = \int_{\Omega^k} v_i \eta_k^{i,j}(\mathbf{x}) \frac{\partial w_k}{\partial x_j} z_k \; d\Omega, \qquad F_s^{q(k)}(\mathbf{w}) = \int_{\Omega^k} \psi_k(\mathbf{x}) \mathbf{f}\cdot\mathbf{w} d\Omega, \tag{3.23}$$

with $\Theta_a^q(\boldsymbol{\mu}) = \beta_k^{i,j}(\boldsymbol{\mu})$, $\Theta_b^q(\boldsymbol{\mu}) = \Theta_c^q(\boldsymbol{\mu}) = \gamma_k^{i,j}(\boldsymbol{\mu})$, $\Theta_s^q(\boldsymbol{\mu}) = \delta_k(\boldsymbol{\mu})$.
In the end, in order to obtain the affine expansion for the problem (3.11) fitting the general form (1.7)-(1.8), it is sufficient to sum the affine expansions corresponding to each term appearing in the expressions (3.12)-(3.14). We remark that if the output (3.15) can be expressed through a similar affine expansion, it can be evaluated in a very cheap way; its treatment is very similar to the linear and bilinear forms already introduced and is omitted here.

In the next subsections we analyze the well-posedness of the parametrized abstract formulation (3.11) starting from the simpler Stokes case. We also discuss the stability of the truth approximation problem obtained by the finite element method, highlighting some features which turn to be important also for the proper formulation of the RB approximation.

### 3.2.1 Well-posedness of parametrized Stokes equations

The analysis of the well-posedness for the Stokes problem is usually recast in the framework of *saddle point problems*. Let us suppose that the bilinear forms $a(\cdot,\cdot;\boldsymbol{\mu}) : V \times V \to \mathbb{R}$ and $b(\cdot,\cdot;\boldsymbol{\mu}) : V \times V \to \mathbb{R}$ are continuous:

$$\gamma_a(\boldsymbol{\mu}) = \sup_{\mathbf{v}\in V}\sup_{\mathbf{w}\in V} \frac{a(\mathbf{v},\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{v}\|_V\|\mathbf{w}\|_V} < +\infty, \quad \forall\, \boldsymbol{\mu} \in \mathcal{D}, \tag{3.24}$$

$$\gamma_b(\boldsymbol{\mu}) = \sup_{q\in Q}\sup_{\mathbf{w}\in V} \frac{b(q,\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{w}\|_V\|q\|_Q} < +\infty, \quad \forall\, \boldsymbol{\mu} \in \mathcal{D}. \tag{3.25}$$

Under a suitable stability condition over the two bilinear forms, the Stokes problem (3.4, with $\delta = 0$) is well-posed, as stated by the following theorem (Brezzi [47]):

**Theorem 3.1.** *Given the Hilbert spaces $V$ and $Q$ and the linear functionals $F(\cdot;\boldsymbol{\mu}) \in V'$, $G(\cdot;\boldsymbol{\mu}) \in Q'$, let us assume that the bilinear forms $a(\cdot,\cdot;\boldsymbol{\mu}) : V\times V \to \mathbb{R}$ and $b(\cdot,\cdot;\boldsymbol{\mu}) : Q\times V \to \mathbb{R}$ are continuous, and that $F(\cdot;\boldsymbol{\mu})$ and $G(\cdot;\boldsymbol{\mu})$ are continuous for any $\boldsymbol{\mu} \in \mathcal{D}$. Moreover, let us assume that the bilinear forms satisfy the following assumptions:*

1. *$a(\cdot,\cdot;\boldsymbol{\mu})$ is weakly coercive over $V_0 = \{\mathbf{w} \in V : b(q,\mathbf{w};\boldsymbol{\mu}) = 0,\ \forall q \in Q\} \subset VX$, i.e.*

$$\exists\, \alpha^{LB}(\boldsymbol{\mu}) > 0 :\ \alpha(\boldsymbol{\mu}) = \inf_{\mathbf{v}\in V_0}\sup_{\mathbf{w}\in V_0} \frac{a(\mathbf{v},\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{v}\|_V\|\mathbf{w}\|_V} = \inf_{\mathbf{w}\in V_0}\sup_{\mathbf{v}\in V_0} \frac{a(\mathbf{v},\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{v}\|_V\|\mathbf{w}\|_V} \geq \alpha^{LB}(\boldsymbol{\mu}),\ \ \forall\, \boldsymbol{\mu} \in \mathcal{D};$$

*2.* $b(\cdot,\cdot;\boldsymbol{\mu})$ *is inf-sup stable over* $X \times Q$*, i.e.*

$$\exists\ \beta_{Br}^{LB}(\boldsymbol{\mu}) > 0 :\ \beta(\boldsymbol{\mu}) = \inf_{q\in Q}\sup_{\mathbf{w}\in X}\frac{b(q,\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{w}\|_V\|q\|_Q} \geq \beta_{Br}^{LB}(\boldsymbol{\mu}),\quad \forall\ \boldsymbol{\mu}\in\mathcal{D}. \tag{3.26}$$

*Then there exists a unique solution* $(\mathbf{v}(\boldsymbol{\mu}), p(\boldsymbol{\mu})) \in V \times Q$ *to problem* (3.4) *(with* $\delta = 0$*), for any* $\boldsymbol{\mu}\in\mathcal{D}$*. Furthermore, the following a priori estimates hold:*

$$\|\mathbf{v}(\boldsymbol{\mu})\|_V \leq \frac{1}{\alpha^{LB}(\boldsymbol{\mu})}\left[\|F(\cdot;\boldsymbol{\mu})\|_{V'} + \frac{\alpha^{LB}(\boldsymbol{\mu})+\gamma_a(\boldsymbol{\mu})}{\beta_{Br}^{LB}(\boldsymbol{\mu})}\|G(\cdot;\boldsymbol{\mu})\|_{Q'}\right], \tag{3.27}$$

$$\|p(\boldsymbol{\mu})\|_Q \leq \frac{1}{\beta_{Br}^{LB}(\boldsymbol{\mu})}\left[\left(1+\frac{\gamma_a(\boldsymbol{\mu})}{\alpha^{LB}(\boldsymbol{\mu})}\right)\|F(\cdot;\boldsymbol{\mu})\|_{V'} + \frac{\gamma_a(\boldsymbol{\mu})(\alpha^{LB}(\boldsymbol{\mu})+\gamma_a(\boldsymbol{\mu}))}{\alpha^{LB}(\boldsymbol{\mu})\beta_{Br}^{LB}(\boldsymbol{\mu})}\|G(\cdot;\boldsymbol{\mu})\|_{Q'}\right].$$

**Remark 3.2.** *In the following, we assume that* $a(\cdot,\cdot;\boldsymbol{\mu})$ *satisfies the (more general) coercivity property over* $V$*:*

$$\exists\ \alpha^{LB}(\boldsymbol{\mu}) > 0 :\ \alpha(\boldsymbol{\mu}) = \inf_{\mathbf{w}\in V}\frac{a(\mathbf{w},\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{w}\|_V^2} \geq \alpha^{LB}(\boldsymbol{\mu}),\quad \forall\ \boldsymbol{\mu}\in\mathcal{D}. \tag{3.28}$$

**Remark 3.3.** *The result stated in Theorem 3.1 provides the classical framework – to which we refer as* Brezzi inf-sup theory *– for mixed variational problems. A second equivalent inf-sup stability theory has been formulated by Babuška [17] directly on the formulation* (3.11)*.*
*Considering the bilinear form* $A(\cdot,\cdot;\boldsymbol{\mu}) : X\times X \to \mathbb{R}$ *defined in* (3.12)*, the* Babuška inf-sup theory *states that the problem* (3.11) *(with* $\delta = 0$*) is well posed if and only if the following continuity condition:*

$$\gamma_A(\boldsymbol{\mu}) = \sup_{V\in X}\sup_{W\in X}\frac{A(V,W;\boldsymbol{\mu})}{\|V\|_X\|W\|_X} < +\infty,\quad \forall\ \boldsymbol{\mu}\in\mathcal{D}, \tag{3.29}$$

*and the following (Babuška) inf-sup condition:*

$$\exists\ \beta_{Ba}^{LB}(\boldsymbol{\mu}) > 0 :\ \beta_{Ba}(\boldsymbol{\mu}) = \inf_{V\in X}\sup_{W\in X}\frac{A(V,W;\boldsymbol{\mu})}{\|V\|_X\|W\|_X} \geq \beta_{Ba}^{LB}(\boldsymbol{\mu}),\quad \forall\ \boldsymbol{\mu}\in\mathcal{D} \tag{3.30}$$

*hold. Furthermore, the unique solution of* (3.11) *satisfies*

$$\|Y(\boldsymbol{\mu})\|_X \leq \frac{\|\tilde{F}(\cdot;\boldsymbol{\mu})\|_{X'}}{\beta_{Ba}^{LB}(\boldsymbol{\mu})}. \tag{3.31}$$

*In this way, the Babuška theory can be seen as a generalization of the Lax-Milgram result for the Galerkin-type formulation to the Petrov-Galerkin case – to which we refer sometimes as* Nečas theorem*; its application to the Stokes problem is just a possible use.*
*We also remark that from the Brezzi estimates* (3.27) *it is possible to derive a global estimate*

$$\|(\mathbf{v},p)(\boldsymbol{\mu})\|_{V\times Q} \leq \mathcal{K}_{Br}(\alpha^{-1}(\boldsymbol{\mu}),\beta_{Br}^{-1}(\boldsymbol{\mu}),\gamma_a(\boldsymbol{\mu}))\|(F(\boldsymbol{\mu}),G(\boldsymbol{\mu}))\|_{V'\times Q'} \tag{3.32}$$

*where* $\mathcal{K}_{Br} = \mathcal{K}_{Br}(\alpha^{-1},\beta_{BR}^{-1},\gamma_a)$ *is such that*

$$\beta_{Ba}(\boldsymbol{\mu}) \geq \frac{1}{\mathcal{K}_{Br}(\alpha^{-1},\beta_{BR}^{-1},\gamma_a)}. \tag{3.33}$$

*Moreover, it is possible to show that* $\beta_{Br}(\boldsymbol{\mu}) \geq \beta_{Ba}(\boldsymbol{\mu})$ *and* $\alpha(\boldsymbol{\mu}) \geq \beta_{Ba}(\boldsymbol{\mu})$ *(see e.g. [324, 77]).*

We exploit these results not only for stating the well-posedness of the parametrized Stokes problem, but also in order to create a general framework to compute error bounds for noncoercive problems under mixed variational form solved by the RB method, as illustrated in Sect. 3.4.1. In particular, our stability analysis will be based on the *Brezzi stability theory*, while error analysis and certification will be based on the *Babuška stability theory*. In particular, *a posteriori* error estimates on velocity and pressure based on the Babuška stability theory require the evaluations of a smaller amount of factors (see remark ).
In the end, concerning the well-posedness of the output evaluation problem, we require that the linear forms $L_{\mathbf{v}}(\cdot;\boldsymbol{\mu}) \in V'$ and $L_p(\cdot;\boldsymbol{\mu}) \in Q'$ are continuous on $V$ and $Q$ for any $\boldsymbol{\mu} \in \mathcal{D}$, respectively, and that the bilinear form $Q(\cdot,\cdot;\boldsymbol{\mu}) : V \times V \to \mathbb{R}$ is coercive over $V$, in the sense stated by (3.28). For the most common quadratic outputs of the velocity field, such as energy, vorticity, etc., this bilinear form turns to be also symmetric.

### 3.2.2 Well-posedness of parametrized Navier-Stokes equations

The proof of existence and uniqueness of solutions of Navier-Stokes problem is much more involved, due to its nonlinear nature. Here we recall the main results necessary to set our RB approximation and corresponding *a posteriori* error estimation, following the two paradigms introduced for the analysis of the Stokes problem (*Brezzi theory* vs. *Babuška theory*). General results about the well-posedness of Navier-Stokes equations can be found in the books by Ladyzhenskaya [178], Temam [307], Girault and Raviart [115]. A common assumption that enables theoretical analysis concerns the existence of a *local branch of non-singular solutions*, see for example [115],

In addition to assumptions (3.24)-(3.25), thanks to *Hölder* inequality and *Sobolev* embedding theorems (assuring that $H^1(\Omega) \subset L^4(\Omega)$ for $d = 2, 3$; see e.g. [307], Chapter 2, Sect. 1.1), the trilinear form $c(\cdot,\cdot,\cdot;\boldsymbol{\mu}) : V \times V \times V \to \mathbb{R}$ satisfies the following continuity condition:

$$\gamma_c(\boldsymbol{\mu}) = \sup_{\mathbf{v}\in V} \sup_{\mathbf{w}\in V} \sup_{\mathbf{z}\in V} \frac{c(\mathbf{v},\mathbf{w},\mathbf{z};\boldsymbol{\mu})}{\|\mathbf{v}\|_V \|\mathbf{w}\|_V \|\mathbf{z}\|_V} < +\infty, \quad \forall\ \boldsymbol{\mu} \in \mathcal{D}. \tag{3.34}$$

In particular the continuity constant $\gamma_c(\boldsymbol{\mu})$ can be expressed as:

$$\gamma_c(\boldsymbol{\mu}) = \sqrt{2}\rho^2 \max_{q=1,\dots,Q_c} \|\Theta_q^c(\boldsymbol{\mu})\|_{L^\infty(\mathcal{D})} \max_{q=1,\dots,Q_c} \|\eta^q\|_{L^\infty(\Omega)} \tag{3.35}$$

where $\rho = \rho(\Omega)$ is the Sobolev embedding constant

$$\rho^2 = \sup_{v\in V} \frac{\|v\|^2_{L^4(\Omega)}}{(v,v)_H}. \tag{3.36}$$

for any $v \in H$ such that $H_0^1(\Omega) \subset H \subset H^1(\Omega)$ and $\|w\|_{L^p(\Omega)} = \left(\int_\Omega |w|^p\right)^{1/p}$. The proof of this relationship is a slight variation of the usual continuity proof for the trilinear form of Navier-Stokes operator (see e.g. [307], Chapter 2, Sect. 1.2), taking into account the additional terms arising from the parametrized formulation. In particular, $\eta^q$ is defined in (3.20) – $q$ is a condensed index of $i, j, k$ – and it is such that $\|\eta^q\|_{L^\infty(\Omega)} = 1$ in the affine case.

In order to state an existence result, let us remark that the trilinear form is such that $c(\mathbf{v};\mathbf{v};\mathbf{v};\boldsymbol{\mu}) = 0$ for any $\mathbf{v} \in (H_0^1(\Omega))^2$. In this case, provided that $a(\cdot,\cdot;\boldsymbol{\mu})$ is coercive over $V$ and $b(\cdot,\cdot;\boldsymbol{\mu})$ is inf-sup stable over $V \times Q$, the problem admits a solution (see e.g. [307], Chapter 2, Thm. 1.2). In the more general case of non-homogeneous problems, recalling that $d(\mathbf{v},\mathbf{v};\boldsymbol{\mu}) = c(\mathbf{v},L_{\mathbf{g}},\mathbf{v};\boldsymbol{\mu}) + c(L_{\mathbf{g}},\mathbf{v},\mathbf{v};\boldsymbol{\mu})$, an additional assumption on $c(\cdot,\cdot,\cdot;\boldsymbol{\mu})$ has to be fulfilled, as stated by the following result, based on a classical stability and uniqueness result (for the proof, see e.g. [307], Chapter 2, Thm. 1.6):

**Theorem 3.4.** *Given the Hilbert spaces $V$ and $Q$ and the linear functionals $F(\cdot;\boldsymbol{\mu}) \in V'$, $G(\cdot;\boldsymbol{\mu}) \in Q'$, let us assume that the bilinear forms $a(\cdot,\cdot;\boldsymbol{\mu}) : V \times V \to \mathbb{R}$ and $b(\cdot,\cdot;\boldsymbol{\mu}) : Q \times V \to \mathbb{R}$ are continuous in the sense of* (3.24)-(3.25) *and stable in the sense of* (3.28)-(3.26)*, and that $F(\cdot;\boldsymbol{\mu})$ and $G(\cdot;\boldsymbol{\mu})$ are continuous for any $\boldsymbol{\mu} \in \mathcal{D}$. Moreover, assume that the trilinear form $c(\cdot,\cdot,\cdot;\boldsymbol{\mu})$ is continuous in the sense of* (3.34) *and such that, for any $\mathbf{w} \in V$, $\boldsymbol{\mu} \in \mathcal{D}$,*

$$|c(\mathbf{w}, L_{\mathbf{g}}, \mathbf{w}; \boldsymbol{\mu})| \leq \frac{\alpha(\boldsymbol{\mu})}{2}\|\mathbf{w}\|_V^2, \quad |c(L_{\mathbf{g}}, \mathbf{w}, \mathbf{w}; \boldsymbol{\mu})| \leq \frac{\alpha(\boldsymbol{\mu})}{2}\|\mathbf{w}\|_V^2, \tag{3.37}$$

*so that the following coercivity property is satisfied:*

$$\exists\ \tilde{\alpha}^{LB}(\boldsymbol{\mu}) > 0 :\ \tilde{\alpha}(\boldsymbol{\mu}) = \inf_{\mathbf{w}\in V} \frac{a(\mathbf{w},\mathbf{w};\boldsymbol{\mu}) + d(\mathbf{w},\mathbf{w};\boldsymbol{\mu}) + c(\mathbf{w},\mathbf{w},\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{w}\|_V^2} \geq \tilde{\alpha}^{LB}(\boldsymbol{\mu}),\ \ \forall\ \boldsymbol{\mu} \in \mathcal{D}. \tag{3.38}$$

*Then* (3.4) *with $\delta = 1$ admits a solution $(\mathbf{v}(\boldsymbol{\mu}), p(\boldsymbol{\mu})) \in V \times Q$, for any $\boldsymbol{\mu} \in \mathcal{D}$. Moreover, if*

$$\frac{4\,\gamma_c(\boldsymbol{\mu})}{\tilde{\alpha}^{LB}(\boldsymbol{\mu})}\|F(\cdot;\boldsymbol{\mu})\|_{(H^{-1}(\Omega))^2} < 1, \tag{3.39}$$

*then there exists a unique solution $(\mathbf{v}(\boldsymbol{\mu}), p(\boldsymbol{\mu})) \in V \times Q$ to* (3.4) *(with $\delta = 1$), for any $\boldsymbol{\mu} \in \mathcal{D}$.*

**Remark 3.5.** *The result stated in Theorem 3.4 provides a possible framework for the analysis of steady Navier-Stokes equations under the* small data *assumption. The more general Brezzi-Rappaz-Raviart (BRR) theory – see e.g. [49, 56] – provides the required tools for the analysis of a wider class of nonlinear equations. In the same way as the* Babuška theory *for mixed variational problem, it is based on the global operator $\tilde{A}(\cdot,\cdot;\boldsymbol{\mu}) : X \times X \to \mathbb{R}$*

$$\tilde{A}(V, W; \boldsymbol{\mu}) = A(V, W; \boldsymbol{\mu}) + C(V, V, W; \boldsymbol{\mu}), \tag{3.40}$$

*and its Fréchet derivative with respect to the first variable; at $V \in X$, it can be expressed as*

$$d\tilde{A}(Y; \boldsymbol{\mu})(V, W) = A(V, W; \boldsymbol{\mu}) + C(Y, V, W; \boldsymbol{\mu}) + C(V, Y, W; \boldsymbol{\mu}), \tag{3.41}$$

*and it is such that*

$$\tilde{A}(V + Z, W; \boldsymbol{\mu}) = \tilde{A}(V, W; \boldsymbol{\mu}) + d\tilde{A}(Z; \boldsymbol{\mu})(V, W) + C(Z, Z, W; \boldsymbol{\mu}), \qquad \forall\ V,\ W,\ Z \in X,$$

*being $A(\cdot,\cdot;\boldsymbol{\mu})$ and $C(\cdot,\cdot,\cdot;\boldsymbol{\mu})$ the operators defined by* (3.12)-(3.13)*. In particular, the problem* (3.11) *(with $\delta = 1$) is well posed if and only if the following continuity condition:*

$$\gamma_{\tilde{A}}(\boldsymbol{\mu}) = \sup_{V\in X}\sup_{W\in X} \frac{d\tilde{A}(Y(\boldsymbol{\mu}); \boldsymbol{\mu})(V, W)}{\|V\|_X \|W\|_X} < +\infty, \quad \forall\ \boldsymbol{\mu} \in \mathcal{D}, \tag{3.42}$$

*and the following (Babuška) inf-sup condition:*

$$\exists\ \beta_{\tilde{A}}^{LB}(\boldsymbol{\mu}) > 0 :\ \beta_{\tilde{A}}(\boldsymbol{\mu}) = \inf_{V\in X}\sup_{W\in X} \frac{d\tilde{A}(Y(\boldsymbol{\mu}); \boldsymbol{\mu})(V, W)}{\|V\|_X \|W\|_X} \geq \beta_{\tilde{A}}^{LB}(\boldsymbol{\mu}), \quad \forall\ \boldsymbol{\mu} \in \mathcal{D} \tag{3.43}$$

*holds. Under these assumptions, the solution in the neighborhood of $Y(\boldsymbol{\mu})$ is unique. Further details about this framework will be provided in Sect. 3.4.4, where it will be exploited to derive an a posteriori error bound for the RB approximation.*

### 3.2.3 Stability of the FE truth approximation

We briefly recall the conditions ensuring the stability of the FE truth approximation lying over our RB approximation. Let $V^{\mathcal{N}} \subset V$ and $Q^{\mathcal{N}} \subset Q$ be two subspaces of $V$ and $Q$, of dimension

$\mathcal{N}_V < +\infty$ and $\mathcal{N}_Q < +\infty$, for velocity and pressure, respectively. The dimension of the FE spaces is thus taken large enough in order to neglect the differences $\|\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}) - \mathbf{v}(\boldsymbol{\mu})\|_V$ and $\|p^{\mathcal{N}}(\boldsymbol{\mu}) - p(\boldsymbol{\mu})\|_Q$, so that it can be effectively considered as a "truth" approximation. The truth FE approximation reads as follows: given $\boldsymbol{\mu} \in \mathcal{D}$, evaluate the scalar output of interest $s^{\mathcal{N}}(\boldsymbol{\mu}) = S(Y^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu})$ where $Y^{\mathcal{N}}(\boldsymbol{\mu}) = (\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}), p^{\mathcal{N}}(\boldsymbol{\mu})) \in X^{\mathcal{N}} := V^{\mathcal{N}} \times Q^{\mathcal{N}}$ is such that

$$A(Y^{\mathcal{N}}(\boldsymbol{\mu}), W^{\mathcal{N}}; \boldsymbol{\mu}) + \delta C(Y^{\mathcal{N}}(\boldsymbol{\mu}), Y^{\mathcal{N}}(\boldsymbol{\mu}), W^{\mathcal{N}}; \boldsymbol{\mu}) = \tilde{F}(W^{\mathcal{N}}; \boldsymbol{\mu}), \quad \forall\, W^{\mathcal{N}} \in X^{\mathcal{N}} \tag{3.44}$$

or, equivalently,

$$\begin{cases} a(\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}), \mathbf{w}^{\mathcal{N}}; \boldsymbol{\mu}) + \delta d(\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}), \mathbf{w}^{\mathcal{N}}; \boldsymbol{\mu}) + b(p^{\mathcal{N}}(\boldsymbol{\mu}), \mathbf{w}^{\mathcal{N}}; \boldsymbol{\mu}) & \\ \qquad\qquad + \delta c(\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}), \mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}), \mathbf{w}^{\mathcal{N}}; \boldsymbol{\mu}) = F(\mathbf{w}^{\mathcal{N}}; \boldsymbol{\mu}), & \forall \mathbf{w}^{\mathcal{N}} \in V^{\mathcal{N}} \\ b(q^{\mathcal{N}}, \mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu}) = G(q^{\mathcal{N}}; \boldsymbol{\mu}), & \forall q \in Q^{\mathcal{N}}. \end{cases} \tag{3.45}$$

All the forms are continuous over the discrete spaces $V^{\mathcal{N}}$ and $Q^{\mathcal{N}}$. In particular, a discrete version of the Sobolev embedding result can be stated (see e.g. [307]), where

$$\rho^2_{\mathcal{N}} = \sup_{v \in V^{\mathcal{N}}} \frac{\|v\|^2_{L^4(\Omega)}}{(v,v)_H}. \tag{3.46}$$

is the discrete version of (3.36), such that $\rho^2_{\mathcal{N}} \leq \rho^2$. Instead, approximation stability is ensured by imposing that the coercivity and inf-sup conditions are still valid at the discrete level.

In the Stokes case, (3.45) admits a unique solution provided that $a(\cdot,\cdot;\boldsymbol{\mu})$ is coercive over $V^{\mathcal{N}}$:

$$\exists\ \alpha^{LB}_{\mathcal{N}}(\boldsymbol{\mu}) > 0 :\ \alpha^{\mathcal{N}}(\boldsymbol{\mu}) = \inf_{\mathbf{v} \in V^{\mathcal{N}}} \frac{a(\mathbf{v},\mathbf{v};\boldsymbol{\mu})}{\|\mathbf{v}\|^2_V} \geq \alpha^{LB}_{\mathcal{N}}(\boldsymbol{\mu}), \quad \forall\ \boldsymbol{\mu} \in \mathcal{D}, \tag{3.47}$$

and $b(\cdot,\cdot;\boldsymbol{\mu})$ is inf-sup stable over $V^{\mathcal{N}} \times Q^{\mathcal{N}}$, so that the following discrete Brezzi inf-sup condition [47, 48]:

$$\exists\ \beta^{LB}_{Br,\mathcal{N}}(\boldsymbol{\mu}) > 0 :\ \beta_{Br,\mathcal{N}}(\boldsymbol{\mu}) = \inf_{q \in Q^{\mathcal{N}}} \sup_{\mathbf{w} \in V^{\mathcal{N}}} \frac{b(q,\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{w}\|_V \|q\|_Q} \geq \beta^{LB}_{Br,\mathcal{N}}(\boldsymbol{\mu}), \quad \forall\ \boldsymbol{\mu} \in \mathcal{D} \tag{3.48}$$

is satisfied. Moreover, stability estimates under the form (3.27) hold for the truth approximation $(\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}), p^{\mathcal{N}}(\boldsymbol{\mu}))$, replacing all the discrete counterparts $\alpha^{LB}_{\mathcal{N}}(\boldsymbol{\mu})$, $\gamma^{\mathcal{N}}_a(\boldsymbol{\mu})$, $\beta^{LB}_{Br,\mathcal{N}}(\boldsymbol{\mu})$ to the continuity/stability constants. In our case $V^{\mathcal{N}} \times Q^{\mathcal{N}}$ is the space of Taylor-Hood $\mathbb{P}_2 - \mathbb{P}_1$ elements [125]; however, this choice is not restrictive, the whole construction keeps holding for other spaces combinations as well. In order to verify the Brezzi inf-sup condition (3.48) let us introduce the following (inner, pressure) supremizer operator $T^{\boldsymbol{\mu}}_p$: $Q^{\mathcal{N}} \to V^{\mathcal{N}}$, defined as

$$(T^{\boldsymbol{\mu}}_p q, \mathbf{w})_V = b(q,\mathbf{w};\boldsymbol{\mu}), \quad \forall\ \mathbf{w} \in V^{\mathcal{N}}. \tag{3.49}$$

From this definition it is straightforward to prove that[2]

[2] By definition (3.49), $\|T^{\boldsymbol{\mu}}_p q\|^2_V = (T^{\boldsymbol{\mu}}_p q, T^{\boldsymbol{\mu}}_p q)_V = b(q, T^{\boldsymbol{\mu}} q; \boldsymbol{\mu})$, but by Cauchy-Schwarz inequality

$$\frac{b(q,\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{w}\|_V} = \frac{(T^{\boldsymbol{\mu}}_p q, \mathbf{w})_V}{\|\mathbf{w}\|_V} \leq \frac{\|T^{\boldsymbol{\mu}}_p q\|_V \|\mathbf{w}\|_V}{\|\mathbf{w}\|_V} \leq \|T^{\boldsymbol{\mu}}_p q\|_V, \qquad \forall \mathbf{w} \in V^{\mathcal{N}}$$

Hence, the supremizer role property gives

$$\beta_{Br,\mathcal{N}}(\boldsymbol{\mu}) = \inf_{q \in Q^{\mathcal{N}}} \left( \frac{1}{\|q\|_Q} \left( \sup_{\mathbf{w} \in V^{\mathcal{N}}} \frac{b(q,\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{w}\|_V} \right) \right) \overset{\mathbf{w}=T^{\boldsymbol{\mu}}_p q}{=} \inf_{q \in Q^{\mathcal{N}}} \frac{\|T^{\boldsymbol{\mu}}_p q\|_V}{\|q\|_Q} \Leftrightarrow (\beta_{Br,\mathcal{N}}(\boldsymbol{\mu}))^2 = \inf_{q \in Q^{\mathcal{N}}} \frac{(T^{\boldsymbol{\mu}}_p q, T^{\boldsymbol{\mu}}_p q)_V}{\|q\|^2_Q}.$$

$$T_p^{\boldsymbol{\mu}} q = \arg \sup_{\mathbf{w} \in V^{\mathcal{N}}} \frac{b(q, \mathbf{w}; \boldsymbol{\mu})}{\|\mathbf{w}\|_V} \qquad \text{and} \qquad (\beta_{Br,\mathcal{N}}(\boldsymbol{\mu}))^2 = \inf_{q \in Q^{\mathcal{N}}} \frac{(T_p^{\boldsymbol{\mu}} q, T_p^{\boldsymbol{\mu}} q)_V}{\|q\|_Q^2}. \tag{3.50}$$

Note from our affine assumption it follows that, for any $\varphi \in Q^{\mathcal{N}}$, the (inner, pressure) supremizer operator can be expressed as

$$T_p^{\boldsymbol{\mu}} \varphi = \sum_{q=1}^{Q_b} \Theta_b^q(\boldsymbol{\mu}) T_p^q \varphi, \quad \text{where} \quad (T_p^q \varphi, \mathbf{v})_V = b^q(\varphi, \mathbf{v}), \quad \forall \mathbf{v} \in V^{\mathcal{N}}, \quad 1 \le q \le Q_b. \tag{3.51}$$

**Remark 3.6.** *Following the Babuška stability theory, an alternative to* (3.47)-(3.48) *ensuring the well posedness of* (3.45) *is the following Babuška inf-sup stability condition:*

$$\exists\ \beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu}) > 0 : \ \beta_{Ba,\mathcal{N}}(\boldsymbol{\mu}) = \inf_{V \in X^{\mathcal{N}}} \sup_{W \in X^{\mathcal{N}}} \frac{A(V, W; \boldsymbol{\mu})}{\|V\|_X \|W\|_X} \ge \beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu}), \quad \forall\ \boldsymbol{\mu} \in \mathcal{D} \tag{3.52}$$

The Navier-Stokes case poses of course much more difficulties. We simply remark that, assuming to consider a branch of nonsingular solutions, (3.45) admits a solution under suitable (discrete) stability conditions, which is also unique provided the *small data* assumption holds. However, since the trilinear form is (in general) no longer such that $c(\mathbf{v}^{\mathcal{N}}, \mathbf{v}^{\mathcal{N}}, \mathbf{v}^{\mathcal{N}}; \boldsymbol{\mu}) = 0$ for any $\mathbf{v}^{\mathcal{N}} \in V^{\mathcal{N}}$, in order to fulfill a stability condition under the form (3.38) also at the discrete level, it has to be replaced by a slightly different version, given for instance (see e.g. [307]) by

$$\tilde{c}(\mathbf{v}, \mathbf{w}, \mathbf{z}; \boldsymbol{\mu}) = \sum_{k=1}^{K_{\text{dom}}} \left( \frac{1}{2} \int_{\Omega^k} v_i\, \chi_{ij}^k(\cdot; \boldsymbol{\mu})\, \frac{\partial w_m}{\partial x_j} z_m \, d\Omega - \frac{1}{2} \int_{\Omega^k} v_i\, \chi_{ij}^k(\cdot; \boldsymbol{\mu})\, w_m \frac{\partial z_m}{\partial x_j} \, d\Omega \right),$$

which is consistent with (3.6) since $\tilde{c}(\mathbf{v}, \mathbf{v}, \mathbf{z}; \boldsymbol{\mu}) = c(\mathbf{v}, \mathbf{v}, \mathbf{z}; \boldsymbol{\mu})$, for any $\mathbf{w} \in V$ and any $\mathbf{v} \in V$ such that $\nabla \cdot \mathbf{v} = 0$, but is such that $\tilde{c}(\mathbf{v}^{\mathcal{N}}, \mathbf{v}^{\mathcal{N}}, \mathbf{v}^{\mathcal{N}}; \boldsymbol{\mu}) = 0$. Thus, if for any $\mathbf{w}^{\mathcal{N}} \in V^{\mathcal{N}}$, $\boldsymbol{\mu} \in \mathcal{D}$,

$$|\tilde{c}(\mathbf{w}^{\mathcal{N}}, \mathbf{v}_D^{\mathcal{N}}, \mathbf{w}^{\mathcal{N}}; \boldsymbol{\mu})| \le \frac{\alpha_{\mathcal{N}}(\boldsymbol{\mu})}{2} \|\mathbf{w}\|_V^2, \ |c(\mathbf{v}_D^{\mathcal{N}}, \mathbf{w}^{\mathcal{N}}, \mathbf{w}^{\mathcal{N}}; \boldsymbol{\mu})| \le \frac{\alpha_{\mathcal{N}}(\boldsymbol{\mu})}{2} \|\mathbf{w}^{\mathcal{N}}\|_V^2,$$

where $\mathbf{v}_D^{\mathcal{N}}$ denotes the FE interpolant of velocity boundary conditions imposed on $\Gamma_D$, the discrete counterpart of the coercivity property (3.38) holds with $\tilde{\alpha}_{\mathcal{N}}^{LB}(\boldsymbol{\mu})$, so that problem (3.45) admits a solution, provided that the pressure/divergence term is inf-sup stable in the sense of (3.48). As before, in order to satisfy the inf-sup condition, we can introduce the (inner, pressure) supremizer operator (3.49). Moreover, a uniqueness result follows under the *small data* assumption:

$$\frac{4\, \gamma_c^{\mathcal{N}}(\boldsymbol{\mu})}{\tilde{\alpha}_{\mathcal{N}}^{LB}(\boldsymbol{\mu})} \|F(\cdot; \boldsymbol{\mu})\|_{(H^{-1}(\Omega))^2} < 1, \quad \gamma_c^{\mathcal{N}}(\boldsymbol{\mu}) = \sqrt{2} \rho_{\mathcal{N}}^2 \max_{q=1,\dots,Q_c} \|\Theta_q^c(\boldsymbol{\mu})\|_{L^\infty(\mathcal{D})} \max_{q=1,\dots,Q_c} \|\eta^q\|_{L^\infty(\Omega)};$$

$\gamma_c^{\mathcal{N}}(\boldsymbol{\mu})$ denotes the continuity constant of $\tilde{c}(\cdot, \cdot, \cdot; \boldsymbol{\mu})$ over $V^{\mathcal{N}}$. As for well-posedness, the BRR theory [49, 56] provides a more general framework for the finite dimensional approximation of branches of nonsingular solutions. We remark that the discrete version of continuity and inf-sup stability conditions (3.42)-(3.43) read as follow:

$$\gamma_{\tilde{A}}^{\mathcal{N}}(\boldsymbol{\mu}) = \sup_{V \in X^{\mathcal{N}}} \sup_{W \in X^{\mathcal{N}}} \frac{d\tilde{A}(Y^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu})(V, W)}{\|V\|_X \|W\|_X} < +\infty, \quad \forall\ \boldsymbol{\mu} \in \mathcal{D}, \tag{3.53}$$

$$\exists\ \beta_{\tilde{A},\mathcal{N}}^{LB}(\boldsymbol{\mu}) > 0 : \ \beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}) = \inf_{V \in X^{\mathcal{N}}} \sup_{W \in X^{\mathcal{N}}} \frac{d\tilde{A}(Y^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu})(V, W)}{\|V\|_X \|W\|_X} \ge \beta_{\tilde{A},\mathcal{N}}^{LB}(\boldsymbol{\mu}), \quad \forall\ \boldsymbol{\mu} \in \mathcal{D}. \tag{3.54}$$

### 3.2.4 Algebraic formulation of the FE truth approximation

Starting from the Galerkin approximation (3.45), we can easily derive the matrix formulation of the Stokes and the Navier-Stokes problems, whose structures are needed to assemble the reduced basis approximation as well. Let us denote by $\{\boldsymbol{\phi}_i^{\mathbf{v}}\}_{i=1}^{\mathcal{N}_V}$ and $\{\phi_i^p\}_{i=1}^{\mathcal{N}_Q}$ the basis of the FE spaces $V^{\mathcal{N}}$ and $Q^{\mathcal{N}}$, respectively, so that we express[3] the FE velocity and pressure as

$$\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}) = \sum_{i=1}^{\mathcal{N}_V} u_i^{\mathcal{N}}(\boldsymbol{\mu})\boldsymbol{\phi}_i^{\mathbf{v}}, \qquad p^{\mathcal{N}}(\boldsymbol{\mu}) = \sum_{i=1}^{\mathcal{N}_Q} p_i^{\mathcal{N}}(\boldsymbol{\mu})\phi_i^p.$$

Consequently $\underline{\mathbf{v}}_{\mathcal{N}}(\boldsymbol{\mu}) = (u_1^{\mathcal{N}}(\boldsymbol{\mu}), \dots, u_{\mathcal{N}_V}^{\mathcal{N}}(\boldsymbol{\mu}))^T \in \mathbb{R}^{\mathcal{N}_V}$ and $\mathbf{p}_{\mathcal{N}}(\boldsymbol{\mu}) = (p_1^{\mathcal{N}}(\boldsymbol{\mu}), \dots, p_{\mathcal{N}_Q}^{\mathcal{N}}(\boldsymbol{\mu}))^T \in \mathbb{R}^{\mathcal{N}_Q}$ denote the vectors of the degrees of freedom related to the FE velocity and pressure, respectively. The problem (3.45) can thus be rewritten under matrix form as follows:

$$\begin{bmatrix} \mathbb{A}_{\mathcal{N}}(\boldsymbol{\mu}) + \delta\mathbb{D}_{\mathcal{N}}(\boldsymbol{\mu}) + \delta\mathbb{C}_{\mathcal{N}}(\underline{\mathbf{v}}_{\mathcal{N}}(\boldsymbol{\mu});\boldsymbol{\mu}) & \mathbb{B}_{\mathcal{N}}^T(\boldsymbol{\mu}) \\ \mathbb{B}_{\mathcal{N}}(\boldsymbol{\mu}) & 0 \end{bmatrix} \begin{bmatrix} \underline{\mathbf{v}}_{\mathcal{N}}(\boldsymbol{\mu}) \\ \mathbf{p}_{\mathcal{N}}(\boldsymbol{\mu}) \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{\mathcal{N}}(\boldsymbol{\mu}) \\ \mathbf{g}_{\mathcal{N}}(\boldsymbol{\mu}) \end{bmatrix} \tag{3.55}$$

where, for $1 \le i, j \le \mathcal{N}_V$ and $1 \le k \le \mathcal{N}_Q$,

$$\begin{aligned} &(\mathbb{A}_{\mathcal{N}}(\boldsymbol{\mu}))_{ij} = a(\boldsymbol{\phi}_j^{\mathbf{v}}, \boldsymbol{\phi}_i^{\mathbf{v}}; \boldsymbol{\mu}), \qquad (\mathbb{D}_{\mathcal{N}}(\boldsymbol{\mu}))_{ij} = c(\mathbf{v}_D^{\mathcal{N}}, \boldsymbol{\phi}_j^{\mathbf{v}}, \boldsymbol{\phi}_i^{\mathbf{v}}; \boldsymbol{\mu}) + c(\boldsymbol{\phi}_j^{\mathbf{v}}, \mathbf{v}_D^{\mathcal{N}}, \boldsymbol{\phi}_i^{\mathbf{v}}\boldsymbol{\mu}), \\ &(\mathbb{C}_{\mathcal{N}}(\underline{\mathbf{w}};\boldsymbol{\mu})_{\mathcal{N}})_{ij} = \sum_{m=1}^{\mathcal{N}_V} w_m^{\mathcal{N}} c\left(\boldsymbol{\phi}_m^{\mathbf{v}}, \boldsymbol{\phi}_j^{\mathbf{v}}, \boldsymbol{\phi}_i^{\mathbf{v}}; \boldsymbol{\mu}\right), \quad (\mathbb{B}_{\mathcal{N}}(\boldsymbol{\mu}))_{ki} = b(\phi_k^p, \boldsymbol{\phi}_i^{\mathbf{v}}; \boldsymbol{\mu}), \\ &(\mathbf{f}_{\mathcal{N}}(\boldsymbol{\mu}))_i = F_s(\boldsymbol{\phi}_i^{\mathbf{v}}; \boldsymbol{\mu}) - a(\mathbf{v}_D^{\mathcal{N}}, \boldsymbol{\phi}_i^{\mathbf{v}}; \boldsymbol{\mu}) - \delta c(\mathbf{v}_D^{\mathcal{N}}, \mathbf{v}_D^{\mathcal{N}}, \boldsymbol{\phi}_i^{\mathbf{v}}; \boldsymbol{\mu}), \quad (\mathbf{g}_{\mathcal{N}}(\boldsymbol{\mu}))_k = -b(\phi_k^p, \mathbf{v}_D^{\mathcal{N}}; \boldsymbol{\mu}). \end{aligned} \tag{3.56}$$

Here, $\mathbb{A}_{\mathcal{N}}$ is the stiffness matrix, $\mathbb{B}_{\mathcal{N}}$ the gradient matrix (whose transpose $\mathbb{B}_{\mathcal{N}}^T$ is the divergence matrix), while $\mathbb{C}_{\mathcal{N}}(\mathbf{v})$ is the nonlinear transport matrix; $\mathbb{D}_{\mathcal{N}}$ is the Newton derivative of the nonlinear transport term, evaluated at $\mathbf{v}_D^{\mathcal{N}}$; the right-hand sides $\mathbf{f}_{\mathcal{N}}$ and $\mathbf{g}_{\mathcal{N}}$ are related to source terms and lifting of Dirichlet conditions. Moreover, we introduce here two other ($\boldsymbol{\mu}$-independent) matrices $\mathbb{X}_V \in \mathbb{R}^{\mathcal{N}_V \times \mathcal{N}_V}$ and $\mathbb{X}_Q \in \mathbb{R}^{\mathcal{N}_Q \times \mathcal{N}_Q}$ – associated with the discrete inner products:

$$(\mathbb{X}_V)_{ij} = (\boldsymbol{\phi}_i^{\mathbf{v}}, \boldsymbol{\phi}_j^{\mathbf{v}})_V, \qquad (\mathbb{X}_Q)_{ij} = (\phi_i^p, \phi_j^p)_Q \tag{3.57}$$

so that the inner product of elements of $V^{\mathcal{N}}$ (resp. $Q^{\mathcal{N}}$) can be calculated as

$$(\mathbf{w}^{\mathcal{N}}, \mathbf{v}^{\mathcal{N}})_X = \underline{\mathbf{w}}_{\mathcal{N}}^T \mathbb{X}_V \underline{\mathbf{v}}_{\mathcal{N}}, \qquad (p^{\mathcal{N}}, q^{\mathcal{N}})_X = \mathbf{p}_{\mathcal{N}}^T \mathbb{X}_Q \mathbf{q}_{\mathcal{N}}. \tag{3.58}$$

In the Stokes case ($\delta = 0$), (3.55) becomes a linear system under the usual form of a saddle-point problem, which can be solved e.g. by means of a *sparse* LU factorization. Other options based on preconditioned iterative solvers (such as the so-called *pressure matrix* method) are also available; see e.g. [260, 92].

In the Navier-Stokes case, the nonlinear saddle-point problem (3.55) can be solved for instance by means of a *fixed-point iteration*[4]. We remark that – see e.g. [92], Chapter 7.2 and references therein – the advantage of the fixed point (or *Picard*) iteration is that, relative to Newton

[3] We remark that the solution of the problem (3.44) is vanishing on the whole Dirichlet boundary, so that the corresponding FE velocity approximation fulfilling the boundary conditions is given by $\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}) + \tilde{\mathbf{g}}_D^{\mathcal{N}}$, where $\tilde{\mathbf{g}}_D^{\mathcal{N}}$ is a discrete function defined just on the set of Dirichlet degrees of freedom, interpolating the Dirichlet data.

[4] By considering $c(\mathbf{w}, \mathbf{v}, \mathbf{z}; \boldsymbol{\mu})$ instead of $c(\mathbf{v}, \mathbf{v}, \mathbf{z}; \boldsymbol{\mu})$ in (3.4), where $\mathbf{w}$ is a known, divergence-free velocity field, we obtain the so-called *Oseen problem*, which is well-posed for any $\mathbf{w}$. Denoting its solution by $\mathbf{u} = T\mathbf{w}$, for an operator $T : V \to V$, the solution of the Navier-Stokes problem results a fixed point of $T$; it is unique under the *small data* assumption, since in this case $T$ is a contractive mapping (Banach fixed-point theorem); see e.g. [92, 115].

iteration, it has a *huge* ball of convergence. Moreover, it can be shown that fixed-point iteration is *globally* convergent if the *small data* uniqueness condition is satisfied; that is, it will converge to the uniquely defined solution for any initial velocity. Thus, starting from an initial guess $(\underline{\mathbf{v}}_{\mathcal{N}}^{(0)}(\boldsymbol{\mu}), \mathbf{p}_{\mathcal{N}}^{(0)}(\boldsymbol{\mu}))$, for $k \geq 1$, we compute $(\underline{\mathbf{v}}_{\mathcal{N}}^{(k)}(\boldsymbol{\mu}), \mathbf{p}_{\mathcal{N}}^{(k)}(\boldsymbol{\mu}))$ as the solution of the problem

$$\begin{bmatrix} \mathbb{A}_{\mathcal{N}}(\boldsymbol{\mu}) + \mathbb{D}_{\mathcal{N}}(\boldsymbol{\mu}) + \mathbb{C}_{\mathcal{N}}(\underline{\mathbf{v}}_{\mathcal{N}}^{(k-1)}(\boldsymbol{\mu}); \boldsymbol{\mu}) & \mathbb{B}_{\mathcal{N}}^{T}(\boldsymbol{\mu}) \\ \mathbb{B}_{\mathcal{N}}(\boldsymbol{\mu}) & 0 \end{bmatrix} \begin{bmatrix} \underline{\mathbf{v}}_{\mathcal{N}}^{(k)}(\boldsymbol{\mu}) \\ \mathbf{p}_{\mathcal{N}}^{(k)}(\boldsymbol{\mu}) \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{\mathcal{N}}(\boldsymbol{\mu}) \\ \mathbf{g}_{\mathcal{N}}(\boldsymbol{\mu}) \end{bmatrix}, \tag{3.59}$$

until a given stopping criterium is satisfied; for instance, we may require that

$$\|\underline{\mathbf{v}}_{\mathcal{N}}^{(k)}(\boldsymbol{\mu}) - \underline{\mathbf{v}}_{\mathcal{N}}^{(k-1)}(\boldsymbol{\mu})\|_X \leq \epsilon_{\text{tol}}^{NS},$$

given a small positive tolerance $\epsilon_{\text{tol}}^{NS}$, where the discrete norm has to be considered in the sense of (3.58). As initial guess, we usually consider the Stokes solution of (3.55). Each Oseen system can be solved e.g. by means of a *sparse* LU factorization, like in the Stokes case. With respect to the Stokes problem, we remark that in the Oseen case the velocity matrix is no longer symmetric, due to the nonconvective term. Different approaches for solving the FE problem are based on the use of *homotopy/continuation* with respect to the parameters, whenever interested in one solution branch, as shown in [78], is .

We remark that exploiting the affine parametric expansions (1.7)-(1.8), we can write the matrices occurring in the FE formulation by decoupling the $\boldsymbol{\mu}$-dependent and $\boldsymbol{\mu}$-independent parts: for the linear terms we have

$$\begin{aligned} &\mathbb{A}_{\mathcal{N}}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_a} \Theta_a^q(\boldsymbol{\mu}) \mathbb{A}_{\mathcal{N}}^q, \qquad (\mathbb{A}_{\mathcal{N}}^q)_{ij} = a^q(\boldsymbol{\phi}_j^{\mathbf{v}}, \boldsymbol{\phi}_i^{\mathbf{v}}), \qquad 1 \leq q \leq Q_a \\ &\mathbb{B}_{\mathcal{N}}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_b} \Theta_b^q(\boldsymbol{\mu}) \mathbb{B}_{\mathcal{N}}^q, \qquad (\mathbb{B}_{\mathcal{N}}^q)_{ki} = b^q(\phi_k^p, \boldsymbol{\phi}_i^{\mathbf{v}}), \qquad 1 \leq q \leq Q_b \end{aligned} \tag{3.60}$$

while, for the nonlinear terms, with $1 \leq q \leq Q_c$, we have

$$\mathbb{C}_{\mathcal{N}}(\underline{\mathbf{w}}_{\mathcal{N}}; \boldsymbol{\mu}) = \sum_{q=1}^{Q_c} \Theta_c^q(\boldsymbol{\mu}) \mathbb{C}_{\mathcal{N}}^q(\mathbf{w}_{\mathcal{N}}), \quad (\mathbb{C}_{\mathcal{N}}^q(\mathbf{w}_{\mathcal{N}}))_{ij} = c^q\left(\sum_{k=1}^{\mathcal{N}_V} w_k^{\mathcal{N}} \boldsymbol{\phi}_k^{\mathbf{v}}, \boldsymbol{\phi}_j^{\mathbf{v}}, \boldsymbol{\phi}_i^{\mathbf{v}}\right). \tag{3.61}$$

A similar expansion can be obtained for the matrix $\mathbb{D}_{\mathcal{N}}(\boldsymbol{\mu})$ and the right-hand sides:

$$\mathbb{D}_{\mathcal{N}}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_d} \Theta_d^q(\boldsymbol{\mu}) \mathbb{D}_{\mathcal{N}}^q, \qquad \mathbf{f}_{\mathcal{N}}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_F} \Theta_F^q(\boldsymbol{\mu}) \mathbf{f}_{\mathcal{N}}^q, \qquad \mathbf{g}_{\mathcal{N}}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_G} \Theta_G^q(\boldsymbol{\mu}) \mathbf{g}_{\mathcal{N}}^q, \tag{3.62}$$

where

$$\begin{aligned} &\Theta_d^q(\boldsymbol{\mu}) = \Theta_c^q(\boldsymbol{\mu}), && (\mathbb{D}_{\mathcal{N}}^q)_{ij} = c^q(\mathbf{v}_D^{\mathcal{N}}, \boldsymbol{\phi}_i^{\mathbf{v}}, \boldsymbol{\phi}_j^{\mathbf{v}}) + c^q(\boldsymbol{\phi}_i^{\mathbf{v}}, \mathbf{v}_D^{\mathcal{N}}, \boldsymbol{\phi}_j^{\mathbf{v}}), && q = 1, \ldots, Q_d \equiv Q_c \\ &\Theta_F^q(\boldsymbol{\mu}) = \Theta_s^q(\boldsymbol{\mu}), && (\mathbf{f}_{\mathcal{N}}^q)_i = F_s^q(\boldsymbol{\phi}_i^{\mathbf{v}}), && q = 1, \ldots, Q_s \\ &\Theta_F^q(\boldsymbol{\mu}) = \Theta_a^q(\boldsymbol{\mu}), && (\mathbf{f}_{\mathcal{N}}^q)_i = -a^q(\mathbf{v}_D^{\mathcal{N}}, \boldsymbol{\phi}_i^{\mathbf{v}}), && q = Q_s + 1, \ldots, Q_s + Q_a \\ &\Theta_F^q(\boldsymbol{\mu}) = \delta\Theta_c^q(\boldsymbol{\mu}), && (\mathbf{f}_{\mathcal{N}}^q)_i = c^q(\mathbf{v}_D^{\mathcal{N}}, \mathbf{v}_D^{\mathcal{N}}, \boldsymbol{\phi}_i^{\mathbf{v}}), && q = Q_s + Q_a + \delta, \ldots, Q_s + Q_a + \delta Q_c \\ &\Theta_G^q(\boldsymbol{\mu}) = \Theta_b^q(\boldsymbol{\mu}), && (\mathbf{g}_{\mathcal{N}}^q)_k = b^q(\phi_k^p, \mathbf{v}_D^{\mathcal{N}}), && q = 1, \ldots, Q_G \equiv Q_b \end{aligned}$$

We point out that *(i)* the RB approximation will be built upon, and the error in our RB approximation will be measured with respect to, this truth FE approximation, and that *(ii)* the Offline/Online procedure is based on relationships (3.60)-(3.62) valid at the FE level.

## 3.3 Reduced basis approximation

We are now in the position to apply the reduced basis framework presented in Sect. 2.2 to the Stokes and the Navier-Stokes problems. Former contributions dealing with the RB approximation of Stokes problems have been addressed by Rozza [278, 276] and Veroy [282], dealing with both affine and nonaffine geometrical parametrizations, even if rigorous *a posteriori* error bounds were still under development. In the Navier-Stokes case, after the pioneering works by Peterson [242], Ito and Ravindran [153, 152], a general framework for both RB approximation and *a posteriori* error estimation has been provided by Patera, Veroy [312] and Nguyen [224], without including the treatment of the pressure field within the equation in the isothermal case. Both RB approximation and error estimation have been further developed by Deparis [78] and Rozza [79], by considering the more general case involving the temperature field and a complete *natural norm* framework. Moreover, the extension to more complex geometries dealing with nonaffine geometrical maps is due to Quarteroni and Rozza [255].

Concerning the Stokes problem, the original contribution of this thesis work – based on a recent work with Rozza and Huynh [279] – is a joint stability study based on the role of the Brezzi's inf-sup constant in the RB context and an error analysis and certification of results based on the estimation of the Babuška's inf-sup constant in the framework of general noncoercive problems. The approximation of the lower bounds of the global Babuška stability factors is performed by extending the *natural norm* Successive Constraint Method (SCM) to the Stokes case.
In the Navier-Stokes case, instead, the original contribution of this thesis is the setting of *a posteriori* error estimation for velocity and pressure fields jointly, following the BRR theory, whereas the stability of the RB approximation is still based on the Brezzi theory. Moreover, we extend the approximation of lower bounds of stability factors based on the SCM to the nonlinear case, providing also a different surrogate method to estimate this quantity (see Sect. A.3). In the end, we also deal with a fully decoupled Offline/Online procedure and with the treatment of more complex nonaffine geometrical mappings also in the Navier-Stokes case.

### 3.3.1 Reduced basis spaces and their approximation stability

The RB method efficiently computes an approximation of the solution $(\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}), p^{\mathcal{N}}(\boldsymbol{\mu}))$ of problem (3.45) by using global approximation spaces made up of well-chosen solutions of this problem, corresponding to specific choices of the parameter values.
We address in this section all the details concerning the RB approximation of Stokes problems – which are in fact still valid in the Navier-Stokes case – leaving to the following section the details concerning the treatment of the nonlinear terms. In both cases, following the general framework described in Sect. 2.1, we select (iteratively, through the greedy algorithm) a set of parameter values $S_N = \{\boldsymbol{\mu}^1, \dots, \boldsymbol{\mu}^N\}$ and consider the corresponding FE solutions $(\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}^1), p^{\mathcal{N}}(\boldsymbol{\mu}^1)), \dots, (\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}^N), p^{\mathcal{N}}(\boldsymbol{\mu}^N))$ of problem (3.45) (with $\delta = 0$ in the Stokes case, $\delta = 1$ in the Navier-Stokes case, respectively), where typically $N \ll \mathcal{N}$. We thus define the *reduced basis pressure space* $Q_N^{\mathcal{N}} \subset Q^{\mathcal{N}}$ as the space engendered by the *pressure snapshots*:

$$Q_N^{\mathcal{N}} = \text{span}\{\tilde{\zeta}_n^p := p^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n = 1, \dots, N\}, \qquad N = 1, \dots N_{\max}. \tag{3.63}$$

The *reduced basis velocity space* $V_N^{\mathcal{N},\boldsymbol{\mu}} \subset V^{\mathcal{N}}$ can be built as

$$V_N^{\mathcal{N},\boldsymbol{\mu}} = \text{span}\{\tilde{\boldsymbol{\zeta}}_n^{\mathbf{v}} := \mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}^n),\ T_p^{\boldsymbol{\mu}}\tilde{\zeta}_n^p,\ n = 1, \dots, N\}, \qquad N = 1, \dots N_{\max}, \tag{3.64}$$

thus enriching the space of velocity *snapshots* with the inner (pressure) supremizers. Details about the Gram-Schmidt orthonormalization procedure are reported in Sect. A.1 of Appendix A.

The RB approximation $(\mathbf{v}_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu}))$ can be obtained by computing a Galerkin projection onto the reduced spaces $V_N^{\mathcal{N},\boldsymbol{\mu}} \times Q_N^{\mathcal{N}}$ as follows: given $\boldsymbol{\mu} \in \mathcal{D}$, evaluate the scalar output of interest $s_N(\boldsymbol{\mu}) = S(Y_N(\boldsymbol{\mu}); \boldsymbol{\mu})$ where $Y_N(\boldsymbol{\mu}) = (\mathbf{v}_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu})) \in X_N := V_N^{\mathcal{N},\boldsymbol{\mu}} \times Q_N^{\mathcal{N}}$ is such that

$$A(Y_N(\boldsymbol{\mu}), W_N; \boldsymbol{\mu}) + \delta C(Y_N(\boldsymbol{\mu}), Y_N(\boldsymbol{\mu}), W_N; \boldsymbol{\mu}) = \tilde{F}(W_N; \boldsymbol{\mu}), \quad \forall\, W_N \in X_N \tag{3.65}$$

or, equivalently,

$$\begin{cases} a(\mathbf{v}_N(\boldsymbol{\mu}), \mathbf{w}_N; \boldsymbol{\mu}) + \delta d(\mathbf{v}_N(\boldsymbol{\mu}), \mathbf{w}_N; \boldsymbol{\mu}) + b(p_N(\boldsymbol{\mu}), \mathbf{w}_N; \boldsymbol{\mu}) \\ \qquad\qquad +\delta c(\mathbf{v}_N(\boldsymbol{\mu}), \mathbf{v}_N(\boldsymbol{\mu}), \mathbf{w}_N; \boldsymbol{\mu}) = F(\mathbf{w}_N; \boldsymbol{\mu})\,, & \forall \mathbf{w}_N \in V_N^{\mathcal{N},\boldsymbol{\mu}} \\ b(q_N, \mathbf{v}_N(\boldsymbol{\mu}); \boldsymbol{\mu}) = G(q_N; \boldsymbol{\mu})\,, & \forall q_N \in Q_N^{\mathcal{N}}. \end{cases} \tag{3.66}$$

We can easily show that problem (3.66) fulfills an equivalent Brezzi RB inf-sup condition. By defining the reduced spaces as in (3.63)-(3.64) and the RB stability factor as

$$\beta_N(\boldsymbol{\mu}) = \inf_{q \in Q_N^{\mathcal{N}}} \sup_{\mathbf{w} \in V_N^{\mathcal{N},\boldsymbol{\mu}}} \frac{b(q, \mathbf{w}; \boldsymbol{\mu})}{\|\mathbf{w}\|_V \|q\|_Q} \tag{3.67}$$

the following inequalities hold [276, 282]:

$$\beta_N(\boldsymbol{\mu}) \geq \beta^{\mathcal{N}}(\boldsymbol{\mu}) > 0, \qquad \forall\, \boldsymbol{\mu} \in \mathcal{D},$$

where $\beta^{\mathcal{N}}(\boldsymbol{\mu})$ is defined in (3.26). In fact, we have that

$$\begin{aligned} \beta^{\mathcal{N}}(\boldsymbol{\mu}) = \inf_{q \in Q^{\mathcal{N}}} \sup_{\mathbf{w} \in V^{\mathcal{N}}} \frac{b(q, \mathbf{w}; \boldsymbol{\mu})}{\|\mathbf{w}\|_V \|q\|_Q} \leq \inf_{q \in Q_N^{\mathcal{N}}} \sup_{\mathbf{w} \in V^{\mathcal{N}}} \frac{b(q, \mathbf{w}; \boldsymbol{\mu})}{\|\mathbf{w}\|_V \|q\|_Q} = \\ \inf_{q \in Q_N^{\mathcal{N}}} \frac{b(q, T_p^{\boldsymbol{\mu}} q; \boldsymbol{\mu})}{\|T_p^{\boldsymbol{\mu}} q\|_V \|q\|_Q} \leq \inf_{q \in Q_N^{\mathcal{N}}} \sup_{\mathbf{w} \in V_N^{\mathcal{N},\boldsymbol{\mu}}} \frac{b(q, \mathbf{w}; \boldsymbol{\mu})}{\|\mathbf{w}\|_V \|q\|_Q} = \beta_N(\boldsymbol{\mu}), \end{aligned} \tag{3.68}$$

where we have applied the fact that $Q_N^{\mathcal{N}} \subset Q^{\mathcal{N}}$, the definition of the (inner, pressure) supremizer operator and the fact that the RB velocity space $V_N^{\mathcal{N},\boldsymbol{\mu}}$ is enriched by supremizers, respectively. We remark that, by enriching the velocity space with the supremizers $T_p^{\boldsymbol{\mu}} \zeta_n^p$, for $n = 1, \dots N$, the RB velocity space (3.64) has dimension $2N$, the double of the dimension $N$ of the RB pressure space. We detail the construction of the supremizer operator in Sect. A.1 of Appendix A.
From now on, we denote the RB velocity space by $V_N^{\mathcal{N}} \equiv V_N^{\mathcal{N},\boldsymbol{\mu}}$ for the sake of simplicity. We introduce two orthonormal basis of the RB spaces $V_N^{\mathcal{N}}$ and $Q_N$, denoted by $\{\zeta_n^p\}_{n=1}^N$ and $\{\boldsymbol{\zeta}_n^{\mathbf{v}}\}_{n=1}^{2N}$, respectively. To highlight the relationship with the truth FE approximation, here we do not consider the additional orthonormalization of velocity and pressure basis functions, as well as the $\boldsymbol{\mu}$-dependence in the supremizer solutions; the rigorous formulation is detailed in Sect. A.1 of Appendix A. We just remark that, in order to exploit a suitable Offline/Online procedure,we need to express the velocity RB space (3.64) in a more viable way. In fact, we want to completely assemble/store the basis functions only once during the Offline stage, while for each new Online evaluation, given a parameter value $\boldsymbol{\mu}$, we do not want to assemble the supremizer solution as combination of previously computed solutions. Since the definition of the RB velocity space (3.64) still depends on $\boldsymbol{\mu}$, we need a different way to express it.
In this way, by expressing the RB approximation as RB solution can be written as a combination of previously computed stored solutions as basis functions:

$$\mathbf{v}_N(\boldsymbol{\mu}) = \sum_{j=1}^{2N} v_{Nj}(\boldsymbol{\mu}) \boldsymbol{\zeta}_j^{\mathbf{v}}, \qquad p_N(\boldsymbol{\mu}) = \sum_{l=1}^{N} p_{Nl}(\boldsymbol{\mu}) \zeta_l^p,$$

the weights $\underline{\mathbf{v}}_N(\boldsymbol{\mu}) = (v_{N\,1}(\boldsymbol{\mu}), \ldots, v_{N\,2N}(\boldsymbol{\mu}))^T \in \mathbb{R}^{2N}$ and $\mathbf{p}_N(\boldsymbol{\mu}) = (p_{N\,1}(\boldsymbol{\mu}), \ldots, p_{N\,N}(\boldsymbol{\mu}))^T \in \mathbb{R}^N$ are obtained by solving the following RB system:

$$\begin{bmatrix} \mathbb{A}_N(\boldsymbol{\mu}) + \delta\mathbb{D}_N(\boldsymbol{\mu}) + \delta\mathbb{C}_N(\underline{\mathbf{v}}_N(\boldsymbol{\mu}); \boldsymbol{\mu}) & \mathbb{B}_N^T(\boldsymbol{\mu}) \\ \mathbb{B}_N(\boldsymbol{\mu}) & 0 \end{bmatrix} \begin{bmatrix} \underline{\mathbf{v}}_N(\boldsymbol{\mu}) \\ \mathbf{p}_N(\boldsymbol{\mu}) \end{bmatrix} = \begin{bmatrix} \mathbf{f}_N(\boldsymbol{\mu}) \\ \mathbf{g}_N(\boldsymbol{\mu}) \end{bmatrix} \tag{3.69}$$

where, for $1 \le m, n \le 2N$ and $1 \le l \le N$,

$$\begin{aligned} &(\mathbb{A}_N(\boldsymbol{\mu}))_{mn} = a(\boldsymbol{\zeta}_n^{\mathbf{v}}, \boldsymbol{\zeta}_m^{\mathbf{v}}; \boldsymbol{\mu}), \quad (\mathbb{D}_N(\boldsymbol{\mu}))_{mn} = c(\mathbf{v}_D^{\mathcal{N}}, \boldsymbol{\zeta}_n^{\mathbf{v}}, \boldsymbol{\zeta}_m^{\mathbf{v}}; \boldsymbol{\mu}) + c(\boldsymbol{\zeta}_n^{\mathbf{v}}, \mathbf{v}_D^{\mathcal{N}}, \boldsymbol{\zeta}_m^{\mathbf{v}}; \boldsymbol{\mu}), \\ &(\mathbb{C}_N(\underline{\mathbf{w}}_N; \boldsymbol{\mu}))_{mn} = \sum_{s=1}^{2N} w_{N\,s}\, c(\boldsymbol{\zeta}_s^{\mathbf{v}}, \boldsymbol{\zeta}_n^{\mathbf{v}}, \boldsymbol{\zeta}_m^{\mathbf{v}}; \boldsymbol{\mu}), \qquad (\mathbb{B}_N(\boldsymbol{\mu}))_{lm} = b(\zeta_l^p, \boldsymbol{\zeta}_m^{\mathbf{v}}; \boldsymbol{\mu}), \\ &(\mathbf{f}_N(\boldsymbol{\mu}))_m = F_s(\boldsymbol{\phi}_m^{\mathbf{v}}; \boldsymbol{\mu}) - a(\mathbf{v}_D^{\mathcal{N}}, \boldsymbol{\phi}_m^{\mathbf{v}}; \boldsymbol{\mu}) - \delta c(\mathbf{v}_D^{\mathcal{N}}, \mathbf{v}_D^{\mathcal{N}}, \boldsymbol{\phi}_m^{\mathbf{v}}; \boldsymbol{\mu}), \quad (\mathbf{g}_N(\boldsymbol{\mu}))_l = -b(\phi_l^p, \mathbf{v}_D^{\mathcal{N}}; \boldsymbol{\mu}). \end{aligned} \tag{3.70}$$

In the Stokes case, (3.69) is a linear system showing the same saddle-point structure of the FE approximation (3.55). Hence, using reduced basis we deal with a matrix of considerably smaller dimension (of order of $N \ll \mathcal{N}$) but with full matrices (instead of sparse ones). In the Navier-Stokes case, the nonlinear saddle-point problem (3.55) can be solved for instance by means of a *fixed-point iteration*, like in the FE case: starting from an initial guess $(\underline{\mathbf{v}}_N^{(0)}(\boldsymbol{\mu}), \mathbf{p}_N^{(0)}(\boldsymbol{\mu}))$ (e.g. the Stokes solution), for $k \ge 1$, we compute $(\underline{\mathbf{v}}_N^{(k)}(\boldsymbol{\mu}), \mathbf{p}_N^{(k)}(\boldsymbol{\mu}))$ as the solution of the problem

$$\begin{bmatrix} \mathbb{A}_N(\boldsymbol{\mu}) + \mathbb{D}_N(\boldsymbol{\mu}) + \mathbb{C}_N(\underline{\mathbf{v}}_N^{(k-1)}(\boldsymbol{\mu}); \boldsymbol{\mu}) & \mathbb{B}_N^T(\boldsymbol{\mu}) \\ \mathbb{B}_N(\boldsymbol{\mu}) & 0 \end{bmatrix} \begin{bmatrix} \underline{\mathbf{v}}_N^{(k)}(\boldsymbol{\mu}) \\ \mathbf{p}_N^{(k)}(\boldsymbol{\mu}) \end{bmatrix} = \begin{bmatrix} \mathbf{f}_N(\boldsymbol{\mu}) \\ \mathbf{g}_N(\boldsymbol{\mu}) \end{bmatrix}, \tag{3.71}$$

until convergence. Despite its lower convergence rate with respect e.g. to a Newton iteration, we adopt a fixed point algorithm because it has a huge ball of convergence, as already remarked in the FE case. We point out that the linearized term can be computed as

$$\sum_{n=1}^{2N} (\mathbb{C}_N(\mathbf{v}_N^{(k-1)}(\boldsymbol{\mu}); \boldsymbol{\mu}))_{mn} v_{N\,n}^{(k)} = \sum_{s=1}^{2N} \sum_{n=1}^{2N} v_{N\,s}^{(k-1)} c(\boldsymbol{\zeta}_s^{\mathbf{v}}, \boldsymbol{\zeta}_n^{\mathbf{v}}, \boldsymbol{\zeta}_m^{\mathbf{v}}; \boldsymbol{\mu}) v_{N\,n}^{(k)}, \qquad 1 \le m \le 2N$$

and thus only the matrices $\mathbb{C}_N(\underline{\boldsymbol{\zeta}}_s^{\mathbf{v}}; \boldsymbol{\mu})$, for $1 \le s \le 2N$, $1 \le N \le N_{max}$, defined by

$$(\mathbb{C}_N(\underline{\zeta}_s^{\mathbf{v}}; \boldsymbol{\mu}))_{mn} = c(\boldsymbol{\zeta}_s^{\mathbf{v}}, \boldsymbol{\zeta}_n^{\mathbf{v}}, \boldsymbol{\zeta}_m^{\mathbf{v}}; \boldsymbol{\mu}),$$

have to be stored. In view of an Offline/Online decomposition, it shall prove convenient to express the RB matrices and vectors appearing in (3.71) in terms of the corresponding FE matrices and vectors of (3.59): the former are linked to the latter via the basis matrices $\mathbb{Z}_{2N} \in \mathbb{R}^{\mathcal{N}_X \times 2N}$ and $\mathbb{Z}_N \in \mathbb{R}^{\mathcal{N}_Q \times N}$ for the velocity and the pressure RB space, respectively, given by

$$(\mathbb{Z}_N)_{il} = \zeta_{p\,l}^i, \qquad (\mathbb{Z}_{2N})_{jm} = \zeta_{\mathbf{v}\,m}^j, \qquad 1 \le i \le \mathcal{N}_Q, \;\; 1 \le j \le \mathcal{N}_V, \;\; 1 \le l \le N, \;\; 1 \le m \le 2N$$

for $1 \le N \le N_{max}$, where the RB basis functions can be expressed as

$$\zeta_l^p(\mathbf{x}) = \sum_{i=1}^{\mathcal{N}_Q} \zeta_{p\,l}^i \phi_i^p(\mathbf{x}), \qquad \boldsymbol{\zeta}_m^{\mathbf{v}}(\mathbf{x}) = \sum_{j=1}^{\mathcal{N}_V} \zeta_{\mathbf{v}\,m}^j \boldsymbol{\phi}_j^{\mathbf{v}}(\mathbf{x}), \qquad 1 \le l \le N, \;\; 1 \le m \le 2N,$$

since they are elements of the FE spaces $Q^{\mathcal{N}}$ and $X^{\mathcal{N}}$, respectively. In particular, from (3.56)-(3.70) we obtain, for $1 \le m, n \le 2N$, $1 \le l \le N$, $1 \le s \le 2N$:

$$\mathbb{A}_N(\boldsymbol{\mu}) = \mathbb{Z}_{2N}^T\ \mathbb{A}_{\mathcal{N}}(\boldsymbol{\mu})\ \mathbb{Z}_{2N}, \qquad a(\boldsymbol{\zeta}_n^{\mathbf{v}}, \boldsymbol{\zeta}_m^{\mathbf{v}}; \boldsymbol{\mu}) = \sum_{i=1}^{\mathcal{N}_X}\sum_{j=1}^{\mathcal{N}_X} \zeta_{\mathbf{v}\,n}^j a(\boldsymbol{\phi}_j^{\mathbf{v}}, \boldsymbol{\phi}_i^{\mathbf{v}}; \boldsymbol{\mu}) \zeta_{\mathbf{v}\,m}^i,$$

$$\mathbb{B}_N(\boldsymbol{\mu}) = \mathbb{Z}_N^T\ \mathbb{B}_{\mathcal{N}}(\boldsymbol{\mu})\ \mathbb{Z}_{2N}, \qquad b(\zeta_l^p, \boldsymbol{\zeta}_n^{\mathbf{v}}; \boldsymbol{\mu}) = \sum_{i=1}^{\mathcal{N}_Q}\sum_{j=1}^{\mathcal{N}_X} \zeta_{p\,l}^i b(\phi_i^p, \boldsymbol{\phi}_j^{\mathbf{v}}; \boldsymbol{\mu}) \zeta_{\mathbf{v}\,n}^j, \tag{3.72}$$

$$\mathbb{C}_N(\boldsymbol{\zeta}_s^{\mathbf{v}}; \boldsymbol{\mu}) = \mathbb{Z}_{2N}^T \mathbb{C}_{\mathcal{N}}(\boldsymbol{\zeta}_s^{\mathbf{v}}; \boldsymbol{\mu}) \mathbb{Z}_{2N} \quad c(\boldsymbol{\zeta}_l^{\mathbf{v}}, \boldsymbol{\zeta}_n^{\mathbf{v}}, \boldsymbol{\zeta}_m^{\mathbf{v}}; \boldsymbol{\mu}) = \sum_{i=1}^{\mathcal{N}_X}\sum_{j=1}^{\mathcal{N}_X} \zeta_{\mathbf{v}\,n}^j c(\boldsymbol{\zeta}_s^{\mathbf{v}}, \boldsymbol{\phi}_j^{\mathbf{v}}, \boldsymbol{\phi}_i^{\mathbf{v}}; \boldsymbol{\mu}) \zeta_{\mathbf{v}\,m}^i,$$

and, in the same way for the source and lifting terms (we omit the explicit expression):

$$\mathbb{D}_N(\boldsymbol{\mu}) = \mathbb{Z}_{2N}^T\ \mathbb{D}_{\mathcal{N}}(\boldsymbol{\mu})\ \mathbb{Z}_{2N}, \qquad \mathbf{f}_N(\boldsymbol{\mu}) = \mathbb{Z}_{2N}^T\ \mathbf{f}_{\mathcal{N}}(\boldsymbol{\mu}), \qquad \mathbf{g}_N(\boldsymbol{\mu}) = \mathbb{Z}_N^T\ \mathbf{g}_{\mathcal{N}}(\boldsymbol{\mu}). \tag{3.73}$$

Furthermore, thanks to the affine decomposition of FE structures (3.60)-(3.62), we obtain

$$\mathbb{A}_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_a} \Theta_q^a(\boldsymbol{\mu}) \mathbb{A}_N^q, \quad \mathbb{B}_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_b} \Theta_q^b(\boldsymbol{\mu}) \mathbb{B}_N^q, \quad \mathbb{C}_N(\boldsymbol{\zeta}_s^{\mathbf{v}}; \boldsymbol{\mu}) = \sum_{q=1}^{Q_c} \Theta_q^c(\boldsymbol{\mu}) \mathbb{C}_N^q(\boldsymbol{\zeta}_s^{\mathbf{v}}) \tag{3.74}$$

$$\mathbb{A}_N^q = \mathbb{Z}_{2N}^T\ \mathbb{A}_{\mathcal{N}}^q\ \mathbb{Z}_{2N}, \qquad \mathbb{B}_N^q = \mathbb{Z}_N^T\ \mathbb{A}_{\mathcal{N}}^q\ \mathbb{Z}_{2N}, \qquad \mathbb{C}_N^q(\boldsymbol{\zeta}_s^{\mathbf{v}}) = \mathbb{Z}_{2N}^T\ \mathbb{C}_{\mathcal{N}}^q(\boldsymbol{\zeta}_s^{\mathbf{v}})\ \mathbb{Z}_{2N}\ ;$$

in the same way, we can express also the RB structures appearing in (3.73).

### 3.3.2 Offline-Online computational strategy

The linear system (3.69) has normally a very small size (and a full structure) compared to the one arising from the FE discretization of (3.55), since it consists of a set of $3N$ linear algebraic equations in $3N$ unknowns – the FE discretization lead to a set of $\mathcal{N}_V + \mathcal{N}_Q$ equations in $\mathcal{N}_V + \mathcal{N}_Q$ unknowns. Nevertheless, the elements of $V_N^{\mathcal{N}}$ and $Q_N^{\mathcal{N}}$ are associated with the underlying FE space and thus are depending on $\mathcal{N}$. To eliminate this $\mathcal{N}$-dependency, which would lead to a very poor online performance, we construct a very efficient Offline/Online procedure (see Sect. 2.2.4): thanks to (3.74), all $\boldsymbol{\mu}$-independent terms can be precomputed and stored in the Offline stage, thus yielding to Online computations which are completely independent of $\mathcal{N}$.
In the *Offline stage* we first compute and store the basis functions $\{\boldsymbol{\zeta}_n^{\mathbf{v}}\}_{n=1}^{2N}$, $\{\zeta_l^p\}_{l=1}^{N}$, and form the RB structures. This requires:

1. $\mathcal{O}((Q_a + Q_b + Q_D)\nu_E)$ operations for the assembly of the FE structures $\mathbb{A}_{\mathcal{N}}^q$, $\mathbb{B}_{\mathcal{N}}^q$, $\mathbb{D}_{\mathcal{N}}^q$, $\mathcal{O}(2N(Q_c)\nu_E)$ for the assembly of the FE structures $\mathbb{C}_{\mathcal{N}}^q(\boldsymbol{\zeta}_n^{\mathbf{v}})$, as well as $\mathcal{O}((Q_F + Q_G)\nu_E)$ for the vectors $\mathbf{f}_{\mathcal{N}}^q$, $\mathbf{g}_{\mathcal{N}}^q$; here $\nu_E$ denotes the number of elements in the FE computational mesh;
2. $\mathcal{O}(N(\mathcal{N}_V + \mathcal{N}_Q)^3(1 + \delta K_{\mathcal{N}}))$ operations for the basis computations, where $K_{\mathcal{N}}$ is the number of the fixed-point iterations required at the FE level;
3. $\mathcal{O}(Q_a 4N^2 + Q_b 2N^2 + \delta Q_d 4N^2)$ operations for the assembly of the linear RB sub-matrices $\mathbb{A}_N^q$, $\mathbb{B}_N^q$ and $\mathbb{D}_N^q$, $\mathcal{O}(\delta 2NQ_c 4N^2)$ for the nonlinear RB sub-matrices $\mathbb{C}_N^q(\boldsymbol{\zeta}_n^{\mathbf{v}})$ and $\mathcal{O}(2NQ_F + NQ_G)$ for the right-hand sides RB vectors $\mathbf{f}_N^q$ and $\mathbf{g}_N^q$.

In the *Online stage*, for each new value of $\boldsymbol{\mu}$ we use the precomputed RB structures to assemble the (full) $3N \times 3N$ system (3.69), whose solution gives $\underline{\mathbf{v}}_N(\boldsymbol{\mu}) \in \mathbb{R}^{2N}$ and $\mathbf{p}_N(\boldsymbol{\mu}) \in \mathbb{R}^N$; then, we evaluate the output approximation. The operation count for the Online stage is then:

1. $\mathcal{O}((Q_a + Q_b + \delta(2NQ_c + Q_d))N^2 + (Q_F + Q_G)N)$ to assemble the RB system;
2. $\mathcal{O}(27N^3(1 + \delta K_N))$ to invert the RB system, where $K_N$ is the number of the fixed-point iterations required at the RB level, and $\mathcal{O}(N + N^2)$ to evaluate the output.

The crucial point is that our Online costs are dependent on $Q_\bullet$ and $N$, but independent of $\mathcal{N}_X + \mathcal{N}_Q$. Since $3N \ll \mathcal{N}_X + \mathcal{N}_Q$, we can expect significant (orders of magnitude) speedup in the Online stage compared to the pure FE approach. This implies also that we may choose $\mathcal{N}$ very large in order to eliminate the error between the exact solution and the FE predictions without affecting the RB Online efficiency – this may be required when dealing with Navier-Stokes equations at moderately large Reynolds numbers, for instance. In fact, the bigger the underlying FE system and thus $\mathcal{N}$ is chosen, the bigger the speedup by the use of the RB method in the Online stage will be. However, we should keep in mind that the Offline stage is still $\mathcal{N}$-dependent.

## 3.4 A posteriori error estimation

In this section we deal with *a posteriori* error estimation in the RB context for parametrized Stokes and Navier-Stokes equations. The approach we address takes advantage in the former case of the Babuška stability theory – in the framework of general noncoercive problems – involving the global Stokes operator, in the latter of the BRR general theory for approximation of branches of nonsingular solutions in nonlinear problems.

### 3.4.1 Stokes equations

As in the model problem of Sect. 2.3, also for the Stokes problem our *a posteriori* error estimation takes advantage of two quantities: the dual norm of residuals and an effective lower bound of the (parametric) stability factor, given in this case by the Babuška inf-sup constant $\beta_{Ba,\mathcal{N}}(\boldsymbol{\mu})$ defined in (3.52). Let us define the residuals $r_{\mathbf{v}}(\cdot\,;\boldsymbol{\mu})$ and $r_p(\cdot\,;\boldsymbol{\mu})$ by

$$\begin{aligned} r_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu}) &:= F(\mathbf{w};\boldsymbol{\mu}) - a(\mathbf{v}_N(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}) - b(p_N(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}), \quad \forall\, \mathbf{w} \in V^{\mathcal{N}}, \\ r_p(q;\boldsymbol{\mu}) &:= G(q;\boldsymbol{\mu}) - b(q,\mathbf{v}_N(\boldsymbol{\mu});\boldsymbol{\mu}), \qquad\qquad\qquad \forall\, q \in Q^{\mathcal{N}}. \end{aligned} \tag{3.75}$$

Note that

$$\begin{aligned} r_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu}) &= a(\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}) + b(e_p(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}) \qquad \forall\, \mathbf{w} \in V^{\mathcal{N}}, \\ r_p(q;\boldsymbol{\mu}) &= b(q,\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu});\boldsymbol{\mu}) \qquad\qquad\qquad\qquad\quad \forall\, q \in Q^{\mathcal{N}}, \end{aligned} \tag{3.76}$$

where $\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu}) = \mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}) - \mathbf{v}_N(\boldsymbol{\mu})$ and $e_p(\boldsymbol{\mu}) = p^{\mathcal{N}}(\boldsymbol{\mu}) - p_N(\boldsymbol{\mu})$. Equivalently, we can write

$$r(W;\boldsymbol{\mu}) = A(Y^{\mathcal{N}}(\boldsymbol{\mu}) - Y_N(\boldsymbol{\mu}), W;\boldsymbol{\mu}) \qquad\qquad \forall\, W \in X^{\mathcal{N}}, \tag{3.77}$$

where $r(W;\boldsymbol{\mu}) := r_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu}) + r_p(q;\boldsymbol{\mu})$. Using the inf-sup condition (3.52), we have

$$\beta_{Ba,\mathcal{N}}(\boldsymbol{\mu})\|Y^{\mathcal{N}}(\boldsymbol{\mu}) - Y_N^{\mathcal{N}}(\boldsymbol{\mu})\|_X \leq \sup_{W \in X^{\mathcal{N}}} \frac{A(Y^{\mathcal{N}}(\boldsymbol{\mu}) - Y_N^{\mathcal{N}}(\boldsymbol{\mu}), W;\boldsymbol{\mu})}{\|W\|_X},$$

so that the following result holds:

**Proposition 3.7.** *Let us denote by $Y^{\mathcal{N}}(\boldsymbol{\mu})$ and $Y_N(\boldsymbol{\mu})$ the truth approximation* (3.45) *and the reduced basis approximation* (3.66) *in the Stokes case $\delta = 0$, respectively. Then, the following residual-based estimation holds:*

$$\|Y^{\mathcal{N}}(\boldsymbol{\mu}) - Y_N(\boldsymbol{\mu})\|_X \leq \frac{\|r(\cdot;\boldsymbol{\mu})\|_{X'}}{\beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu})} =: \Delta_N(\boldsymbol{\mu}), \qquad \forall\, \boldsymbol{\mu} \in \mathcal{D}, \tag{3.78}$$

*where $\|r(\cdot;\boldsymbol{\mu})\|_{X'} = \sup_{W \in X^{\mathcal{N}}} r(W;\boldsymbol{\mu})/\|W\|_X$ is the dual norm of the residual and $\beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu})$ is a computable lower bound for $\beta_{Ba,\mathcal{N}}(\boldsymbol{\mu})$.*

An alternative expression of (3.78) is

$$\|\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}) - \mathbf{v}_N(\boldsymbol{\mu})\|_V^2 + \|p^{\mathcal{N}}(\boldsymbol{\mu}) - p_N(\boldsymbol{\mu})\|_Q^2 \leq \frac{1}{(\beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu}))^2} \left(\|r_{\mathbf{v}}(\cdot\,;\boldsymbol{\mu})\|_{V'}^2 + \|r_p(\cdot\,;\boldsymbol{\mu})\|_{Q'}^2\right)$$

where

$$\|r_{\mathbf{v}}(\cdot\,;\boldsymbol{\mu})\|_{V'} = \sup_{\mathbf{w}\in V^{\mathcal{N}}} \frac{r_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{v}\|_V}, \qquad \|r_p(\cdot\,;\boldsymbol{\mu})\|_{Q'} = \sup_{q\in Q^{\mathcal{N}}} \frac{r_p(q;\boldsymbol{\mu})}{\|q\|_Q}$$

are the dual norms of the residuals for the velocity and the pressure variables, respectively, such that $\|r(\cdot;\boldsymbol{\mu})\|_{X'}^2 = \|r_{\mathbf{v}}(\cdot\,;\boldsymbol{\mu})\|_{V'}^2 + \|r_p(\cdot\,;\boldsymbol{\mu})\|_{Q'}^2$.
We point out that these error bounds (as well as the ones we are going to introduce in the following sections) are of no utility without an accompanying Offline-Online computational approach. In order to be computed in a very rapid and efficient way, both the dual norm of the residuals and the lower bounds $\beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu})$ have to be based on such a procedure. The *natural norm* version of the SCM algorithm for the computation of $\beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu})$ – applied for the first time this algorithm to saddle point Stokes problems, as shown in the work with Rozza and Huynh [279] – is analyzed in Sect. A.3 of Appendix A. Computation of dual norms of residuals is based instead on the Riesz representation of $r_{\mathbf{v}}(\cdot\,;\boldsymbol{\mu})$ and $r_p(\cdot\,;\boldsymbol{\mu})$ and on the affine decomposition of the parametric operators; see Sect. A.2.1 in the Appendix A for further details.

**Remark 3.8.** *Using the estimations* (3.27) *on the residual equations* (3.76) *it is possible to derive analogous error estimates*[5] *for the velocity and the pressure errors, separately:*

$$\|\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}) - \mathbf{v}_N(\boldsymbol{\mu})\|_V \leq \Delta_{N;\mathbf{v}}^{\mathcal{N}}(\boldsymbol{\mu}), \qquad \|p^{\mathcal{N}}(\boldsymbol{\mu}) - p_N(\boldsymbol{\mu})\|_Q \leq \Delta_{N;p}^{\mathcal{N}}(\boldsymbol{\mu}), \qquad \forall \boldsymbol{\mu}\in\mathcal{D},$$

*where*

$$\Delta_{N;\mathbf{v}}(\boldsymbol{\mu}) := \frac{\|r_{\mathbf{v}}(\cdot\,;\boldsymbol{\mu})\|_{V'}}{\alpha_{\mathcal{N}}^{LB}(\boldsymbol{\mu})} + \left(1 + \frac{\gamma_{a,\mathcal{N}}^{UB}(\boldsymbol{\mu})}{\alpha_{\mathcal{N}}^{LB}(\boldsymbol{\mu})}\right) \frac{\|r_p(\cdot\,;\boldsymbol{\mu})\|_{Q'}}{\beta_{Br,\mathcal{N}}^{LB}(\boldsymbol{\mu})} \tag{3.79}$$

$$\Delta_{N;p}(\boldsymbol{\mu}) := \frac{\|r_{\mathbf{v}}(\cdot\,;\boldsymbol{\mu})\|_{V'}}{\beta_{Br,\mathcal{N}}^{LB}(\boldsymbol{\mu})} + \frac{\gamma_{a,\mathcal{N}}^{UB}(\boldsymbol{\mu})}{\beta_{Br,\mathcal{N}}^{LB}(\boldsymbol{\mu})} \left(\frac{\|r_{\mathbf{v}}(\cdot\,;\boldsymbol{\mu})\|_{V'}}{\alpha_{\mathcal{N}}^{LB}(\boldsymbol{\mu})} + \frac{\alpha_{\mathcal{N}}^{LB}(\boldsymbol{\mu}) + \gamma_{a,\mathcal{N}}^{UB}(\boldsymbol{\mu})}{\alpha_{a,\mathcal{N}}^{LB}(\boldsymbol{\mu})\beta_{Br,\mathcal{N}}^{LB}(\boldsymbol{\mu})} \|r_p(\cdot\,;\boldsymbol{\mu})\|_{Q'}\right),$$

$\alpha_{\mathcal{N}}^{LB}(\boldsymbol{\mu})$ *and* $\gamma_{a,\mathcal{N}}^{UB}(\boldsymbol{\mu})$ *the lower bound of the coercivity constant* (3.47) *and the upper bound of the continuity constant* $\gamma_a^{\mathcal{N}}(\boldsymbol{\mu})$*, respectively, of* $a(\cdot,\cdot;\boldsymbol{\mu})$*. A detailed analysis of these error estimates is provided for instance in [111]. Moreover, if we compare the Babuška error bound* (3.78) *and the one obtained by summing the two separate bounds* (3.79) *as*

$$\|Y^{\mathcal{N}}(\boldsymbol{\mu}) - Y_N(\boldsymbol{\mu})\|_X \leq ((\Delta_{N;\mathbf{v}}^{\mathcal{N}}(\boldsymbol{\mu}))^2 + (\Delta_{N;p}(\boldsymbol{\mu}))^2)^{1/2} =: \tilde{\Delta}_N(\boldsymbol{\mu}),$$

*effectivities* $\eta_N(\boldsymbol{\mu}) = \Delta_N(\boldsymbol{\mu})/\|\mathbf{e}_N(\boldsymbol{\mu})\|_X$ *and* $\tilde{\eta}_N(\boldsymbol{\mu}) = \tilde{\Delta}_N(\boldsymbol{\mu})/\|\mathbf{e}_N(\boldsymbol{\mu})\|_X$ *are such that*

$$1 \leq \eta_N(\boldsymbol{\mu}) \leq \frac{\gamma_{a,\mathcal{N}}(\boldsymbol{\mu}) + \gamma_{b,\mathcal{N}}(\boldsymbol{\mu})}{\beta_{Ba,\mathcal{N}}(\boldsymbol{\mu})}, \qquad 1 \leq \tilde{\eta}_N(\boldsymbol{\mu}) \leq \mathcal{K}_{Br}^{\mathcal{N}}(\gamma_{a,\mathcal{N}}(\boldsymbol{\mu}) + \gamma_{b,\mathcal{N}}(\boldsymbol{\mu})), \qquad \forall \boldsymbol{\mu}\in\mathcal{D},$$

*where* $\mathcal{K}_{Br}^{\mathcal{N}} = \mathcal{K}_{Br}^{\mathcal{N}}(\alpha_{\mathcal{N}}^{-1}, \beta_{Br,\mathcal{N}}^{-1}, \gamma_{a,\mathcal{N}})$ *is the discrete counterpart of the factor* $\mathcal{K}_{Br}$ *introduced in* (3.32)*. Thanks to relationship* (3.33)*, by considering aggregate error estimates for both RB velocity and pressure of type* (3.31) *or* (3.32)*, we argue that the Babuška-based estimate is sharper than the Brezzi-based one since the "safety factor"* $\beta_{BA}$ *is in any case greater than* $1/\mathcal{K}_{Br}$*.*

[5] A former contribution on *a posteriori* error bounds for Stokes problem based on the splitting between viscous and pressure-divergence terms was provided by Rovas [272]. A recent approach has been proposed by Veroy and Gerner [110] based on a penalty method, thus reporting the problem in the coercive case.

*Numerical analysis supporting this conjecture are still in progress. In any case, the advantage of the Babuška-based error estimator is that only the (lower bound $\beta_{BA}^{\mathcal{N}}(\boldsymbol{\mu})$ of a) stability constant needs to be evaluated to get the error bound. Instead, the Brezzi-based error estimators would require the evaluation of the coercivity (lower bound) and continuity (upper bound) constants of $a(\cdot,\cdot;\boldsymbol{\mu})$ and the Brezzi inf-sup constant of $b(\cdot,\cdot;\boldsymbol{\mu})$.*
*Clearly, if we are interested in error bounds for outputs depending on either velocity or pressure, separate bounds like* (3.79) *can be more effective, provided that all the quantities appearing in the expression of the error bounds are computed.*

### 3.4.2 Linear outputs

We now build *a posteriori* error bounds for our outputs of interest (3.15), given by

$$s(\boldsymbol{\mu}) = L(Y(\boldsymbol{\mu});\boldsymbol{\mu}) + M(Y(\boldsymbol{\mu});\boldsymbol{\mu}) = L_{\mathbf{v}}(\mathbf{v}(\boldsymbol{\mu});\boldsymbol{\mu}) + L_p(p(\boldsymbol{\mu});\boldsymbol{\mu}), \tag{3.80}$$

dealing formerly with the linear case, and distinguishing between compliant and noncompliant cases. The quadratic case will be addressed in the next section.

#### Compliant case

Given the solution $(\mathbf{v}(\boldsymbol{\mu}), p(\boldsymbol{\mu}))$ to Stokes problem (3.4), in the compliant case we have $L_{\mathbf{v}}(\cdot;\boldsymbol{\mu}) = F(\cdot;\boldsymbol{\mu})$, $L_p(\cdot;\boldsymbol{\mu}) = G(\cdot;\boldsymbol{\mu})$. Correspondingly, the FE approximation of the output is given by

$$s^{\mathcal{N}}(\boldsymbol{\mu}) = L(Y^{\mathcal{N}}(\boldsymbol{\mu});\boldsymbol{\mu}) = F(\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu});\boldsymbol{\mu}) + G(p^{\mathcal{N}}(\boldsymbol{\mu});\boldsymbol{\mu}), \tag{3.81}$$

while the RB approximation of the output, considering a suitable correction as proposed in [243, 244] in order to improve the output accuracy, is given by

$$s_N(\boldsymbol{\mu}) = L(Y_N(\boldsymbol{\mu});\boldsymbol{\mu}) + r(Y_N(\boldsymbol{\mu});\boldsymbol{\mu}), \tag{3.82}$$

and thus $s^{\mathcal{N}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu}) = r_{\mathbf{v}}(\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu});\boldsymbol{\mu}) + r_p(e_p(\boldsymbol{\mu});\boldsymbol{\mu})$. Thanks to the relationship

$$\begin{aligned} |s^{\mathcal{N}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu})| &\leq \sup_{\mathbf{w}\in V} \frac{r_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{w}\|_V}\|\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu})\|_V + \sup_{q\in Q}\frac{r_p(q;\boldsymbol{\mu})}{\|q\|_Q}\|e_p(\boldsymbol{\mu})\|_Q \\ &= \|r_{\mathbf{v}}(\cdot;\boldsymbol{\mu})\|_{V'}\|\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu})\|_V + \|r_p(\cdot;\boldsymbol{\mu})\|_{Q'}\|e_p(\boldsymbol{\mu})\|_Q, \end{aligned}$$

and to the error bound (3.78) on velocity and pressure variables, the following result holds, yielding to the quadratic convergence effect (see e.g. [280]) also in the Stokes case:

**Proposition 3.9.** *Let us denote by $s^{\mathcal{N}}(\boldsymbol{\mu})$ and $s_N(\boldsymbol{\mu})$ the finite element and the reduced basis approximation, defined by* (3.81) *and* (3.82)*, respectively, of a linear output* (3.80) *in the compliant case. Then, the following error estimation holds:*

$$|s^{\mathcal{N}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu})| \leq 2\left(\frac{\|r_{\mathbf{v}}(\cdot;\boldsymbol{\mu})\|_{V'}^2 + \|r_p(\cdot;\boldsymbol{\mu})\|_{Q'}^2}{\beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu})}\right) := \Delta_N^{s,\,c}(\boldsymbol{\mu}), \qquad \forall\,\boldsymbol{\mu}\in\mathcal{D}, \tag{3.83}$$

#### Noncompliant case

Let us now consider the more general case where the output of interest can be written as a linear functional of $(\mathbf{v}(\boldsymbol{\mu}), p(\boldsymbol{\mu}))$ as $s(\boldsymbol{\mu}) = L(Y(\boldsymbol{\mu});\boldsymbol{\mu}) = L_{\mathbf{v}}(\mathbf{v}(\boldsymbol{\mu});\boldsymbol{\mu}) + L_p(p(\boldsymbol{\mu});\boldsymbol{\mu})$, where $L_{\mathbf{v}}(\cdot;\boldsymbol{\mu}) \in V'$ and $L_p(\cdot;\boldsymbol{\mu}) \in Q'$ for all $\boldsymbol{\mu}\in\mathcal{D}$. In this case, we can readily develop an *a posteriori*

error bound for $s_N(\boldsymbol{\mu})$: by standard arguments [280, 237] we have that

$$|s^{\mathcal{N}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu})| \leq ||L(\cdot;\boldsymbol{\mu})||_{(X^{\mathcal{N}})'} \Delta_N(\boldsymbol{\mu}),$$

where $\Delta_N(\boldsymbol{\mu})$ is given by (3.78). In case of outputs depending just on velocity or pressure fields, error bounds $\Delta_{N;\mathbf{v}}(\boldsymbol{\mu})$, $\Delta_{N;p}(\boldsymbol{\mu})$ defined by (3.79) can be used. We denote this method as *"primal-only"*. Although for many outputs this is perhaps the best approach (each additional output, and associated error bound, is a simple "add-on"), it has two deficiencies: *(i)* we lose the "quadratic convergence" effect (2.45) for outputs, and *(ii)* the effectivities may be unbounded (see e.g. [280]).

To overcome these limitations, we introduce the *dual problem* associated with $l(\cdot;\boldsymbol{\mu})$: find $(\boldsymbol{\psi}(\boldsymbol{\mu}), \lambda(\boldsymbol{\mu})) \in V \times Q$ such that

$$\begin{cases} a(\mathbf{w}, \boldsymbol{\psi}(\boldsymbol{\mu}); \boldsymbol{\mu}) + b(\lambda(\boldsymbol{\mu}), \mathbf{w}; \boldsymbol{\mu}) = -L_{\mathbf{v}}(\mathbf{w}; \boldsymbol{\mu}) & \forall\ \mathbf{w} \in V \\ b(q, \boldsymbol{\psi}(\boldsymbol{\mu}); \boldsymbol{\mu}) = -L_p(q; \boldsymbol{\mu}) & \forall\ \ q \in Q, \end{cases} \tag{3.84}$$

where $\Psi(\boldsymbol{\mu}) = (\boldsymbol{\psi}(\boldsymbol{\mu}), \lambda(\boldsymbol{\mu}))$ is denoted the dual (or adjoint) field. The corresponding FE approximation of the output is $s^{\mathcal{N}}(\boldsymbol{\mu}) = L(\Psi^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu}) = L_{\mathbf{v}}(\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu}) + L_p(p^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu})$, where $\Psi^{\mathcal{N}}(\boldsymbol{\mu}) = (\boldsymbol{\psi}^{\mathcal{N}}(\boldsymbol{\mu}), \lambda^{\mathcal{N}}(\boldsymbol{\mu})) \in X^{\mathcal{N}}$ is the solution of:

$$\begin{cases} a(\mathbf{w}^{\mathcal{N}}, \boldsymbol{\psi}^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu}) + b(\lambda^{\mathcal{N}}(\boldsymbol{\mu}), \mathbf{w}^{\mathcal{N}}; \boldsymbol{\mu}) = -L_{\mathbf{v}}(\mathbf{w}^{\mathcal{N}}; \boldsymbol{\mu}) & \forall\ \mathbf{w}^{\mathcal{N}} \in V^{\mathcal{N}} \\ b(q^{\mathcal{N}}, \boldsymbol{\psi}^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu}) = -L_p(q^{\mathcal{N}}; \boldsymbol{\mu}) & \forall\ \ q^{\mathcal{N}} \in Q^{\mathcal{N}}. \end{cases} \tag{3.85}$$

The dual problem is subject to the same Brezzi inf-sup condition of the primal problem. Its RB approximation is as follows: find $\Psi_M(\boldsymbol{\mu}) = (\boldsymbol{\psi}_M(\boldsymbol{\mu}), \lambda_M(\boldsymbol{\mu})) \in V_M^{\mathcal{N}} \times Q_M^{\mathcal{N}}$ such that

$$\begin{cases} a(\mathbf{w}_M, \boldsymbol{\psi}_M(\boldsymbol{\mu}); \boldsymbol{\mu}) + b(\lambda_M(\boldsymbol{\mu}), \mathbf{w}_M; \boldsymbol{\mu}) = -L_{\mathbf{v}}(\mathbf{w}_M; \boldsymbol{\mu}) & \forall\ \mathbf{w}_M \in V_M^{\mathcal{N}} \\ b(q_M, \boldsymbol{\psi}_M(\boldsymbol{\mu}); \boldsymbol{\mu}) = -L_p(q_M; \boldsymbol{\mu}) & \forall\ \ q_M \in Q_M^{\mathcal{N}}. \end{cases} \tag{3.86}$$

where the RB dual spaces $V_M^{\mathcal{N}}$, $Q_M^{\mathcal{N}}$ are built by means of a greedy algorithm and the dimension $M \ll \mathcal{N}$ is *a priori* different from the dimension $N$ of the primal RB spaces. Reduced spaces and approximation for the dual problem are denoted in the same way as for the primal problem, where $M$ (instead of $N$) is the dimension of the dual space.

Similarly to (3.75), we define the errors $\mathbf{e}_v^{du}(\boldsymbol{\mu}) = \boldsymbol{\psi}^{\mathcal{N}}(\boldsymbol{\mu}) - \boldsymbol{\psi}_M(\boldsymbol{\mu})$ and $e_p^{du}(\boldsymbol{\mu}) = \lambda^{\mathcal{N}}(\boldsymbol{\mu}) - \lambda_M(\boldsymbol{\mu})$ and the residuals $r_{\mathbf{v}}^{du}(\cdot;\boldsymbol{\mu}) \in V'$ and $r_p^{du}(\cdot;\boldsymbol{\mu}) \in Q'$ as follows:

$$\begin{aligned} r_{\mathbf{v}}^{du}(\mathbf{w};\boldsymbol{\mu}) &:= -L_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu}) - a(\mathbf{w}, \boldsymbol{\psi}_M(\boldsymbol{\mu});\boldsymbol{\mu}) - b(\lambda_M(\boldsymbol{\mu}), \mathbf{w};\boldsymbol{\mu}), \\ r_p^{du}(q;\boldsymbol{\mu}) &:= -L_p(q;\boldsymbol{\mu}) - b(q, \boldsymbol{\psi}_M(\boldsymbol{\mu});\boldsymbol{\mu}). \end{aligned} \tag{3.87}$$

Note that

$$\begin{aligned} r_{\mathbf{v}}^{du}(\mathbf{w};\boldsymbol{\mu}) &= a(\mathbf{e}_{\mathbf{u}}^{du}(\boldsymbol{\mu}), \mathbf{w};\boldsymbol{\mu}) + b(e_p^{du}(\boldsymbol{\mu}), \mathbf{w};\boldsymbol{\mu}) \qquad \forall\ \mathbf{w} \in V^{\mathcal{N}}, \\ r_p^{du}(q;\boldsymbol{\mu}) &= b(q, \mathbf{e}_{\mathbf{u}}^{du}(\boldsymbol{\mu});\boldsymbol{\mu}) \qquad\qquad\qquad\quad \forall\ q \in Q^{\mathcal{N}}; \end{aligned} \tag{3.88}$$

equivalently,

$$r^{du}(W;\boldsymbol{\mu}) = A(\Psi^{\mathcal{N}}(\boldsymbol{\mu}) - \Psi_M(\boldsymbol{\mu}), W;\boldsymbol{\mu}), \qquad \forall\ W \in X^{\mathcal{N}} \equiv V^{\mathcal{N}} \times Q^{\mathcal{N}},$$

where $r^{du}(W;\boldsymbol{\mu}) := r^{du}_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu}) + r^{du}_{p}(q;\boldsymbol{\mu})$. The RB approximation of the output is thus given by

$$s_N(\boldsymbol{\mu}) = L(Y_N(\boldsymbol{\mu});\boldsymbol{\mu}) - r(\Psi_M(\boldsymbol{\mu});\boldsymbol{\mu}) \tag{3.89}$$

where the correction helps improving the accuracy of the approximation. Hence, we have

$$\begin{aligned} s^{\mathcal{N}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu}) &= L(Y^{\mathcal{N}}(\boldsymbol{\mu});\boldsymbol{\mu}) - L(Y_N(\boldsymbol{\mu});\boldsymbol{\mu}) + r(\Psi_M(\boldsymbol{\mu});\boldsymbol{\mu}) \\ &= L_{\mathbf{v}}(\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu});\boldsymbol{\mu}) + L_p(e_p(\boldsymbol{\mu});\boldsymbol{\mu}) + r(\Psi_M(\boldsymbol{\mu});\boldsymbol{\mu}); \end{aligned}$$

thanks to (3.84) and (3.75), this expression can be also written as

$$\begin{aligned} s^{\mathcal{N}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu}) =& - a(\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu}), \boldsymbol{\psi}^{\mathcal{N}}(\boldsymbol{\mu});\boldsymbol{\mu}) - b(\lambda^{\mathcal{N}}(\boldsymbol{\mu}), \mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu});\boldsymbol{\mu}) - b(e_p(\boldsymbol{\mu}), \boldsymbol{\psi}^{\mathcal{N}}(\boldsymbol{\mu});\boldsymbol{\mu}) \\ &+ a(\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu}), \boldsymbol{\psi}_M(\boldsymbol{\mu});\boldsymbol{\mu}) + b(\lambda_M(\boldsymbol{\mu}), \mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu});\boldsymbol{\mu}) + b(e_p(\boldsymbol{\mu}), \boldsymbol{\psi}_M(\boldsymbol{\mu});\boldsymbol{\mu}) \\ =& - a(\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu}), \mathbf{e}^{du}_{\mathbf{v}}(\boldsymbol{\mu});\boldsymbol{\mu}) - b(e^{du}_p(\boldsymbol{\mu}), \mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu});\boldsymbol{\mu}) - b(e_p(\boldsymbol{\mu}), \mathbf{e}^{du}_{\mathbf{v}}(\boldsymbol{\mu});\boldsymbol{\mu}) \\ =& - r^{du}_{\mathbf{v}}(\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu});\boldsymbol{\mu}) - r^{du}_{p}(e_p(\boldsymbol{\mu});\boldsymbol{\mu}). \end{aligned}$$

Using the same procedure exploited in the compliant case, we obtain the following expression:

$$\begin{aligned} |s^{\mathcal{N}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu})| &\le \sup_{\mathbf{w}\in V} \frac{r^{du}_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{w}\|_V} \|\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu})\|_V + \sup_{q\in Q} \frac{r^{du}_{p}(q;\boldsymbol{\mu})}{\|\phi\|_Q} \|e_p(\boldsymbol{\mu})\|_Q \\ &= \|r^{du}_{\mathbf{v}}(\cdot;\boldsymbol{\mu})\|_{V'} \|\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu})\|_V + \|r^{du}_{p}(\cdot;\boldsymbol{\mu})\|_{Q'} \|e_p(\boldsymbol{\mu})\|_Q, \end{aligned}$$

so that the error bound is given by a combination of the dual norms of the dual residuals and the error on the primal variables. We have thus shown the following proposition:

**Proposition 3.10.** *Let us denote by $s^{\mathcal{N}}(\boldsymbol{\mu})$ and $s_N(\boldsymbol{\mu})$ the finite element and the reduced basis approximation, defined by* (3.81) *and* (3.89)*, respectively, of a linear output* (3.80) *in the noncompliant case. Then, the following error estimation holds:*

$$|s^{\mathcal{N}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu})| \le \Delta_N^{s,n}(\boldsymbol{\mu}), \qquad \forall\, \boldsymbol{\mu} \in \mathcal{D},$$

*where*

$$2 \left( \frac{\|r^{du}_{\mathbf{v}}(\cdot;\boldsymbol{\mu})\|^2_{V'} + \|r^{du}_{p}(\cdot;\boldsymbol{\mu})\|^2_{Q'}}{\beta^{LB}_{Ba,\mathcal{N}}(\boldsymbol{\mu})} \right)^{1/2} \left( \frac{\|r_{\mathbf{v}}(\cdot;\boldsymbol{\mu})\|^2_{V'} + \|r_{p}(\cdot;\boldsymbol{\mu})\|^2_{Q'}}{\beta^{LB}_{Ba,\mathcal{N}}(\boldsymbol{\mu})} \right)^{1/2}. \tag{3.90}$$

This result is the noncompliant version of (3.83): in fact, it extends the estimation obtained for the compliant case, since in the latter case, choosing $V_N^{\mathcal{N}} \equiv V_M^{\mathcal{N}}$ and $Q_N^{\mathcal{N}} \equiv Q_M^{\mathcal{N}}$, we have $\psi_M(\boldsymbol{\mu}) \equiv -\mathbf{v}_N(\boldsymbol{\mu})$, $\lambda_M(\boldsymbol{\mu}) = -p_N(\boldsymbol{\mu})$ and the same expression for primal and dual residuals.

### 3.4.3 Quadratic outputs

Let us move on to the case of a quadratic output

$$s(\boldsymbol{\mu}) = Q(\mathbf{v}(\boldsymbol{\mu}), \mathbf{v}(\boldsymbol{\mu});\boldsymbol{\mu}), \tag{3.91}$$

expressed as a quadratic bilinear form over the velocity space $V$. Many examples of outputs of physical interest, such as energy dissipation, drag forces (conveniently expressed), vorticity, can be recast in this framework. In order to derive a suitable *a posteriori* error estimation, we exploit a dual-based strategy as in the noncompliant case of linear outputs.

However, the dual problem now involves the derivative (with respect to the primal variables) of the output, which is a linear functional of the (primal) velocity itself. The case of a quadratic output depending on the pressure field, not addressed here, is straightforward and can be treated in the same way.

Let us first consider a preliminary error bound which does not show the so-called *quadratic effect*. In fact, by using a continuity argument, we have directly

$$\begin{aligned} |s^{\mathcal{N}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu})| = & \left|Q(\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}), \mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu}) - Q(\mathbf{v}_N(\boldsymbol{\mu}), \mathbf{v}_N(\boldsymbol{\mu}); \boldsymbol{\mu})\right| \\ \leq & \left|Q(\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}), \mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}) - \mathbf{v}_N(\boldsymbol{\mu}); \boldsymbol{\mu}) + Q(\mathbf{v}_N(\boldsymbol{\mu}), \mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}) - \mathbf{v}_N(\boldsymbol{\mu}); \boldsymbol{\mu})\right| \\ \leq & \xi_Q(\boldsymbol{\mu}) \left(\|\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu})\|_V + \|\mathbf{v}_N(\boldsymbol{\mu})\|_V\right) \|\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}) - \mathbf{v}_N(\boldsymbol{\mu})\|_V, \end{aligned}$$

where $\xi_Q(\boldsymbol{\mu}) > 0$ is the continuity constant of $Q(\cdot, \cdot; \boldsymbol{\mu})$; the norm of the FE velocity approximation can be bounded by means of stability estimates introduced in Sect. 3.2.3 and similar estimates can be derived also for the RB approximation. As already pointed out in the previous section, the primal-only approach has some deficiencies: moreover, in the quadratic case, the estimate would be affected by a factor bounding the norm of the truth FE approximation.

We thus introduce a *primal-dual* alternative as for the case of a noncompliant linear output, exploiting a general approach based on duality principles [24]. This approach is widely used in optimal control and error estimation for adaptivity (leading to the so-called *dual-weighted-residual* method) – and has been already used for *a posteriori* error estimation in optimal control problems for advection-diffusion equations solved within the RB framework [73].

Within this context, the estimation problem is recast in a more general optimal control framework, by introducing a suitable Lagrangian functional for a fictitious minimization problem, which in our case reads as follows:

$$\min_{\mathbf{v} \in V} Q(\mathbf{v}(\boldsymbol{\mu}), \mathbf{v}(\boldsymbol{\mu}); \boldsymbol{\mu}) \quad \text{s.t.} \quad \begin{cases} a(\mathbf{v}(\boldsymbol{\mu}), \boldsymbol{\Phi}; \boldsymbol{\mu}) + b(p(\boldsymbol{\mu}), \boldsymbol{\Phi}; \boldsymbol{\mu}) = F(\boldsymbol{\Phi}; \boldsymbol{\mu}) & \forall\ \boldsymbol{\Phi} \in V \\ b(\phi, \mathbf{v}(\boldsymbol{\mu}); \boldsymbol{\mu}) = G(\phi; \boldsymbol{\mu}) & \forall\ \ \phi \in Q, \end{cases}$$

which is in fact equivalent to the evaluation of the output $s(\boldsymbol{\mu})$ from the solution of the (primal) Stokes problem. For the sake of notation, the explicit $\boldsymbol{\mu}$-dependence is often omitted in this section. By introducing the following Lagrangian functional:

$$L((\mathbf{v}, p), (\boldsymbol{\psi}, \lambda)) := Q(\mathbf{v}, \mathbf{v}) + F(\boldsymbol{\psi}) + G(\lambda) - a(\mathbf{v}, \boldsymbol{\psi}) - b(p, \boldsymbol{\psi}) - b(\lambda, \mathbf{v}), \qquad \forall (\boldsymbol{\psi}, \lambda) \in V \times Q,$$

the solution $(\mathbf{v}, p)$ of the previous minimization problem is characterized as the first component of a stationary point of $L((\mathbf{v}, p), (\boldsymbol{\psi}, \lambda))$. Hence, we seek solutions $(Y, \Psi) \equiv ((\mathbf{v}, p), (\boldsymbol{\psi}, \lambda)) \in (V \times Q)^2$ to the following Euler-Lagrange system (first-order necessary conditions):

$$\begin{cases} L'_{(\boldsymbol{\psi}, \lambda)}[\boldsymbol{\Phi}, \phi] = 0 \quad \Leftrightarrow \quad \begin{cases} a(\mathbf{v}, \boldsymbol{\Phi}; \boldsymbol{\mu}) + b(p, \boldsymbol{\Phi}; \boldsymbol{\mu}) = F(\boldsymbol{\Phi}; \boldsymbol{\mu}) & \forall\ (\boldsymbol{\Phi}, \phi) \in V \times Q \\ b(\phi, \mathbf{v}; \boldsymbol{\mu}) = G(\phi; \boldsymbol{\mu}) \end{cases} \\ L'_{(\mathbf{v}, p)}[\boldsymbol{\Phi}, \phi] = 0 \quad \Leftrightarrow \quad \begin{cases} a(\boldsymbol{\Phi}, \boldsymbol{\psi}; \boldsymbol{\mu}) + b(\lambda, \boldsymbol{\Phi}; \boldsymbol{\mu}) = Q(\mathbf{v}, \boldsymbol{\Phi}; \boldsymbol{\mu}) & \forall\ (\boldsymbol{\Phi}, \phi) \in V \times Q. \\ b(\phi, \boldsymbol{\psi}; \boldsymbol{\mu}) = 0 \end{cases} \end{cases} \tag{3.92}$$

We recall that these equations are stated for, and our *a posteriori* error estimation for RB solution is computed with respect to, the truth FE solution. Following the so-called *optimize-then-discretize* approach, we can introduce the truth FE approximation of (3.92) and then a subsequent RB approximation, which is given by:

$$\begin{cases} L'_{(\boldsymbol{\psi}_M,\lambda_M)}[\boldsymbol{\Phi},\phi]=0 \Leftrightarrow & \begin{cases} a(\mathbf{v}_N,\boldsymbol{\Phi};\boldsymbol{\mu})+b(p_N,\boldsymbol{\Phi};\boldsymbol{\mu})=F(\boldsymbol{\Phi};\boldsymbol{\mu}) \\ b(\phi,\mathbf{v}_N;\boldsymbol{\mu})=G(\phi;\boldsymbol{\mu}) \end{cases} & \forall\ (\boldsymbol{\Phi},\phi)\in V_N^{\mathcal{N}}\times Q_N^{\mathcal{N}} \\ L'_{(\mathbf{v}_N,p_N)}[\boldsymbol{\Phi},\phi]=0 \Leftrightarrow & \begin{cases} a(\boldsymbol{\Phi},\boldsymbol{\psi}_M;\boldsymbol{\mu})+b(\lambda_M,\boldsymbol{\Phi};\boldsymbol{\mu})=Q(\mathbf{v}_N,\boldsymbol{\Phi};\boldsymbol{\mu}) \\ b(\phi,\boldsymbol{\psi}_M;\boldsymbol{\mu})=0 \end{cases} & \forall\ (\boldsymbol{\Phi},\phi)\in V_M^{\mathcal{N}}\times Q_M^{\mathcal{N}}, \end{cases} \tag{3.93}$$

where, as in the previous section, $\Psi_M(\boldsymbol{\mu})=(\boldsymbol{\psi}_M(\boldsymbol{\mu}),\lambda_M(\boldsymbol{\mu}))\in V_M^{\mathcal{N}}\times Q_M^{\mathcal{N}}$ is the RB dual approximation. As in the linear case, primal errors and residuals satisfy equation (3.76), here rewritten for the reader's convenience (expliciting the dependence on the primal variables):

$$\begin{aligned} &r_{\mathbf{v}}^{pr}(\boldsymbol{\Phi};Y_N(\boldsymbol{\mu});\boldsymbol{\mu}))=a(\mathbf{e}_{\mathbf{v}}^{pr}(\boldsymbol{\mu}),\boldsymbol{\Phi};\boldsymbol{\mu})+b(e_p^{pr}(\boldsymbol{\mu}),\boldsymbol{\Phi};\boldsymbol{\mu}),\\ &r_p^{pr}(\phi;Y_N(\boldsymbol{\mu});\boldsymbol{\mu})=b(\mathbf{e}_{\mathbf{v}}^{pr}(\boldsymbol{\mu}),\mathbf{v}_N(\boldsymbol{\mu});\boldsymbol{\mu}). \end{aligned} \tag{3.94}$$

In the same way, dual errors and residuals satisfy the following relationship (such as in (3.88)):

$$\begin{aligned} r_{\mathbf{v}}^{du}(\boldsymbol{\Phi};\Psi_M(\boldsymbol{\mu});\boldsymbol{\mu}) &= Q(\mathbf{v}_N(\boldsymbol{\mu});\boldsymbol{\Phi};\boldsymbol{\mu})-a(\boldsymbol{\Phi},\boldsymbol{\psi}_M(\boldsymbol{\mu});\boldsymbol{\mu})-b(\lambda_M(\boldsymbol{\mu}),\boldsymbol{\Phi};\boldsymbol{\mu})\\ &= a(\boldsymbol{\Phi},\mathbf{e}_{\mathbf{v}}^{du}(\boldsymbol{\mu});\boldsymbol{\mu})+b(e_p^{du}(\boldsymbol{\mu}),\boldsymbol{\Phi};\boldsymbol{\mu}),\\ r_p^{du}(\phi;\Psi_M(\boldsymbol{\mu});\boldsymbol{\mu}) &= -b(\phi,\boldsymbol{\psi}_M(\boldsymbol{\mu});\boldsymbol{\mu})=b(\phi,\mathbf{e}_{\mathbf{v}}^{du}(\boldsymbol{\mu});\boldsymbol{\mu}). \end{aligned} \tag{3.95}$$

In this case, we can show the following results:

**Proposition 3.11.** *Let us denote by $s^{\mathcal{N}}(\boldsymbol{\mu})$ and $s_N(\boldsymbol{\mu})$ the finite element and the reduced basis approximation of a quadratic output* (3.91) *of the velocity fields. Then, the following error estimation holds:*

$$|s^{\mathcal{N}}(\boldsymbol{\mu})-s_N(\boldsymbol{\mu})|\le\Delta_N^{s,q}(\boldsymbol{\mu}),\qquad\forall\,\boldsymbol{\mu}\in\mathcal{D},$$

*where*

$$\Delta_N^{s,q}(\boldsymbol{\mu})=\frac{\|r_{\mathbf{v}}^{pr}(\cdot;\boldsymbol{\mu})\|_{V'}^2+\|r_p^{pr}(\cdot;\boldsymbol{\mu})\|_{Q'}^2}{\beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu})}+\frac{\|r_{\mathbf{v}}^{du}(\cdot;\boldsymbol{\mu})\|_{V'}^2+\|r_p^{du}(\cdot;\boldsymbol{\mu})\|_{Q'}^2}{\beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu})} \tag{3.96}$$

*Proof.* Let us start by considering the following relationship:

$$s^{\mathcal{N}}(\boldsymbol{\mu})-s_N(\boldsymbol{\mu})=L(Y^{\mathcal{N}}(\boldsymbol{\mu}),\Psi^{\mathcal{N}}(\boldsymbol{\mu}))-L(Y_N(\boldsymbol{\mu}),\Psi_M(\boldsymbol{\mu})).$$

By indicating with $\boldsymbol{\xi}^{\mathcal{N}}(\boldsymbol{\mu})=(Y^{\mathcal{N}}(\boldsymbol{\mu}),\Psi^{\mathcal{N}}(\boldsymbol{\mu}))$ and $\boldsymbol{\xi}_{N,M}(\boldsymbol{\mu})=(Y_N(\boldsymbol{\mu}),\Psi_M(\boldsymbol{\mu}))$, the previous difference can be rewritten as

$$s^{\mathcal{N}}(\boldsymbol{\mu})-s_N(\boldsymbol{\mu})=L(\boldsymbol{\xi}^{\mathcal{N}}(\boldsymbol{\mu}))-L(\boldsymbol{\xi}_{N,M}(\boldsymbol{\mu}))=\int_0^1 L'(\boldsymbol{\xi}_{N,M}+t\boldsymbol{\eta};\boldsymbol{\eta})dt$$

where $\boldsymbol{\eta}=\boldsymbol{\eta}(\boldsymbol{\mu})=(Y^{\mathcal{N}}(\boldsymbol{\mu})-Y_N(\boldsymbol{\mu}),\Psi^{\mathcal{N}}(\boldsymbol{\mu})-\Psi_M(\boldsymbol{\mu}))$ is a compact expression of the primal and dual errors $(\mathbf{e}_{\mathbf{v}}^{pr}(\boldsymbol{\mu}),e_p^{pr}(\boldsymbol{\mu}))$, $(\mathbf{e}_{\mathbf{v}}^{du}(\boldsymbol{\mu}),e_p^{du}(\boldsymbol{\mu}))$, and the condensed notation $L'(\xi;\boldsymbol{\Pi})$ denotes

$$L'(\boldsymbol{\xi};\boldsymbol{\Pi}):=L'_{(\mathbf{v},p)}[\boldsymbol{\Phi},\phi]+L'_{(\mathbf{z},q)}[\tilde{\boldsymbol{\Phi}},\tilde{\phi}],\quad\boldsymbol{\xi}=(Y,\Psi)=((\mathbf{v},p),(\boldsymbol{\psi},\lambda)),\quad\boldsymbol{\Pi}=((\boldsymbol{\Phi},\phi),(\tilde{\boldsymbol{\Phi}},\tilde{\phi})).$$

By using the trapezoidal rule for approximating the integral, we obtain

$$s^{\mathcal{N}}(\boldsymbol{\mu})-s_N(\boldsymbol{\mu})=\frac{1}{2}L'(\boldsymbol{\xi}^{\mathcal{N}}(\boldsymbol{\mu});\boldsymbol{\eta})+\frac{1}{2}L'(\boldsymbol{\xi}_{N,M}(\boldsymbol{\mu});\boldsymbol{\eta})+R$$

where the remainder term $R$ is given by

$$R = \frac{1}{2}\int_0^1 L'''(\boldsymbol{\xi}_{N,M}(\boldsymbol{\mu}) + t\boldsymbol{\eta}; \boldsymbol{\eta}, \boldsymbol{\eta}, \boldsymbol{\eta}) t(t-1) dt$$

and vanishes if $a(\cdot,\cdot;\boldsymbol{\mu})$, $b(\cdot,\cdot;\boldsymbol{\mu})$ are bilinear and $s(\cdot)$ is quadratic (see e.g. [24]). Since $L'(\boldsymbol{\xi}^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\eta}) = 0$ for all $\boldsymbol{\eta} \in X^{\mathcal{N}} \times X^{\mathcal{N}}$ due to the Euler-Lagrange system (3.92) and

$$\begin{aligned} &L'(\boldsymbol{\xi}_{N,M}(\boldsymbol{\mu}); \boldsymbol{\eta}) = L'_{(\mathbf{v}_N, p_N)}[\mathbf{e}^{pr}_{\mathbf{v}}(\boldsymbol{\mu}), e^{pr}_p(\boldsymbol{\mu})] + L'_{(\psi_M, \lambda_M)}[\mathbf{e}^{du}_{\mathbf{v}}(\boldsymbol{\mu}), e^{du}_p(\boldsymbol{\mu})] = \\ &= r^{pr}_{\mathbf{v}}(\mathbf{e}^{pr}_{\mathbf{v}}(\boldsymbol{\mu}); Y_N(\boldsymbol{\mu})) + r^{pr}_p(e^{pr}_p(\boldsymbol{\mu}); Y_N(\boldsymbol{\mu})) + r^{du}_{\mathbf{v}}(\mathbf{e}^{du}_{\mathbf{v}}(\boldsymbol{\mu}); \Psi_M(\boldsymbol{\mu})) + r^{du}_p(e^{du}_p(\boldsymbol{\mu}); \Psi_M(\boldsymbol{\mu})), \end{aligned}$$

we end up with

$$\begin{aligned} s^{\mathcal{N}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu}) = \frac{1}{2}\left(r^{pr}_{\mathbf{v}}(\mathbf{e}^{pr}_{\mathbf{v}}(\boldsymbol{\mu}); Y_N(\boldsymbol{\mu})) + r^{pr}_p(e^{pr}_p(\boldsymbol{\mu}); Y_N(\boldsymbol{\mu}))\right) \\ + \frac{1}{2}\left(r^{du}_{\mathbf{v}}(\mathbf{e}^{du}_{\mathbf{v}}(\boldsymbol{\mu}); \Psi_M(\boldsymbol{\mu})) + r^{du}_p(e^{du}_p(\boldsymbol{\mu}); \Psi_M(\boldsymbol{\mu}))\right) \end{aligned}$$

so that the following *a posteriori* error bound on the output holds:

$$\begin{aligned} |s^{\mathcal{N}}(\boldsymbol{\mu}) - s_N(\boldsymbol{\mu})| \leq \; & \frac{1}{2}\left(\|r^{pr}_{\mathbf{v}}(\cdot; Y_N(\boldsymbol{\mu}))\|_{V'} \|\mathbf{e}^{pr}_{\mathbf{v}}(\boldsymbol{\mu})\|_V + \|r^{pr}_p(\cdot; Y_N(\boldsymbol{\mu}))\|_{Q'} \|e^{pr}_p(\boldsymbol{\mu})\|_Q\right) \\ + \; & \frac{1}{2}\left(\|r^{du}_{\mathbf{v}}(\cdot; \Psi_M(\boldsymbol{\mu}))\|_{V'} \|\mathbf{e}^{du}_{\mathbf{v}}(\boldsymbol{\mu})\|_X + \|r^{du}_p(\cdot; \Psi_M(\boldsymbol{\mu}))\|_{Q'} \|e^{du}_p(\boldsymbol{\mu})\|_Q\right), \end{aligned}$$

which has a similar form of estimate (3.90). The result easily follows by using the error estimates on velocity and pressure fields. □

The evaluation of (both primal and dual) residual norms to compute the output error bounds (3.83)-(3.90) in the case of linear outputs, and (3.96) in the case of quadratic outputs, is based on the same Offline-Online procedure already introduced for error bounds on velocity and pressure fields, and detailed in Sect. A.2.1 of Appendix A.

### 3.4.4 Navier-Stokes equations

We move on to the Navier-Stokes case, describing an error estimation procedure for velocity and pressure fields jointly, based on the BRR theory. We require some slight modifications with respect to the linear preliminaries: also for the Navier-Stokes problem the *a posteriori* error estimation takes advantage of the dual norm of residuals and of an effective lower bound of the (parametric) stability factor, given in this case by the Babuška inf-sup constant $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$ referred to the *Fréchet* derivative of the global operator $\tilde{A}(\cdot,\cdot;\boldsymbol{\mu})$, defined in (3.54). We first define the residuals $r_{\mathbf{v}}(\cdot\,;\boldsymbol{\mu})$ and $r_p(\cdot\,;\boldsymbol{\mu})$ by

$$\begin{aligned} &r_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu}) := F(\mathbf{w};\boldsymbol{\mu}) - \tilde{a}(\mathbf{v}_N(\boldsymbol{\mu}), \mathbf{w};\boldsymbol{\mu}) - b(p_N(\boldsymbol{\mu}), \mathbf{w};\boldsymbol{\mu}) - c(\mathbf{v}_N(\boldsymbol{\mu}), \mathbf{v}_N(\boldsymbol{\mu}), \mathbf{w};\boldsymbol{\mu}), \\ &r_p(q;\boldsymbol{\mu}) := G(q;\boldsymbol{\mu}) - b(q, \mathbf{v}_N(\boldsymbol{\mu});\boldsymbol{\mu}), \end{aligned} \tag{3.97}$$

for any $\mathbf{w} \in V^{\mathcal{N}}$, $q \in Q^{\mathcal{N}}$, respectively, where $\tilde{a}(\cdot,\cdot;\boldsymbol{\mu}) = a(\cdot,\cdot;\boldsymbol{\mu}) + d(\cdot,\cdot;\boldsymbol{\mu})$. Equivalently, $r(W;\boldsymbol{\mu}) := r_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu}) + r_p(q;\boldsymbol{\mu})$ is such that

$$r(W;\boldsymbol{\mu}) = \tilde{F}(W;\boldsymbol{\mu}) - \tilde{A}(Y_N(\boldsymbol{\mu}), W;\boldsymbol{\mu}) \qquad \forall\, W \in X^{\mathcal{N}}. \tag{3.98}$$

Moreover, let us denote by $\|r(\cdot;\boldsymbol{\mu})\|_{X'} = \sup_{W\in Y^{\mathcal{N}}} r(W;\boldsymbol{\mu})/\|W\|_X$ the dual norm of the residual and $\beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$ a computable lower bound for $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}) = \beta_{\tilde{A},\mathcal{N}}(Y_N(\boldsymbol{\mu});\boldsymbol{\mu})$, recalling that

$$\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}) \equiv \beta_{\tilde{A},\mathcal{N}}(Y_N(\boldsymbol{\mu})(\boldsymbol{\mu});\boldsymbol{\mu}) = \inf_{V\in X^{\mathcal{N}}} \sup_{W\in X^{\mathcal{N}}} \frac{d\tilde{A}(Y_N(\boldsymbol{\mu})(\boldsymbol{\mu});\boldsymbol{\mu})(V,W)}{\|V\|_X\|W\|_X},$$

where we now highlight the dependence on the RB solution $Y_N(\boldsymbol{\mu})$ where the derivative is evaluated. Since the error bound is related to the RB solution $Y_N(\boldsymbol{\mu})$, the derivative of the global operator, as well as the stability factor, have to be evaluated with respect to $Y_N(\boldsymbol{\mu})$. We next introduce a *proximity* indicator (a *non-dimensional* measure of the the residual) required by the BRR theory:

$$\tau_N(\boldsymbol{\mu}) = \frac{4\gamma(\rho;\boldsymbol{\mu})\|r(\cdot;\boldsymbol{\mu})\|_{X'}}{(\beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}))^2},$$

where $\gamma(\rho;\boldsymbol{\mu}) \equiv \gamma_c^{\mathcal{N}}(\boldsymbol{\mu})$ is the (discrete) continuity constant of the trilinear form $c(\cdot,\cdot,\cdot;\boldsymbol{\mu})$, depending on the Sobolev embedding constant $\rho_{\mathcal{N}}$ defined in (3.46). Here in the continuity constant we highlight the dependence by $\rho = \rho_{\mathcal{N}}$ (omitting the subscript $_{\mathcal{N}}$ for simplicity), since the Sobolev embedding constant is a further, requested ingredient to be computed for the sake of error estimation. Finally, we define $N^*(\boldsymbol{\mu})$ such that $\tau_N(\boldsymbol{\mu}) < 1$ for $N \geq N^*(\boldsymbol{\mu})$; we require that $N^*(\boldsymbol{\mu}) \leq N_{\max}$, for any $\boldsymbol{\mu} \in \mathcal{D}$. Then, we can state the following theorem:

**Theorem 3.12.** *Let us denote by $Y^{\mathcal{N}}(\boldsymbol{\mu})$ and by $Y_N(\boldsymbol{\mu})$ the truth approximation* (3.45) *and the reduced basis approximation* (3.66)*, respectively. If $N \geq N^*(\boldsymbol{\mu})$, there exists a unique solution $Y^{\mathcal{N}}(\boldsymbol{\mu})$ to* (3.45) *in the open ball*

$$\mathcal{B}_X\left(Y_N(\boldsymbol{\mu}); \frac{\beta^{LB}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})}{2\gamma(\rho;\boldsymbol{\mu})}\right) = \left\{Y \in X \,:\, \|Y - Y_N(\boldsymbol{\mu})\|_X \leq \frac{\beta^{LB}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})}{2\gamma(\rho;\boldsymbol{\mu})}\right\}.$$

*Furthermore, the following* a posteriori *error estimation holds:*

$$\|Y^{\mathcal{N}}(\boldsymbol{\mu}) - Y_N(\boldsymbol{\mu})\|_X \leq \frac{\beta^{LB}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})}{2\gamma(\rho;\boldsymbol{\mu})}\left(1 - \sqrt{1-\tau_N(\boldsymbol{\mu})}\right) =: \Delta_N(\boldsymbol{\mu}), \qquad \forall\, \boldsymbol{\mu} \in \mathcal{D}. \quad (3.99)$$

*Proof.* This result is a slight variation of the result shown by Veroy and Patera in [312], that in turn derives from the abstract results of the BRR theory stated in Thm. 2.1 of [56]. For this reason, we sketch only the main points, reminding to the original references for further details. First of all, let us denote by $g(\cdot,\cdot;\boldsymbol{\mu}) : X \times X \to \mathbb{R}$ the following operator:

$$g(W,V;\boldsymbol{\mu}) = \tilde{A}(W,V;\boldsymbol{\mu}) - \tilde{F}(V;\boldsymbol{\mu})$$

and remark that $dg(Z;W,V;\boldsymbol{\mu}) = d\tilde{A}(Z;W,V;\boldsymbol{\mu})$, recalling that

$$d\tilde{A}(Z;W,V;\boldsymbol{\mu}) = A(W,V;\boldsymbol{\mu}) + C(Z,W,V;\boldsymbol{\mu}) + C(W,Z,V;\boldsymbol{\mu}).$$

Moreover, let us remark that, by Taylor expansion,

$$\tilde{A}(Y_2^{\mathcal{N}},W^{\mathcal{N}};\boldsymbol{\mu}) - \tilde{A}(Y_1^{\mathcal{N}},W^{\mathcal{N}};\boldsymbol{\mu}) = \int_0^1 d\tilde{A}(Y_1^{\mathcal{N}} + t(Y_2^{\mathcal{N}} - Y_1^{\mathcal{N}}); Y_2^{\mathcal{N}} - Y_1^{\mathcal{N}}, V;\boldsymbol{\mu}), \quad (3.100)$$

whereas from the continuity of $c(\cdot,\cdot,\cdot;\boldsymbol{\mu})$ we have

$$|d\tilde{A}(Y_2^{\mathcal{N}};W^{\mathcal{N}},V^{\mathcal{N}}) - d\tilde{A}(Y_1^{\mathcal{N}};W^{\mathcal{N}},V^{\mathcal{N}})| \leq 2\gamma(\rho;\boldsymbol{\mu})\|W^{\mathcal{N}}\|_X\|V^{\mathcal{N}}\|_X\|Y_2^{\mathcal{N}} - Y_1^{\mathcal{N}}\|_X. \quad (3.101)$$

The central ingredient is the construction of a suitable *contraction mapping* which underlies a standard *fixed-point* argument. To this aim, we define the operator $H^{\boldsymbol{\mu}} : X^{\mathcal{N}} \to X^{\mathcal{N}}$ as

$$dg(Y_N(\boldsymbol{\mu}); T^{\boldsymbol{\mu}}(W^{\mathcal{N}}), V^{\mathcal{N}}; \boldsymbol{\mu}) = dg(Y_N(\boldsymbol{\mu}); W^{\mathcal{N}}, V^{\mathcal{N}}; \boldsymbol{\mu}) - g(W^{\mathcal{N}}, V^{\mathcal{N}}; \boldsymbol{\mu}), \quad \forall V^{\mathcal{N}} \in X^{\mathcal{N}}; \tag{3.102}$$

it is evident that a fixed point of $H^{\boldsymbol{\mu}}(W^{\mathcal{N}})$, such that $H^{\boldsymbol{\mu}}(\overline{W}^{\mathcal{N}}) = \overline{W}^{\mathcal{N}}$, is a zero of $g$, i.e. $g(\overline{W}^{\mathcal{N}}, V; \boldsymbol{\mu}) = 0$ for any $V \in X^{\mathcal{N}}$, that is $\overline{W}^{\mathcal{N}}$ solves the truth approximation problem (3.45). Let us now consider $W_1^{\mathcal{N}} \in \overline{\mathcal{B}}(Y_N(\boldsymbol{\mu}); \alpha)$, $W_2^{\mathcal{N}} \in \overline{\mathcal{B}}(Y_N(\boldsymbol{\mu}); \alpha)$, for a given $\alpha > 0$; from (3.100)–(3.102) and the definition of $\beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$, it follows that

$$\|H^{\boldsymbol{\mu}}(W_2^{\mathcal{N}}) - H^{\boldsymbol{\mu}}(W_1^{\mathcal{N}})\|_X \leq \frac{2\gamma(\rho;\boldsymbol{\mu})\alpha}{\beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})} \|W_2^{\mathcal{N}} - W_1^{\mathcal{N}}\|_X,$$

so that $\|H^{\boldsymbol{\mu}}(W_2^{\mathcal{N}}) - H^{\boldsymbol{\mu}}(W_1^{\mathcal{N}})\|_X \leq \|W_2^{\mathcal{N}} - W_1^{\mathcal{N}}\|_X$, i.e. $H^{\boldsymbol{\mu}}(\cdot)$ is a contraction mapping, for any $\alpha \in [0, \beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})/2\gamma(\rho;\boldsymbol{\mu}))$. Furthermore, from the same relationships, we also have that

$$\|H^{\boldsymbol{\mu}}(W^{\mathcal{N}}) - Y_N(\boldsymbol{\mu})\|_X \leq \frac{\|r(\cdot;\boldsymbol{\mu})\|_{X'} + \alpha^2\gamma(\rho;\boldsymbol{\mu})}{\beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})}, \qquad \forall\, W^{\mathcal{N}} \in \overline{\mathcal{B}}(Y_N(\boldsymbol{\mu}), \alpha),$$

where the residual $r(W;\boldsymbol{\mu})$ is given by $r(W;\boldsymbol{\mu}) = -g(Y_N(\boldsymbol{\mu}), W; \boldsymbol{\mu})$ and its dual norm such that $|g(Y_N(\boldsymbol{\mu}), W^{\mathcal{N}}; \boldsymbol{\mu})| \leq \|r(\cdot;\boldsymbol{\mu})\|_{X'} \|W^{\mathcal{N}}\|_X$, for any $W^{\mathcal{N}} \in X^{\mathcal{N}}$. Thus, since

$$\frac{\|r(\cdot;\boldsymbol{\mu})\|_{X'} + \alpha^2\gamma(\rho;\boldsymbol{\mu})}{\beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})} < \alpha \Leftrightarrow \frac{\beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})}{2\gamma(\rho;\boldsymbol{\mu})} \left(1 - \sqrt{1 - \tau_N(\boldsymbol{\mu})}\right) < \alpha < \frac{\beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})}{2\gamma(\rho;\boldsymbol{\mu})} \left(1 + \sqrt{1 - \tau_N(\boldsymbol{\mu})}\right),$$

we obtain that for any $\alpha \in [\Delta_N(\boldsymbol{\mu}), (\beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})/2\gamma(\rho;\boldsymbol{\mu}))(1 + \sqrt{1 - \tau_N(\boldsymbol{\mu})})]$ the operator $H^{\boldsymbol{\mu}}(\cdot)$ maps $\overline{\mathcal{B}}(Y_N(\boldsymbol{\mu}), \alpha)$ into itself. In conclusion, for any $\alpha \in [\Delta_N(\boldsymbol{\mu}), \beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})/2\gamma(\rho;\boldsymbol{\mu})]$ $H^{\boldsymbol{\mu}}(\cdot)$ is a contraction that maps $\overline{\mathcal{B}}(Y_N(\boldsymbol{\mu}), \alpha)$ into itself, and thus it admits a unique fixed points, i.e. there exists a unique function $Y^{\mathcal{N}}(\boldsymbol{\mu}) \in \overline{\mathcal{B}}(Y_N(\boldsymbol{\mu}), \alpha)$ which is a solution of the truth approximation (3.45).

□

**Remark 3.13.** *This theorem generalizes available versions of the* a posteriori *error bounds for Navier-Stokes equations in two ways:* (i) *by developing the computations on the global Navier-Stokes operator $\tilde{A}(\cdot,\cdot;\boldsymbol{\mu})$ (including also the pressure terms) and involving the lower bound $\beta^{LB}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$, we provide an aggregate error estimate for both RB velocity and pressure;* (ii) *by considering the continuity factor $\gamma(\rho;\boldsymbol{\mu})$ of the trilinear form instead of just the Sobolev embedding constant $\rho$, we can deal with problems involving parametrized trilinear terms – entering, for instance, in geometrical parametrizations. Moreover, in nonaffine problems the correction factor $\max_{q=1,\ldots,Q_c} \|\eta^q\|_{L^\infty(\Omega)}$ affects the definition of $\gamma(\rho;\boldsymbol{\mu})$, whereas it reduces to 1 in the affine case. In any case, we presume that in the Navier-Stokes case the parameter-dependent solution lays on an isolated branch.*

As for the Stokes case, the computational challenge is the development of suitable and efficient procedures for the calculation of the quantities appearing in the error bound (3.99): dual norms of residuals, embedding constants and lower bounds $\beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$ of the stability factor. Details about the evaluation of these quantities are provided in the Appendix A – see Sect. A.2.2, A.3.2 and A.4, respectively. Error bounds for linear and quadratic outputs related to Navier-Stokes equations have been treated for instance in [78], but not employed. In our framework, based on error estimation for velocity and pressure jointly, their analysis and implementation is another topic of forthcoming research [209].

## 3.5 Further remarks and open issues

Before presenting some numerical examples dealing with RB approximation of both Stokes and Navier-Stokes problems, we point out major numerical difficulties and bottlenecks. First of all, let us remark that the most challenging computational aspects are *(i)* the estimation of lower bounds for stability factors and *(ii)* the calculation of dual norms of residuals, more than the RB approximation itself.
Concerning the latter point, computation of dual norms of residuals is based instead on the Riesz representation of $r_{\mathbf{v}}(\cdot\,;\boldsymbol{\mu})$ and $r_p(\cdot\,;\boldsymbol{\mu})$ and on the affine decomposition of the parametric operators. We thus need to store during the Offline stage several structures representing the scalar products of parameter-independent terms, so that during the Online stage we just need to evaluate the $\Theta_\bullet(\boldsymbol{\mu})$ functions and combine the former contributions. This procedure requires the storage of up to $\mathcal{O}(Q_c^2N^2 + Q_c(Q_aN))$ full matrices of dimension $\mathcal{O}(N)$ in the nonlinear case (see Sect. A.2.2), which may lead to some difficulties in case of affine expansions made by large number of terms. This case occurs whenever the affine structure of the operators is recovered through the EIM procedure, featuring sometimes $Q_\bullet = \mathcal{O}(10^2)$ terms in each operator.

Regarding instead the estimation of lower bounds for stability factors, we rely on the (*natural norm* version of the) Successive Constraint Method (SCM), which is based on the successive solution of suitable linear optimization problems and features an efficient Offline-Online strategy; see Sect. A.3 of Appendix A. As shown in the forthcoming section, the natural norm SCM algorithm has been successfully applied to several Stokes problems, dealing with both (physical and geometrical) affine and (geometrical) nonaffine parametrizations, and $p \leq 3$ parameters. Some preliminary results are available also for a simple Navier-Stokes problem dealing with a physical affine parametrization.
From our current experience, the natural norm SCM algorithm starts converging very slowly when the number of parameters is getting larger and larger, at least for $p \geq 4$. However, we observe that for free-shape nonaffine geometrical parametrizations such as in the FFD case, adding more geometrical parameters does not have a huge impact on the range of stability factors (see also [184]). We were thus able to circumvent the difficulty of poor convergence of the SCM algorithm by proposing several alternatives, such as: *(i)* global constant lower bound in case of free-shape nonaffine parametrizations; *(ii)* parametric lower bounds obtained by activating up to $\tilde{p} \leq 3$ parameters among the ones defining the parametrization; *(iii)* surrogate models, based e.g. on RBF interpolation, as sketched in Sect. A.3.2 of Appendix A.
More rigorous alternatives capable to overcome also the difficulty of large affine expansions (e.g. in nonaffine problems treated through EIM procedures) take advantage of the possibility to approximate a nonaffine operator by means of two affine operators at different levels of accuracy. In this way, SCM is applied on the coarse level to obtain a lower bound for the stability factors, and this bound is extended to the fine level by adding a proper correction term. This procedure has been implemented for simple scalar coercive and noncoercive problems in [183] and its extension to Stokes and (possibly Navier-Stokes) operators is currently under investigation.

## 3.6 Numerical examples

Flows in pipes and channels or around bodies are of great interest in fluid mechanics applications [234], especially when they can be studied in a parametrized geometrical configuration. We present in the following some numerical results dealing with RB approximation and error estimation for steady parametrized Stokes and Navier-Stokes flows. We consider both physical and geometrical parametrizations: the former involve boundary conditions and physical coefficients, while the latter consist of both (automatic, piecewise) affine and global nonaffine mappings.

### 3.6.1 Stokes equations

The following examples consider low Reynolds viscous flows described by 2D steady Stokes equations in different geometries; they can be seen as examples for the design of parametrized fluidic devices or considered as elements of more complex modular fluidic systems. By increasing complexity, we deal with classic Poiseuille and Couette flows, a flow in a channel contraction, around a curved bluff body and in a bifurcation. The following subsections are devoted to the description of these problems with results – submitted in a recent paper [279] – showing the application of the RB methodology. All numerical details concerning the construction of RB spaces and the computation of lower bounds $\beta^{\mathrm{LB}}_{Ba,\mathcal{N}}(\boldsymbol{\mu})$ for the inf-sup stability constants are reported in Tables 3.1 and 3.2.

#### Case 1: Poiseuille and Couette flows

This first example deals with a classical flow in straight pipes of uniform cross-section, known as Hagen-Poiseuille and Couette flows [234]. In the former a parabolic velocity profile is imposed at the inflow, while in the latter we deal with a flow in the space between two parallels sections, one of which is moving relative to the other.
For the Poiseuille case, we consider the physical domain $\Omega_o(\boldsymbol{\mu})$ shown in Fig. 3.1 and $P = 2$ parameters. Here $\mu_1 = \nu$, while $\mu_2$ is a geometrical parameter representing the length of the right narrow channel. In this case a dimensional formulation for the equations has been considered. The parameter domain is given by $\mathcal{D} = [0.25, 0.75] \times [1.5, 2.5]$. The forcing term is $\mathbf{f} = (1, 0)$.

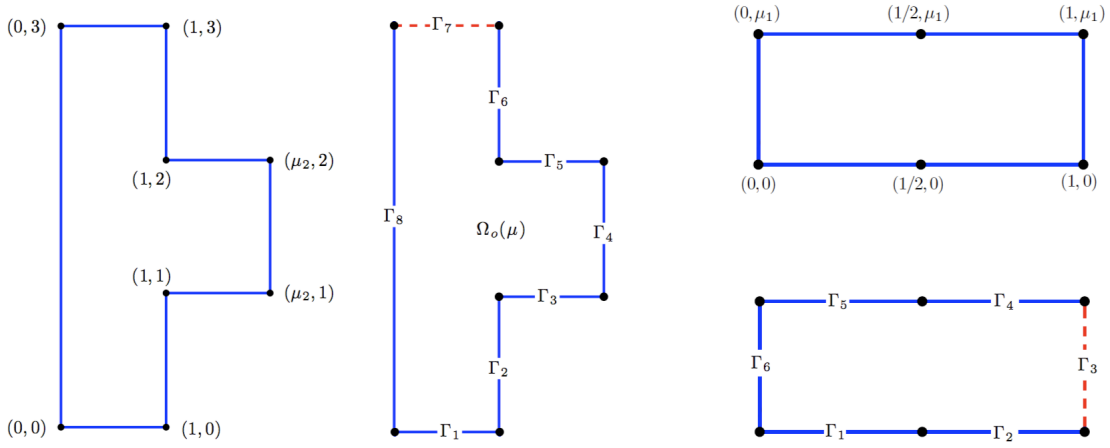



Figure 3.1: Parametrized geometry and domain boundaries for the Poiseuille (left) and the Couette (right) case.

We impose the following boundary conditions (with $\Gamma_D \equiv \Gamma_{D_0} = \partial\Omega \setminus (\Gamma_1 \cup \Gamma_7)$):

$$\begin{aligned} &\mathbf{v} = \mathbf{0} && \text{on} \quad \Gamma_D \\ &v_1 = 0, \quad v_2 = 4x_1(1 - x_1) && \text{on} \quad \Gamma_1 \\ &v_1 = 0, \quad -pn_2 + \mu_1 \frac{\partial v_2}{\partial x_2} n_2 = 0 && \text{on} \quad \Gamma_7 \end{aligned} \qquad \text{(Poiseuille case)}$$

where $\mathbf{n} = (n_1, n_2)^T$ denotes the normal unit vector. For the Couette case, we consider the physical domain $\Omega_o(\boldsymbol{\mu})$ shown in Fig. 3.1 (right side) and $P = 1$ parameter, $\mu_1 \in [0.5, 2]$, being both the height of the channel and the maximum value of the linear profile of inlet velocity prescribed. The forcing term is $\mathbf{f} = (0, -1)$. Denoting $\Gamma_D = \partial\Omega \setminus \Gamma_3$, we impose the following

boundary conditions:

$$\begin{aligned} &v_1 = x_2, \quad v_2 = 0 && \text{on} \quad \Gamma_D \\ &v_2 = 0, \quad -pn_1 + \frac{\partial v_1}{\partial x_1} n_1 = 0 && \text{on} \quad \Gamma_3 \end{aligned} \qquad \text{(Couette case)}$$

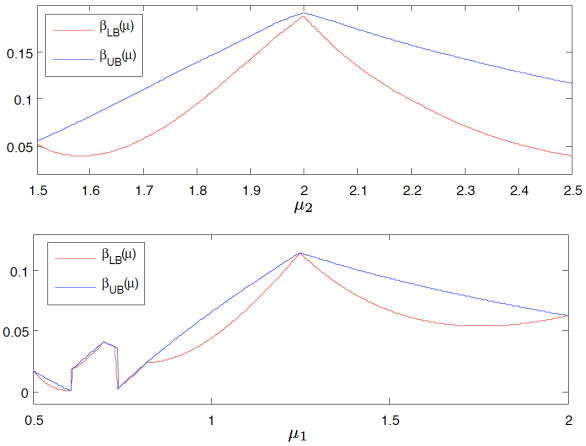



Figure 3.2: Poiseuille (top) and Couette (bottom) cases: lower and upper bounds for the Babuška inf-sup constant; here $\Xi_{train}$ is a uniform sample of size $n_{\text{train}} = 1000$: $\beta^{LB}_{Ba,\mathcal{N}}(\boldsymbol{\mu})$ (red curve) and $\beta^{UB}_{Ba,\mathcal{N}}(\boldsymbol{\mu})$ (blue curve) as a function of $\mu_2$ for the Poiseuille case (being $\mu_1 = 0.5$ fixed) and of $\mu_1$ for the Couette case after 27 and 15 iterations of the SCM greedy algorithm, respectively.

With a fixed tolerance $\varepsilon^{\text{RB}}_{\text{tol}} = 10^{-2}$, $N_{max} = 7$ and $N_{max} = 6$ basis functions have been selected for the Poiseuille and the Couette cases, respectively, through the greedy procedure. We also plot in Fig. 3.2 the SCM lower and upper bounds for the Babuška inf-sup constant (e.g. for a selected value of $\mu_1$ in the Poiseuille case, using for both the cases in the Online evaluation a uniform test sample of 1000 parameter values). For these cases the output of interest is provided by the visualization of velocity and pressure contour fields; two examples are reported in Fig. 3.3-3.4.

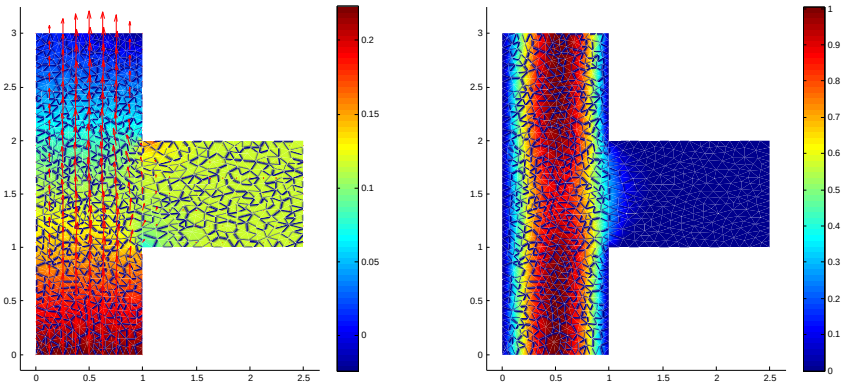


Figure 3.3: Poiseuille case: representative solution for pressure with streamlines (left) and velocity (right) for $\boldsymbol{\mu} = [0.25, 2.5]$.

We plot in Fig. 3.5 the errors between the "truth" FE solution and the RB approximation, for $N = 1, \ldots, N_{max}$, and the corresponding error bounds. We remark both the rigor and the sharpness of the error bounds, being the effectivity $\eta_N(\boldsymbol{\mu}) := \Delta_N(\boldsymbol{\mu}) / \|Y^{\mathcal{N}}(\boldsymbol{\mu}) - Y_N(\boldsymbol{\mu})\|_Y$ larger than 1 (rigor) and not far from unity (sharpness).

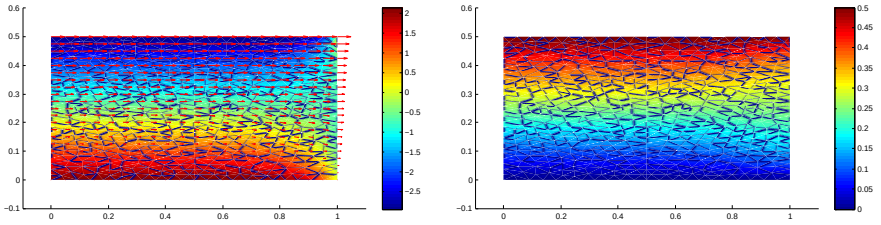


Figure 3.4: Couette case: representative solution for pressure with streamlines (left) and velocity (right) for $\boldsymbol{\mu} = 0.5$.

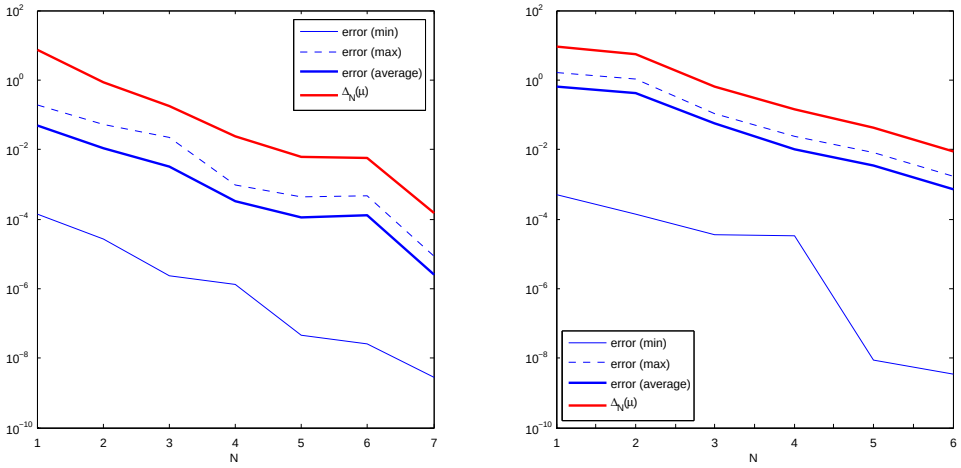



Figure 3.5: Poiseuille (left) and Couette (right) cases: *a posteriori* error bounds and (minimum, maximum and average over $\Xi_{\text{test}}$) computed errors between the "truth" FE solution and the RB approximation, for $N = 1, \ldots, N_{max}$. Here $\Xi_{\text{test}}$ is a uniform sample of size $n_{\text{test}} = 1000$.

### Case 2: A channel contraction

The problem of the change of a sectional area characterizes many engineering problems dealing with internal flows. The physical phenomena observed in the channel at the change of the sectional area are based on the continuity equation; another important aspect is the calculation of flow rates at a selected section of the channel. We consider the physical domain $\Omega_o(\boldsymbol{\mu})$ shown in Fig. 3.6; we identify the regions $\mathcal{R}_\ell$, $1 \leq \ell \leq 2$, which represent the portions of the channel with different sectional area.
We consider $P = 3$ parameters; here $\mu_1$, $\mu_2$, $\mu_3$ are geometrical parameters defined in Fig. 3.6: $\mu_1$ is the length of the larger zone of the channel before the contraction, $\mu_2$ is the length of the narrow zone of the channel (just before the outflow) and $\mu_3$ is the diameter of the channel at the inflow. The parameter domain is given by $\mathcal{D} = [3, 5] \times [3, 5] \times [2.5, 3]$. The forcing term is $\mathbf{f} = (0, 0)$. We impose the following boundary conditions:

$$
\begin{aligned}
&\mathbf{v} = \mathbf{0} && \text{on} \quad \Gamma_1, \Gamma_2, \Gamma_4, \Gamma_6, \Gamma_7, \Gamma_8 \\
&v_2 = 0, \quad -pn_1 + \frac{\partial v_1}{\partial x_1} n_1 = 1 && \text{on} \quad \Gamma_5. \\
&v_2 = 0, \quad -pn_1 + \frac{\partial v_1}{\partial x_1} n_1 = -1 && \text{on} \quad \Gamma_9, \Gamma_{10}, \Gamma_{11}.
\end{aligned}
$$

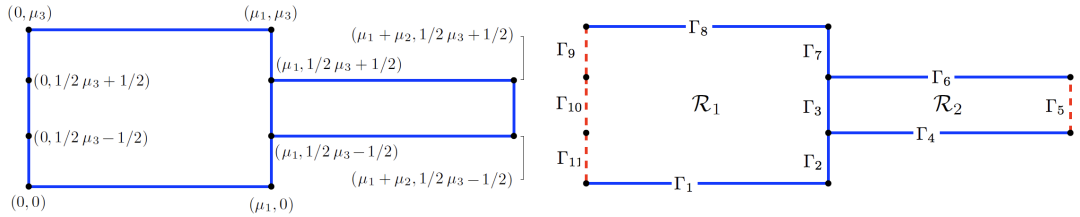



Figure 3.6: Parametrized geometry and domain boundaries for the channel contraction case.

The output of interest is the flowrate on $\Gamma_3$ (internal boundary at the interface, on which the continuity of velocity and stresses is assured), given by

$$s(\boldsymbol{\mu}) = \int_{\Gamma_3} u_1(\boldsymbol{\mu}) d\Gamma.$$

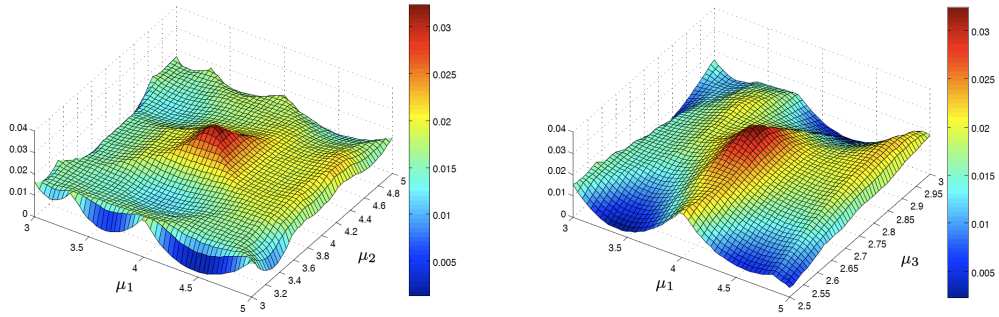



Figure 3.7: Channel contraction case: lower bounds for the Babuška inf-sup constant; here $\Xi_{train}$ is a uniform sample of size $n_{\text{train}} = 2500$: $\beta^{LB}_{Ba,\mathcal{N}}(\boldsymbol{\mu})$as a function of $\mu_1, \mu_2$ (left, being $\mu_3 = 2.75$ fixed) and of $\mu_1, \mu_3$ (right, being $\mu_2 = 4$ fixed) after 49 iterations of the SCM greedy algorithm.

The greedy procedure for the construction of the primal and dual RB spaces selects $N^{pr}_{max} = 11$ and $N^{du}_{max} = 17$ basis functions for the primal and the dual problem, respectively, with a fixed tolerance $\varepsilon^{\text{RB}}_{\text{tol}} = 10^{-2}$. We plot in Fig. 3.7 the SCM lower bounds for the Babuška inf-sup constant, using in the Online evaluation a uniform test sample of size $n_{\text{test}} = 2500$. In Fig. 3.8 we report some representative solutions for selected values of the parameters.

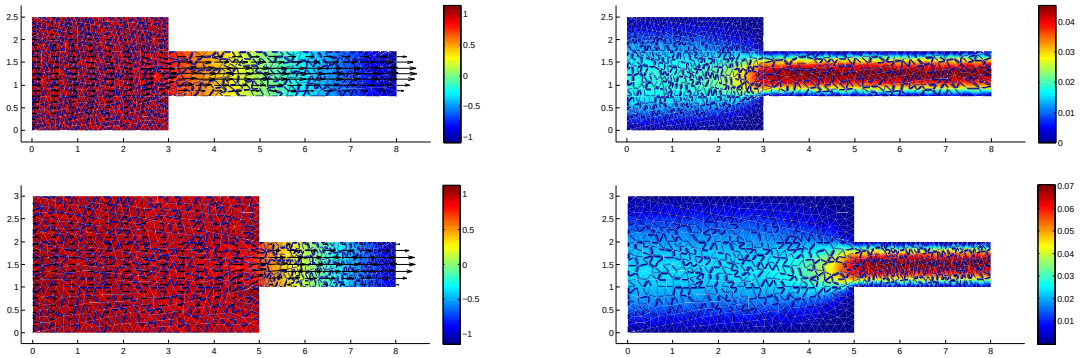


Figure 3.8: Channel contraction case: representative solutions for pressure with streamlines (left) and velocity (right) for $\boldsymbol{\mu} = [3, 5, 2.5]$ and $\boldsymbol{\mu} = [5, 3, 3]$.

We plot in Fig. 3.9 the errors between the "truth" FE solution and the RB approximation, for $N = 1, \ldots, N_{max}$, and the corresponding error bounds.

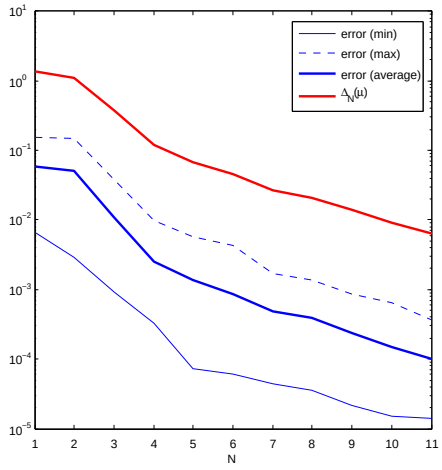



Figure 3.9: Channel contraction case: *a posteriori* error bounds and (minimum, maximum and average over $\Xi_{\text{test}}$) computed errors between the "truth" FE solution and the RB approximation, for $N = 1, \ldots, N_{max}$. Here $\Xi_{\text{test}}$ is a uniform sample of size $n_{\text{test}} = 500$.

In Fig. 3.10 we plot the computed output, together with the related error bound, as functions of $\mu_1$ and $\mu_2$, being $\mu_3$ fixed to its intermediate value. We recall the quadratic effect recovered by introducing and solving the dual problem in the case of a noncompliant output.

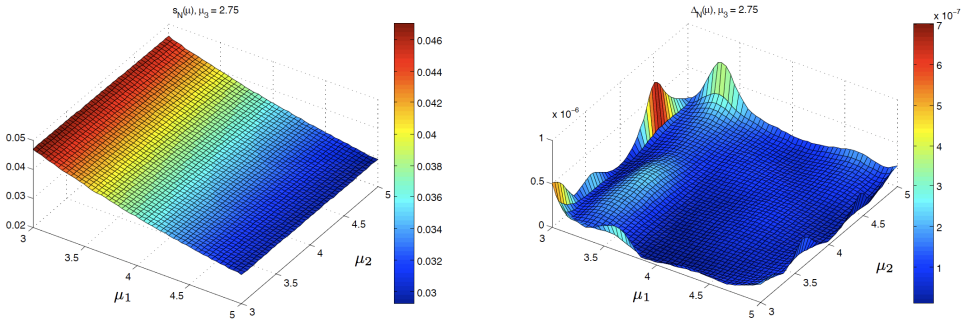



Figure 3.10: Channel contraction case: output $s_N(\boldsymbol{\mu})$ (left) and related error bound $\Delta_N^s(\boldsymbol{\mu})$ (right) as functions of $\mu_1, \mu_2$, with $\mu_3 = 2.75$. Average time for output evaluation is $0.148s$.

### Case 3: A curved bluff body

A common problem in fluid dynamics is the drag minimization around a body which is in relative motion in a fluid (see Sect. 1.5). Here we consider a simplified version of the drag minimization problem addressed in [71], in which drag forces are minimized by controlling the velocity through the body boundary. Our goal is to compute the Stokes flow and related drag forces around a profile in relative motion with a laminar viscous fluid, with respect to simple parametric variations. We consider the geometrical setting depicted in Fig. 3.11: here $\mu_1 \in [0.1, 0.25]$ is a geometrical parameter representing the body length, while $\mu_2 \in [-25, 25]$ is the Neumann datum prescribed on the boundaries $\Gamma_9 \cup \Gamma_{11}$: as in [71], this corresponds to regulate the aspiration or the blowing of the boundary layer for reducing the effects of the vortices coming off from the rear of the body.

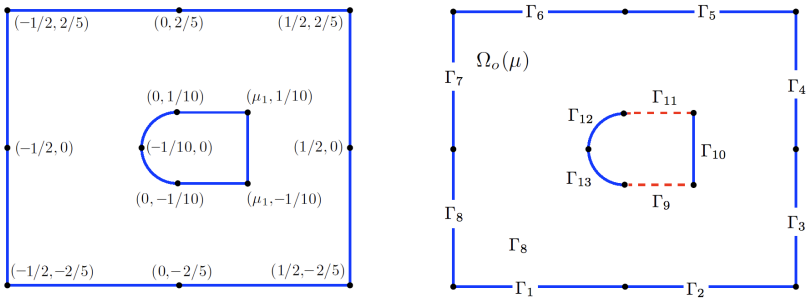



Figure 3.11: Parametrized geometry and domain boundaries for the curved bluff body case.

The forcing term is $\mathbf{f} = (0, 0)$. A parabolic flow is imposed at the inlet $\Gamma_7 \cup \Gamma_8$, while a free-stress condition is imposed at the outflow $\Gamma_3 \cup \Gamma_4$. Thus, we impose the following boundary conditions:

$$
\begin{aligned}
&\mathbf{v} = \mathbf{0} && \text{on} \quad \Gamma_1, \Gamma_2, \Gamma_5, \Gamma_6, \Gamma_{10}, \Gamma_{12}, \Gamma_{13}, \\
&v_1 = \alpha(x_2 - 0.4)(x_2 + 0.4), \quad v_2 = 0 && \text{on} \quad \Gamma_7, \Gamma_8, \\
&v_1 = 0, \quad -pn_2 + \frac{\partial v_2}{\partial x_2} n_2 = \mu_2 && \text{on} \quad \Gamma_9, \\
&v_1 = 0, \quad -pn_2 + \frac{\partial v_2}{\partial x_2} n_2 = \mu_2 && \text{on} \quad \Gamma_{11}, \\
&-p\mathbf{n} + \nu \frac{\partial \mathbf{v}}{\partial \mathbf{n}} = \mathbf{0} && \text{on} \quad \Gamma_3, \Gamma_4.
\end{aligned}
$$

where $\alpha = 0.16$ in order to have a maximum inlet velocity equal to 1. A first output of interest is the drag force acting on the Dirichlet boundary of the body $\Gamma_B = \Gamma_{10} \cup \Gamma_{12} \cup \Gamma_{13}$, given by

$$
s(\boldsymbol{\mu}) = \int_{\Gamma_B} \left( p\mathbf{n} - \frac{\partial \mathbf{v}}{\partial \mathbf{n}} \right) \cdot \hat{\mathbf{v}}_D d\Gamma,
$$

where $\hat{\mathbf{v}}_D = (1, 0)$ is the direction of the inflow velocity. By means of the greedy procedure (with a fixed tolerance $\varepsilon_{\text{tol}}^{\text{RB}} = 10^{-2}$) $N_{max}^{pr} = 12$ and $N_{max}^{du} = 6$ basis functions have been selected for the primal and the dual problem, respectively. We also plot in Fig. 3.12 the SCM lower bounds $\beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu})$ for the inf-sup constant; clearly, they do not depend on $\mu_2$, which does not affect the left-hand-side of the Stokes operator. In Fig. 3.13 we report some representative solutions for selected values of the parameters.

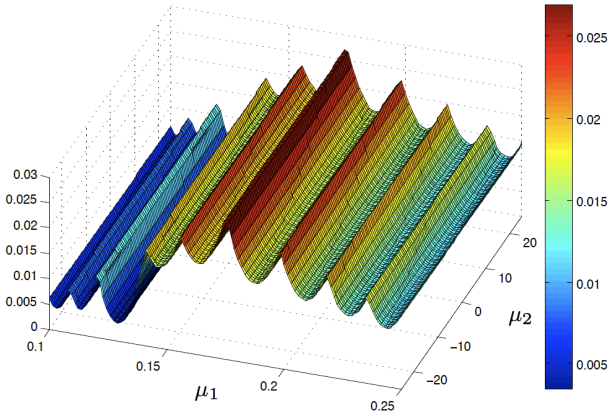



Figure 3.12: Curved bluff body case: lower bounds $\beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu})$ for the Babuška inf-sup constant as a function of $\mu_1, \mu_2$; here $\Xi_{train}$ is a uniform sample of size $n_{\text{train}} = 2500$.

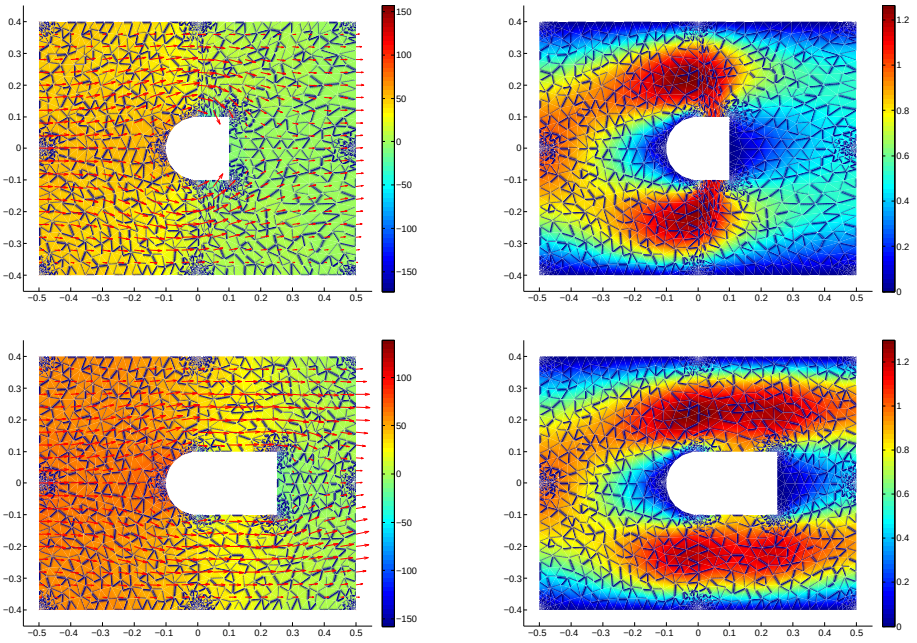


Figure 3.13: Curved bluff body case: representative solutions for pressure with streamlines (left) and velocity (right) for $\boldsymbol{\mu} = [0.1, -25]$ (top) and $\boldsymbol{\mu} = [0.25, 25]$ (bottom).

We can underline a strong sensitivity of the flow with respect to geometrical variations and, clearly, also on the aspiration/blowing of the fluid across the body. In Fig. 3.14 we plot the computed output, together with the related error bound. The output behaves as a non-monotonic function w.r.t. the two parameters. There is a different influence of the bluff body geometry (i.e. short or long body) w.r.t. the shear layers and the separation.
We plot in Fig. 3.15 the errors between the "truth" FE solution and the RB approximation, for $N = 1, \dots, N_{max}$, and the corresponding error bounds (we remark that $\Delta_N^s(\boldsymbol{\mu}) \approx 10^{-7} \div 10^{-8}$).

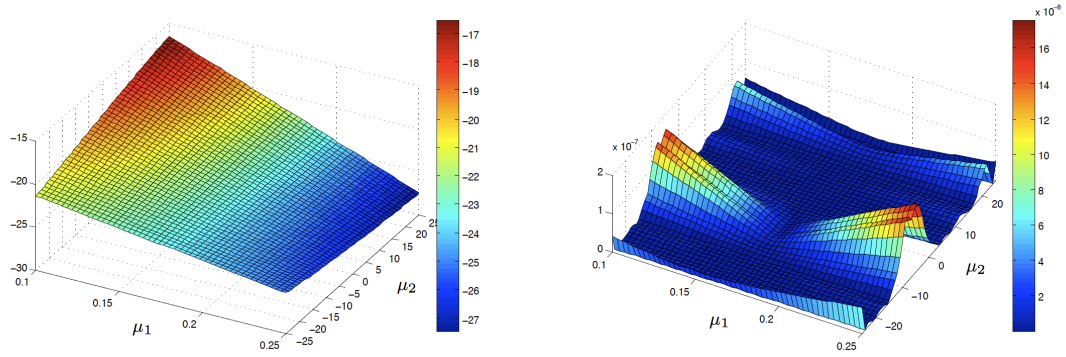



Figure 3.14: Curved bluff body case: computed output $s_N(\boldsymbol{\mu})$ (left) and related error bound $\Delta_N^s(\boldsymbol{\mu})$ (right) as a function of $\mu_1, \mu_2$. The average time for Online output evaluation is $0.087s$.

A second output of interest is the viscous energy dissipation over the domain, given by

$$s(\boldsymbol{\mu}) = \nu \int_\Omega |\nabla \mathbf{v}(\boldsymbol{\mu})|^2 d\Omega \equiv a(\mathbf{v}, \mathbf{v}; \boldsymbol{\mu}),$$

thus fitting the quadratic case of Sect. 3.4.3 with $Q(\cdot, \cdot; \boldsymbol{\mu}) = a(\cdot, \cdot; \boldsymbol{\mu})$. In the case of a full Dirichlet problem, this output is in fact equivalent to the *drag* output considered above [27].

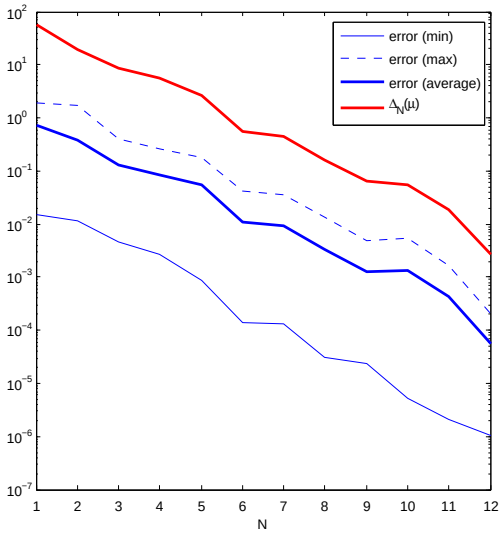



Figure 3.15: Curved bluff body case: *a posteriori* error bounds and (minimum, maximum and average over $\Xi_{\text{test}}$) computed errors between the "truth" FE solution and the RB approximation, for $N = 1, \ldots, N_{max}$. Here $\Xi_{\text{test}}$ is a uniform sample of size $n_{\text{test}} = 500$.

By means of the greedy procedure (with a fixed tolerance $\varepsilon_{\text{tol}}^{\text{RB}} = 10^{-2}$) $N_{max}^{pr} = N_{max}^{du} = 12$ basis functions have been selected for both the primal and the dual problem. We show in Fig. 3.16 the convergence of the simultaneous[6] *primal-dual* greedy algorithm, as well as the (averages computed over over a uniform sample $\Xi_{\text{test}}$ of size $n_{\text{test}} = 2500$ of the) online evaluations for the error bounds on both primal and dual solutions.

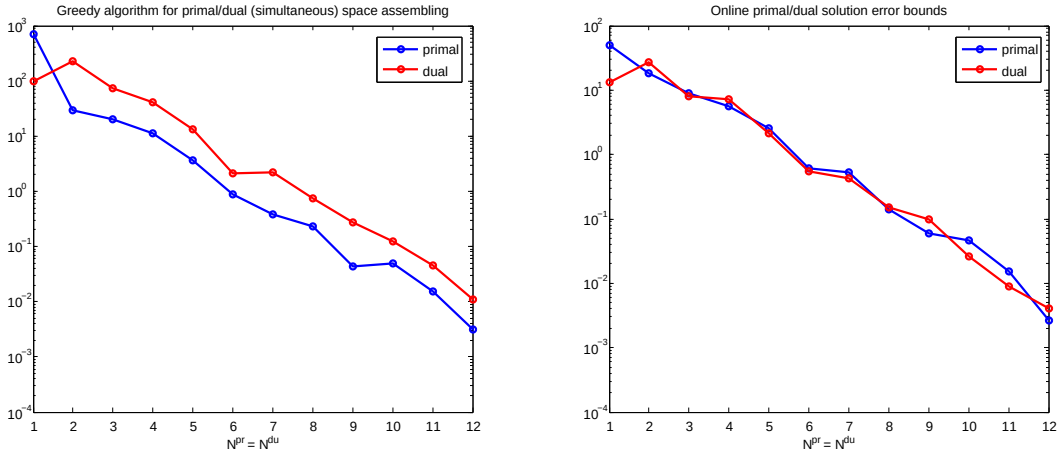



Figure 3.16: Curved bluff body case. Left: $\max_{\boldsymbol{\mu} \in \Xi_{train}} \Delta_N(\boldsymbol{\mu})$ and $\max_{\boldsymbol{\mu} \in \Xi_{train}} \Delta_M(\boldsymbol{\mu})$ as a function of $N = N^{pr}$ and $M = N^{du}$ computed during the greedy procedure, for the primal (blue) and the dual (red) problem, respectively. Here $\Xi_{train}$ is a uniform random sample of size $n_{\text{train}} = 1000$ and RB tolerance is $\varepsilon_{\text{tol}}^{\text{RB}} = 10^{-2}$. Right: online evaluations of error bounds on primal and dual variables (average over $n_{\text{test}} = 2500$ sampled $\boldsymbol{\mu}$ values).

[6] For a noncompliant linear output, the dual problem does not depend on the primal solution, so the assembling procedure for the dual space can be executed one the primal RB space has been built. Instead, when dealing with quadratic outputs, the (right-hand side of the) dual problem depends on the primal solution, so that we can exploit a simultaneous primal-dual greedy for constructing both spaces at the same time. Stopping criterium is fulfilled when both primal and dual error bounds are below a given tolerance, for any $\boldsymbol{\mu} \in \mathcal{D}$. This explains why $N_{max}^{pr} = N_{max}^{du}$ after the selection procedure. See e.g. [72] for alternative greedy strategies in this context.

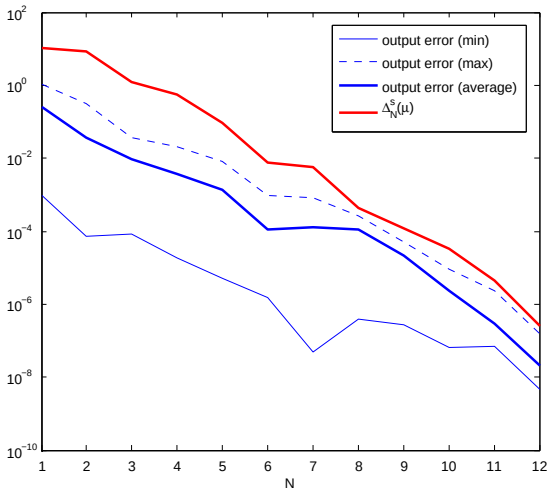



Figure 3.17: Curved bluff body case: *a posteriori* error bounds on energy output and (minimum, maximum and average over $\Xi_{\text{test}}$) computed errors. Here $\Xi_{\text{test}}$ is a uniform sample of size $n_{\text{test}} = 2500$. The average time for Online output evaluation is $0.11s$.

In the end, we plot in Fig. 3.17 the errors between the "truth" FE output $s^{\mathcal{N}}(muv)$ and the RB output $s_N(\boldsymbol{\mu})$, for $N = 1, \ldots, N_{max}$, and the corresponding error bounds. We can remark that we are able to recover the *quadratic effect* also in the case of a quadratic output.

### Case 4: A parametrized bifurcation

We now consider the case of Stokes flows through a nonaffinely parametrized bifurcation, obtained by a global FFD mapping. This geometrical configuration can be considered as a simplified version of a carotid bifurcation, which represents a well-studied case of interaction between blood flows and (possibly occluded) arterial vessels, potentially able to affect the development of arterial diseases. See Sect. 6.3 for more details and other study cases related to these features. Here we are interested in the RB approximation and certification of a simpler steady Stokes flow in the domain represented in Fig. 3.18. A $10 \times 5$ lattice of control points is placed over the bifurcation and the two control points represented in red are allowed to move in the $x_2$-direction, giving $p = 2$ geometrical parameters $\boldsymbol{\mu} = (\mu_1, \mu_2) \in (-0.5, 0.5)^2$, adjusting the width of the bifurcation. A parabolic Poiseuille profile is imposed at the inflow $\Gamma_{\text{in}}$, no-slip conditions are imposed along the wall $\Gamma_{\text{w}}$, while free-stress conditions are imposed at the two outflows $\Gamma_{\text{out}}$.

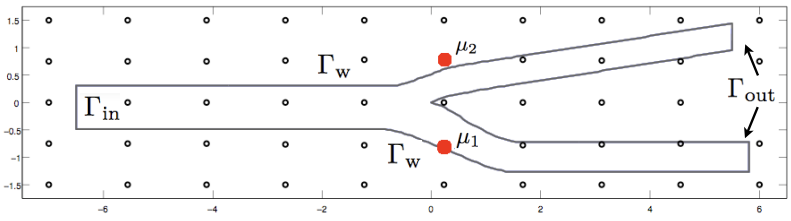



Figure 3.18: Parametrized geometry and domain boundaries for the bifurcation case.

By means of the greedy procedure (with $\varepsilon_{\text{tol}}^{\text{RB}} = 10^{-3}$) $N_{max} = 16$ basis functions have been selected. We also plot in Fig. 3.19 the SCM lower bounds $\beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu})$ for the inf-sup constant; we can remark that the pattern is nearly symmetric with respect to both variables, and does not

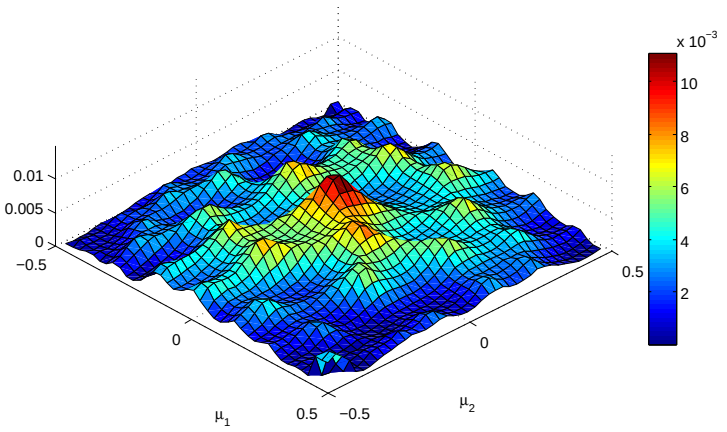



Figure 3.19: Bifurcation case: lower bounds $\beta^{LB}_{Ba,\mathcal{N}}(\boldsymbol{\mu})$ for the Babuška inf-sup constant as a function of $\mu_1, \mu_2$ after 329 iterations of the SCM greedy algorithm. Here $\Xi_{train}$ is a uniform sample of size $n_{\text{train}} = 2500$.

show strong variations over the parametric range. We always experienced these features when dealing with FFD shape parametrizations (see e.g. also case 3 of the following section).
We plot in Fig. 3.20 both the convergence of the greedy algorithm and the errors between the "truth" FE solutions and the RB approximations, with the corresponding error bounds.

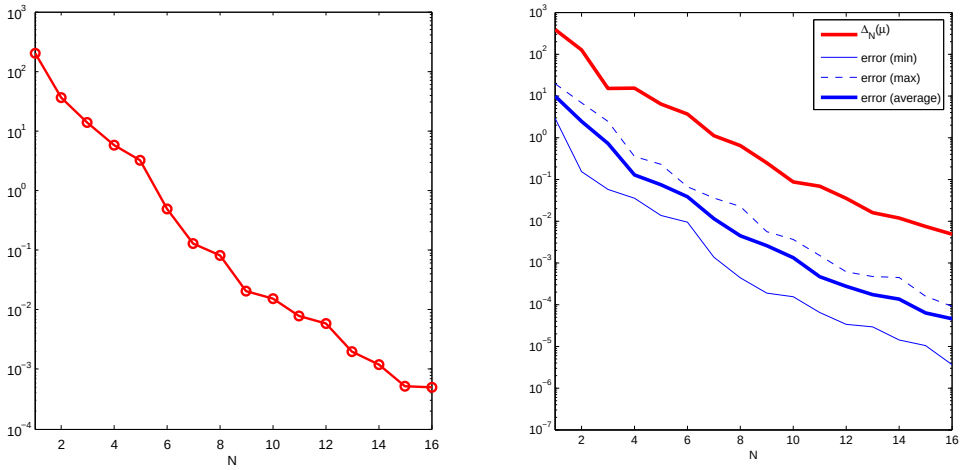



Figure 3.20: Bifurcation case. Left: $\max_{\boldsymbol{\mu}\in\Xi_{train}} \Delta_N(\boldsymbol{\mu})$ as a function of $N$ for the RB approximations computed during the greedy procedure. Here $\Xi_{train}$ is a uniform random sample of size $n_{\text{train}} = 1000$ and the RB tolerance is $\varepsilon^{\text{RB}}_{\text{tol}} = 10^{-2}$. Right: *a posteriori* error bounds and (minimum, maximum and average over $\Xi_{\text{test}}$) computed errors between the "truth" FE solution and the RB approximation, for $N = 1, \ldots, N_{max}$. Here $\Xi_{\text{test}}$ is a uniform sample of size $n_{\text{test}} = 2500$.

In Fig. 3.21 we report some representative solutions for selected values of the parameters. We can underline a strong sensitivity of the flow with respect to geometrical variations, which represents one of the most interesting features when dealing with this kind of problems.

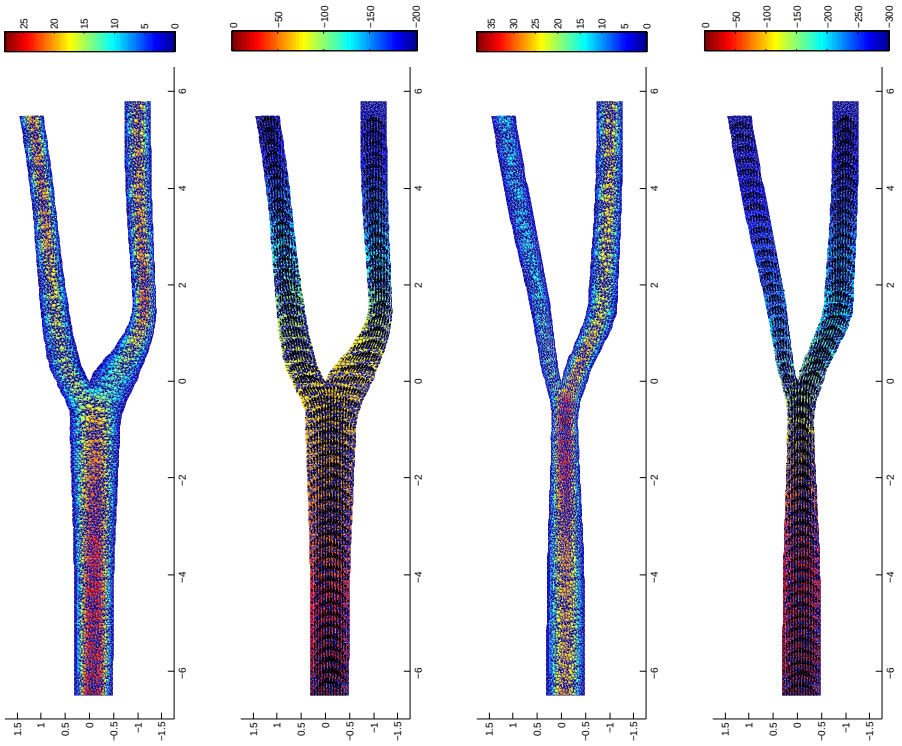


Figure 3.21: Bifurcation case: representative solutions for velocity and pressure with streamlines for $\boldsymbol{\mu} = [-0.5, 0.5]$ (left) and $\boldsymbol{\mu} = [0.5, -0.5]$ (right).

### Summary of results

As shown by the results obtained with the previous test cases, the greedy algorithm enables to build reduced spaces of contained dimension performing an exponential convergence. We report all the details of the numerical simulations related to the discussed test cases in Table 3.1. We remark the very small dimension $N$ of the RB approximation problems with respect to the FE approximation space dimension $\mathcal{N}$, which leads to effective computational economies, necessary when dealing with numerical simulations in both real time and many query context. The reduction in linear systems dimension is between 200 and 400 times, depending on the test cases, while the computational speedup is of order $10^2$, varying from 98 to 442. Computational time for Online evaluation is of order $10^{-2}$ seconds. The number of affine operator components $Q_a + 2Q_b$ is larger in the Case 4 compared to other cases because of the EIM procedure employed to recover the affine parametric dependence in this nonaffine case. By setting $\varepsilon_{tol}^{EIM} = 2.5 \times 10^{-4}$, we obtain an affine expansion of $\sum_{i,j} K_{ij}^a = 117$ terms for $\nu_{ij}$ components and of $\sum_{i,j} K_{ij}^b = 7$ terms for $\chi_{ij}$ components. This has a strong impact on the running of the SCM algorithm, which in this case has taken about 300 iterations to converge in about 20 hours, by using the *Lanczos* method to solve the eigenproblems required by SCM.

In Table 3.2 we report all the details related to the SCM procedure for the approximation of lower bounds $\beta_{Ba,\mathcal{N}}^{LB}(\boldsymbol{\mu})$ used in the test cases presented. We point out that the natural norm SCM algorithm enables to contain the computational costs arising from the computation of the lower bound of the inf-sup constant, also in the cases of larger parameter spaces $\mathcal{D}$, as for the curved bluff body case, at least in the case $p \leq 3$ – the algorithm starts converging very slowly when the number of parameters is getting larger and larger.

| Approximation data | Case 1 *Poiseuille* | Case 1 *Couette* | Case 2 *Channel* | Case 3 *Body* | Case 4 *Bifurcation* |
|---|---|---|---|---|---|
| Number of parameters $P$ | 2 | 1 | 3 | 2 | 2 |
| Affine op. components $Q_a + 2Q_b$ | 5 | 3 | 8 | 8 | 135 |
| Affine rhs components $Q_F + Q_G$ | 2 | 4 | 2 | 9 | 140 |
| FE space dim. $\mathcal{N}$ | 8354 | 5093 | 6490 | 13216 | 24046 |
| RB primal space dim. $N_{max}^{pr}$ | 7 | 6 | 11 | 12 | 16 |
| RB dual space dim. $N_{max}^{du}$ | - | - | 17 | 6 | - |
| FE evaluation $t_{FE}^{online}$ (s) | 3.987 | 2.005 | 3.464 | 10.483 | 14.563 |
| RB evaluation $t_{RB}^{online}$ (s) | 0.0101 | 0.0205 | 0.0212 | 0.0237 | 0.0289 |
| Computational speedup | 395 | 98 | 163 | 442 | 504 |

Table 3.1: Numerical details for the test cases presented. RB spaces have been built by means of the greedy procedure, using a tolerance $\varepsilon_{\text{tol}}^{\text{RB}} = 10^{-2}$ and a uniform RB greedy train sample of size $n_{\text{train}}$. A comparison of the computational times between the Online RB evaluations and the corresponding FE simulations is reported. Here $t_{RB}^{online}$ is the time of an Online RB computation, while $t_{FE}^{online}$ is the time for a FE computation, once FE matrices are built.

| Approximation data | *Poiseuille* | *Couette* | *Channel* | *Body* | *Bifurcation* |
|---|---|---|---|---|---|
| # of selected $\bar{\boldsymbol{\mu}}$ | 2 | 3 | 1 | 1 | 6 |
| # of selected $\hat{\boldsymbol{\mu}}$ ($\forall \bar{\boldsymbol{\mu}}$) | 22; 5 | 7; 5; 3 | 49 | 10 | 215; 33; 21; 38; 17; 5 |
| # of eigenproblems | 39 | 24 | 66 | 27 | 605 |

Table 3.2: Numerical details for the test cases presented. The lower and upper bounds of the Babuška inf-sup constants have been computed by means of the natural norm SCM algorithm detailed in Sect. A.3, using a tolerance $\varepsilon_{\text{tol}}^{SCM} = 0.85$ and a uniform train sample of size $n_{\text{train}}$. SCM requires the solution of $\#\bar{\boldsymbol{\mu}} + \#\hat{\boldsymbol{\mu}} + 2(Q_a + 2Q_b)$ eigenproblems.

### 3.6.2 Navier-Stokes equations

In the following we consider moderate Reynolds viscous flows described by 2D steady Navier-Stokes equations in different geometries. By increasing complexity, we deal with a classic cavity flow parametrized by the Reynolds number, with a double elbow pipeline parametrized by the Reynolds number and the aspect ratio, and with a flow around an airfoil profile, whose shape is parametrized through a FFD map. In any case, we are interested to the approximation of velocity and pressure fields, as well as to error estimation; no outputs (and related dual problem will be considered at this level. All numerical details concerning the construction of RB spaces are reported in Table 3.3, as well as some remarks concerning the computation of lower bounds $\beta_{\hat{A},\mathcal{N}}^{\text{LB}}(\boldsymbol{\mu})$ for the inf-sup stability constants.

#### Case 1: A channel expansion

In this first example we deal with a parametrized flow in the fixed geometrical configuration depicted in Fig. 3.22, representing a channel expansion (sometimes assimilated to the classical *cavity* benchmarks). We identify the regions $\mathcal{R}_\ell$, $1 \leq \ell \leq 2$, which represent the portions of the channel with different sectional area. We consider $p = 1$ parameter, defined as $\mu_1 = 1/\nu$; the parameter domain is given by $\mathcal{D} = [10, 250]$. The forcing term is $\mathbf{f} = (0, 0)$. A parabolic flow is imposed at the inlet $\Gamma_1$, while a free-stress condition is imposed at the outflow $\Gamma_5$. Thus, we

impose the following boundary conditions:

$$\begin{aligned} &\mathbf{v} = \mathbf{0} && \text{on} \quad \Gamma_2, \Gamma_3, \Gamma_4, \Gamma_6, \Gamma_7, \\ &v_1 = \alpha(x_2 - 1)(2 - x_2), \quad v_2 = 0 && \text{on} \quad \Gamma_1, \\ &-p\mathbf{n} + \frac{1}{\mu_1}\frac{\partial \mathbf{v}}{\partial \mathbf{n}} = \mathbf{0} && \text{on} \quad \Gamma_5. \end{aligned}$$

where $\alpha > 0$ is a positive constant chosen such that $|\bar{\mathbf{v}}| = \max v_1|_{\Gamma_1} = 0.5$. Equivalently, the flow is parametrized with respect to the Reynolds number, which can be defined in this case by $\mathrm{Re} = \mu_1/2$, where $L = 1$ is the characteristic length – the width of the channel at the inflow – and $|\bar{\mathbf{v}}|$ the characteristic velocity.

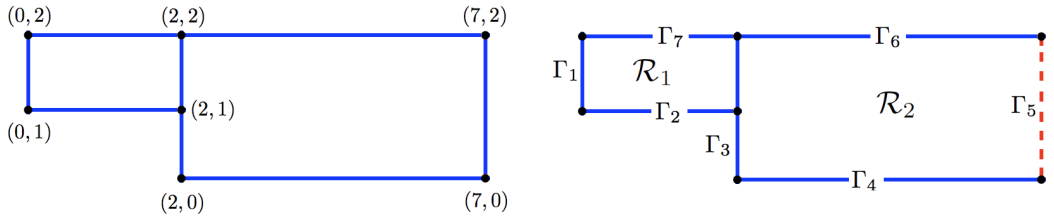



Figure 3.22: Parametrized geometry and domain boundaries for the channel expansion case.

The greedy procedure for the construction of the RB space selects $N_{max} = 11$ basis functions with a fixed tolerance $\varepsilon_{\text{tol}}^{\text{RB}} = 10^{-4}$. We report in Fig. 3.23 the convergence of the greedy procedure, as well as the (online evaluation over a train sample $\Xi_{\text{test}}$ of size $n_{\text{test}} = 100$) errors between the "truth" FE solution and the RB approximation, for $N = 1, \ldots, N_{max}$, and the corresponding error bounds. In this case we have $N^*(\boldsymbol{\mu}) = 3$ ($N^*(\boldsymbol{\mu})$ is such that $\tau_N(\boldsymbol{\mu}) < 1$ for $N \geq N^*(\boldsymbol{\mu})$); for $N < N^*(\boldsymbol{\mu})$ we can keep $\tau_N(\boldsymbol{\mu})$ as error bound (this choice is employed by the greedy algorithm, which thus selects the first snapshots as those which maximize $\tau_N(\boldsymbol{\mu})$, and also during the online evaluations). We also remark that the difference between $\Delta_N(\boldsymbol{\mu})$ and $\tau_N(\boldsymbol{\mu})$ is not so large for $N \geq N^*(\boldsymbol{\mu})$, even if a more general mathematical connection between $\Delta_N(\boldsymbol{\mu})$ and $\tau_N(\boldsymbol{\mu})$ is less obvious – e.g., there could be coincidences related to embedding constants and *inf-sup* constants.

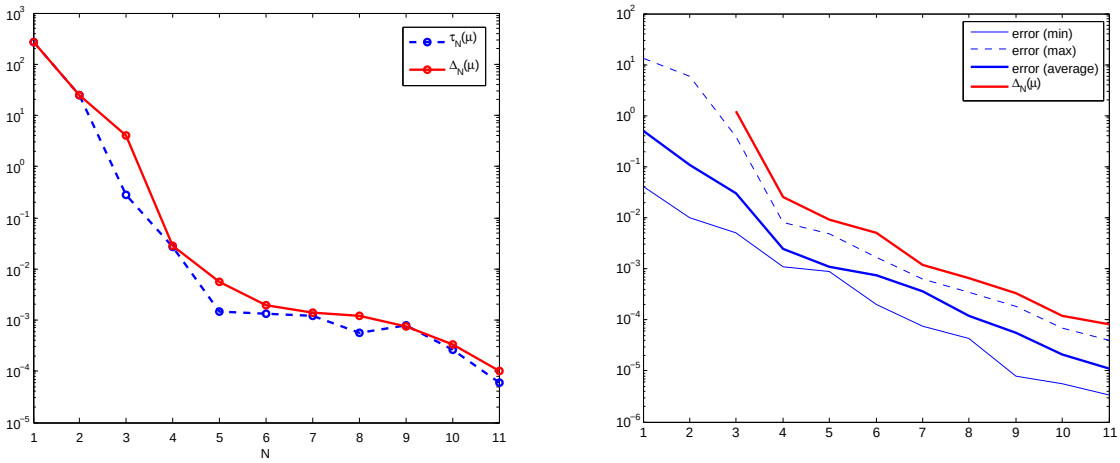



Figure 3.23: Channel expansion case. Left: $\max_{\boldsymbol{\mu} \in \Xi_{train}} \Delta_N(\boldsymbol{\mu})$ (in red) and $\max_{\boldsymbol{\mu} \in \Xi_{train}} \tau_N(\boldsymbol{\mu})$ (in blue) as a function of $N$, computed during the greedy procedure. Here $\Xi_{train}$ is a uniform random sample of size $n_{\text{train}} = 500$ and RB tolerance is $\varepsilon_{\text{tol}}^{\text{RB}} = 10^{-4}$. Right: online evaluations of error bounds (average over $n_{\text{test}} = 100$ sampled $\boldsymbol{\mu}$ values) and (minimum, average and maximum over $\Xi_{\text{test}}$) error between FE and RB approximations.

We plot in Fig. 3.24 the lower bounds for the Babuška inf-sup constant $\beta^{LB}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$ computed by the SCM algorithm (in red), as well as the values (blue dots) of the true stability factor $\beta(\boldsymbol{\mu}^{j*})$, for the points in $\mathcal{S} = \{\boldsymbol{\mu}^{j*}, j = 1, \ldots, J\}$ selected through the SCM-greedy procedure. As in the linear case (see e.g. [146], Section 3.2), the method interpolates $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$ at each $\boldsymbol{\mu}^* \in \mathcal{S}$, i.e. $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}^*) = \beta^{LB}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}^*)$ for each $\boldsymbol{\mu}^* \in \mathcal{S}$. Moreover, we show (in blue) the *surrogate* lower bounds computed[7] through a RBF interpolation of the values $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$, for $\boldsymbol{\mu}^* \in \mathcal{S}$, by using thin plate splines, as well as (in black) a set of $n_{\text{test}} = 10$ computed *true* values $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}_i)$, with $\boldsymbol{\mu}_i \notin \mathcal{S}$. Although the surrogate fails to bound $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}_i)$ in at least one case ($\boldsymbol{\mu}_i \in (200, 250)$), it can be globally considered a good surrogate of the SCM lower bound $\beta^{LB}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$, avoiding the lower peaks usually shown by SCM outputs and due to the imposition of the successive constraints. Nevertheless, the RBF surrogate (as well as a constant lower bound) might fail in particular in the case of separated flows also at moderate Reynolds number, because of possible undergoing bifurcation phenomena.

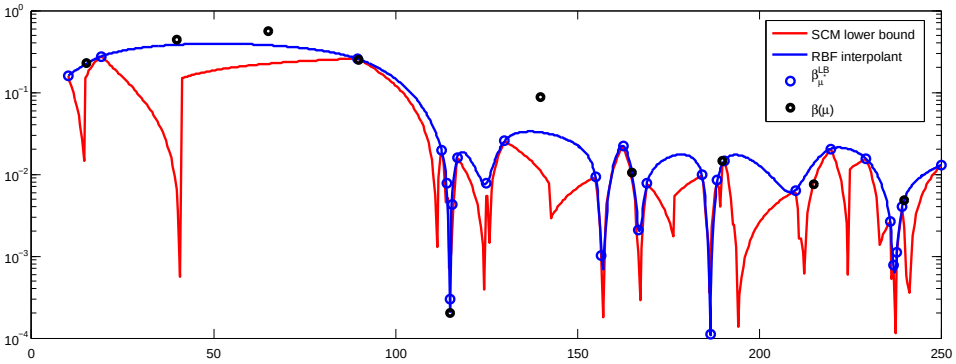



Figure 3.24: Channel expansion case. Lower bounds $\beta^{LB}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$ for the Babuška inf-sup constant (in red) and RBF surrogate lower bound (in blue) as a function of $\boldsymbol{\mu}$; here $\Xi_{train}$ is a uniform sample of size $n_{\text{train}} = 400$; Values of the stability factor $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}^*)$ for each selected $\boldsymbol{\mu}^* \in \mathcal{S}$ (blue dots) and stability factors $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}_i)$ for 10 parameter values $\{\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{10}\}$ (in black).

In Fig. 3.25 we report some representative solutions. We can underline that the size of the vortex core located in the bottom left corner increases as $\mu_1$ increases, indicating the presence of a larger and larger recirculation with increasing Reynolds numbers, as expected.

### Case 2: A double elbow pipeline

We now consider a parametrized flow in the geometrical configuration depicted in Fig. 3.26, representing a double elbow pipeline (sometimes assimilated to a *double cavity* benchmark). We identify the regions $\mathcal{R}_\ell$, $1 \le \ell \le 3$, which represent the three portions of the channel with constant sizes ($\ell = 1, 3$) and variable sizes ($\ell = 2$). We consider $p = 2$ parameter, defined as $\mu_1 = 1/\nu$ and the aspect ratio $\mu_2$ of the vertical portion of the pipeline; the parameter domain is given by $\mathcal{D} = [50, 350] \times [2, 5]$. The forcing term is $\mathbf{f} = (0, 0)$. A parabolic flow is imposed at the inlet $\Gamma_1$, while a free-stress condition is imposed at the outflow $\Gamma_6$:

$$
\begin{aligned}
&\mathbf{v} = \mathbf{0} && \text{on} \quad \partial\Omega_o(\mu_2) \setminus (\Gamma_1 \cup \Gamma_6),\\
&v_1 = \alpha x_2(1 - x_2), \quad v_2 = 0 && \text{on} \quad \Gamma_1,\\
&-p\mathbf{n} + \frac{1}{\mu_1}\frac{\partial \mathbf{v}}{\partial \mathbf{n}} = \mathbf{0} && \text{on} \quad \Gamma_6.
\end{aligned}
$$

[7] See Sect. A.3.2 in Appendix A for some remarks about the construction of surrogate lower bounds.

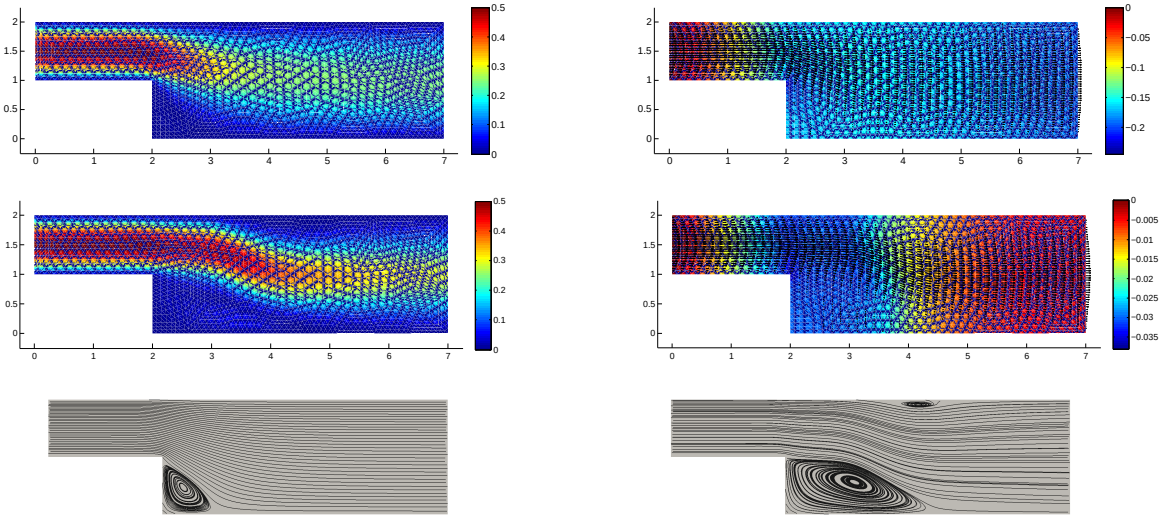


Figure 3.25: Channel expansion case: representative solutions for velocity and pressure with streamlines for $\boldsymbol{\mu} = 50$ (left) and $\boldsymbol{\mu} = 250$ (right).

where $\alpha > 0$ is a positive constant chosen such that $|\bar{\mathbf{v}}| = \max v_1|_{\Gamma_1} = 0.5$. Equivalently, the flow is parametrized with respect to the Reynolds number, given by $\mathrm{Re} = \mu_1/2$, where $L = 1$ is the characteristic length – the width of the channel – and $|\bar{\mathbf{v}}|$ the characteristic velocity.
The greedy procedure for the construction of the RB space selects $N_{max} = 29$ basis functions with a fixed tolerance $\varepsilon_{\mathrm{tol}}^{\mathrm{RB}} = 10^{-2}$. We report in Fig. 3.27 the convergence of the greedy procedure, as well as the (online evaluation over a train sample $\Xi_{\mathrm{test}}$ of size $n_{\mathrm{test}} = 200$) errors between the FE solution and the RB approximation, for $N = 1, \ldots, N_{max}$, and the corresponding error bounds. In this case we have $N^*(\boldsymbol{\mu}) = 2$; for $N < N^*(\boldsymbol{\mu})$ we can keep $\tau_N(\boldsymbol{\mu})$ as error bound (this choice is employed by the greedy algorithm, which thus selects the first snapshots as those which maximize $\tau_N(\boldsymbol{\mu})$).

We plot in Fig. 3.28 the surrogate lower bounds for the Babuška inf-sup constant $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$ computed through a RBF interpolation (by using thin plate splines) over a set of $J = 50$ values $\beta(\boldsymbol{\mu}^{j*})$ corresponding to a set of *a priori* selected points $\mathcal{S} = \{\boldsymbol{\mu}^{j*}, j = 1, \ldots, J\}$, represented as blue dots. In this case we did not rely on the SCM algorithm for the construction of the lower bounds since the (current implementation of the) procedure showed a very poor convergence, leading to computational times sensibly larger than required by the construction of the RB space itself.

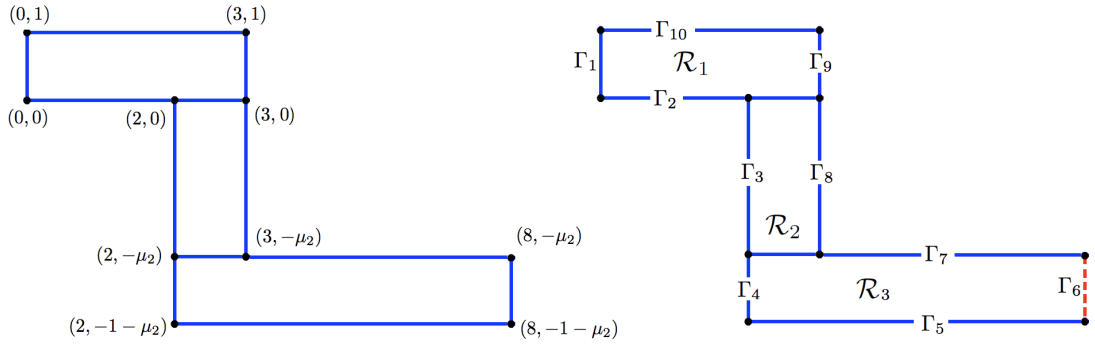



Figure 3.26: Parametrized geometry and domain boundaries for the double elbow pipeline case.

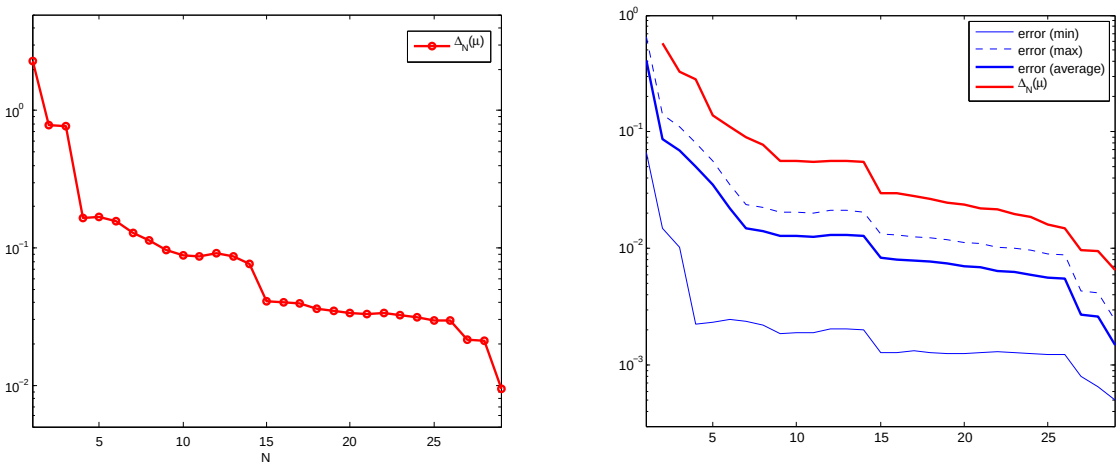



Figure 3.27: Double elbow pipeline case. Left: $\max_{\boldsymbol{\mu}\in\Xi_{train}} \Delta_N(\boldsymbol{\mu})$ as a function of $N$, computed during the greedy procedure. Here $\Xi_{train}$ is a uniform random sample of size $n_{\text{train}} = 500$ and RB tolerance is $\varepsilon_{\text{tol}}^{\text{RB}} = 10^{-2}$. Right: online evaluations of error bounds (average over $n_{\text{test}} = 200$ sampled $\boldsymbol{\mu}$ values) and (minimum, average and maximum over $\Xi_{\text{test}}$) error between FE and RB approximations.

We point out that in this case the method interpolates $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$ at each $\boldsymbol{\mu}^* \in \mathcal{S}$, but it may lack of rigor for values of $\boldsymbol{\mu}$ not close to $\mathcal{S}$. See Sect. A.3.2 in Appendix A for further details.
In Fig. 3.29 we report some representative solutions for selected values of the parameters. The pattern of the vortex cores becomes more and more complex when $\mu_1$ (and Reynolds number as well) increases – despite we are dealing with steady flows – showing the usual recirculation cells close to the corners of the channel. In case of shorter configurations at larger Reynolds numbers (e.g. the first case in Fig. 3.29) the recirculation fills up the whole channel.

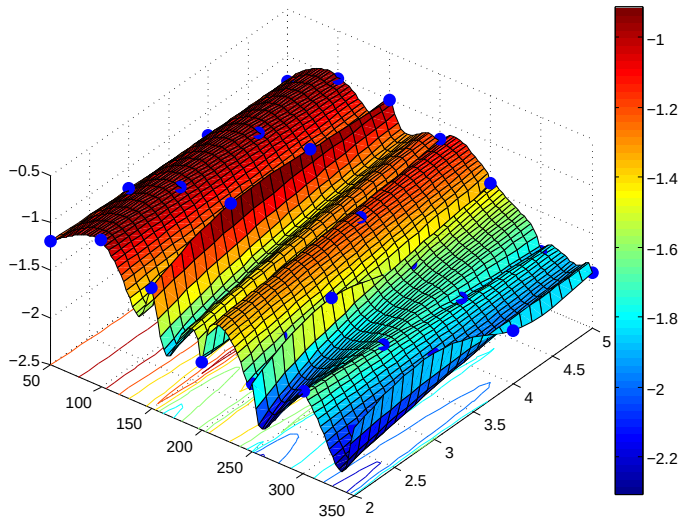


Figure 3.28: Double elbow pipeline case. Surrogate RBF lower bounds (*log* scale) and a set of $j = 1, \ldots, J = 50$ computed values $\beta(\boldsymbol{\mu}^{j*})$ (blue dots) for the Babuška inf-sup constant $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$.

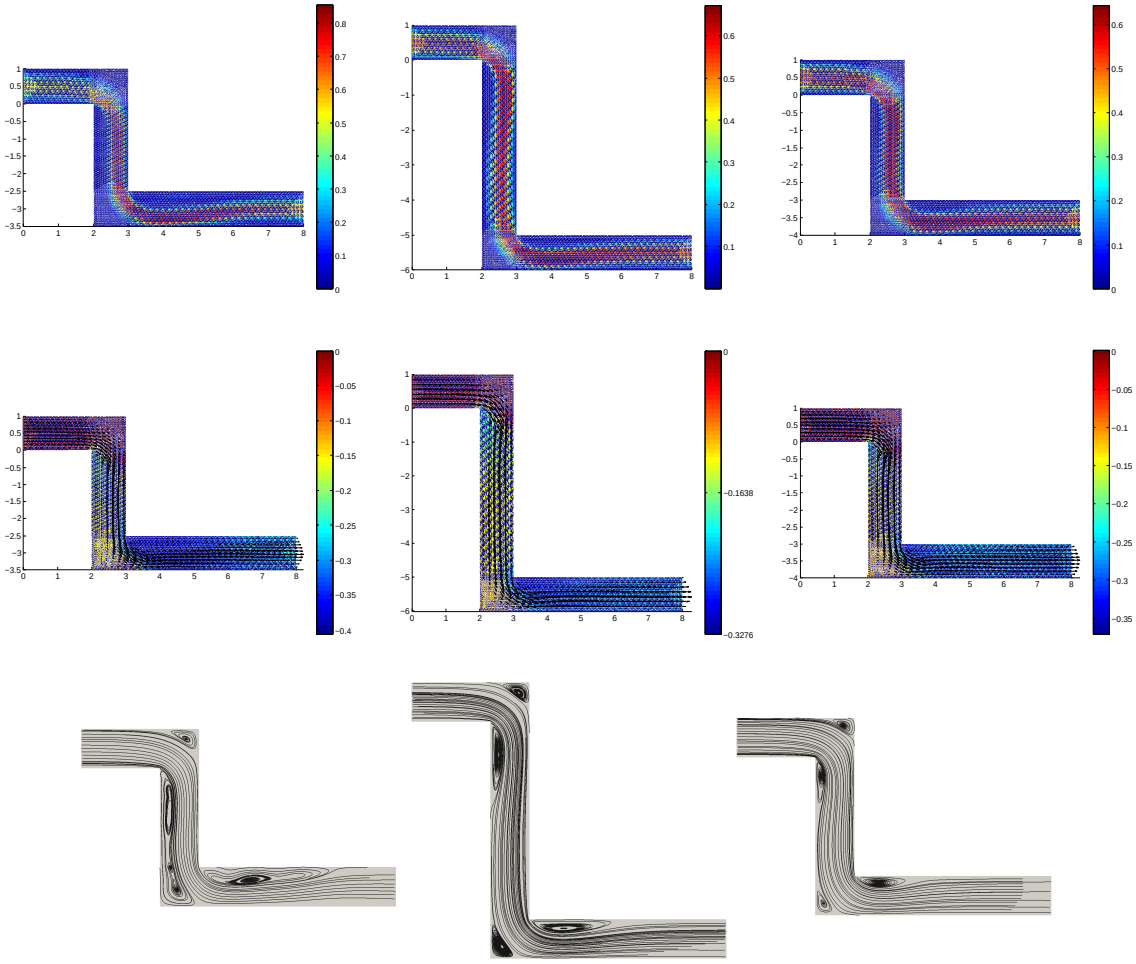


Figure 3.29: Double elbow pipeline case: representative solutions for velocity and pressure for $\boldsymbol{\mu} = (500, 2.5)$ (left), $\boldsymbol{\mu} = (400, 5)$ (center) and $\boldsymbol{\mu} = (250, 3)$ (right).

### Case 3: A NACA airfoil

In this last example we consider the approximation of a steady Navier-Stokes flow around a family of NACA0012 airfoil profiles (see Fig. 3.30). We parametrize the geometry around the airfoil using a FFD map: a $6 \times 6$ lattice of control points is placed around the airfoil and the closest four control points (represented in red) are allowed to move in the $x_2$-direction. We thus consider $p = 4$ geometrical parameters, given by the vertical displacements of the four control points; the parameter domain is given by $\mathcal{D} = [-0.4, 0.4]^4$.

The forcing term is $\mathbf{f} = (0, 0)$, whereas $\nu = 1/150$. A parabolic horizontal flow $\mathbf{v}_{\text{in}} = (\alpha(x_2 - 2.5)(2.5 - x_2), 0)^T$ is imposed at the inlet $\Gamma_{\text{in}}$, where $\alpha > 0$ is a positive constant chosen such that $|\bar{\mathbf{v}}| = \max v_1|_{\Gamma_1} = 1$, while a free-stress condition is imposed at the outflow $\Gamma_{\text{out}}$.
The greedy procedure for the construction of the RB space selects $N_{max} = 17$ basis functions with a fixed tolerance $\varepsilon_{\text{tol}}^{\text{RB}} = 2.5 \times 10^{-3}$. We report in Fig. 3.31 the convergence of the greedy procedure, as well as the (online evaluation over a train sample $\Xi_{\text{test}}$ of size $n_{\text{test}} = 500$) errors between the "truth" FE solution and the RB approximation, for $N = 1, \ldots, N_{max}$, and the corresponding error bounds. In this case $N^*(\boldsymbol{\mu}) = 1$, so that the error bound $\Delta_N(\boldsymbol{\mu})$ can be employed from the very beginning both for greedy selection procedure and the online error evaluation.

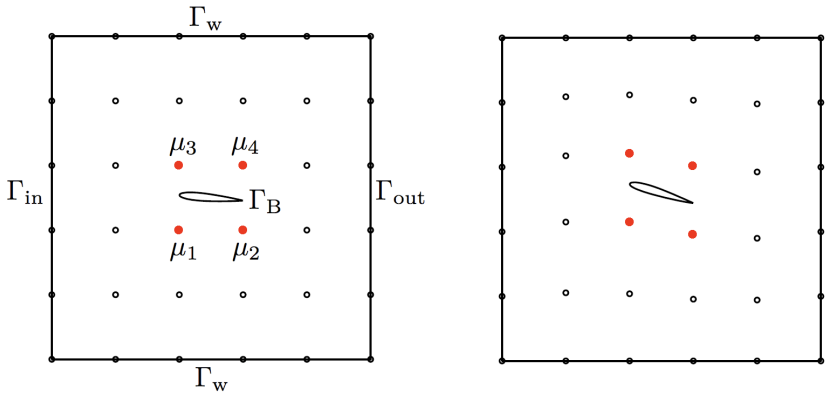



Figure 3.30: Parametrized geometry and domain boundaries for the NACA airfoil case.

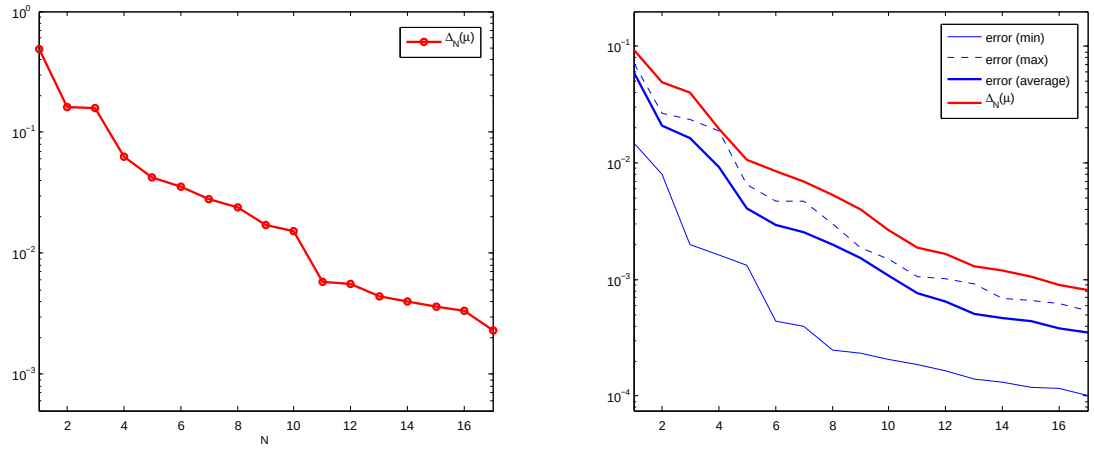



Figure 3.31: NACA Airfoil case. Left: $\max_{\boldsymbol{\mu}\in\Xi_{train}} \Delta_N(\boldsymbol{\mu})$ as a function of $N$, computed during the greedy procedure. Here $\Xi_{train}$ is a uniform random sample of size $n_{\text{train}} = 1000$ and RB tolerance is $\varepsilon_{\text{tol}}^{\text{RB}} = 2.5 \times 10^{-3}$. Right: online evaluations of error bounds (average over $n_{\text{test}} = 500$ sampled $\boldsymbol{\mu}$ values) and (minimum, average and maximum over $\Xi_{\text{test}}$) error between FE and RB approximations.

We plot in Fig. 3.32 (bottom) the surrogate lower bounds for the Babuška inf-sup constant $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$ computed through a RBF interpolation (by using thin plate splines) over a set of $J = 100$ values $\beta(\boldsymbol{\mu}^{j*})$ corresponding to a set of *a priori* selected points $\mathcal{S} = \{\boldsymbol{\mu}^{j*}, j = 1, \ldots, J\}$ (top). As already remarked for the Case 4 of Sect. 3.6.1 dealing with a Stokes flow in a nonaffinely (FFD) parametrized bifurcation, the range of variation of the stability factor over the parametric domain is rather contained, so that also a constant lower bound can provide good results. See Sect. A.3.2 in Appendix A for further details. In the end, we report in Fig. 3.33 some representative solutions for selected values of the parameters.

### Summary of results

As in the case of Stokes flows, also for Navier-Stokes flows the greedy algorithm enables to build reduced spaces of contained dimension performing an exponential convergence, as shown by the results of previous test cases.

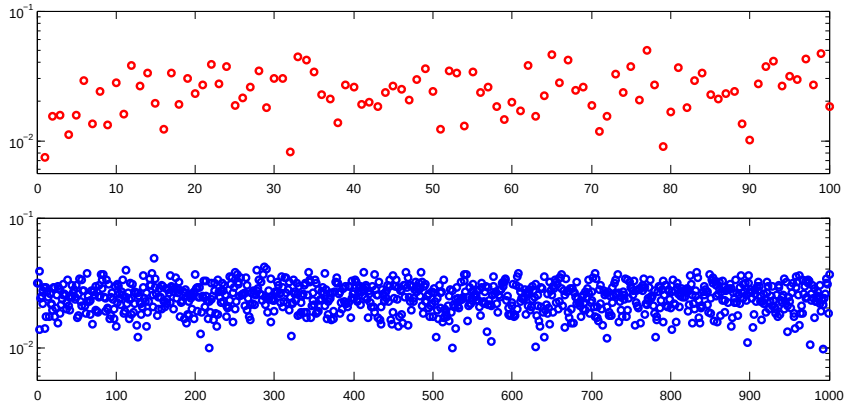


Figure 3.32: NACA Airfoil case. Top: set of $J = 100$ computed values $\beta(\boldsymbol{\mu}^{j*})$, $j = 1, \dots, J$ for the Babuška inf-sup constant $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$. Bottom: surrogate RBF lower bounds computed on a train sample $\Xi_{\text{train}}$ of size $n_{\text{train}} = 1000$.

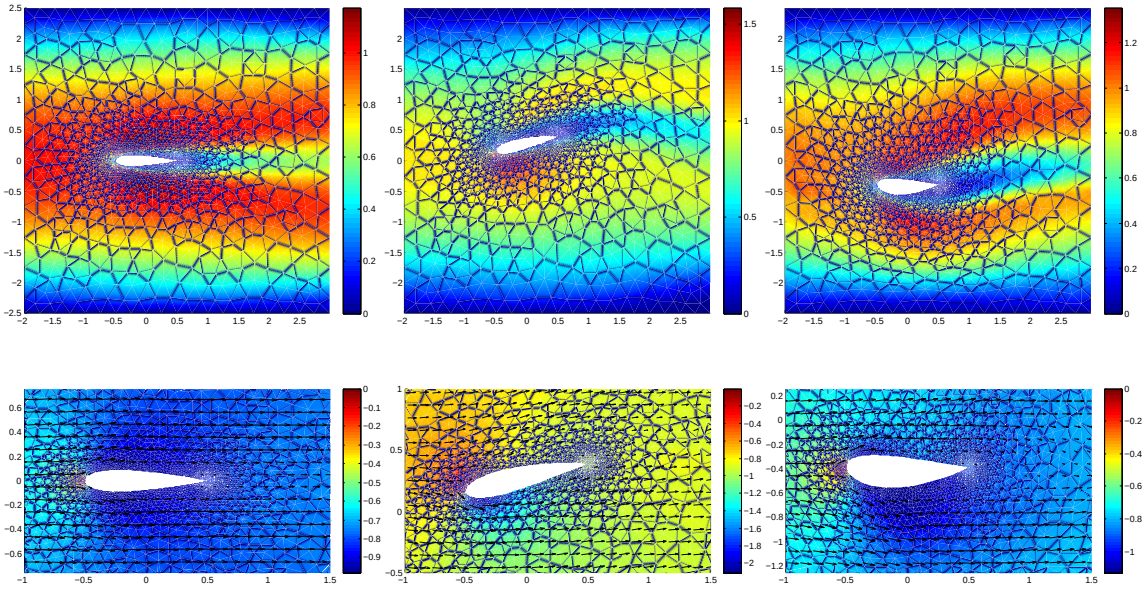


Figure 3.33: NACA airfoil case: representative solutions for velocity and pressure for different values of $\boldsymbol{\mu} \in \mathcal{D}$.

Some weak points, mainly concerned with the estimation of parametric lower bounds to the stability factor $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$, are shared with the Stokes cases; moreover, other critical aspects arise within the approximation of Navier-Stokes equations, such as the efficient computation of dual norms of residuals and the solution of nonlinear equations – both during basis computation and online RB approximation. We report all the details of the numerical simulations related to the discussed test cases in the following Table 3.3. The number of affine operator components $Q_a + 2Q_b + Q_c$ is larger in the Case 3 compared to other cases because of the EIM procedure employed to recover the affine parametric dependence in this nonaffine case.

We remark the very small dimension $N$ of the RB approximation problems with respect to the FE approximation space dimension $\mathcal{N}$, although we are currently dealing with RB tolerances of

order $10^{-2} \div 10^{-4}$. The reduction in linear systems dimension ranges between 357 and 699 times, depending on the test cases.

Computational speedup is of order $10^2$, varying from 377 to 564. Computational time for each RB Online evaluation is of order 1 second and is almost constant, also when dealing with problems parametrized w.r.t. Reynolds number. On the other hand, FE solutions require increasing CPU times (and number of iterations) in these cases (see e.g. Fig. 3.34, concerning the test Case 1); both for FE and RB solutions of nonlinear problems we rely on a fixed-point iteration with a tolerance $\varepsilon_{\text{tol}}^{NS} = 10^{-5}$.

| Approximation data | Case 1 *Channel* | Case 2 *Pipeline* | Case 3 *Airfoil* |
|---|---|---|---|
| Number of parameters $P$ | 1 | 2 | 4 |
| Affine op. components $Q_a + 2Q_b + Q_c$ | 9 | 18 | 77 |
| Affine rhs components $Q_F + Q_G$ | 9 | 13 | 67 |
| FE space dim. $\mathcal{N}$ | 23077 | 31093 | 32538 |
| RB space dim. $N_{max}$ | 11 | 29 | 17 |
| RB greedy tolerance $\varepsilon_{\text{tol}}^{\text{RB}}$ | $10^{-4}$ | $10^{-2}$ | $2.5 \times 10^{-3}$ |
| RB greedy train sample size $n_{\text{train}}$ | 500 | 500 | 1000 |
| FE evaluation $t_{FE}^{online}$ (s) | 137.47 | 2070.25 | 460.28 |
| RB evaluation $t_{RB}^{online}$ (s) | 1.3707 | 3.667 | 1.2201 |
| Computational speedup | 100 | 564 | 377 |

Table 3.3: Numerical details for the test cases presented. RB spaces have been built by means of the greedy procedure, using a tolerance $\varepsilon_{\text{tol}}^{\text{RB}}$ and a uniform RB greedy train sample of size $n_{\text{train}}$. A comparison of the computational times between the Online RB evaluations and the corresponding FE simulations is reported. Here $t_{RB}^{online}$ is the time of an Online RB computation, while $t_{FE}^{online}$ is the time for a FE computation, once FE matrices (linear operators) are built.

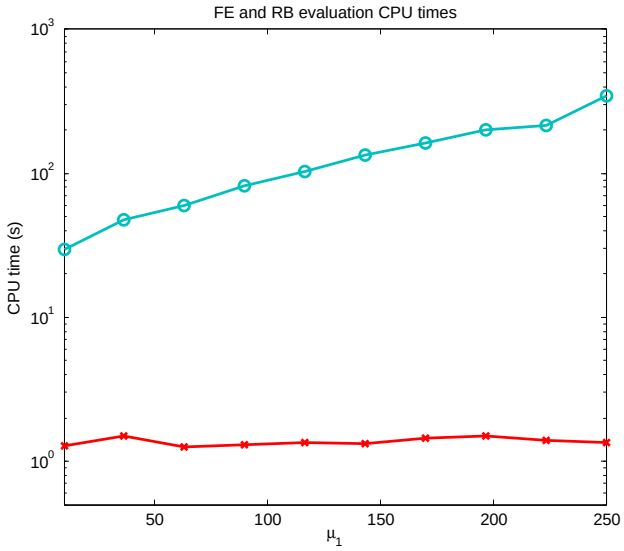



Figure 3.34: Channel expansion case. Comparison between FE and RB evaluation CPU times. Here the problem is parametrized w.r.t. $\mu_1 = 1/\nu \in [10, 250]$, proportional to Reynolds number.

# 4 A reduced framework for optimization and inverse problems

In this chapter we analyze the reduced framework developed for the efficient solution of optimal control, shape optimization and other inverse problems governed by parametrized PDEs, which is based on the integration of RB methods for computational reduction and suitable shape parametrizations for geometrical complexity reduction.
We first provide some general results for optimal control and shape optimization problems, which are applied to some problems dealing with fluid flows. Then, we present a reduced order framework for parametrized optimal control and shape optimization problems, built over the reduced basis approximation of the state flow problems. In particular, we recast FFD and RBF parametrization techniques in the general class of perturbation of the identity map (which assure suitable regularity conditions on admissible shapes). We extend this framework to other inverse problems related with both variable shapes and uncertainty, for which we characterize a deterministic and a Bayesian framework, recasting them in a many-query context. Further, we propose a parameter coupling algorithm for steady fluid-structure interaction problems. Finally, we describe a numerical procedure for solving optimization problems governed by parametrized PDEs. Throughout the chapter, we omit the proofs of well-known results from literature.

## 4.1 Optimal control problems

In the following we recall a general well-posedness result for some optimal control (OC) problems, which will then be specified for the case of velocity boundary control problems governed by Navier-Stokes equations, relevant to the applications presented in Part III.

### 4.1.1 Abstract formulation

As seen in Sect. 1.3.1, an optimal control problem can be expressed as a constrained optimization problem in the general form (1.10), here rewritten for the reader's convenience:

$$\hat{u} = \arg\min_{u \in \mathcal{U}_{\text{ad}}} \mathcal{J}(u) \quad \text{s.t.} \quad \tilde{A}(y(u), w) = F(w; u), \quad \forall w \in X; \tag{4.1}$$

here $y = y(u) \in X$ is the state variable, $X$ the state space, $\tilde{A} : X \times X \to \mathbb{R}$ the state operator. Moreover, $F : X \to \mathbb{R}$ is a linear operator depending on the control function $u \in \mathcal{U}_{\text{ad}}$ and $\mathcal{U}_{\text{ad}} \subseteq \mathcal{U}$ is the space of admissible controls; $X$ and $\mathcal{U}$ are reflexive Banach spaces. The constraint is the PDE modelling the state system, whereas $\mathcal{J} : \mathcal{U}_{\text{ad}} \to \mathbb{R}$ is the cost functional. The well-posedness of the abstract OC problem (4.1) is assured by the following result:

**Theorem 4.1.** *Let $\mathcal{U}$ be a reflexive Banach space and $\mathcal{U}_{ad} \subseteq \mathcal{U}$ a nonempty, closed, bounded and convex set. Then, if $\mathcal{J} : \mathcal{U}_{ad} \to \mathbb{R}$ is a weakly lower semicontinuous functional, such that $\min_{u \in \mathcal{U}_{ad}} \mathcal{J}(u) > -\infty$, problem (4.1) admits a solution.*

**Remark 4.2.** *This classical result in* calculus of variations *states that a bounded lower semicontinuous functional attains its minimum in a weakly sequentially compact set (see e.g. [311], Chapter 2, Thm. 2.14 for the proof). We remark that:* (i) *every bounded subset of a reflexive Banach space is weakly sequentially compact, and* (ii) *every convex and closed subset of a Banach space $\mathcal{U}$ is weakly sequentially closed, so that if the space $\mathcal{U}$ is reflexive and the set is in addition bounded, then it is weakly sequentially compact. The first result is the main reason why the concept of weak convergence is fundamental: the notion of weak sequential compactness takes over the role of compactness used in the finite-dimensional case to assure the existence of a minimizer for a lower semicontinuous function (Weierstrass theorem).*

**Remark 4.3.** *In order to fulfill the hypotheses of Theorem 4.1, let us recall that a continuous and convex functional $J : \mathcal{U} \to \mathbb{R}$ on a Banach space $\mathcal{U}$ is weakly lower semicontinuous, i.e.*

$$\liminf_{n \to \infty} \mathcal{J}(u_n) \geq \mathcal{J}(u)$$

*for any sequence $\{u_n\}_{n=1}^{\infty} \subset \mathcal{U}$ such that $u_n \rightharpoonup u$ as $n \to \infty$. We recall that we can express*

$$\mathcal{J}(u) = \frac{1}{2}\mathcal{J}_1(u) + \frac{\varepsilon}{2}\mathcal{J}_2(u) = \frac{1}{2}\|Su - z_d\|_Z + \frac{\varepsilon}{2}(Nu, u)_{\mathcal{U}},$$

*where $S : \mathcal{U} \to Z$ is the operator $S = E_X G$, where $G : \mathcal{U} \to X$ is the control-to-state operator and $E_X : X \to Z$ the embedding operator of $X$ into $Z$. If $\varepsilon > 0$, the second summand is strictly convex, where $N \in \mathcal{L}(\mathcal{U},\mathcal{U})$ is a symmetric and coercive operator, whereas if $\varepsilon = 0$ the strict convexity is implied by the injectivity of $S : \mathcal{U} \to Z$.*

In order to find the solution to the OC problem (4.1) we can use the *Lagrangian functional* framework (see e.g. [128], Chapter 6), for which $(\hat{u}, \hat{y}(\hat{u}), \hat{p}(\hat{u})) \in \mathcal{U}_{ad} \times X \times X$, where $\hat{u} \in \mathcal{U}_{\text{ad}}$ is the solution to (4.1), is a stationary point of the Lagrangian functional

$$\mathcal{L}(y, z, u) := \tilde{\mathcal{J}}(y, u) + F(z; u) - \tilde{A}(y, z),$$

where $z \in X$ is the Lagrangian multiplier and $\mathcal{J}(u) = \tilde{\mathcal{J}}(y(u), u)$. Then, $(\hat{u}, \hat{y}(\hat{u}), \hat{p}(\hat{u}))$ fulfills the following (first order) optimality conditions:

$$\begin{aligned}
&\partial_p \mathcal{L}(y, z, u)[\phi] = 0 && \forall\ \phi \in X && \textit{state equation},\\
&\partial_y \mathcal{L}(y, z, u)[\psi] = 0 && \forall\ \psi \in X && \textit{adjoint equation},\\
&\partial_u \mathcal{L}(y, z, u)[\theta] \geq 0 && \forall\ \theta \in \mathcal{U}_{\text{ad}} && \textit{optimality condition}.
\end{aligned} \tag{4.2}$$

Hence, by differentiating $\mathcal{L}$ w.r.t. $z$, we find again the state equation in weak form:

$$\text{find}\ \ y \in X\ :\ \tilde{A}(y, \phi) = F(\phi; u) \qquad \forall \phi \in X.$$

The second equation of (4.2) is the *adjoint equation* (for which $z \in X$ is also defined as the adjoint variable) and is obtained by differentiating $\mathcal{L}$ w.r.t. $y$; its weak form reads as follows:

$$\text{find}\ \ p \in X\ :\ \tilde{A}_y(y)(z, \psi) = \partial_y \tilde{\mathcal{J}}(y, u)(\psi) \qquad \forall \psi \in X,$$

where $\tilde{A}_y(y)(z, \psi)$ is the Fréchet derivative of $\tilde{A}(y, \psi)$ with respect to the first argument $y \in X$, evaluated at $z \in X$. If $\tilde{A}(\cdot, \cdot)$ is bilinear, then $\tilde{A}_y(y)(z, \psi) = \tilde{A}(z, \psi)$; if $\tilde{A}(\cdot, \cdot)$ is a more general semilinear form (linear in the second argument), the adjoint equation is still linear in $z \in X$. In the end, by differentiating $\mathcal{L}$ w.r.t. $u$, we obtain the third equation of system (4.2):

$$\partial_u \tilde{\mathcal{J}}(y, u)(\theta) + F_u(u)(z; \theta) \geq 0 \qquad \forall \theta \in \mathcal{U}_{\text{ad}},$$

which states the optimality condition and provides the expression of the gradient of the cost functional, needed e.g. for gradient-based numerical optimization procedures.

### 4.1.2 Velocity boundary control of Navier-Stokes flows

We now turn to the analysis of a velocity boundary control problem for the Navier-Stokes equations, which is of particular interest in view of an application to optimal design of bypass grafts (see Sect. 5.2.4). We consider the Lagrangian multiplier method [16] for the treatment of the Dirichlet boundary conditions, which includes the control function itself. Let us consider a state system similar to (1.17), given by:

$$\begin{cases} -\nu\Delta\mathbf{v} + (\mathbf{v}\cdot\nabla)\mathbf{v} + \nabla p = \mathbf{0} & \text{in } \Omega \\ \nabla\cdot\mathbf{v} = 0 & \text{in } \Omega \\ \mathbf{v} = \mathbf{u}_{\mathrm{c}} & \text{on } \Gamma_{\mathrm{c}} \\ \mathbf{v} = \mathbf{v}_{\mathrm{in}} & \text{on } \Gamma_{\mathrm{in}} \\ \mathbf{v} = \mathbf{0} & \text{on } \Gamma_{\mathrm{w}} \\ -p\mathbf{n} + \nu\dfrac{\partial\mathbf{v}}{\partial\mathbf{n}} = \mathbf{0} & \text{on } \Gamma_{\mathrm{out}}, \end{cases} \tag{4.3}$$

where $\mathbf{u}_{\mathrm{c}} \in \mathcal{U}_{\mathrm{c}} := (H^{1/2}(\Gamma_{\mathrm{c}}))^2$ is the boundary control function and $\mathbf{u}_{\mathrm{in}} \in \mathcal{U}_{\mathrm{in}}(H^{1/2}(\Gamma_{\mathrm{in}}))^2$ a second inflow condition. The domain $\Omega \subset \mathbb{R}^d$ for $d = 2, 3$ is assumed to be piecewise $C^2$ with convex corners, while the Dirichlet portion of $\partial\Omega$ is $\Gamma_D = \Gamma_{\mathrm{c}} \cup \Gamma_{\mathrm{in}} \cup \Gamma_{\mathrm{w}}$. Moreover, let us define

$$\mathbf{u}_{\mathrm{D}} \in (H^{1/2}(\Gamma_{\mathrm{D}}))^d \ : \ \mathbf{u}_{\mathrm{D}} = \begin{cases} \mathbf{u}_{\mathrm{c}} & \text{on } \Gamma_{\mathrm{c}} \\ \mathbf{v}_{\mathrm{in}} & \text{on } \Gamma_{\mathrm{in}} \\ \mathbf{0} & \text{on } \Gamma_{\mathrm{w}}. \end{cases}$$

Let us consider a cost functional $\tilde{\mathcal{J}} : V \times \mathcal{U}_{\mathrm{c}} \times \mathcal{U}_{\mathrm{in}} \to \mathbb{R}_+$ which represents the objective of the optimization. It depends both on velocity and inflow conditions, and will be chosen according to the specific applications (see Sect. 5.2.4). Let us denote by $Y = (\mathbf{v}, p, \boldsymbol{\xi}) \in X = V \times Q \times \mathcal{G}$ the state variable, given by the velocity, the pressure and the Lagrange multiplier accounting for the Dirichlet conditions, with $V = (H^1(\Omega))^d$, $Q = L^2(\Omega)$ and $\mathcal{G} = (H^{-1/2}(\Gamma_D))^d$, respectively. If the function $\mathbf{u}_{\mathrm{in}} \in \mathcal{U}_{\mathrm{in}}$ is known, the first problem of interest – denoted as *deterministic design (OC) problem* – can be formulated as follows: given $\mathbf{v}_{\mathrm{in}}$, find the boundary control function $\hat{\mathbf{u}}_{\mathrm{c}}$ solving

$$\begin{aligned} &\hat{\mathbf{u}}_c = \arg\min_{\mathbf{u}_c \in \mathcal{U}_{\mathrm{c,\,ad}}} \mathcal{J}(\mathbf{u}_c) \quad \text{s.t.} \quad Y = (\mathbf{v}, p, \boldsymbol{\xi}) \in X \ : \\ &\tilde{A}(Y, W) = F(W; \mathbf{u}_{\mathrm{D}}), \ \forall W = (\mathbf{z}, q, \boldsymbol{\lambda}) \in X \end{aligned} \tag{DD-OC}$$

where $\mathcal{U}_{\mathrm{c,\,ad}} \subseteq \mathcal{U}_{\mathrm{c}}$, $\mathcal{J}(\mathbf{u}_c) = \tilde{\mathcal{J}}(\mathbf{v}, \mathbf{u}_{\mathrm{c}}, \mathbf{v}_{\mathrm{in}})$ and the operators appearing in the state problem is

$$\begin{aligned} &\tilde{A}(Y, W) = a(\mathbf{v}, \mathbf{z}) + b(p, \mathbf{z}) + b(q, \mathbf{y}) + c(\mathbf{v}, \mathbf{v}, \mathbf{z}) + g_D(\boldsymbol{\xi}, \mathbf{z}) + g_D(\mathbf{v}, \boldsymbol{\lambda}), \\ &F(w; \mathbf{u}_{\mathrm{D}}) = -g_D(\mathbf{u}_{\mathrm{D}}, \boldsymbol{\lambda}), \end{aligned} \tag{4.4}$$

where (using the same notation of Sect. 3.1)

$$a(\mathbf{v}, \mathbf{z}) := \nu \int_\Omega \nabla\mathbf{u} : \nabla\mathbf{z} \, d\Omega, \quad b(p, \mathbf{z}) := \int_\Omega p\nabla\cdot\mathbf{z} \, d\Omega, \quad c(\mathbf{v}, \mathbf{w}, \mathbf{z}) := \int_\Omega (\mathbf{v}\cdot\nabla)\mathbf{w}\cdot\mathbf{z} \, d\Omega,$$

and the Dirichlet conditions on $\Gamma_D$ with boundary data $\mathbf{u}_D$ are enforced through the bilinear form

$$g_D(\mathbf{v}, \mathbf{w}) := \int_{\Gamma_D} \mathbf{v}\cdot\mathbf{w} \, d\Gamma.$$

We recall that the state problem (4.4) is well-posed under the conditions stated in Theorem 3.4. Moreover, the velocity field satisfies the following stability estimate:

$$\|\mathbf{v}\|_V \leq \frac{2}{\nu}\|F(\cdot;\mathbf{u}_{\mathrm{D}})\|_{(H^{-1}(\Omega))^d},$$

where $\|F(\cdot;\mathbf{u}_{\mathrm{D}})\|_{(H^{-1}(\Omega))^d}$ is defined like in (3.39) and depends on the Dirichlet boundary conditions $\mathbf{u}_{\mathrm{D}}$. In particular, the solution $\mathbf{v}$ is unique and depends continuously on the data $\mathbf{u}_{\mathrm{D}}$.

Let us consider $\tilde{\mathcal{J}} : V \times \mathcal{U}_{\mathrm{c}} \times \mathcal{U}_{\mathrm{in}} \to \mathbb{R}_+$ as a functional of the velocity field $\mathbf{v} \in V$, and denote $J(\mathbf{v}) = \mathcal{J}(\mathbf{v}, \mathbf{u}_{\mathrm{c}}, \mathbf{v}_{\mathrm{in}})$. A general existence result for the first optimality problem (DD-OC) can be found in [130] (see Lemma 2.1 and the related proof):

**Theorem 4.4.** *Let us assume that the cost functional $J(\mathbf{v})$:*

1. *is bounded, i.e. there exists $C_0 > 0$ s.t. $J(\mathbf{v}) \leq C_0\|\mathbf{v}\|_V^2$;*
2. *is convex, i.e. for any $\mathbf{u}_1, \mathbf{u}_2 \in V$ and $\gamma \in [0,1]$ it holds that $(1-\gamma)J(\mathbf{u}) + \gamma J(\mathbf{u}) \geq J((1-\gamma)\mathbf{u} + \gamma\mathbf{u})$;*
3. *satisfies for some constants $C_1, C_2, C_3 > 0$ the weak coercivity inequality*

$$J(\mathbf{v}) \geq C_1\|\mathbf{v}\|_V^2 - C_2\|\mathbf{v}\|_V - C_3 \quad \text{for all } \mathbf{v} \in \mathcal{V}. \tag{4.5}$$

*Let the admissible set $\mathcal{U}_{c,\,ad}$ for the control function be a closed and convex subset of $\mathcal{U}_c = (H^{1/2}(\Gamma_c))^d$. Then the problem* (DD-OC) *admits at least one solution.*

The proof of Theorem 4.4 is a direct application of the abstract result stated in Theorem 4.1. A second optimal flow control problem of remarkable interest for the sake of applications is a worst-case OC problem, where we are interested to seek for the optimal control $\hat{\mathbf{u}}_c$ that minimizes the worst-case value of $\mathcal{J}(\mathbf{u}_c, \mathbf{v}_{\mathrm{in}}) = \tilde{\mathcal{J}}(\mathbf{v}, \mathbf{u}_{\mathrm{c}}, \mathbf{v}_{\mathrm{in}})$ over all admissible values of the function $\mathbf{v}_{\mathrm{in}}$. Hence, we are interested in studying the so-called *robust design problem*: find the boundary control function $\hat{\mathbf{u}}_{\mathrm{c}}$ solving the worst-case optimization problem

$$\begin{aligned} &\hat{\mathbf{u}}_{\mathrm{c}} = \arg \min_{\mathbf{u}_{\mathrm{c}} \in \mathcal{U}_{\mathrm{c,\,ad}}} \max_{\mathbf{v}_{\mathrm{in}} \in \mathcal{U}_{\mathrm{in,\,ad}}} \mathcal{J}(\mathbf{u}_c, \mathbf{v}_{\mathrm{in}}) \quad \text{s.t.} \quad Y = (\mathbf{v}, p, \boldsymbol{\xi}) \in X \ : \\ &\tilde{A}(Y, W) = F(W; \mathbf{u}_{\mathrm{c}}, \mathbf{v}_{\mathrm{in}}), \ \forall W = (\mathbf{z}, q, \boldsymbol{\lambda}) \in X \end{aligned} \tag{RD-OC}$$

The robust design problem (RD-OC) can be understood as a one-shot game, where the optimizer plays first and chooses the control function $\mathbf{u}_{\mathrm{c}}$ to minimize the cost functional $\mathcal{J}$. The second player then follows by choosing the function $\mathbf{v}_{\mathrm{in}}$ to maximize the cost function $\mathcal{J}$. The payoff for the designer is $-\mathcal{J}$ and for the second player $\mathcal{J}$. Thus the optimal strategy for the optimizer is given as the solution of a min-max type of strategy obtained by solving (RD-OC), while the second player will choose his response by solving another problem. We call this the *complementary uncertainty problem*, and it is defined as: given a known boundary control function $\mathbf{u}_{\mathrm{c}}$, find the function $\hat{\mathbf{v}}_{\mathrm{in}}$ solving

$$\begin{aligned} &\hat{\mathbf{v}}_{\mathrm{in}} = \arg \max_{\mathbf{v}_{\mathrm{in}} \in \mathcal{U}_{\mathrm{in,\,ad}}} \mathcal{J}(\mathbf{u}_c, \mathbf{v}_{\mathrm{in}}) \quad \text{s.t.} \quad Y = (\mathbf{v}, p, \boldsymbol{\xi}) \in X \ : \\ &\tilde{A}(Y, W) = F(W; \mathbf{v}_{\mathrm{in}}), \ \forall W = (\mathbf{z}, q, \boldsymbol{\lambda}) \in X \end{aligned} \tag{CU}$$

The well-posedness of problem (CU) is ensured by the following result:

**Theorem 4.5.** *Let us assume that $\Gamma_{in}$ is an open and connected, non-empty subset of $\partial\Omega$, and that the cost functional $J(\mathbf{v})$:*

1. *is bounded (see (i), Theorem. 4.4);*
2. *is upper semicontinuous, i.e. $\limsup_{\mathbf{v}\to\mathbf{v}^*} J(\mathbf{v}) \leq J(\mathbf{v}^*)$ for all $\mathbf{v}^* \in V$.*

*Let the admissible set $\mathcal{U}_{in,\,ad} \subseteq \mathcal{U}_{C_4}$ be a closed subset of*

$$\mathcal{U}_{C_4} := \left\{ \mathbf{u} \in (H^2(\Gamma_{in}))^d \ : \ \|\mathbf{u}\|_{(H^2(\Gamma_{in}))^d} \leq C_4 \right\}, \tag{4.6}$$

*for some $C_4 > 0$ small enough such that* (3.37) *is satisfied, and furthermore that the viscosity is large enough to satisfy* (3.39)*. Then the problem* (CU) *admits at least one optimal solution.*

*Proof.* Since $\Gamma_{\text{in}}$ is bounded, the embedding $H^2(\Gamma_{\text{in}}) \hookrightarrow H^1(\Gamma_{\text{in}})$ is compact by Rellich's theorem and $\mathcal{U}_{\text{in}}$ is compact in $(H^1(\Gamma_{\text{in}}))^d$. According to Theorem 3.4, the solution map $\mathbf{v}_{\text{in}} \mapsto \mathbf{v}(\mathbf{v}_{\text{in}})$ is continuous in the $H^1$-topology under our assumptions. Thus the image of $\mathcal{U}_{C_4}$ through the resolvent operator (4.4), restricted to velocity field, is a compact set in $V$. By Theorem 4.1, a bounded upper semicontinuous functional attains its maximum in a compact set. □

**Remark 4.6.** *For coercive cost functionals satisfying (4.5) the maximizer of* (CU) *will be found on the boundary of the set of admissible functions $\mathcal{U}_{in,\,ad}$. Thus we expect to find maximizers that become increasingly singular as we increase $C_4$ in (4.6).*

Existence of solutions for the worst-case problem (RD-OC) in the infinite-dimensional case has not been extensively studied. In a recent paper [140], the concept of weak lower semi-continuity for set-valued functions is used to prove existence results for OC problems governed by PDEs for functionals of the min-max type. In the case that the admissible set $\mathcal{U}_{\text{in, ad}}$ does not depend on the control variable $\mathbf{u}_{\text{c}}$, and therefore a sufficient condition for the weak lower semi-continuity of

$$\widehat{\mathcal{J}}(\mathbf{u}_{\text{c}}) := \sup_{\mathbf{v}_{\text{in}} \in \mathcal{U}_{\text{in, ad}}} \tilde{\mathcal{J}}(\mathbf{v}, \mathbf{u}_{\text{c}}, \mathbf{v}_{\text{in}})$$

is that $\tilde{\mathcal{J}}(\,\cdot\,,\,\cdot\,,\mathbf{v}_{\text{in}})$ is weakly lower semi-continuous for all admissible $\mathbf{v}_{\text{in}} \in \mathcal{U}_{\text{in, ad}}$ (Theorem 2.5 of [140]). This assumption of independence does not strictly hold for problems discussed in chapter 6 (see eq. (5.3)), so further study on well-posedness of the min-max formulation would be needed.

## 4.2 Shape optimization problems

We discuss in this section the main results assuring the well-posedness of the shape optimization (SO) problems relevant to the applications presented in Part III. These problems are more difficult than the optimal control case, because of the shape dependence and the need of a suitable definition of shape variations. After treating the general case of an abstract SO problem, we characterize the case of problems governed by Navier-Stokes equations.

### 4.2.1 Abstract formulation

Recalling Sect. 1.3.2, a shape optimization problem can be expressed as a constrained optimization problem in the general form (1.11), here rewritten for the reader's convenience:

$$\hat{\Omega}_o = \arg\min_{\Omega_o \in \mathcal{O}_{\text{ad}}} \mathcal{J}(\Omega_o) \quad \text{s.t.} \quad \tilde{A}(y(\Omega_o), w; \Omega_o) = F(w; \Omega_o), \quad \forall w \in X(\Omega_o); \tag{4.7}$$

$y(\Omega_o) \in X(\Omega_o)$ is the state variable, $X(\Omega_o)$ a reflexive Banach space defined on $\Omega_o$, $\mathcal{O}_{\text{ad}} \subseteq \mathcal{O}$ the set of admissible shapes, where $\mathcal{O}$ is a generic class of shapes (to be specified later). Here the state equation representing the constraint is defined over the domain $\Omega_o$ – which for the sake of convenience is identified with the original domain of the parametrized framework – whereas $\mathcal{J} : \mathcal{O}_{\text{ad}} \to \mathbb{R}$ is the cost functional, depending on the state variable as well, i.e. $\mathcal{J}(\Omega_o) = \tilde{\mathcal{J}}(y(\Omega_o), \Omega_o)$, where $\tilde{\mathcal{J}} : X \times \mathcal{O}_{\text{ad}} \to \mathbb{R}$. Let us recall that we denote by $\Gamma_{\text{w}}$ the *free-boundary* which can be displaced in order to optimize the objective.

Provided that the state problem admits a unique solution in any domain $\Omega$, we can introduce a mapping $y$ that with any $\Omega \in \mathcal{O}$ associates the state solution $y(\Omega) \in X(\Omega)$, i.e. $y : \Omega \mapsto y(\Omega) \in X(\Omega)$. Moreover, let $\{\Omega_n\}_{n=1}^{\infty} \subset \mathcal{O}$ be a sequence of shapes and $\{y_n\}_{n=1}^{\infty}$ a sequence of state solutions, being $y_n \equiv y(\Omega_n) \in X(\Omega_n)$. Let us denote $\Omega_n \xrightarrow{\tau} \hat{\Omega}$ and $y_n \rightsquigarrow y$ two suitable notions of convergence[1], of the sequence of shapes and of state solutions, respectively; remark that in the latter case convergence involves different functional spaces, defined on the sequence $\{\Omega_n\}_{n=1}^{\infty}$. Then, the following abstract existence result holds (see for example [137], Theorem 2.10):

**Theorem 4.7.** *Let $\mathcal{E} = \{(\Omega, y(\Omega)), \forall\, \Omega \in \mathcal{O}_{ad}\}$ be the graph of the mapping $y(\cdot)$ restricted to $\mathcal{O}_{ad}$. Assume that*

*i) $\mathcal{E}$ is compact, i.e. for any sequence $\{(\Omega_n, y(\Omega_n)) \in \mathcal{E}\}_{n=1}^{\infty}$, there exists a subsequence $\{(\Omega_{n_k}, y(\Omega_{n_k})) \in \mathcal{E}\}_{k=1}^{\infty}$ and $(\hat{\Omega}, y(\hat{\Omega})) \in \mathcal{E}$, such that $\Omega_{n_k} \xrightarrow{\tau} \hat{\Omega}$, $y(\Omega_{n_k}) \rightsquigarrow y(\hat{\Omega})$ for $k \to \infty$;*

*ii) the cost functional $\mathcal{J}(\Omega)$ is lower semicontinuous, i.e. if $\Omega_n \xrightarrow{\tau} \hat{\Omega}$ and $y_n \rightsquigarrow \hat{y}$, then $\liminf_{n\to\infty} \tilde{\mathcal{J}}(y_n, \Omega_n) \geq \tilde{\mathcal{J}}(\hat{y}, \hat{\Omega})$.*

*Then, if $\tilde{\mathcal{J}}$ is bounded from below, problem* (4.7) *has at least one solution.*

**Remark 4.8.** *The usual way to verity the first hypothesis stands on the two following assumptions:*

1. *compactness of $\mathcal{O}_{ad}$ in $\mathcal{O}$: for any sequence $\{\Omega_n\}_{n\geq 1} \subset \mathcal{O}_{ad}$ there exists a subsequence $\{\Omega_{n_k}\}_{f\geq 1} \subset \mathcal{O}_{ad}$ such that $\Omega_{n_k} \xrightarrow{\tau} \Omega$ when $k \to \infty$;*
2. *continuity of $y(\Omega)$ with respect to the domain: for any sequence $\{\Omega_n\}_{n\geq 1} \subset \mathcal{O}_{ad}$, $\Omega \subset \mathcal{O}_{ad}$, the following implication holds: $\Omega_n \xrightarrow{\tau} \Omega \Rightarrow y(\Omega_n) \rightsquigarrow y(\Omega)$.*

With respect to OC problems, verifying the well-posedness of SO problems involves additional assumptions of regularity on admissible shapes and continuity of state solution with respect to shape deformations. Thanks to Remark 4.8, we can *decouple* the dependence of the shape on the state solution: the first assumption is in fact just related to the set of admissible shapes. Usually, further regularity assumptions on the set of admissible shapes or admissible deformations ensure the compactness of $\mathcal{O}_{\text{ad}}$. For instance, Lipschitz domains (or, equivalently, domains satisfying the so-called *uniform cone condition* [139]) yield to compact sets of admissible shapes; additional constraints, for instance on the volume of the admissible domains might also be imposed. Another class of domains yielding this property are the ones obtained through a *perturbation of the identity* map; this concept is useful not only to frame the analysis of problem (4.7), but also to express optimality conditions as well.

Let us denote

$$\mathcal{O} \equiv \mathcal{O}(\Omega) := \{\Omega_o = T(\Omega) \text{ for some } T \in \mathcal{T}\}, \tag{4.8}$$

[1] A possible choice for the topology of $\mathcal{O}$ is the set distance topology of Hausdorff or the pseudo-distance (4.9).

the space of shapes obtained by deforming $\Omega$ through the mapping $T$, where $\mathcal{T} = \{T : \mathbb{R}^d \to \mathbb{R}^d,\ (T-I) \in W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d), (T^{-1}-I) \in W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)\}$ is a space of diffeomorphisms[2] in $\mathbb{R}^d$; in other words, $T \in \mathcal{T}$ is a differentiable mapping whose inverse is also differentiable. It can be shown that such a map is also continuous, so that the topology of the shapes obtained by deformation through $T \in \mathcal{T}$ is the same as the one of $\Omega$.

As a set $\mathcal{O}_{\rm ad}$ of admissible shapes for the SO problem (4.7), we can thus define

$$\mathcal{O}_{\rm ad} = \{\Omega_o \in \mathcal{O}(\Omega) \,:\, \partial\Omega_o \setminus \Gamma_{\rm w} \text{ is fixed}, |\Omega_o| \le V\}.$$

Nevertheless, to ensure a compactness property, we need to enforce a *uniform regularity* condition – exploited e.g. in the forthcoming analysis of the FFD mappings – by introducing the following (pseudo) distance over the set of shapes $\mathcal{O}$:

$$d(\Omega^1,\Omega^2) = \inf_{T\in\mathcal{T}\,:\,T(\Omega^1)=\Omega^2} \left(\|T-I\|_{W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)} + \|T^{-1}-I\|_{W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)}\right). \tag{4.9}$$

Thus, it can be shown (see e.g. [221]) that the set of admissible shapes defined by

$$\mathcal{O}_{\rm ad} = \{\Omega_o \in \mathcal{O}(\Omega) \,:\, d(\Omega,\Omega_o) \le R \,:\, \partial\Omega_o \setminus \Gamma_{\rm w} \text{ is fixed}, |\Omega_o| \le V\}$$

verifies a suitable compactness[3] property. This ensures the existence of (at least) one optimal shape, provided the other conditions required by Theorem 4.7 are verified.

Thanks to previous results, we thus can focus on perturbation of the identity maps, given by

$$T(\cdot;\boldsymbol{\theta}) = I + \boldsymbol{\theta}, \quad \text{with} \quad \boldsymbol{\theta} \in W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d);$$

$\boldsymbol{\theta} : \mathbb{R}^d \to \mathbb{R}^d$ can be regarded as a vectorial field which deforms the reference domain $\Omega$ to obtain the original domain $\Omega_o$. In particular, *(i)* $T(\Omega;\boldsymbol{\theta}) \in \mathcal{O}(\Omega)$ if the displacement field $\boldsymbol{\theta}$ is sufficiently *small* and *(ii)* the set of shapes obtained through $T(\Omega;\boldsymbol{\theta})$ is compact, as stated by the following Lemma (see e.g. [7], Lemma 6.13 and [221], Theorem 2.4):

**Lemma 4.9.** *Let $\boldsymbol{\theta} \in W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)$ be a vectorial field such that $\|\boldsymbol{\theta}\|_{W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)} \le 1$. Then $T(\cdot;\boldsymbol{\theta}) = I + \boldsymbol{\theta} \in \mathcal{T}$, i.e. it is a perturbation of the identity. Moreover:*

1. *if $\|\boldsymbol{\theta}\|_{W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)} \le 1-\alpha$ for some $\alpha \in (0,1)$, then $\|T^{-1}(\cdot;\boldsymbol{\theta}) - I\|_{W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)} \le 1/\alpha$;*
2. *the following family of shapes is compact (with respect to the pseudo-distance (4.9)):*

$$\mathcal{O}_T(\Omega) := \{\Omega_o = T(\Omega;\boldsymbol{\theta}) \,:\, \|\boldsymbol{\theta}\|_{W^{1,\infty}} \le 1-\alpha, \quad |\Omega_o| \le V\}. \tag{4.10}$$

Not only, we can also define the notion of *shape derivative*, unavoidable to formulate the optimality conditions, by using the derivative operation with respect to $\boldsymbol{\theta}$. We just point out that, as in the OC case, also for a shape optimization problem the optimality conditions can be expressed by means of the *Lagrangian functional* framework (see e.g. [7], Chapter 6). Following the same notation of Sect. 4.1.1, let us define the Lagrangian functional associated to (4.7) as

$$\mathcal{L}(y,z,\Omega) := \tilde{\mathcal{J}}(y,\Omega) + F(z;\Omega) - \tilde{A}(y,z;\Omega),$$

[2]Recall that $W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)$ is the space of Lipschitz functions $\boldsymbol{\varphi} : \mathbb{R}^d \to \mathbb{R}^d$ such that $\boldsymbol{\varphi}$ and $\nabla\boldsymbol{\varphi}$ are bounded in $\mathbb{R}^d$, which can be equipped with the norm $\|\boldsymbol{\varphi}\|_{W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)} = \sup_{\mathbf{x}\in\mathbb{R}^d}(\|\boldsymbol{\varphi}\|_2 + |||\nabla\boldsymbol{\varphi}|||_2)$ where $\|\cdot\|_2$ (resp. $|||\cdot|||_2$) is the Euclidean norm over $\mathbb{R}^2$ (resp. the induced Euclidean matrix norm over $\mathbb{R}^{d\times d}$).

[3]This is another important topology which can be used in conjunction with mapping techniques. Whenever domain perturbations are described by a family of bijective mappings having some regularities (for example, $T \in \mathcal{T}$), convergence of domains can be defined using the pseudo-distance (4.9) among the mappings $T$ s.t. $T(\Omega) \in \mathcal{T}$.

where $z \in X(\Omega)$ is the Lagrangian multiplier. The optimum $(\hat{\Omega}, \hat{y}(\hat{\Omega}), \hat{z}(\hat{\Omega})) \in \mathcal{O}_{ad} \times X(\hat{\Omega}) \times X(\hat{\Omega})$ is a stationary point of the Lagrangian functional and fulfills the following optimality conditions[4]:

$$\begin{array}{lll} \partial_p \mathcal{L}(\Omega, y(\Omega), z(\Omega))[\phi] = 0 & \forall\ \phi \in V & \textit{state equation}, \\ \partial_y \mathcal{L}(\Omega, y(\Omega), z(\Omega))[\psi] = 0 & \forall\ \psi \in V & \textit{adjoint equation}, \\ \partial_\Omega \mathcal{L}(\Omega, y(\Omega), z(\Omega))[\boldsymbol{\theta}] \geq 0 & \forall\ \boldsymbol{\theta} \in W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d), & \textit{optimality condition}. \end{array} \tag{4.11}$$

As in the OC case, by differentiating $\mathcal{L}(y, z, \Omega)$ with respect to $z$ we find the state equation, while by differentiating $\mathcal{L}(y, z, \Omega)$ with respect to $y$ we find the adjoint equation:

$$\text{find}\ \ p(\Omega) \in V(\Omega)\ \ :\ \ a'(y(\Omega))(z(\Omega), \psi) = \partial_y \tilde{\mathcal{J}}(\Omega, y(\Omega))(\psi) \qquad \forall \psi \in V(\Omega). \tag{4.12}$$

Finally, by differentiating $\mathcal{L}(y, z, \Omega)$ with respect to the shape we obtain

$$\partial_\Omega \mathcal{L}(\Omega, y(\Omega), z(\Omega))[\boldsymbol{\theta}] = d\mathcal{J}(\Omega; \boldsymbol{\theta}), \tag{4.13}$$

where $d\mathcal{J}(\Omega; \boldsymbol{\theta}) = \lim_{t \to 0} (\mathcal{J}(\Omega_o^t) - \mathcal{J}(\Omega))/t$ is the derivative of $\mathcal{J}$ in $\Omega$ and direction $\boldsymbol{\theta}$, and $\Omega_o^t = T(\Omega; t\boldsymbol{\theta})$; in the same way, we can define the derivative of the Lagrangian. Last equation of system (4.11) thus provides the expression of the gradient of the cost functional, needed e.g. for gradient-based numerical optimization procedures.

### 4.2.2 Shape optimization for Stokes/Navier-Stokes flows

We now turn to the analysis of a shape optimization problem for the Navier-Stokes equations, in view of the applications presented in Sect. 5.4. Let us consider the state system (5.2), where the condition $\mathbf{v} = \mathbf{v}_{\rm c}$ on the inflow boundary $\Gamma_{\rm c}$ is now set and does not represent a control function anymore. Furthermore, let us consider a cost functional $\tilde{\mathcal{J}} : V \times \mathcal{O}_{\rm ad} \times \mathcal{U}_{\rm in} \to \mathbb{R}_+$ which represents the objective of the optimization, depending on velocity, on the shape of the domain $\Omega_o \subset \mathbb{R}^d$ for $d = 2, 3$ (which is assumed to be piecewise $C^2$ with possibly convex corners) and on the boundary data $\mathbf{v}_{\rm in} \in \mathcal{U}_{\rm in}$ for reasons that will be clarified later.

Let us denote by $Y = (\mathbf{v}, p, \boldsymbol{\xi}) \in X = V \times Q \times \mathcal{G}$ the state variable (see Sect. 4.1.2 for the functional setting). If the function $\mathbf{v}_{\rm in} \in \mathcal{U}_{\rm in}$ is known, the first shape optimization problem we are interested in is the following *deterministic design (SO) problem*: given $\mathbf{v}_{\rm in}$, find the optimal shape $\hat{\Omega}_o$ solving

$$\begin{array}{l} \hat{\Omega}_o = \arg \min_{\Omega_o \in \mathcal{O}_{\rm ad}} \mathcal{J}(\Omega_o) \quad \text{s.t.} \quad Y(\Omega_o) = (\mathbf{v}(\Omega_o), p(\Omega_o), \boldsymbol{\xi}(\Omega_o)) \in X(\Omega_o)\ : \\ \tilde{A}(Y(\Omega_o), W; \Omega_o) = F(W; \Omega_o),\ \forall W = (\mathbf{z}, q, \boldsymbol{\lambda}) \in X(\Omega_o) \end{array} \tag{DD-SO}$$

where the operators appearing in the state problem are defined by (4.4) – here we have made explicit the dependence of the differential operators on the shape of the domain $\Omega_o$.

Analysis of this SO problem exploits the general Theorem 4.7 and its proof is mainly based on the observations sketched in the previous section. To this aim, let us denote

$$\mathcal{O}_{\rm ad} = \{\Omega_o \in \mathcal{O}_T(\Omega)\ :\ \Omega_o \subset D,\ \partial\Omega_o \setminus \Gamma_{\rm w} \text{ is fixed}\}, \tag{4.14}$$

where $\mathcal{O}(\Omega)$ is the family of shapes defined by (4.10) and $D \subset \mathbb{R}^d$ a fixed box of volume $V = |D|$ such that $\partial D \cap \partial\Omega = \partial\Omega_o \setminus \Gamma_{\rm w}$, i.e. the free boundary $\Gamma_{\rm w}$ lies inside the box, whereas $\partial\Omega \setminus \Gamma_{\rm w}$ is

[4] This abstract formulation is valid if $X$ does not depend on $\Omega$; the same result stands also for problems in which $X = X(\Omega)$ introducing suitable Lagrange multipliers for treating Dirichlet boundary conditions.

a subset of the boundary $\partial D$. Moreover, let us consider $\tilde{\mathcal{J}} : V \times \mathcal{O}_{\rm ad} \to \mathbb{R}_+$ as a functional of the velocity field $\mathbf{v} \in V$ and denote $J(\mathbf{v}) = \tilde{\mathcal{J}}(\mathbf{v}, \Omega_o)$. An existence result for the problem (DD-SO) can be found e.g. in [132] (see Theorem 5.1 and related proof):

**Theorem 4.10.** *Assume that $\mathcal{O}_{ad}$ is given by* (4.14) *and the cost functional $J(\mathbf{v})$ is such that*

$$J(\mathbf{v}) = \int_{\Omega_o} \phi(\nabla \mathbf{v})\, d\Omega_o,$$

*where $\phi(s)$ is a continuous, nonnegative and convex function of $s \in \mathbb{R}$. Then, if the state problem* (5.2) *is well posed, problem* (DD-SO) *admits at least one solution.*

*Proof.* (*sketch*) Continuity of the state solution with respect to shape variations can be proved by introducing a sequence of shapes $\{\Omega_{o,n}\}_{n=1}^{\infty} \subset \mathcal{O}_{\rm ad}$ such that $\Omega_{o,n} \xrightarrow{\tau} \hat{\Omega}_o$ and extending by zero the velocity solution in $B \setminus \Omega_{o,n}$. Extended solutions on $B$ verify stability estimates of the Navier-Stokes problem and thus result uniformly bounded; from this sequence we can extract a subsequence that is weakly convergent (thanks to Banach-Alaoglu theorem) and converges to an element fulfilling the state equations over $\hat{\Omega}_o$. To verify this, it is sufficient to pass to the limit in the weak formulation over $\Omega_{o,n}$. The lower semicontinuity of the cost functional follows directly from the assumption made in the statement (see e.g. [103]). □

As in the optimal control case, a second interesting problem we wish to deal with is the following *robust design (SO)* problem, where we aim at finding an optimal shape of the domain $\Omega_o$ which is *robust* with respect to the boundary data $\mathbf{v}_{\rm in} \in \mathcal{U}_{\rm in}$:

$$\begin{aligned} &\hat{\Omega}_o = \arg \min_{\Omega_o \in \mathcal{O}_{\rm ad}} \max_{\mathbf{v}_{\rm in} \in \mathcal{U}_{\rm in,\,ad}} \mathcal{J}(\Omega_o, \mathbf{v}_{\rm in}) \quad \text{s.t.} \quad Y(\Omega_o) = (\mathbf{v}(\Omega_o), p(\Omega_o), \boldsymbol{\xi}(\Omega_o)) \in X(\Omega_o) \ : \\ &\tilde{A}(Y(\Omega_o), W; \Omega_o) = F(W; \Omega_o, \mathbf{v}_{\rm in}),\ \forall\, W = (\mathbf{z}, q, \boldsymbol{\lambda}) \in X(\Omega_o) \end{aligned} \tag{RD-SO}$$

being for simplicity $\mathcal{J}(\Omega_o, \mathbf{v}_{\rm in}) = \tilde{\mathcal{J}}(\mathbf{v}, \Omega_o, \mathbf{v}_{\rm in})$. Existence of solutions for the worst-case shape optimization problem (RD-SO) in the infinite-dimensional case has not been extensively studied yet. We argue that the concept of weak lower semi-continuity for set-valued functions analyzed in [140] could be employed also in the present case. Nevertheless, as in the OC case, further study of the well-posedness of this min-max formulation is needed.

## 4.3 Reduced order modelling for optimization problems governed by parametrized PDEs

Computational and geometrical reduction tools come into play in order to solve optimal control and shape optimization problems, by providing a general and *modular* framework for their efficient numerical approximation and recasting them into the *many-query* context. In practice, for both optimal control and shape optimization problems, the standard adjoint-based approach addressed in the previous sections is very often too expensive. Substantial computational saving becomes possible thanks to a *reduced order model* (ROM) which relies on two reduction steps: *(i)* parametrization of control functions and/or admissible shapes and *(ii)* substitution of the full-order truth solution of state equations with a reduced solution obtained by the reduced basis method. Here we focus on the formulation of the parametrized version of OC and SO problems, while the resulting numerical procedure – based on the RB method for parametrized PDEs – is addressed in Sect. 4.7. The extension to other inverse identification problems will be addressed in Sect. 4.5, by focusing on the a parametric coupling technique for fluid-structure interaction problems in Sect. 4.6.

The key factor is provided by the description of either control functions or admissible shapes through a set of *input* parameters. In particular, we assume that the control functions (in OC case) and admissible shapes (in SO case) are described through a *control* parameter $\boldsymbol{\pi} \in \mathcal{D}_{\boldsymbol{\pi}} \subset \mathbb{R}^{p_\pi}$, while other *uncertain* factors such as boundary conditions, sources or physical properties are described through an uncertainty parameter $\boldsymbol{\omega} \in \mathcal{D}_{\boldsymbol{\omega}} \subset \mathbb{R}^{p_\omega}$. In this way, the state problem is described by a parametrized PDE, where the parameter vector $\boldsymbol{\mu} = (\boldsymbol{\pi}, \boldsymbol{\omega}) \in \mathcal{D} \subset \mathbb{R}^p$ consists of two parts and the parametric set is given by $\mathcal{D} = \mathcal{D}_{\boldsymbol{\pi}} \times \mathcal{D}_{\boldsymbol{\omega}}$, being $p = p_{\boldsymbol{\pi}} + p_{\boldsymbol{\omega}}$. Concerning the two classes of OC and SO problems, we have that:

- in the OC case, representing the control functions by means of a set of parametric inputs is straightforward: for instance, in the boundary velocity control case, we need to provide an expression $\mathbf{u}_\mathrm{c} = \mathbf{u}_\mathrm{c}(\boldsymbol{\pi})$ of the boundary data. By formulating the state problem over a fixed geometrical domain – which can be taken as the reference configuration $\Omega$ in the parametrized PDE context – the deterministic OC problem (DD-OC) can be rephrased as the following *parametrized optimization* problem:

$$\begin{aligned} &\hat{\boldsymbol{\pi}} = \arg\min_{\boldsymbol{\pi} \in \mathcal{D}_{\boldsymbol{\pi}}} \mathcal{J}(\boldsymbol{\pi}) \quad \text{s.t.} \quad Y(\boldsymbol{\mu}) \in X \ : \\ &\tilde{A}(Y(\boldsymbol{\mu}), W; \boldsymbol{\mu}) = F(W; \boldsymbol{\mu}), \ \forall W \in X, \end{aligned} \tag{DD-OC$\boldsymbol{\mu}$}$$

  where $\mathcal{J}(\boldsymbol{\pi})$ is the corresponding parametrized version of the cost functional $\mathcal{J}(\mathbf{u}_\mathrm{c})$ appearing in (DD-OC). On the other hand, by assuming that uncertain elements are parametrized as well, also a robust OC problem can be formulated in the same way. For instance, by denoting $\mathbf{v}_\mathrm{in} = \mathbf{v}_\mathrm{in}(\boldsymbol{\omega})$ the unknown boundary condition in (5.2), the robust OC problem (RD-OC) can be expressed as the following *parametrized robust optimization* problem:

$$\begin{aligned} &\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\pi} \in \mathcal{D}_{\boldsymbol{\pi}}} \max_{\boldsymbol{\omega} \in \mathcal{D}_{\boldsymbol{\omega}}} \mathcal{J}(\boldsymbol{\mu}) \quad \text{s.t.} \quad Y(\boldsymbol{\mu}) \in X \ : \\ &\tilde{A}(Y(\boldsymbol{\mu}), W; \boldsymbol{\mu}) = F(W; \boldsymbol{\mu}), \ \forall W \in X. \end{aligned} \tag{RD-OC$\boldsymbol{\mu}$}$$

  where $\mathcal{J}(\boldsymbol{\mu}) = \mathcal{J}(\boldsymbol{\pi}, \boldsymbol{\omega})$ is the analogous of $\mathcal{J}(\mathbf{u}_c, \mathbf{v}_\mathrm{in})$ in (RD-OC). Equivalently, we denote by $\tilde{\mathcal{J}}(Y(\boldsymbol{\mu}); \boldsymbol{\mu})$ the parametrized equivalent of $\mathcal{J}(\mathbf{v}, \mathbf{u}_c, \mathbf{v}_\mathrm{in})$.

- In the SO case, a parametrized expression of the admissible shapes is achieved through a suitable shape parametrization techniques, so that $\Omega_o = \Omega_o(\boldsymbol{\pi})$. In this way, problem (DD-SO), for instance, reduces to

$$\begin{aligned} &\hat{\boldsymbol{\pi}} = \arg\min_{\boldsymbol{\pi} \in \mathcal{D}_{\boldsymbol{\pi}}} \mathcal{J}_o(\boldsymbol{\pi}) \quad \text{s.t.} \quad Y_o(\boldsymbol{\mu}) \in X(\Omega_o(\boldsymbol{\pi})) \ : \\ &\tilde{A}_o(Y_o(\boldsymbol{\mu}), W; \boldsymbol{\mu}) = F_o(W; \boldsymbol{\mu}), \ \forall W \in X(\Omega_o(\boldsymbol{\mu})), \end{aligned} \tag{4.15}$$

  Whenever we make use of a parametrized representation of the admissible shapes, we end up with a problem under this form. However, it is still defined over the original $\boldsymbol{\mu}$-dependent domain, so that numerical optimization procedures at this stage cannot avoid shape deformations and related mesh motion at each step. Then, in order to exploit fully the parametric framework and the reduced basis methodology, we need to formulate the state problem in the reference domain $\Omega$, taking advantage of the parametric mapping $T(\cdot; \boldsymbol{\mu}) : \Omega \to \Omega_o(\boldsymbol{\mu})$ and the related change of variables. Problem (4.15) can be thus rewritten as follows:

$$\begin{aligned} &\hat{\boldsymbol{\pi}} = \arg\min_{\boldsymbol{\pi} \in \mathcal{D}_{\boldsymbol{\pi}}} \mathcal{J}(\boldsymbol{\pi}) \quad \text{s.t.} \quad Y(\boldsymbol{\mu}) \in X(\Omega) \ : \\ &\tilde{A}(Y(\boldsymbol{\mu}), W; \boldsymbol{\mu}) = F(W; \boldsymbol{\mu}), \ \forall W \in X(\Omega), \end{aligned} \tag{DD-SO$\boldsymbol{\mu}$}$$

Furthermore, by assuming that uncertain elements are parametrized as in the OC case, the robust SO problem (RD-SO) can be expressed as the following *parametrized robust optimization* problem:

$$\begin{aligned}&\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\pi}\in\mathcal{D}_{\boldsymbol{\pi}}}\max_{\boldsymbol{\omega}\in\mathcal{D}_{\boldsymbol{\omega}}} \mathcal{J}(\boldsymbol{\mu}) \quad \text{s.t.} \quad Y(\boldsymbol{\mu})\in X\ :\\ &\tilde{A}(Y(\boldsymbol{\mu}),W;\boldsymbol{\mu}) = F(W;\boldsymbol{\mu}),\ \forall W\in X(\Omega).\end{aligned} \tag{RD-SO$\boldsymbol{\mu}$}$$

Moreover, based on the introduction of a suitable control/shape parametrization, the analysis of the well-posedness of problem (DD-OC$\boldsymbol{\mu}$) (likewise, (DD-SO$\boldsymbol{\mu}$)) is simpler than the non-parametrized cases addressed in the previous sections, thanks to the assumption that input parameter $\boldsymbol{\mu}$ belongs to a finite dimensional space. In fact, the following existence result, based on the classical result for continuous real functions over finite-dimensional compact spaces (see e.g. [283], Thm. 4.16):

**Theorem 4.11.** *Let $\mathcal{D}_{\boldsymbol{\pi},ad}\subseteq\mathcal{D}_{\boldsymbol{\pi}}$ a nonempty, compact and convex subset of $\mathcal{D}\subseteq\mathbb{R}^p$. Then, if $\mathcal{J}:\mathcal{D}_{\boldsymbol{\pi},ad}\to\mathbb{R}$ is a lower semicontinuous functional, problem* (DD-OC$\boldsymbol{\mu}$) *(respectively,* (DD-SO$\boldsymbol{\mu}$)*) admits a solution. Moreover, if $\mathcal{J}$ is also convex, the solution is unique.*

In both cases, we denote $\tilde{A}(Y,W;\boldsymbol{\mu}) = A(Y,W;\boldsymbol{\mu}) + C(Y,Y,W;\boldsymbol{\mu})$ a nonlinear stationary parametrized operator (keeping always in mind the Navier-Stokes case) as already done in the previous chapter. We remark that, thanks to the *(i)* introduction of a suitable parametrization for either control functions or admissible shapes, and to the *(ii)* parametrized PDEs framework, extensively analyzed in the previous chapters, optimal control and shape optimization problems can be recast in a general parametrized optimization context.
We will detail the numerical strategy for solving the parametrized optimization problems introduced so far in Sect. 4.7. It relies on the reduced basis method for parametrized PDEs and a suitable numerical iterative optimization procedure.

Let us close this section by pointing out two facts about the parametrized formulation and the optimality conditions. For sake of simplicity, although these considerations are absolutely general, we consider a state problem given by the parametrized Navier-Stokes equations (3.11)-(3.14). First of all, we observe that, by deriving formally with respect to $\mu_1,\dots,\mu_p$ the expression of the state problem (3.11)-(3.14), we obtain the following *sensitivity equations* for the unknowns $\partial Y(\boldsymbol{\mu})/\partial\boldsymbol{\mu} = (\partial\mathbf{v}(\boldsymbol{\mu})/\partial\boldsymbol{\mu}, \partial p(\boldsymbol{\mu})/\partial\boldsymbol{\mu})\in X = V\times Q$:

$$\begin{aligned}&a(\frac{\partial\mathbf{v}(\boldsymbol{\mu})}{\partial\boldsymbol{\mu}},\mathbf{w};\boldsymbol{\mu}) + d(\frac{\partial\mathbf{v}(\boldsymbol{\mu})}{\partial\boldsymbol{\mu}},\mathbf{w};\boldsymbol{\mu}) + b(\frac{\partial p(\boldsymbol{\mu})}{\partial\boldsymbol{\mu}},\mathbf{w};\boldsymbol{\mu}) + c(\frac{\partial\mathbf{v}(\boldsymbol{\mu})}{\partial\boldsymbol{\mu}},\mathbf{v}(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu})\\ &\quad +c(\mathbf{v}(\boldsymbol{\mu}),\frac{\partial\mathbf{v}(\boldsymbol{\mu})}{\partial\boldsymbol{\mu}},\mathbf{w};\boldsymbol{\mu}) = \frac{\partial F}{\partial\boldsymbol{\mu}}(\mathbf{w};\boldsymbol{\mu}) - \frac{\partial a}{\partial\boldsymbol{\mu}}(\mathbf{v}(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}) - \frac{\partial d}{\partial\boldsymbol{\mu}}(\mathbf{v}(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu})\\ &\quad -\frac{\partial b}{\partial\boldsymbol{\mu}}(p(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}) - \frac{\partial c}{\partial\boldsymbol{\mu}}(\mathbf{v}(\boldsymbol{\mu}),\mathbf{v}(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}), \qquad \forall\mathbf{w}\in V;\\ &b(r,\frac{\partial\mathbf{v}(\boldsymbol{\mu})}{\partial\boldsymbol{\mu}};\boldsymbol{\mu}) = \frac{\partial G}{\partial\boldsymbol{\mu}}(r;\boldsymbol{\mu}) - \frac{\partial b}{\partial\boldsymbol{\mu}}(r,\mathbf{v}(\boldsymbol{\mu});\boldsymbol{\mu}), \qquad \forall r\in Q,\end{aligned} \tag{4.16}$$

(we adopt a compact notation to express $p$ equations, one for each derivative $\partial/\partial\mu_i$, with $i=1,\dots,p$, in $p$ unknowns, $\partial Y(\boldsymbol{\mu})/\partial\mu_i = (\partial\mathbf{v}(\boldsymbol{\mu})/\partial\mu_i, \partial p(\boldsymbol{\mu})/\partial\mu_i)$). By introducing the global Navier-Stokes operators $\tilde{A}(V,W;\boldsymbol{\mu}) = A(V,W;\boldsymbol{\mu}) + C(V,V,W;\boldsymbol{\mu})$ and $\tilde{F}(W;\boldsymbol{\mu})$ defined by (3.40)–(3.14), as well as the Fréchet derivative $d\tilde{A}(Y;\boldsymbol{\mu})(V,W)$ defined by (3.41), problem (4.16) can be rewritten as follows:

$$d\tilde{A}(Y(\boldsymbol{\mu});\boldsymbol{\mu})(\frac{\partial Y(\boldsymbol{\mu})}{\partial\boldsymbol{\mu}},W) = \frac{\partial\tilde{F}}{\partial\boldsymbol{\mu}}(W;\boldsymbol{\mu}) - \frac{\partial\tilde{A}}{\partial\boldsymbol{\mu}}(V,W;\boldsymbol{\mu}), \qquad \forall W\in X,$$

where the right-hand side is computed, thanks to the affinity assumptions (1.7)-(1.8), by taking the derivatives of the $\Theta_\bullet^q$ with respect to $\boldsymbol{\mu}$ – provided that the functions $\Theta_\bullet^q \in \mathcal{C}^1(\mathcal{D};\mathbb{R})$. This assumption is automatically verified in case of affine parametrizations, and also in nonaffine parametrizations treated by EIM if the tensors appearing in the parametrized formulation are at least $\mathcal{C}^1(\mathcal{D}\times\mathbb{R}^d;\mathbb{R}^{d\times d})$ functions. For instance, in the case of the viscous term $a(\cdot,\cdot;\boldsymbol{\mu})$ we obtain:

$$\frac{\partial a}{\partial \mu_i}(\mathbf{v},\mathbf{w};\boldsymbol{\mu}) = \sum_{q=1}^{Q_a} \frac{\partial \Theta_a^q(\boldsymbol{\mu})}{\partial \mu_i} a^q(\mathbf{v},\mathbf{w}), \qquad 1 \le i \le p;$$

in the same way we can compute the parametric derivatives of remaining terms. It is straightforward to show that, under the assumption (3.43), problem (4.16) is well-posed. Moreover, from (3.43) we can easily draw a continuity result for the solution $Y(\boldsymbol{\mu})$ with respect to the parameter $\boldsymbol{\mu}$. In fact, we end up with the following relationship:

$$\left|\left|\frac{\partial Y(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right|\right|_X \le \frac{C_A}{\beta_{BA}(\boldsymbol{\mu})} C_{\boldsymbol{\mu}} \left( \left|\left| \frac{\partial \Theta_\bullet(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}} \right|\right|_{\mathcal{C}^0(\mathcal{D})} \right),$$

where $C_A > 0$ is a constant originating from the continuity of the $\boldsymbol{\mu}$-independent linear/bilinear/trilinear forms appearing at the right-hand side of (4.16) and $C_{\boldsymbol{\mu}}(\cdot)$ is a suitable combination of the $\infty$-norms of the parametric derivatives of the $\Theta_\bullet(\boldsymbol{\mu})$ functions. In the case of a parametrized shape optimization problem, such a relationship plays the same role as the continuity of the state solution with respect to shape variations, required for instance by the well-posedness analysis.

The second fact is about optimality conditions for a parametrized optimization problem like (DD-OC), which we rewrite here by considering a quadratic cost functional of the velocity field (this is a case occurring very often in our applications):

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}\in\mathcal{D}} \mathcal{J}(\boldsymbol{\mu}) \quad \text{s.t.} \quad Y(\boldsymbol{\mu}) \in X : \quad \tilde{A}(Y(\boldsymbol{\mu}),W;\boldsymbol{\mu}) = F(W;\boldsymbol{\mu}),\ \forall W \in X, \tag{4.17}$$

being

$$\mathcal{J}(\boldsymbol{\mu}) = \tilde{\mathcal{J}}(\mathbf{v}(\boldsymbol{\mu}),\boldsymbol{\mu}) = \frac{1}{2} Q(\mathbf{v}(\boldsymbol{\mu}),\mathbf{v}(\boldsymbol{\mu});\boldsymbol{\mu}).$$

Also in the parametric case the approach based on the Lagrangian functional allows to derive the optimality conditions. By introducing the Lagrangian functional

$$\begin{aligned}\mathcal{L}(\mathbf{v},p,\mathbf{z},q,\boldsymbol{\mu}) &= \tilde{\mathcal{J}}(\mathbf{v}(\boldsymbol{\mu}),\boldsymbol{\mu}) + F(z;\boldsymbol{\mu}) + G(q;\boldsymbol{\mu}) \\ &\quad - a(\mathbf{v},\mathbf{z};\boldsymbol{\mu}) - d(\mathbf{v},\mathbf{z};\boldsymbol{\mu}) - b(p,\mathbf{z};\boldsymbol{\mu}) - b(r,\mathbf{u};\boldsymbol{\mu}) - c(\mathbf{v},\mathbf{v},\mathbf{z};\boldsymbol{\mu})\end{aligned} \tag{4.18}$$

where $(\mathbf{z},q) \in V\times Q$ is the Lagrange multipliers (or adjoint variables). Then, $(\hat{\mathbf{v}}(\hat{\boldsymbol{\mu}}),\hat{p}(\hat{\boldsymbol{\mu}}),\hat{\mathbf{z}}(\hat{\boldsymbol{\mu}}),\hat{q}(\hat{\boldsymbol{\mu}}),\hat{\boldsymbol{\mu}}) \in X\times X\times\mathcal{D}$ fulfills the following optimality conditions:

$$\begin{aligned}
&\partial_{\mathbf{z}}\mathcal{L}(\mathbf{v},p,\mathbf{z},q,\boldsymbol{\mu})[\mathbf{w}] = 0 && \forall\ \mathbf{w}\in V && \textit{velocity state equation},\\
&\partial_{q}\mathcal{L}(\mathbf{v},p,\mathbf{z},q,\boldsymbol{\mu})[r] = 0 && \forall\ r\in Q && \textit{pressure state equation},\\
&\partial_{\mathbf{v}}\mathcal{L}(\mathbf{v},p,\mathbf{z},q,\boldsymbol{\mu})[\mathbf{w}] = 0 && \forall\ \mathbf{w}\in V && \textit{velocity adjoint equation},\\
&\partial_{p}\mathcal{L}(\mathbf{v},p,\mathbf{z},q,\boldsymbol{\mu})[r] = 0 && \forall\ r\in Q && \textit{pressure adjoint equation},\\
&\partial_{\boldsymbol{\mu}}\mathcal{L}(\mathbf{v},p,\mathbf{z},q,\boldsymbol{\mu})[\tilde{\boldsymbol{\mu}}] \ge 0 && \forall\ \tilde{\boldsymbol{\mu}}\in\mathcal{D}_{\text{ad}} && \textit{optimality equation}.
\end{aligned} \tag{4.19}$$

It is straightforward to see that by deriving the Lagrangian with respect to the adjoint variables we obtain again the parametrized Navier-Stokes equations representing the state problem. By deriving the Lagrangian with respect to the state variables we obtain instead the following *adjoint* equations:

$$\begin{aligned} &a(\mathbf{w},\mathbf{z}(\boldsymbol{\mu});\boldsymbol{\mu}) + d(\mathbf{w},\mathbf{z}(\boldsymbol{\mu});\boldsymbol{\mu}) + b(q(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}) \\ &\qquad +c(\mathbf{w},\mathbf{u}(\boldsymbol{\mu}),\mathbf{z}(\boldsymbol{\mu});\boldsymbol{\mu}) + c(\mathbf{u}(\boldsymbol{\mu}),\mathbf{w},\mathbf{z}(\boldsymbol{\mu});\boldsymbol{\mu}) = Q(\mathbf{v}(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}), \quad \forall \mathbf{w}\in V; \\ &b(r,\mathbf{z}(\boldsymbol{\mu});\boldsymbol{\mu}) = 0, \qquad\qquad \forall r \in Q. \end{aligned} \tag{4.20}$$

In the end, by deriving the Lagrangian with respect to the parameter $\boldsymbol{\mu}$, we obtain (using the same compact notation as in (4.16)) :

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \boldsymbol{\mu}}(\mathbf{v}(\boldsymbol{\mu}),p(\boldsymbol{\mu}),\mathbf{z}(\boldsymbol{\mu}),q(\boldsymbol{\mu}),\boldsymbol{\mu}) &= \frac{1}{2}\frac{\partial Q}{\partial \boldsymbol{\mu}}(\mathbf{v}(\boldsymbol{\mu}),\mathbf{v}(\boldsymbol{\mu});\boldsymbol{\mu}) + \frac{\partial F}{\partial \boldsymbol{\mu}}(\mathbf{z}(\boldsymbol{\mu});\boldsymbol{\mu}) + \frac{\partial G}{\partial \boldsymbol{\mu}}(q(\boldsymbol{\mu});\boldsymbol{\mu}) \\ -\frac{\partial a}{\partial \boldsymbol{\mu}}(\mathbf{v}(\boldsymbol{\mu}),\mathbf{z}(\boldsymbol{\mu});\boldsymbol{\mu}) - \frac{\partial d}{\partial \boldsymbol{\mu}}(\mathbf{v}(\boldsymbol{\mu}),\mathbf{z}(\boldsymbol{\mu});\boldsymbol{\mu}) - \frac{\partial b}{\partial \boldsymbol{\mu}}(p(\boldsymbol{\mu}),\mathbf{z}(\boldsymbol{\mu});\boldsymbol{\mu}) - \frac{\partial b}{\partial \boldsymbol{\mu}}(q(\boldsymbol{\mu}),\mathbf{v}(\boldsymbol{\mu});\boldsymbol{\mu}) \\ -\frac{\partial c}{\partial \boldsymbol{\mu}}(\mathbf{v}(\boldsymbol{\mu}),\mathbf{v}(\boldsymbol{\mu}),\mathbf{z}(\boldsymbol{\mu});\boldsymbol{\mu}). \end{aligned}$$

This expression turns to be equal to the derivative of the cost functional $\partial\mathcal{J}(\boldsymbol{\mu})/\partial\boldsymbol{\mu}$, with respect to $\boldsymbol{\mu}$, i.e.

$$\frac{\partial \mathcal{J}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}) = \frac{1}{2}\frac{\partial Q}{\partial \boldsymbol{\mu}}(\mathbf{v}(\boldsymbol{\mu}),\mathbf{v}(\boldsymbol{\mu});\boldsymbol{\mu}) + Q(\mathbf{v}(\boldsymbol{\mu}),\frac{\partial \mathbf{v}(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}};\boldsymbol{\mu}) = \frac{\partial \mathcal{L}}{\partial \boldsymbol{\mu}}(\mathbf{v}(\boldsymbol{\mu}),p(\boldsymbol{\mu}),\mathbf{z}(\boldsymbol{\mu}),q(\boldsymbol{\mu}),\boldsymbol{\mu}),$$

thanks to the expression of the adjoint problem (4.20) with $\mathbf{w} = \partial\mathbf{v}(\boldsymbol{\mu})/\partial\boldsymbol{\mu}$ and the sensitivity equations (4.16). Thus, also in the parametric case we can exploit a gradient-based procedure, based on the previous optimality conditions, for the numerical solution of parametrized optimization problems. In any case, the solution of parametrized state problems relies on the reduced basis method presented in the previous chapter.
At the moment, we rely on direct method of nonlinear programming to solve the parametrized optimization problems addressed in this section. In this way, we avoid the explicit solution of the adjoint equations and the computation of the parametric derivatives $\partial\Theta^p_\bullet(\boldsymbol{\mu})/\partial\boldsymbol{\mu}$ – entailing for each operator $Q_\bullet p$ derivatives – for which the empirical interpolation method could potentially be exploited in the most expensive nonaffine cases, as shown in [91]. We will come back on these points in the last section of the chapter 4.7, devoted to the numerical implementation.

## 4.4 Further analysis of FFD and RBF mappings

We provide in this section a deeper analysis of the two *volume-based* shape parametrization techniques – Free-Form Deformation (FFD) and Radial Basis Functions (RBF) techniques – we have introduced in order to deal with complex shapes, recasting them in the class of *perturbation of identity* mappings. In particular, since we rely on FFD mappings to parametrize our shape optimization problems, we provide some conditions ensuring *(i)* a good definition of the mapping and *(ii)* the compactness of the resulting admissible shapes obtained through this parametrization with respect to the parameter $\boldsymbol{\mu}$. Moreover, we provide some results on RBF deformations with respect to the location and the admissible displacements of the control points.

### 4.4.1 Free-Form Deformation mappings

Recalling from Sect. 2.6, a FFD mapping – here denoted as $T_{\text{FFD}} : \mathbb{R}^2 \times \mathcal{D} \to \mathbb{R}^2$ – is obtained as the composition $T_{\text{FFD}}(\mathbf{x};\boldsymbol{\mu}) = \Psi^{-1} \circ \hat{T}_{\text{FFD}} \circ \Psi(\mathbf{x};\boldsymbol{\mu})$, where $\Psi : D \to \hat{D} \equiv (0,1)^2$ is a simple affine mapping,

$$\hat{T}_{\text{FFD}}(\hat{\mathbf{x}};\boldsymbol{\mu}) = \sum_{l=0}^{L}\sum_{m=0}^{M} b_{l,m}^{L,M}(\hat{\mathbf{x}})(\mathbf{P}_{l,m} + \boldsymbol{\mu}_{l,m}), \tag{4.21}$$

and $b_{l,m}^{L,M}(\hat{\mathbf{x}}) = b_l^L(\hat{x}_1)b_m^M(\hat{x}_2)$ are tensor products of the 1D *Bernstein* basis polynomials (2.57). For sake of simplicity, we consider the two-dimensional case ($d = 2$), although the extension of these results to the case $d > 2$ is straightforward.
Thanks to the *partition of unity* property (2.59) of the Bernstein polynomials, to the equispaced distribution of the control points over the lattice, and exploiting the composition with the mapping $\Psi$, we can easily characterize the FFD mapping as a perturbation of identity, under the form

$$T_{\text{FFD}}(\mathbf{x};\boldsymbol{\mu}) = \mathbf{x} + \boldsymbol{\theta}_{\text{FFD}}(\mathbf{x};\boldsymbol{\mu}), \quad \text{being} \quad \boldsymbol{\theta}_{\text{FFD}}(\mathbf{x};\boldsymbol{\mu}) = \sum_{l=0}^{L}\sum_{m=0}^{M} b_{l,m}^{L,M}(\Psi(\mathbf{x}))\mathbb{D}\boldsymbol{\mu}_{l,m}, \tag{4.22}$$

where $\mathbb{D} = diag(D_1, D_2)$ and $D_1, D_2$ are the dimensions of the box $D$ such that $\Omega \subset D$. Let us show some interesting properties of FFD mappings. First of all, owing to the expression of *Bernstein* polynomials, $T_{\text{FFD}}(\cdot;\boldsymbol{\mu}) \in \mathcal{C}^\infty(\mathbb{R}^2;\mathbb{R}^2)$. Moreover, we can prove the following result:

**Proposition 4.12.** *Let $\alpha \in (0,1)$ be an arbitrary constant and let $R \in \mathbb{R}_+$ be defined as*

$$R = R(D, L, M, \alpha) = \frac{1-\alpha}{\max(D_1, D_2)\left(1 + \dfrac{2\sqrt{L^2+M^2}}{\min(D_1, D_2)}\right)}. \tag{4.23}$$

*Then, if the parameter vectors $\{\boldsymbol{\mu}_{l,m}\}$, $l = 1,\dots,L$, $m = 1,\dots,M$ are such that*

$$\sum_{l=0}^{L}\sum_{m=0}^{M} \|\boldsymbol{\mu}_{l,m}\|_2 \le R,$$

*the displacement field $\boldsymbol{\theta}_{FFD}(\cdot;\boldsymbol{\mu})$ verifies $\|\boldsymbol{\theta}_{FFD}(\cdot;\boldsymbol{\mu})\|_{W^{1,\infty}(\mathbb{R}^2;\mathbb{R}^2)} \le 1-\alpha$, so that the FFD mapping $T_{FFD}(\cdot;\boldsymbol{\mu})$ defined by (4.22) is a perturbation of the identity.*

*Proof.* We report here a simple sketch of the proof (see [18] for more details). We show that the constant $R$ defined by (4.23) is such that

$$\|\boldsymbol{\theta}_{\text{FFD}}(\cdot;\boldsymbol{\mu})\|_{W^{1,\infty}(\mathbb{R}^2;\mathbb{R}^2)} = \sup_{\mathbf{x}\in\mathbb{R}^2} \|\boldsymbol{\theta}_{\text{FFD}}(\mathbf{x};\boldsymbol{\mu})\|_2 + \sup_{\mathbf{x}\in\mathbb{R}^2} |||\nabla\boldsymbol{\theta}_{\text{FFD}}(\mathbf{x};\boldsymbol{\mu})|||_2 < 1-\alpha, \tag{4.24}$$

in order to apply the first result of Lemma 4.9. For the first term, we have that:

$$\|\boldsymbol{\theta}_{\text{FFD}}(\mathbf{x};\boldsymbol{\mu})\|_2 \le \sum_{l=0}^{L}\sum_{m=0}^{M} |b_{l,m}^{L,M}(\Psi(\mathbf{x}))|\|\mathbb{D}\boldsymbol{\mu}_{l,m}\|_2 \le \max(D_1, D_2)\sum_{l=0}^{L}\sum_{m=0}^{M}\|\boldsymbol{\mu}_{l,m}\|_2.$$

For the second term, since (see (2.61))

$$\nabla\boldsymbol{\theta}_{\text{FFD}}(\mathbf{x};\boldsymbol{\mu}) = \mathbb{D}\left(\sum_{l=0}^{L}\sum_{m=0}^{M}\boldsymbol{\mu}_{l,m}\otimes\nabla b_{l,m}^{L,M}(\Psi(\mathbf{x}))\right)\mathbb{D}^{-1}$$

we have, thanks to Cauchy-Schwarz inequality,

$$|||\nabla \boldsymbol{\theta}_{\mathrm{FFD}}(\mathbf{x};\boldsymbol{\mu})|||_2 \leq \|\mathbb{D}\|_2 \sum_{l=0}^{L}\sum_{m=0}^{M} \|\boldsymbol{\mu}_{l,m}\|_2 \|\nabla b_{l,m}^{L,M}(\Psi(\mathbf{x}))\|_2 \|\mathbb{D}^{-1}\|_2,$$

where the bound $\|\nabla b_{l,m}^{L,M}(\Psi(\mathbf{x}))\|_2 \leq 2\sqrt{L^2+M^2}$ is obtained thanks to the representation formula of the gradient (2.60) and to the properties of Bernstein polynomials. In the end, we find

$$\|\boldsymbol{\theta}_{\mathrm{FFD}}(\cdot;\boldsymbol{\mu})\|_{W^{1,\infty}(\mathbb{R}^2;\mathbb{R}^2)} \leq \max(D_1,D_2) \sum_{l=0}^{L}\sum_{m=0}^{M} \|\boldsymbol{\mu}_{l,m}\|_2 \left(1+\frac{2\sqrt{L^2+M^2}}{\min(D_1,D_2)}\right)$$

so that (4.24) is verified because $R$ satisfies (4.23). □

By taking the bound $R$ on the norm of the control points displacements as in (4.23), the second result of Lemma 4.9 implies that $\|T_{\mathrm{FFD}}^{-1} - I\|_{W^{1,\infty}(\mathbb{R}^2;\mathbb{R}^2)} \leq 1/\alpha$, so that the family of shapes $\mathcal{O}_{\mathrm{FFD}}(\Omega;R)$ is compact. In particular, the pseudo-distance (4.9) between the reference and the deformed shape is equal to $d(\Omega,\Omega_o(\boldsymbol{\mu})) = \alpha + 1/\alpha$. Hence, the previous result guarantees that the domains $\Omega_o(\boldsymbol{\mu}) = T_{\mathrm{FFD}}(\Omega;\boldsymbol{\mu})$ generated from $\Omega$ through the FFD mapping $T_{\mathrm{FFD}}(\cdot;\boldsymbol{\mu})$ form a compact set under a suitable *small parameters* condition.

**Remark 4.13.** *We point out that parameter vectors $\boldsymbol{\mu}_{l,m} \in \mathbb{R}^2$ and displacement fields are defined over two different spaces (see definition* (2.56) *of the FFD mapping). This explains the presence of the matrix $\mathbb{D}$ in the definition* (4.22) *of the field $\boldsymbol{\theta}_{FFD}(\mathbf{x};\boldsymbol{\mu})$.*

**Remark 4.14.** *From a practical viewpoint, by defining the actual degrees of freedom as $\mu_1,\ldots,\mu_p$ directly on the domain $\Omega$, each corresponding to the displacement of a control point in either the $x_1$ or the $x_2$ direction (i.e. to one of the components of a vector $\boldsymbol{\mu}_{l,m}$), condition* (4.23) *can be simply rewritten as follows:*

$$\boldsymbol{\mu} = (\mu_1,\ldots,\mu_p) \in \mathbb{R}^p \ : \sum_{k'\in K_1} \frac{\mu_{k'}}{D_1} + \sum_{k''\in K_2} \frac{\mu_{k''}}{D_2} \leq R,$$

*where $K_i$ is the set of indices $1 \leq k \leq p$ related to displacements of a control point in the $x_i$ direction.*

The result stated in Theorem 4.12 provides a practical condition to ensure the right definition of a FFD mapping and, in particular, that the determinant of its Jacobian $\det(J_{T_{\mathrm{FFD}}})$ is always (strictly) positive, so that self-intersections of deformed shapes are avoided. We point out that often this condition is quite restrictive, featuring very small deformations; from a numerical point of view, we experienced FFD parametrizations still preserving $\det(J_{T_{\mathrm{FFD}}}) \geq \delta > 0$ also for larger deformations than required in Theorem 4.12, by checking the sign of $\det(J_{T_{\mathrm{FFD}}})$ for very large train sets of possible deformations. Different set of necessary and sufficient injectivity conditions which are however more difficult to implement are reported e.g. in [105].

### 4.4.2 Radial Basis Functions mappings

We now turn to the analysis of the RBF mappings, introduced in Sect. 2.7. In particular, we state some conditions ensuring the well-posedness of the interpolation problem (2.65), and an *a priori* estimate on the (parametrized) coefficients entering in the expression of the RBF mapping

with respect to the parameter $\boldsymbol{\mu}$ and the set of interpolation points.
Since we do not use RBF mappings in shape optimization problems, we do not derive any condition ensuring that the norm of the displacement field obtained through a RBF mapping is bounded by a suitable norm of the parameter vector $\boldsymbol{\mu}$ as in (4.24). Recalling from Sect. 2.7, a RBF mapping is given by

$$T(\mathbf{x};\boldsymbol{\mu}) = \mathbf{c}(\boldsymbol{\mu}) + \mathbb{A}(\boldsymbol{\mu})\mathbf{x} + \mathbb{W}(\boldsymbol{\mu})^T\mathbf{s}(\mathbf{x}), \tag{4.25}$$

being $\mathbf{c}(\boldsymbol{\mu}) \in \mathbb{R}^2$, $\mathbb{A}(\boldsymbol{\mu}) \in \mathbb{R}^{2\times 2}$, $\mathbb{W}(\boldsymbol{\mu}) \in \mathbb{R}^{k\times 2}$ the coefficients of the mapping. Recall that $s(\mathbf{x}) \in \mathbb{R}^k$ is given by $s(\mathbf{x}) = (\Phi(\|\mathbf{x}-\mathbf{P}_1\|), \dots, \Phi(\|\mathbf{x}-\mathbf{P}_k\|))^T$.

Vector and matrices components $c_m(\boldsymbol{\mu})$, $(\mathbb{A}(\boldsymbol{\mu}))_{mn} = A_{mn}(\boldsymbol{\mu})$ and $(\mathbb{W}(\boldsymbol{\mu}))_{im} = w_{im}(\boldsymbol{\mu})$, for $1 \le m, n \le 2$, $1 \le i \le k$ are obtained by solving the following linear systems:

$$\begin{bmatrix} \mathbb{M} & \mathbb{P} \\ \mathbb{P}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbb{W}_m(\boldsymbol{\mu}) \\ \boldsymbol{\gamma}_m(\boldsymbol{\mu}) \end{bmatrix} = \begin{bmatrix} \mathbf{P}^m + \boldsymbol{\mu}^m \\ 0 \end{bmatrix}, \qquad m = 1, 2 \tag{4.26}$$

where (for sake of analysis, notation slightly differs from (2.71)):

- $\mathbb{W}_m(\boldsymbol{\mu}) = (w_{1m}, (\boldsymbol{\mu}) \dots, w_{km}(\boldsymbol{\mu}))^T$ and $\gamma_m(\boldsymbol{\mu}) = (c_m(\boldsymbol{\mu}), A_{1m}(\boldsymbol{\mu}), A_{2m}(\boldsymbol{\mu}))^T$ are the coefficients related to radial and polynomial components to deformation, for the $m$-th coordinate;
- $\mathbf{P}^m = (P_{1m}, \dots, P_{km})^T$ the vector of the $m$-th coordinates of the $k$ control points in the undeformed, reference configuration; in the same way, $\mathbf{P}_i = (P_{i1}, P_{i2})^T$ for any $1 \le i \le k$;
- $\boldsymbol{\mu}^m = (\mu_{1m}, \mu_{km})$ the displacement of the set of $k$ control points in the $m$-th direction;
- $\mathbb{M} \in \mathbb{R}^{k\times k}$ the interpolation matrix, of components $M_{ij} = \Phi(\|\mathbf{P}_i - \mathbf{P}_j\|)$, for any $1 \le i, j \le k$;
- $\mathbb{P} \in \mathbb{R}^{k\times 3}$ the matrix arising from the side constraints and given by $\mathbb{P} = [\ \mathbf{1} \mid \mathbf{P}^1 \mid \mathbf{P}^2]$.

It is possible to show (see e.g. [50], Proposition 2.1) that the interpolation matrix $\mathbb{M}$ is in general symmetric and positive semi-definite; it is positive definite if $\Phi(r) = \phi(r^2)$, where $\phi : \mathbb{R}_+ \to \mathbb{R}$ is a continuous completely monotonic function, i.e. $\phi : \mathbb{R}_+ \to \mathbb{R}$, $\phi \in \mathcal{C}^\infty(0,\infty)$ is such that $(-1)^l\phi^{(l)}(t) \ge 0$ for $l = 0, 1, 2, \dots$ and for all positive $t$. In this case, the RBF interpolation problem admits a unique solution even without adding a polynomial function to the RBF map, provided that $\{\mathbf{P}_i\}_{i=1}^k \subset \mathbb{R}^2$ is a set of $k > 2$ distinct, non-collinear control points. Then, if we use e.g. Gaussian or Multiquadric RBFs (see table 2.1), the interpolation problem is well posed.

On the other hand, we can recover the possibility to solve the RBF interpolation problem even for function which are not completely monotonic, by relaxing the requirement on $\Phi(\cdot)$ and the definition of positive definiteness. In fact, we can always interpolate uniquely with (strictly) *conditionally* positive definite functions[5] if we add polynomials to the interpolant, and if the only polynomial that vanishes on the set of control points $\{\mathbf{P}_i\}_{i=1}^k \subset \mathbb{R}^2$ sites is zero.

[5] A function $F : \mathbb{R}^d \to \mathbb{R}$ is *conditionally* positive definite (of order $k$) if for all finite subsets $\Xi \subset \mathbb{R}^d$, the quadratic form $\sum_{\xi\in\Xi}\sum_{\zeta\in\Xi}\lambda_\xi\lambda_\zeta F(\xi-\zeta)$ is nonnegative for all $\boldsymbol{\lambda} = \{\lambda_\xi\}_{\xi\in\Xi}$ which satisfies $\sum_{\xi\in\Xi}\lambda_\xi q(\xi) = 0$ for any polynomial functions $q \in \Pi_d^{k-1}$ of total degree at most $k-1$ in $d$ unknowns. In our case, $F(\mathbf{x}) = \Phi(\|\mathbf{x}\|)$ and it can be proved that $\Phi(\|\mathbf{x}\|)$ is strictly conditionally positive definite (of order $k$) if $(-1)^k d^k/dt^k\phi(\sqrt{t})$, $t > 0$ is completely monotonic (see e.g. [50], Theorem 5.1). Common choices of RBFs satisfy this requirement wit $k = 2$, so that adding a polynomial of degree $k-1 = 1$ is sufficient.

The side constraints in (4.26) are used to take up the extra degrees of freedom introduced through the use of the polynomial $\pi(\mathbf{x}) \in \Pi_2^1$. Moreover, we can prove the following result:

**Proposition 4.15.** *The RBF mapping* (4.25) *satisfies the following properties:*

1. $T(\mathbf{x};\boldsymbol{\mu}) = \mathbf{x} + \boldsymbol{\theta}(\mathbf{x};\boldsymbol{\mu})$, *where* $\boldsymbol{\theta}(\mathbf{x};\boldsymbol{\mu}) = \mathbf{0}$ *if and only if* $\boldsymbol{\mu} = \mathbf{0}$;

2. *the following* a priori *estimate holds:*

$$\frac{1}{\kappa_2(\mathbb{K})}\frac{\|\mathbf{P}^m\|_2}{\|\boldsymbol{\mu}^m\|_2} \leq (\|\boldsymbol{\gamma}_m(\boldsymbol{\mu})\|^2 + \|\mathbb{W}_m(\boldsymbol{\mu})\|^2)^{1/2} \leq \kappa_2(\mathbb{K})\frac{\|\boldsymbol{\mu}^m\|_2}{\|\mathbf{P}^m\|_2}, \tag{4.27}$$

*where* $\kappa_2(\mathbb{K}) = \dfrac{\max_i \sigma_i(\mathbb{K})}{\min_i \sigma_i(\mathbb{K})}$ *is the spectral condition number of* $\mathbb{K} = \begin{bmatrix} \mathbb{M} & \mathbb{P} \\ \mathbb{P}^T & 0 \end{bmatrix}$.

*Proof.* First of all, we can rewrite expression (4.25) as follows:

$$T(\mathbf{x};\boldsymbol{\mu}) = \mathbf{x} + \mathbf{c}(\boldsymbol{\mu}) + (\mathbb{A}(\boldsymbol{\mu}) - \mathbb{I})\mathbf{x} + \mathbb{W}(\boldsymbol{\mu})^T s(\mathbf{x}).$$

It is easy to check that, since $\mathbb{P} = [\ \mathbf{1} \mid \mathbf{P}^1 \mid \mathbf{P}^2]$, the solution of the linear systems (4.26) with $\boldsymbol{\mu}^m = \mathbf{0}$ is given by $\mathbb{W}_m(\mathbf{0}) = \mathbf{0}$ for $m = 1, 2$ and $\boldsymbol{\gamma}_1(\mathbf{0}) = [0, 1, 0]^T$, $\boldsymbol{\gamma}_2(\mathbf{0}) = [0, 0, 1]^T$, i.e. we have $\mathbb{A}(\mathbf{0}) = \mathbb{I}$, $\mathbf{c}(\mathbf{0}) = \mathbf{0}$, $\mathbb{W}(\mathbf{0}) = \mathbf{0}$, provided that $\mathbf{K}$ is nonsingular. To prove this, let us remark that:

- if $\mathbb{M}$ is symmetric positive definite, $\mathbb{K}$ has a *saddle-point* structure and is nonsingular if and only if $\mathbb{P}$ has full column rank, i.e. if and only if $rank(\mathbb{P}) = 3$. In the present case, this condition is verified if and only if the $k > 2$ control points are non-collinear. In fact, $\mathbb{K}$ is congruent to the block diagonal $diag(\mathbb{M}, \mathbb{S})$, where $\mathbb{S} = -\mathbb{P}^T\mathbb{M}\mathbb{P}$ is the Schur complement, so that $\mathbb{K}$ is nonsingular if and only if $\mathbb{S}$ is nonsingular, provided that $\mathbb{M}$ is nonsingular.

- if $\mathbb{M}$ is symmetric but positive semi-definite, we can rely on theorem 5.1 in [50].

To prove the *a priori* estimate (4.27), we simply apply a basic result in numerical linear algebra (see e.g. [258], Theorem 3.2) for the stability analysis of a linear system $\mathbb{K}\mathbf{x} = \mathbf{b}$, $\mathbf{b} \neq \mathbf{0}$:

$$\mathbb{K}(\mathbf{x} + \boldsymbol{\delta}\mathbf{x}) = \mathbf{b} + \boldsymbol{\delta}\mathbf{b} \quad \Rightarrow \quad \frac{1}{\kappa(\mathbb{K})}\frac{\|\boldsymbol{\delta}\mathbf{b}\|}{\|\mathbf{b}\|} \leq \frac{\|\boldsymbol{\delta}\mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa(\mathbb{K})\frac{\|\boldsymbol{\delta}\mathbf{b}\|}{\|\mathbf{b}\|},$$

by taking, for $m = 1, 2$:

$$\mathbf{b} = \begin{bmatrix} \mathbf{P}^m \\ 0 \end{bmatrix}, \quad \boldsymbol{\delta}\mathbf{b} = \begin{bmatrix} \boldsymbol{\mu}^m \\ 0 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} \mathbb{W}_m(\mathbf{0}) \\ \boldsymbol{\gamma}_m(\mathbf{0}) \end{bmatrix}, \qquad \boldsymbol{\delta}\mathbf{x} = \begin{bmatrix} \mathbb{W}_m(\boldsymbol{\mu}) - \mathbb{W}_m(\mathbf{0}) \\ \boldsymbol{\gamma}_m(\boldsymbol{\mu}) - \boldsymbol{\gamma}_m(\mathbf{0}) \end{bmatrix}.$$

In this way, by taking the Euclidean norm $\|\cdot\|_2$ vectors (and the induced matrix norm to get the spectral condition number $\kappa_2(\mathbb{K})$, the estimate directly follows.

□

We point out that some relationships between the spectral condition numbers of the interpolation matrix $\kappa_2(\mathbb{M})$ and of the global matrix $\kappa_2(\mathbb{K})$ can also be derived, by exploiting the theory of *saddle-point* problems (see e.g. [29], Theorem 3.5). In this way, we can characterize the stability estimate (4.27) by means of the spectral condition number $\kappa_2(\mathbb{M})$ of $\mathbb{M}$.
Moreover, we remark that $\kappa_2(\mathbb{M})$ provides an indication about the complexity of the interpolation

problem, related to location/number of control points and radial basis. In particular, upper bounds under the following form can be established:

$$\kappa_2(\mathbb{M}) \leq k\frac{\max_{i\neq j}|\Phi(\|\mathbf{P}_i-\mathbf{P}_j\))|}{G(q)}, \qquad q=\frac{1}{2}\min_{i\neq j}\|\mathbf{P}_i-\mathbf{P}_j\|_2,$$

where $q$ is the *separation radius* of the control points positions and $G(\cdot)$ a real function depending on the radial basis; e.g., for the Gaussian RBF $\Phi(h)=e^{-h^2}$ we have $G(h)=h^{-d}e^{-\gamma/h^2}$, for the spline-type $\Phi(h)=h^\lambda$ we have $G(h)=h^\lambda$. See e.g. [50], Chapter 5 or [291] and references therein for further details and a general setting.

For a small number of control points, as in our approach, linear systems (4.26) can be efficiently solved by a suitable direct method (matrix factorization is not depending on the parameters). When using a large number of control points – as for example in fluid-structure interaction coupled problems or, more generally, when dealing with mesh motion through RBF – the matrix appearing in (4.26) may be badly conditioned and non-sparse (because of the global character of RBF), and some difficulties may arise. In these cases, suitable scaling or preconditioning strategies may help, as discussed for example in [50]. See also [155] for theoretical construction of shape deformations based on RBFs for mesh deformations and the works by Schaback and Wendland (e.g., [321, 293]) for computational aspects related to RBFs interpolation.

## 4.5 Reduced order modelling for parametrized inverse problems

The parametrized framework enables to deal with more general inverse problems beyond optimization problems governed by PDEs, such as optimal control and shape optimization problems. In this section we address both a deterministic framework and a Bayesian framework for the solution of different types of inverse problems.
Given a parameter vector $\boldsymbol{\mu}\in\mathcal{D}$, the forward problem consists in computing the state solution $Y_o(\boldsymbol{\mu})\in X(\Omega_o)$ solving the state equation

$$A_o(Y_o(\boldsymbol{\mu}),W:\boldsymbol{\mu})+C_o(Y_o(\boldsymbol{\mu}),Y_o(\boldsymbol{\mu}),W;\boldsymbol{\mu})=F_o(W;\boldsymbol{\mu}), \qquad \forall\, W\in X(\Omega_o) \tag{4.28}$$

and in evaluating the corresponding output of interest

$$s_o(\boldsymbol{\mu})=S_o(Y_o(\boldsymbol{\mu})) \tag{4.29}$$

being $s_o:\mathcal{D}\to Z$ and $S_o:X(\Omega_o)\to Z$ suitable observation operators, and $Z$ the vector space of observables. Usually the forward problem (4.28) is solved *many times*, each one corresponding to a new value of the parameter vector $\boldsymbol{\mu}$, in order to predict the outcome $s_o(\boldsymbol{\mu})$ – which is thus equivalent to exploring the graph of the observation operator $\boldsymbol{\mu}\mapsto s_o(\boldsymbol{\mu})$. A more challenging situation is that of inverse problems, where observed data or measurements $\hat{s}_o^*\in Z$ are used to estimate unknown parameters $\boldsymbol{\mu}$ of the physical state system.
As in Sect. 4.3, we shall assume that the input parameter vector $\boldsymbol{\mu}=(\boldsymbol{\pi},\boldsymbol{\omega})\in\mathcal{D}\subset\mathbb{R}^p$ consists of two parts – a *control* parameter $\boldsymbol{\pi}\in\mathcal{D}_{\boldsymbol{\pi}}\subset\mathbb{R}^{p_{\boldsymbol{\pi}}}$ and an uncertainty parameter $\boldsymbol{\omega}\in\mathcal{D}_{\boldsymbol{\omega}}\subset\mathbb{R}^{p_{\boldsymbol{\omega}}}$, so that $\mathcal{D}=\mathcal{D}_{\boldsymbol{\pi}}\times\mathcal{D}_{\boldsymbol{\omega}}$, being $p=p_{\boldsymbol{\pi}}+p_{\boldsymbol{\omega}}$. Depending on the single cases, we may be interested in identifying one or both the components $\boldsymbol{\pi}$ and $\boldsymbol{\omega}$; for sake of notation, whenever possible we will deal in this presentation with the *global* input parameter vector $\boldsymbol{\mu}\in\mathcal{D}$. In the following subsection we provide a first deterministic approach to the solution of inverse problems, which are seen as *parameter identification* problems.

### 4.5.1 A deterministic optimization framework

A deterministic approach for the solution of the inverse problems recasts them in a more general optimization framework, and computes the *best-fit* for the parameter value by minimizing the error (usually in the *least squares* sense) between the observation $s_o^*$ and the output prediction given by the model. We can also refer to this class of problems as *parameter identification* problems. In our context based on parametrized PDEs, we define the abstract parametrized *inverse problem* as follows: given an observation operator $s_o : \mathcal{D} \to Z$, and a target observation $s_o^* \in Z$ – being $Z$ the vector space of observables – find the parameter vector $\boldsymbol{\mu}^* \in \mathcal{D}$ that solves the optimization problem

$$\boldsymbol{\mu}^* = \arg\min_{\boldsymbol{\mu}\in\mathcal{D}} \ \mathcal{J}_o(\boldsymbol{\mu}) := \ \tfrac{1}{2}\|s_o^* - s_o(\boldsymbol{\mu})\|_Z^2 \ + \tfrac{1}{2}\|\boldsymbol{\mu} - \boldsymbol{\mu}_p\|_M^2, \tag{4.30}$$

being $\boldsymbol{\mu}_p \in \mathcal{D}$ a given, target parameter value, $M : \mathbb{R}^P \to \mathbb{R}^P$ a positive definite matrix so that $\|\boldsymbol{\mu} - \boldsymbol{\mu}_p\|_M^2 := (M(\boldsymbol{\mu} - \boldsymbol{\mu}_p), \boldsymbol{\mu} - \boldsymbol{\mu}_p)_{\mathbb{R}^P}$ is a regularization term. In the applications addressed in Part III, we deal with scalar outputs $S_o : X(\Omega_o) \to Z \equiv \mathbb{R}$, given by a linear and/or a quadratic functional of the state solution $Y_o(\boldsymbol{\mu})$, under the form

$$S_o(Y_o(\boldsymbol{\mu})) = L_o(Y_o(\boldsymbol{\mu}); \boldsymbol{\mu}) + Q_o(Y_o(\boldsymbol{\mu}), Y_o(\boldsymbol{\mu}); \boldsymbol{\mu}).$$

Nevertheless, the present framework can be extended to more complex outputs $\mathcal{S}_o : X(\Omega_o) \to Z$, being $Z$ a suitable vector or functional space – a first example in this direction is represented by the inverse identification problem arising in the FSI context. The abstract formulation (4.30) thus covers a large number of practical problems, of which we next give a few examples – recasting also in this framework the cases presented in the previous sections:

- *Parameter identification problem.* The most common problem under the form (4.30) is the identification of uncertain parameters $\boldsymbol{\omega}$ (related e.g. to boundary conditions) or to control parameters $\boldsymbol{\pi}$ (e.g. related with shape) so that the misfit between the computed output $s_o(\boldsymbol{\mu})$ and some given observations$s_o^*$ (possibly affected by experimental noise) is minimized.

- *Optimal design problem (without uncertainty)*:

$$\min_{\boldsymbol{\pi}\in\mathcal{D}_{\boldsymbol{\pi}}} \ \mathcal{J}_o(\boldsymbol{\pi}) := \ \tfrac{1}{2} S_o(Y_o(\boldsymbol{\pi})), \tag{4.31}$$

  where the state equations (4.28) do not contain any uncertainty $\boldsymbol{\omega}$ and $S_o : X(\Omega_o(\boldsymbol{\pi})) \to \mathbb{R}$ is a quadratic functional of the state variables. We can also consider a tracking-type functional $S_o(Y_o(\boldsymbol{\pi}) - Y_o^*)$, where $Y_o^*$ is a reference solution with some desirable properties. For example, in a shape optimization problem related to minimizing the vortices generated behind a bluff body, $Y_o^*$ might represent the Stokes solution (see e.g. [163, 174]).

- *Robust optimal design problem (with uncertainty)*:

$$\min_{\boldsymbol{\pi}\in\mathcal{D}_{\boldsymbol{\pi}}} \ \mathcal{J}_o(\boldsymbol{\pi}) := \ \max_{\boldsymbol{\omega}\in\mathcal{D}_{\boldsymbol{\omega}}} \tfrac{1}{2} S_o(Y_o(\boldsymbol{\mu})), \tag{4.32}$$

  where we seek e.g. a single optimal shape represented through a parameter vector $\hat{\boldsymbol{\pi}}$ to cover a range of possible realizations of the uncertainty $\boldsymbol{\omega}$.

- *Parameter identification (coupled) problem.* In the following section we introduce a parametric coupling algorithm for simple FSI problems. In this case, $\boldsymbol{\omega}$ represents a vector of material parameters, whereas $\boldsymbol{\pi}$ a set of geometrical parameters, describing the free-boundary at the interface between the fluid and the structural domain. Suppose we want to identify the material parameters $\hat{\boldsymbol{\omega}}$ such that a given output, depending on the state (fluid) solution $Y_o(\boldsymbol{\pi})$ computed in $\Omega_o(\boldsymbol{\pi})$ gives the best fit to some measurement $s_o^*$.

This entails the solution of the following parameter identification problem:

$$\min_{\boldsymbol{\omega}\in\mathcal{D}_{\boldsymbol{\omega}}} \mathcal{J}_o(\boldsymbol{\mu}) := \tfrac{1}{2}|s_o^* - S_o(Y_o(\boldsymbol{\pi}^*(\boldsymbol{\omega})))|^2 + \tfrac{1}{2}\|\boldsymbol{\omega}-\boldsymbol{\omega}_p\|_M^2, \tag{4.33}$$

where for any realization $\boldsymbol{\omega}\in\mathcal{D}_{\boldsymbol{\omega}}$ in addition to the forward equation

$$A_o(Y_o(\boldsymbol{\pi}^*,\boldsymbol{\omega}),W;\boldsymbol{\mu})+C_o(Y_o(\boldsymbol{\pi}^*,\boldsymbol{\omega}),Y_o(\boldsymbol{\pi}^*,\boldsymbol{\omega}),W;\boldsymbol{\mu}) = \mathcal{F}_o(W;\boldsymbol{\mu}) \quad \forall W\in X(\Omega_o(\boldsymbol{\pi}^*)),$$

we need to specify another equation, $E_o(\boldsymbol{\pi}^*,\boldsymbol{\omega})=0$ (in this case the parametric coupling problem) in order to close the problem and determine $\boldsymbol{\pi}^*=\boldsymbol{\pi}^*(\boldsymbol{\omega})$. Here $E_o(\boldsymbol{\pi}^*,\boldsymbol{\omega})=0$ returns the shape control parameters $\boldsymbol{\pi}^*$ best fitting the structure displacement.

No matter which numerical algorithm ends up being used (see Sect. 4.7), a deterministic setting for inverse problems leads to the repeated evaluation of the state solution $Y_o(\boldsymbol{\mu})$ and the cost functional $\mathcal{J}_o(\boldsymbol{\mu})$, depending on the output of interest $S_o(Y_o(\boldsymbol{\mu}))$, which is a typical *many-query* problem. Hence, the efficient solution of (4.30) hinges on our ability to simulate effectively the forward problem (4.28): reduction techniques are thus mandatory in order to contain the computational costs, since optimization procedures may require up to $\mathcal{O}(10^2)-\mathcal{O}(10^3)$ input/output evaluations.

Of course, further difficulties are exhibited when the parameters $\boldsymbol{\pi}$ control explicitly or implicitly the domain of the forward problem $\Omega_o(\boldsymbol{\pi})$. In this case, by considering a *fixed domain approach* as in the previous sections, we end up with an inverse problem on the fixed domain $\Omega$:

$$\min_{\boldsymbol{\mu}\in\mathcal{D}} \mathcal{J}(\boldsymbol{\mu}) := \tfrac{1}{2}\|s^* - s(\boldsymbol{\mu})\|_Z^2 + \tfrac{1}{2}\|\boldsymbol{\mu}-\boldsymbol{\mu}_p\|_M^2, \tag{4.34}$$

where $s(\boldsymbol{\mu})=S(Y(\boldsymbol{\mu});\boldsymbol{\mu})$ and $Y(\boldsymbol{\mu})$ is the solution of the forward problem on the fixed domain:

$$\mathcal{A}(Y(\boldsymbol{\mu}),W;\boldsymbol{\mu})+\mathcal{C}(Y(\boldsymbol{\mu}),Y(\boldsymbol{\mu}),W;\boldsymbol{\mu}) = \mathcal{F}(W;\boldsymbol{\mu}) \quad \forall\, W\in X(\Omega). \tag{4.35}$$

Once the problem has been rewritten in the proposed parametrized form (4.34)-(4.35), we can exploit the reduced basis method for parametrized PDEs, in order to speed up the solution of the state problem and related output evaluations during the optimization procedure.

### 4.5.2 A reduced Bayesian statistical framework

Parameter estimation problems search for a best choice of parameters relative to (typically noisy) experimental data. This usually entails the solution of an optimization problem, for any realization of the experimental noise affecting the data, thus yielding to extensive numerical simulations which are unaffordable without resorting to computational reduction strategies. Moreover, in the case that $s^*$ is an experimental measure, possibly polluted by measurement error, a deterministic approach is no longer sufficient to describe fully the propagation of errors through the numerical model. We need a method for quantifying the uncertainty contained in the best estimator. A different perspective, presented e.g. in a recent work by Patera et al. [147], is based on statistical *validation* procedures, which assess the consistency of any candidate parameter value – and hence associated proposed PDE model – with (typically noisy) experimental data. In order to facilitate such an approach, we assume that *the uncertainty can be embedded in the parameters of the forward problem* – we do not consider nonparametric approaches here[6] – and that the *parametric dimension is relatively small* to avoid the curse of dimensionality.

[6] As previously mentioned, we call the subset of parameters that describe the uncertainty as *uncertainty parameters*, as opposed to the other types of parameters in our inverse problem called *control parameters* that can be either directly or indirectly influenced.

We can distinguish between Bayesian and frequentistic validation. The former approach builds directly upon Bayesian parameter estimation [219] from experimental data to likelihood through prior to posterior and finally credible regions [304]. This approach is very general and can easily provide several information on the error propagation through the numerical model; however, it strongly hinges on the choice of a prior probability density.
The latter approach avoids the introduction of prior distributions and is able to provide less subjective conclusions, but still some assumptions on the statistical (parametric) model to describe the observed data (e.g. the Gaussian model) may limit its capability. Many other reasons can of course be addressed in order to motivate the choice between these two general approaches; here we decide to follow a Bayesian approach (opting e.g. for non informative prior distributions) mainly because of its intuitive nature, just in order to provide a possible uncertainty quantification framework where computational reduction can be extremely helpful (as detailed in [101,106,191]). We remark, in fact, that a statistical approach to uncertainty in inverse problems, either from the frequentistic [263] or the Bayesian viewpoint [162,305], requires in any case the repeated evaluation of outputs over large sample sets (e.g. in order to compute sample statistics such as expectations, variances, and higher moments). With this respect, the reduced basis framework can be suitable also for uncertainty quantification [147] and more general probabilistic problems [43, 227].

To address our goals, we can identify two approaches to deal with uncertainty. In the first approach the uncertainty can be "explored" thanks to the inexpensive solution algorithm so that some application-specific conclusions can be drawn. Exploring the parametric uncertainty space of possible outcomes leads naturally to a "many-query" context – the forward problem will need to be solved for multiple different values of the uncertainty parameters.
In the second approach a Bayesian viewpoint (see e.g. [106,191]) is adopted, providing a probability distribution function for the parameter $\boldsymbol{\mu}$ that carries on the uncertainties related to measurements. We assume that we have available certain prior information about the uncertainty parameters plus a characterization of the noise. In particular, we model observations $s^*$ and parameters $\boldsymbol{\mu}$ as random variables and we assume that the probability distributions can be written in terms of (measurable) probability density functions.

By applying the well-known *Bayes' theorem*, the probability distributions can be inverted to obtain the posterior distribution of the uncertainty parameters given the information about the observation. In this way, the solution of the inverse problem is given by the *posterior probability* density $p(\boldsymbol{\mu}|s^*) : \mathcal{D}_{\boldsymbol{\mu}} \times Y \to \mathbb{R}_0^+$, i.e. the probability density of the parameter $\boldsymbol{\mu}$ given the prior information and the measured value of $s^*$, which can be obtained as

$$p(\boldsymbol{\mu} \mid s^*) = \frac{p(s^* \mid \boldsymbol{\mu})\, p(\boldsymbol{\mu})}{p(s^*)}, \tag{4.36}$$

where in (4.36) $p(\boldsymbol{\mu}) : \mathcal{D}_{\boldsymbol{\mu}} \to \mathbb{R}_0^+$ is the *prior probability* density, expressing all available information on $\boldsymbol{\mu}$ independently of the particular measurements on $s^*$ that will be considered as data; $p(s^* \mid \boldsymbol{\mu}) : Y \times \mathcal{D}_{\boldsymbol{\mu}} \to \mathbb{R}_0^+$ is the likelihood function, i.e. the conditional probability density of the observation $s^*$ if $\boldsymbol{\mu}$ were known; the denominator $p(s^*)$ is a norming constant, determined by the normalization of the posterior density, and has often little importance.

The simplest probabilistic model that can be used to describe experimental uncertainties is the Gaussian model; in particular, we assume that our data are $n$ surrogate measures $s^* = s(\boldsymbol{\mu}) + \boldsymbol{\varepsilon}$, obtained by perturbing the output value depending on a parameter vector $\boldsymbol{\mu}$ by adding normally distributed errors $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \Sigma_s)$, uncorrelated and equally distributed, i.e. $\Sigma_s = \sigma_s^2 \mathbb{I}_n$, where the variance $\sigma_s^2$ is supposed to be known. Under this assumption, the likelihood function is given by

$$p(s^* \mid \boldsymbol{\mu}) = p_{\text{noise}}(s^* - s(\boldsymbol{\mu})) \propto \exp\left(-\tfrac{1}{2}\|s^* - s(\boldsymbol{\mu})\|^2_{\Sigma_s^{-1}}\right) = \exp\left(-\frac{1}{2\sigma_s^2}\|s^* - s(\boldsymbol{\mu})\|^2\right).$$

If we can assume a Gaussian model also on the *prior* knowledge of the parameters distributions, i.e. $p(\boldsymbol{\mu}) \sim \mathcal{N}(\boldsymbol{\mu}_p, \Sigma_p)$, then the posterior probability density will be normally distributed, under the form

$$p(\boldsymbol{\mu} \mid s^*) \propto \exp\left(-\tfrac{1}{2}\|s^* - s(\boldsymbol{\mu})\|^2_{\Sigma_s^{-1}} - \tfrac{1}{2}\|\boldsymbol{\mu} - \boldsymbol{\mu}_p\|^2_{\Sigma_p^{-1}}\right).$$

In this case, the maximum *a posteriori* estimator $\boldsymbol{\mu}^*$ is then obtained as

$$\boldsymbol{\mu}^* = \arg\min_{\boldsymbol{\mu}\in\mathbb{R}^p}\left(\tfrac{1}{2}\|s^* - s(\boldsymbol{\mu})\|^2_{\Sigma_s^{-1}} + \tfrac{1}{2}\|\boldsymbol{\mu} - \boldsymbol{\mu}_p\|^2_{\Sigma_p^{-1}}\right),$$

i.e. by solving a specific instance of the regularized minimization problem (4.30). We can also remark that in the Bayesian framework the norms $\|\cdot\|_Y$, $\|\cdot\|_M$ and the value $\boldsymbol{\mu}_p$ have a clear interpretation in terms of the statistics of the measurement noise and the prior probability distribution [305].

On the other hand, if we assume that no information is available about the parameter distribution except that it resides in the parameter space $\mathcal{D}$, i.e. $p(\boldsymbol{\mu}) \sim \mathcal{U}(\mathcal{D})$ is a uniform distribution over $\mathcal{D}$, the posterior probability density is given by

$$p(\boldsymbol{\mu} \mid s^*) \propto \begin{cases} \exp\left(-\dfrac{1}{2\sigma_s^2}\|s^* - s(\boldsymbol{\mu})\|^2\right), & \boldsymbol{\mu} \in \mathcal{D}, \\ 0, & \boldsymbol{\mu} \notin \mathcal{D}. \end{cases}$$

If the parameter space $\mathcal{D}$ has a large dimension (e.g. $P > 4$), an exhaustive exploration of this space is not possible, and we have to rely on Monte Carlo sampling techniques to extract information from the posterior probability density; a well-known technique for sampling probability distributions is the *Metropolis-Hastings* algorithm, which proceeds as follows:

```
pick an initial ω^1; set k = 1;
for k = 2 : K
    compute p(ω^k | s*);
    draw ξ ~ N(0, σ²I), set ω̂ = ω^k + ξ, compute p(ω̂ | s*);
    compute α(ω^k, ω̂) = min(1, p(ω̂ | s*)/p(ω^k | s*))(acceptance ratio);
    draw t ~ U([0,1]);
    if t ≤ α(ω^k, ω̂)
        accept: set ω^{k+1} = ω̂, Ξ = Ξ ∪ ω^{k+1};
    else
        reject: keep ω^{k+1} = ω^k;
    end if
    k ← k + 1;
end
```

After the algorithm has taken $K$ steps, the sample set $\Xi$ contains realizations of the probability distribution $p(\boldsymbol{\mu} \mid s^*)$. In particular, the variance $\sigma^2$ affecting the step $\boldsymbol{\xi}$ should be chosen as large as possible while still maintaining a reasonable acceptance rate [162]. The algorithm reported above is the simplest version of the Metropolis-Hastings, where we assume for instance that the so-called *proposal distribution* from which we sample $\boldsymbol{\xi}$ is symmetric (in the present case it is assumed to be Gaussian). See e.g. Chapter 6 of [157] or to more details about Markov Chain Monte Carlo (MCMC) algorithms for sampling (complex and high-dimensional) posterior densities probability distributions.

## 4.6 Reduced order modelling for steady FSI problems

As seen in the previous section, the same framework already exploited for the efficient solution of optimal control and shape optimization problems can be applied also to other inverse problems. A case of remarkable interest is the numerical simulation of steady Fluid-Structure Interaction (FSI) problems [220], where a deformable structure mutually interacts with a fluid flow, which can be either internal (like blood flowing in compliant vessels) or external (like in aircraft wings/surfaces interactions).

### 4.6.1 General formulation

Following the so-called Arbitrary Lagrangian Eulerian (ALE) formulation [220], which is particularly suitable for the numerical approximation of FSI problems, we consider a Lagrangian description for the structure and the usual Eulerian description for the flow. In this way, the computational domain has to comply with the behavior of the (free) boundary, described by a suitable structural model, and through the (free) boundary acts on the fluid flow across it. In the ALE formulation we have then the interplay of (at least) two motions: the motion of the structural medium – governed by e.g. elasticity equations – and the motion of the computational domain, which is *arbitrary* but is subject to the constraint provided by the movement of the free boundary. A overview on the most recent results on the analysis of this type of problems may be found in [199]. Even if several elements provided in the following come from, and may be applied to, time-dependent FSI problems, we focus throughout this section on *steady* state FSI problems, which have been extensively analyzed by Grandmont et al. [118, 119].

Let us denote by $(Y_o, \eta_f, \eta_s)$ the flow variables $Y_o = (\mathbf{v}_o, p_o)$, the displacement of the computational (fluid) domain $\Omega_o$ and the displacement of the structure domain, respectively. Then, we can express a FSI problem by means of three coupled subproblems as follows:

$$\begin{cases} \mathcal{F}(Y_o, \eta_s, \eta_f) = 0 & \textit{Fluid problem} \\ \mathcal{S}(Y_o, \eta_s) = 0 & \textit{Structure problem} \\ \mathcal{G}(\eta_s, \eta_f) = 0 & \textit{Geometry} \end{cases}$$

describing the physics of the fluid (given the solid displacement and the domain displacement), the solid displacement (given the traction exerted by the fluid) and the displacement of the fluid computational domain, respectively. Coupling conditions are expressed through the continuity of the stresses and the continuity of the domain displacement (or geometric adherence) at the interface $\Gamma_w$ between fluid and structure. In this way, we can denote by $\eta := \eta_s \equiv \eta_f$ the structural displacement and by $\Omega_o = \Omega_o(\eta)$ the original domain for the fluid equations, which depend on the displacement $\eta$ of the (free) structural boundary. Then, we consider as fluid problem in the variable, original domain $\Omega_o = \Omega_o(\eta)$ the steady Navier-Stokes equations (3.3), which are rewritten here for the reader's convenience:

$$\begin{cases} a_o(\mathbf{v}, \mathbf{w}) + \delta d_o(\mathbf{v}, \mathbf{w}) + b_o(p, \mathbf{w}) + \delta c_o(\mathbf{v}, \mathbf{v}, \mathbf{z}) = \langle F^o, \mathbf{w} \rangle \,, & \forall \mathbf{w} \in X_o(\Omega_o(\eta)) \\ b_o(q, \mathbf{v}) = \langle G^o, q \rangle \,, & \forall q \in Q_o(\Omega_o(\eta)). \end{cases} \tag{4.37}$$

For sake of simplicity, we assume to deal with variable domains under the form $\Omega_o(\eta) = \{\mathbf{x}_o = (x_{o1}, x_{o2}) \in (0, L) \times (0, R(x_{o1}))\}$ where $R(x_{o1}) = \bar{R} + \eta(x_{o1})$ and $\eta \in D := H_0^2(0, L)$ is the smooth displacement of the outer wall of the domain (free boundary) $\Gamma_{\mathrm{w}}$ from its reference configuration $\Omega = (0, L) \times (0, \bar{R})$ (see Fig. 4.1). This configuration can also be seen as a cylindrical tube of radius $\bar{R}$ which is axisymmetrically deformed.

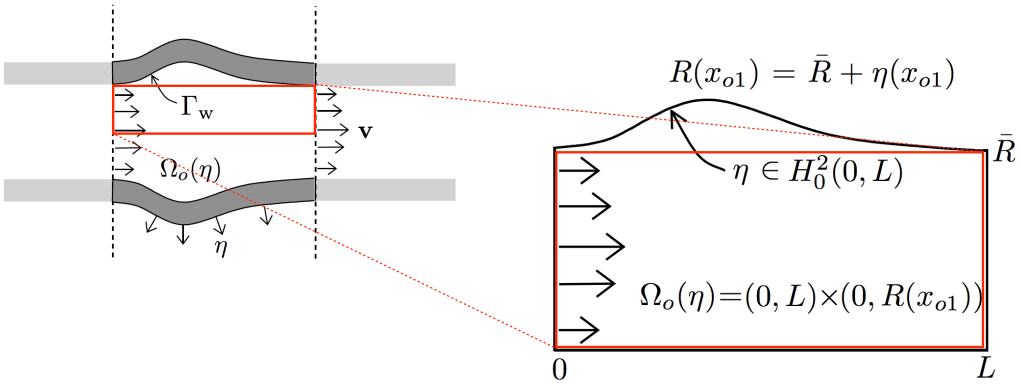



Figure 4.1: Schematic representation of the geometrical configuration for the FSI case.

In order to close the system, we need to provide an equation for solving the displacement $\eta$. This equation is in the Lagrangian form on the undeformed configuration of the structural boundary $\Gamma_w$, which we identify as the interval $(0, L)$ in the current simplified case. Furthermore, we assume to deal with small displacements, which take place in the normal direction of $\Gamma_w$, the tangential displacement being equal to zero. A linear elasticity model for the displacement is chosen, as the following second order equation with a fourth order perturbation (with $\varepsilon > 0$ small)

$$\varepsilon \frac{\partial^4 \eta}{\partial x_1^4} - kGh \frac{\partial^2 \eta}{\partial x_1^2} + \frac{Eh}{1-\nu_P^2} \frac{\eta}{R_0(x_1)^2} = \tau_{\Gamma_w}, \quad x_1 \in (0, L), \tag{4.38}$$

where $h$ is the wall thickness, $k$ is the Timoshenko shear correction factor, $G$ the shear modulus, $E$ the Young modulus, $\nu_P$ the Poisson ratio, $R_0$ the radius of the reference configuration, and $\tau_{\Gamma_w}$ denotes the traction applied to the wall by the fluid inside the domain $\Omega_o$ (see e.g. [259]). We have added a fourth order term (weighted by the quantity $\varepsilon$) in order to allow a better regularity for the displacement. The weak form of (4.38) reads as follows: find the structural displacement in the normal direction $\eta \in D$ such that

$$e(\eta, \phi) = \langle \tau_{\Gamma_w}, \phi \rangle, \qquad \forall \phi \in D, \tag{4.39}$$

being

$$e(\eta, \phi) = \varepsilon \int_0^L \frac{\partial^2 \eta}{\partial x_1^2} \frac{\partial^2 \phi}{\partial x_1^2} \, dx_1 + kGh \int_0^L \frac{\partial \eta}{\partial x_1} \frac{\partial \phi}{\partial x_1} \, dx_1 + \frac{Eh}{1-\nu_P^2} \int_0^L \frac{\eta \phi}{R_0(x_1)^2} \, dx_1$$

and $D := H_0^2(0, L)$ the space of kinematically admissible displacements. The fluid and structure are coupled together by taking the applied traction $\tau_{\Gamma_w}$ on the structure to be the normal component of the normal Cauchy stress of the fluid on $\Gamma_w$, i.e.

$$\tau_{\Gamma_w} = (\boldsymbol{\sigma}(\mathbf{v}_o, p_o)\boldsymbol{n}) \cdot \boldsymbol{n}, \quad \text{on } \Gamma_w. \tag{4.40}$$

We remark that $\tau_{\Gamma_w}$ can be expressed in the weak sense using[7] the residual $r(\mathbf{v}_o; \mathbf{w}) \in V_w'$ of the fluid solution on the interface $\Gamma_w$ (see e.g. [186]), which is defined for any $\mathbf{w} \in V_w = \{\mathbf{v} \in (H^1(\Omega_o))^2 : \mathbf{v} = \mathbf{0} \text{ on } \Gamma_w\}$ as

$$r(\mathbf{v}_o; \mathbf{w}) := \langle F^o, \mathbf{w} \rangle - a_o(\mathbf{v}_o + L_{o,\mathbf{g}}, \mathbf{w}) - b_o(p_o, \mathbf{w}) - c_o(\mathbf{v}_o + L_{o,\mathbf{g}}, \mathbf{v}_o + L_{o,\mathbf{g}}, \mathbf{w}); \tag{4.41}$$

we refer to the notation introduced in Sect. 3.1. In this way, the coupled FSI problem reads as

[7] More precisely, we should introduce the Riesz representant $\tilde{r}(\mathbf{v}_o; \mathbf{w}) \in V_w$ such that $r(\mathbf{v}_o; \mathbf{w}) = {}_{X_w'}\langle \tilde{r}(\mathbf{v}_o), \mathbf{w} \rangle_{X_w}$ for any $\mathbf{w} \in X_w$ and the trace operator $\gamma_{\Gamma_w} : X_w \to (H^{1/2}(\Gamma_w))^2$.

follows:

$$\begin{cases} a_o(\mathbf{v},\mathbf{w}) + \delta d_o(\mathbf{v},\mathbf{w}) + b_o(p,\mathbf{w}) + \delta c_o(\mathbf{v},\mathbf{v},\mathbf{z}) = \langle F^o, \mathbf{w}\rangle\,, & \forall \mathbf{w} \in X_o(\Omega_o(\eta)) \\ b_o(q,\mathbf{v}) = \langle G^o, q\rangle\,, & \forall q \in Q_o(\Omega_o(\eta)) \\ e(\eta,\phi) = {}_{H^{-1/2}(0,L)}\langle \tilde{r}(\mathbf{v}_o)\cdot\mathbf{n}, \phi\rangle_{H^{1/2}(0,L)}\,, & \forall \phi \in D. \end{cases} \tag{4.42}$$

The coupled problem (4.42) is well-posed under the assumptions already made for the well-posedness of the Navier-Stokes equations and the hypotheses about the structural equation discussed in this section. The proof is based on a Schauder fixed point theorem; we refer for instance to the work of Grandmont [118, 119] for the details.

### 4.6.2 Parametric Fluid-Structure coupling

Several numerical strategies have been devised in the last decades (see e.g. [96]) to tackle FSI problems; we can roughly distinguish between *monolithic* [68] and *partitioned* approaches, where the equations describing the fluid flow and the structure displacement are solved simultaneously (in the former) or separately, and the coupling is obtained through an iterative scheme (in the latter). Following the second strategy, a trial configuration of the geometry is used to solve the fluid and structure problems, then the coupling conditions are tested: in case they are not fulfilled (up to a prescribed tolerance), the trial configuration is updated by displacing the free boundary and another step is performed. This entails the use of a recursive mesh updating by moving the boundary nodes (and adjusting the interior mesh points to maintain mesh quality). In order to avoid the reassembly of matrices corresponding to the discretization of the fluid and the structure problems, we formulate the fluid problem[8] on a reference domain, by describing its deformations through a low dimensional shape parametrization, for instance a FFD mapping. The nonlinear coupled problem (4.42) is thus solved through a partitioned algorithm based on a parametric coupling method: fluid and structure are coupled together by parametrizing the computational domain (and consequently the displacement) for the fluid, and measuring the coupling error in the parametric displacement space. This technique, introduced in [184] for a Stokes problem, has been extended to Navier-Stokes equations in a subsequent work [180], on which this presentation is based.

In particular, once the computational fluid domain is parametrized with a FFD mapping $T(\cdot;\boldsymbol{\pi}) : \Omega \to \Omega_o(\boldsymbol{\pi})$, any fluid problem corresponding to an original domain $\Omega_o(\boldsymbol{\pi})$ is traced back on the reference domain $\Omega$, whereas the shape deformation is managed through the input parameters $\pi_1, \ldots, \pi_{p_{\boldsymbol{\pi}}}$ defining the FFD mapping. From the fluid problem we extract the loads representing the data for the structure problem, whose solution gives in turn the updated domain (through the new values of the parameters $\pi_1, \ldots, \pi_{p_{\boldsymbol{\pi}}}$ for the fluid problem. Hence, a fixed point coupling algorithm can be written in the parameter space – rather than the displacement space. Again, an iterative procedure is needed to ensure the coupling conditions are satisfied up to a desired tolerance. In this way, starting from an initial configuration $\Omega_o(\boldsymbol{\pi}^{(0)})$, at each iteration $k \geq 1$ of the algorithm we must solve a least-squares problem

$$\min_{\boldsymbol{\pi}^{(k+1)} \in \mathcal{D}_{\text{ad}}} \mathcal{J}(\boldsymbol{\pi}^{(k+1)}; \boldsymbol{\pi}^{(k)}) := \int_{\Gamma_{\text{w}}} |\eta(\boldsymbol{\pi}^{(k+1)}) - \mathcal{S}\circ\mathcal{L}\circ\mathcal{F}(\boldsymbol{\pi}^{(k)})|^2 \, d\Gamma \tag{4.43}$$

to find the configuration of the fluid domain $\Omega_o(\boldsymbol{\pi}^{k+1})$ at the next iteration, being:

[8] The same approach, based on a suitable shape parametrization and RB methods for computational reduction, can also be applied to the structural equation. However, since in this case we deal with a simple one-dimensional equation for the displacement $\eta$, we simply rely on the truth 1D finite element approximation of the structure problem.

- $\mathcal{F} : \mathcal{D}_{\boldsymbol{\pi}} \to X$ the fluid resolvant operator, returning the fluid solution $Y(\boldsymbol{\pi}^{(k)})$ in $\Omega_o(\boldsymbol{\pi}^{(k)})$;
- $\mathcal{L} : X \to H^{-1/2}(0, L)$ an operator transferring the loads from fluid to structure in a weak way using the fluid residual (4.41);
- $\mathcal{S} : H^{-1/2}(0, L) \to D$ the structure resolvant operator, returning the structural displacement $\eta$ solution of (4.38) assuming the applied loads;
- $\eta(\boldsymbol{\pi})$ the parametric displacement of the fluid domain at configuration $\Omega_o(\boldsymbol{\pi})$, computed through the FFD mapping, without solving any further equation.

During the coupling procedure, seeking for the *best* parameter value (and related parametrized displacement) matching the displacement yielded by the fluid-structure coupling, a potentially large number of parametrized PDEs need to be solved in different parametric configurations – this is the stage where the RB method comes into play. Reduction obtained through this procedure is thus twofold, since *(i)* the fluid equations for any given parametrized domain are solved through the RB method; *(ii)* the displacement $\eta(\boldsymbol{\pi})$ of the fluid domain is described through a FFD parametrization and recovered from the structural displacement $\eta$ by solving a parameter identification problem.

## 4.7 Numerical approximation of optimization problems governed by parametrized PDEs

In this section we present the numerical procedures developed and implemented for the solution of the problems discussed throughout this chapter. They are nonlinear programming problems depending on a relatively small number of optimization variables, but entailing a very large number of (nonlinear) constraints – *a priori*, one equation for each degree of freedom of the truth finite element approximation. By means of a suitable combination of the techniques analyzed in the previous chapters – RB methods, flexible shape parametrizations and EIM for treating non-affinities – the solution of such a problem can be greatly enhanced.

### 4.7.1 A possible *reduce-then-optimize* approach

For the numerical solution of optimization problems under PDE constraints, two different paradigms[9] can be adopted, both relying on iterative procedures : *(i) optimize-then-discretize*, where we first apply the iterative method to the (continuous) system of optimality conditions, then we discretize the various steps of the algorithm, or *(i) discretize-then-optimize*, where we first discretize our optimal control problem, then we apply an iterative algorithm to solve the discrete version of the system of optimality conditions [128, 143]. We will follow a (reduced version) of this latter procedure.

Let us consider the case of a parametrized problem like (4.17) – that we rewrite here for convenience:

$$\begin{aligned} &\hat{\boldsymbol{\mu}} = \arg \min_{\boldsymbol{\mu} \in \mathcal{D}_{\text{ad}}} \mathcal{J}(\boldsymbol{\mu}) \quad \text{s.t.} \quad Y(\boldsymbol{\mu}) \in X \ : \\ &\tilde{A}(Y(\boldsymbol{\mu}), W; \boldsymbol{\mu}) = F(W; \boldsymbol{\mu}), \ \forall W \in X. \end{aligned} \tag{$\boldsymbol{\mu}$-OPT}$$

Then, let us *discretize* the former problem, by introducing a truth approximation FE space $X^{\mathcal{N}}$

[9] Here we do not consider alternatives to iterative procedures, based for instance on the so-called *one-shot* approach [268], which aims at solving the discretized systems arising from the optimality conditions in a *monolithic* way, instead than pursuing a partitioned scheme, like in most common iterative procedures.

(whose dimension will be denoted $\mathcal{N}$ throughout this section), in order to get the following version:

$$\begin{aligned} &\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}\in\mathcal{D}_{\text{ad}}} \mathcal{J}^{\mathcal{N}}(\boldsymbol{\mu}) \quad \text{s.t.} \quad Y^{\mathcal{N}}(\boldsymbol{\mu}) \in X^{\mathcal{N}} \ : \\ &\tilde{A}(Y^{\mathcal{N}}(\boldsymbol{\mu}), W^{\mathcal{N}}; \boldsymbol{\mu}) = F(W^{\mathcal{N}}; \boldsymbol{\mu}), \ \forall\, W^{\mathcal{N}} \in X^{\mathcal{N}}. \end{aligned} \qquad (\boldsymbol{\mu}\text{-OPT-}\mathcal{N})$$

At this stage, the first possibility (also referred as *indirect method*) is to tackle the system of optimality conditions arising from ($\boldsymbol{\mu}$-OPT-$\mathcal{N}$) which involve the repeated solution of the (discrete) state and adjoint equations, as well as the update of the control variables [143, 39]. Due to the fact that the cost functional $\mathcal{J}^{\mathcal{N}}(\boldsymbol{\mu})$ depends on the solution of the state problem, it can be very nonconvex and contain several local minima. This is unavoidable in the case of *non-parametrized* optimization problems; in this case, large-scale trust-region methods can help achieve global convergence of the optimization method. Moreover, as we have already remarked, a standard adjoint-based approach is in practice too expensive [128].

In a parametrized context – i.e. when the control parameter vector $\boldsymbol{\mu} \in \mathcal{D} \subset \mathbb{R}^p$ is finite-dimensional – we can rely on more efficient procedures (also referred as *direct methods*) and have access to many tools from the theory of nonlinear programming [229]. Thus, we solve directly the minimization problem by seeking a sequence $\{\boldsymbol{\mu}^{(k)}\}_{k=1,2,\dots}$ of values of $\boldsymbol{\mu}$, on which the value of the cost functional $\mathcal{J}^{\mathcal{N}}(\boldsymbol{\mu})$ is reduced. Provided that the analytical expression of the first-order sensitivities $d\mathcal{J}^{\mathcal{N}}/d\mu_j$ are available, optimization can then be efficiently performed e.g. using quasi-Newton methods[10] (such as the *sequential quadratic programming* technique); see the forthcoming subsection for a brief sketch of these methods. For pure least-squares type functionals, there also exists the possibility of using the efficient *Levenberg-Marquardt-Fletcher* [99] method. In the case of a min-max optimization problem of the type (RD-OC$\boldsymbol{\mu}$), we either need to solve an inner maximization problem inside each outer minimization problem, or consider a relaxation

$$\min_{\boldsymbol{\pi}\in\mathcal{D}_{\boldsymbol{\pi}}} \max_{i=1,\dots,n} \tfrac{1}{2}\mathcal{J}^{\mathcal{N}}(\boldsymbol{\pi}, \boldsymbol{\omega}^i) \qquad (4.44)$$

for a discrete set of uncertainty values $\boldsymbol{\omega}^1, \dots, \boldsymbol{\omega}^n \in \mathcal{D}_{\boldsymbol{\omega}}$, after which the inner maximization reduces to a procedure of enumeration.

From a purely optimization standpoint – and avoiding for simplicity the inequality constraints given by $\boldsymbol{\mu} \in \mathcal{D}_{\text{ad}}$, problem ($\boldsymbol{\mu}$-OPT-$\mathcal{N}$) can be seen as a particular instance of the following more general problem

$$\min_{\boldsymbol{\mu}} \mathcal{J}^{\mathcal{N}}(\boldsymbol{\mu}) \quad \text{s.t.} \quad g^{\mathcal{N}}(\boldsymbol{\mu}) = 0,$$

where $\mathcal{J}^{\mathcal{N}} : \mathbb{R}^p \to \mathbb{R}$ and $g^{\mathcal{N}} : \mathbb{R}^p \to \mathbb{R}^{\mathcal{N}}$ are the discrete representations of the cost functional and the PDE constraints, respectively. The number of constraints at the discretized level corresponds to the total amount of degrees of freedom related to the truth full-order approximation of the state problem and has a strong impact on the speed of the iterative optimization procedure. Thus, thanks to the reduced basis approximation of the state (parametrized PDE) problem and related output of interest, we can pursue a remarkable reduction at this stage, since the RB version of the state problem entails just $N \ll \mathcal{N}$ degrees of freedom (and the same for the number of constraints in the optimization problem).

Our procedure for numerical optimization is rather based on a *(discretize-then) reduce-then-optimize* approach, where the optimization is performed directly at the reduced order level. The

[10] On the other hand, *derivative-free* method for optimization such as genetic algorithms or surrogate models [211, 300] replacing the exact cost functional with e.g. low-order polynomial interpolants can also be exploited.

corresponding reduced version of the problem we end up with is as follows:

$$\begin{aligned} &\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}\in\mathcal{D}_{\text{ad}}} \mathcal{J}_N(\boldsymbol{\mu}) \quad \text{s.t.} \quad Y_N(\boldsymbol{\mu}) \in X_N \ : \\ &\tilde{A}(Y_N(\boldsymbol{\mu}), W_N; \boldsymbol{\mu}) = F(W_N; \boldsymbol{\mu}), \ \forall\, W_N \in X_N, \end{aligned} \tag{$\boldsymbol{\mu}$-OPT-N}$$

Clearly, error bounds are necessary to certify the reduced approximation also for sake of the optimization procedure. At the moment, we rely on (partial) error estimation procedures for state variables and outputs (as shown e.g. in Sect. 3.4.3) separately, whereas an error estimation procedure on the global optimization procedure – say, to bound the difference $|\hat{\mathcal{J}}^{\mathcal{N}}(\hat{\boldsymbol{\mu}}) - \hat{\mathcal{J}}_N(\hat{\boldsymbol{\mu}})|$ are still under investigation and are available just for simple cases at the moment (see e.g. [223, 121]). Several remarks about e.g. *optimality* and *predictability* errors, as well as about the *general* sub-optimality of the RB optimal solution with respect to the corresponding FE optimal solution, can be found e.g. in [72].
We remark that here we consider a numerical approximation of the gradient of the cost functional based on finite differences, instead of solving the *adjoint* problem. As sketched in the forthcoming subsection, within this approach we just require an efficient tool for evaluating the state solution and the cost functional for many parameter values: this perfectly fits the RB framework, which can thus provide *rapid responses* in terms of input/output evaluations thanks to the reduced dimension $N \ll \mathcal{N}$ of the approximation. The extension of the current framework to optimization procedures accounting for a reduced version of the adjoint problem is currently under investigation.

### 4.7.2 An iterative procedure for constrained optimization

Let us consider the following finite-dimensional optimization problem:

$$\min_{\boldsymbol{\mu}} \mathcal{J}_N(\boldsymbol{\mu}) \quad \text{s.t.} \quad g_N(\boldsymbol{\mu}) = 0, \tag{4.45}$$

where $\mathcal{J}_N : \mathbb{R}^p \to \mathbb{R}$ and $g_N : \mathbb{R}^p \to \mathbb{R}^N$ are the reduced representations of the cost functional and the PDE constraints, respectively. The basic idea for solving this minimization problem is to build an iterative procedure that, starting from a given $\boldsymbol{\mu}^{(0)}$, generates a sequence $\{\boldsymbol{\mu}^{(k)}\}$ typically such that $\mathcal{J}_N(\boldsymbol{\mu}^{(k+1)}) < \mathcal{J}_N(\boldsymbol{\mu}^{(k)})$, until a suitable stopping criterium is fulfilled. In order to move from the current iterate $\boldsymbol{\mu}^{(k)}$ to the new $\boldsymbol{\mu}^{(k+1)}$, we can rely on *(i) line search* type methods, where the algorithm chooses a direction $\mathbf{d}_k$ and searches along this direction from the current $\boldsymbol{\mu}^{(k)}$ for a new $\boldsymbol{\mu}^{(k+1)}$ such that $\mathcal{J}_N(\boldsymbol{\mu}^{(k+1)}) < \mathcal{J}_N(\boldsymbol{\mu}^{(k)})$; *(ii) trust region* type methods, where the information gathered about $\mathcal{J}_N$ is used to construct a model function whose behavior in a given region (the *trust region*) near the current iterate $\boldsymbol{\mu}^{(k)}$ is similar to that of the actual objective function $\mathcal{J}_N(\cdot)$, modifying the size of the region during the iteration process.

Among the former class, we will exploit the so-called *sequential quadratic programming* (SQP) technique, whose idea is to model (4.45) at the current iterate $\boldsymbol{\mu}^{(k)}$ by a quadratic programming subproblem and to use the minimizer of this subproblem to define a new iterate $\boldsymbol{\mu}^{(k+1)}$. Following [229], SQP methods can be seen as an application of the Newton's method to the *Karush-Kuhn-Tucker* (KKT) optimality conditions for (4.45). Denoting by $\mathcal{L}_N(\boldsymbol{\mu}, \boldsymbol{\lambda}) = \mathcal{J}_N(\boldsymbol{\mu}) - \boldsymbol{\lambda}^T g_N(\boldsymbol{\mu})$ the Lagrangian functional for (4.45), where $\boldsymbol{\lambda} \in \mathbb{R}^N$ is the Lagrange multiplier, the Newton-step from $\boldsymbol{\mu}^{(k)}$ is given by the solution $\mathbf{d}^{(k)}$ of the following KKT system:

$$\begin{bmatrix} H_k & -M_k^T \\ M_K & 0 \end{bmatrix} \begin{bmatrix} \mathbf{d}^{(k)} \\ \boldsymbol{\lambda}^{(k+1)} \end{bmatrix} = \begin{bmatrix} -\nabla \mathcal{J}_N(\boldsymbol{\mu}^{(k)}) \\ -g_N(\boldsymbol{\mu}^{(k)}) \end{bmatrix} \tag{4.46}$$

where $H_k = H(\boldsymbol{\mu}^{(k)}, \boldsymbol{\lambda}^{(k)})$ is the Hessian of the cost functional, such that $H(\boldsymbol{\mu}) = \nabla^2_{\boldsymbol{\mu}\boldsymbol{\mu}} \mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\lambda})$, and $M_k = M(\boldsymbol{\mu}^{(k)})$ is the Jacobian matrix of the constraints, such that $M(\boldsymbol{\mu})^T = [\nabla g_1(\boldsymbol{\mu}), \ldots, \nabla g_N(\boldsymbol{\mu})]$.

It is straightforward to prove that the solution $(\mathbf{d}^{(k)}, \boldsymbol{\lambda}^{(k+1)})$ of Newton's problem (4.46) is equivalent to that of the quadratic program

$$\min_{\boldsymbol{\mu}} \frac{1}{2}\mathbf{d}^T H_k \mathbf{d} + \nabla \mathcal{J}_N(\boldsymbol{\mu}^{(k)})^T \mathbf{d} \quad \text{s.t.} \quad M_k \mathbf{d} + g_N(\boldsymbol{\mu}^{(k)}) = 0. \tag{4.47}$$

With respect to the Newton method, *quasi*-Newton methods allow to avoid the calculation of the Hessian, replacing it with a suitable approximation; in this way, an approximated inverse of the Hessian matrix, say $B_k^{-1}$, is used to replace $H_k^{-1}$ in equation (4.46). One of the most popular formulas for updating the Hessian approximation $B_k$ is the so called BFGS (from *Broyden-Fletcher-Goldfarb-Shanno*) formula:

$$B_{k+1} = B_k - \frac{B_k \boldsymbol{s}_k \boldsymbol{s}_k^T B_k}{\boldsymbol{s}_k^T B_k \boldsymbol{s}_k} + \frac{\mathbf{w}_k \mathbf{w}_k^T}{\mathbf{w}_k^T \boldsymbol{s}_k}, \tag{4.48}$$

being

$$\boldsymbol{s}_k = \boldsymbol{\mu}^{(k+1)} - \boldsymbol{\mu}^{(k)}, \qquad \mathbf{w}_k = \nabla_{\boldsymbol{\mu}} \mathcal{L}(\boldsymbol{\mu}^{(k+1)}, \boldsymbol{\lambda}^{(k+1)}) - \nabla_{\boldsymbol{\mu}} \mathcal{L}(\boldsymbol{\mu}^{(k)}, \boldsymbol{\lambda}^{(k)}).$$

In any case, calculation of the quantities involving the gradient of the cost functional relies on a finite difference approximation instead than on the solution of the corresponding adjoint problem, as already mentioned. In this way, a general SQP algorithm can be set as follows:

Given a tolerance $\epsilon > 0$ and an initial guess $(\boldsymbol{\mu}^{(0)}, \boldsymbol{\lambda}^{(0)})$;
repeat
  evaluate $\mathcal{J}_N(\boldsymbol{\mu}^{(k)}), \nabla \mathcal{J}_N(\boldsymbol{\mu}^{(k)}), H_k = H(\boldsymbol{\mu}^{(k)}, \boldsymbol{\lambda}^{(k)}), g_N(\boldsymbol{\mu}^{(k)})$ and $M(\boldsymbol{\mu}^{(k)})$;
  compute the search direction $\mathbf{d}^{(k)}$ by solving (4.46);
  compute the step length $\tau_k$ (line search) and set $\mathbf{s}_k = \tau_k \mathbf{d}^{(k)}$;
  set $\boldsymbol{\mu}^{(k+1)} = \boldsymbol{\mu}^{(k)} + \mathbf{s}_k$;
  update the Lagrange multiplier $\boldsymbol{\lambda}^{(k+1)}$;
  set $\mathbf{w}_k = \nabla_{\boldsymbol{\mu}} \mathcal{L}(\boldsymbol{\mu}^{(k+1)}, \boldsymbol{\lambda}^{(k+1)}) - \nabla_{\boldsymbol{\mu}} \mathcal{L}(\boldsymbol{\mu}^{(k)}, \boldsymbol{\lambda}^{(k)})$
  obtain $B_{k+1}$ by updating $B_k$ (quasi-Newton BFGS formula (4.48));
  set $k = k + 1$;
until $\|\nabla \mathcal{J}(\boldsymbol{\mu}_k)\| < \varepsilon^*_{OPT}$ or $|\mathcal{J}(\boldsymbol{\mu}_{k+1}) - \mathcal{J}(\boldsymbol{\mu}_k)| < \varepsilon^*_{OPT}$ or $\|\boldsymbol{\mu}_{k+1} - \boldsymbol{\mu}_k\| < \varepsilon^*_{OPT}$.

where the step length $\tau_k$ is computed through a *line search* method in order to guarantee that $\tau_k \mathbf{d}^{(k)}$ is a descent direction. Once the direction $\mathbf{s}_k$ has been determined, the expression of the Lagrange multipliers has to be updated. Several strategies for dealing with these issues can be devised – one of the most popular choices of $\tau_k$ consists e.g. in performing an inexact line search with the *Armijo* rule; see e.g. [229] for details about the practical implementation and [143, 39] for a general overview of available approaches to numerical approximation of optimization problems under PDE constraints. Numerical approximation of optimal flow control problems by SQP method is discussed e.g. in [266].

### 4.7.3 Implementation of the reduced basis optimization procedure

We now sum up the numerical procedure which has been implemented for the numerical solution of the optimization problems based on parametrized PDEs presented in this chapter. We rely on the reduced basis approximation of the state flow problems, on suitable shape parametrization such as FFD or RBF techniques for dealing with variable shapes – and eventually on the empirical interpolation method to recover the affinity assumption – as well as on the numerical optimization procedure sketched in the previous subsection.

All these ingredients make up an integrated and modular framework (both shape parametrizations and PDE models are in fact interchangeable), where we can easily distinguish between four main stages, represented in the scheme of Fig. 4.2:

1. *Pre-processing.* After parametrizing the control functions or the admissible shapes, determine a parametric range of variation $\mathcal{D} \subset \mathbb{R}^p$ for the parameter vector $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)$ on the basis of problem-dependent criteria. Shape parametrization entails the computation of the map $T(\cdot; \boldsymbol{\mu})$ and of the parametrized tensors like $\boldsymbol{\nu}(\boldsymbol{\mu})$ and $\boldsymbol{\chi}(\boldsymbol{\mu})$. Their expressions are computed symbolically by means of a Computer Algebra System (CAS).
   Then, compute the affine expansions (1.7)-(1.8) in case of nonaffinely parametrized tensors by applying the EIM procedure to each tensor component.

2. *RB Offline stage.* This is the most extensive stage of the whole computations. It can be divided in four substages:

   2.1 *Finite element structure assembling.* Assemble the finite element structures corresponding to the $\boldsymbol{\mu}$-independent terms of (1.7)-(1.8) using the affine expansions of the tensors previously computed. In particular, assemble the matrices $\mathbb{A}^q \in \mathbb{R}^{\mathcal{N}_V \times \mathcal{N}_V}$ and the right-hand sides $\mathbf{f}^q \in \mathbb{R}^{\mathcal{N}_V}$ (in the same way other linear/bilinear/trilinear terms).

   2.2 *Stability factors lower bound computation.* Construct the (parametric) lower bound of the inf-sup constant $\beta_{LB}^{\mathcal{N}}(\boldsymbol{\mu})$ exploiting the *Successive constraint method* (Sect. A.3).

   2.3 *Reduced basis space assembling.* Perform the *greedy* algorithm (Sect. 3.3.1) in order to build the RB spaces $X_N$ for the state variables (in our case, velocity and pressure). For each selected $\boldsymbol{\mu}^n$, solve the truth approximation problem (fixed point iteration in the nonlinear case), enrich the velocity space by the solution of the supremizer equation and execute a Gram-Schmidt orthonormalization. In this way the basis matrices $\mathbb{Z}_{2N}$ and $\mathbb{Z}_N$ for velocity and pressure RB spaces, respectively, are filled.

   2.4 *RB post-processing.* After the *greedy* basis selection procedure, $\boldsymbol{\mu}$-independent reduced structures (e.g. $\mathbb{A}_N^q$, $\mathbf{f}_N^q, \dots$) are obtained by a pre- and post- multiplication of the corresponding FE structures with the basis matrices $\mathbb{Z}_{2N}$ and $\mathbb{Z}_N$. In the same way, all the $\boldsymbol{\mu}$-independent structures entering in the computation of dual norms of residuals are assembled and stored.

3. *RB Online stage/Numerical Optimization.* Solve the optimization problem as a *many-query* problem, by a suitable iterative subroutine (based e.g. on the sequential quadratic programming techniques), up to a given (very small) tolerance. At each step, an *Online* evaluation of the RB state solution and the cost functional to be minimized are required.

Concerning the software exploited to deal with the numerical simulations, for the RB Offline/Online computations we have used (and extended) the `rbMIT` library [149, 237], while the truth FE approximation and assembly of related FE structures rely on the `MLife` library [285], set in the `Matlab`-PDE toolbox environment.
We remark that the *Offline* stage thus depends on the choice of the FE spaces, the reference domain $\Omega$, the shape parametrization and the EIM approximation in the nonaffine cases. The *Online* stage consists of many evaluations of field variables $Y_N(\boldsymbol{\mu}) = (\mathbf{v}_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu}))$ and parametric outputs $s_N(\boldsymbol{\mu})$. It depends on the dimension $N$ of the RB spaces and the numerical procedure employed to solve the optimization problem.
As extensively shown by the applications provided within next Part, the parametrized framework introduced and analyzed in this chapter enables to obtain remarkable computational savings, above all in complex optimization problems dealing with variable shapes, thanks to the coupling of suitable geometrical parametrizations and the reduced basis method for parametrized PDEs.

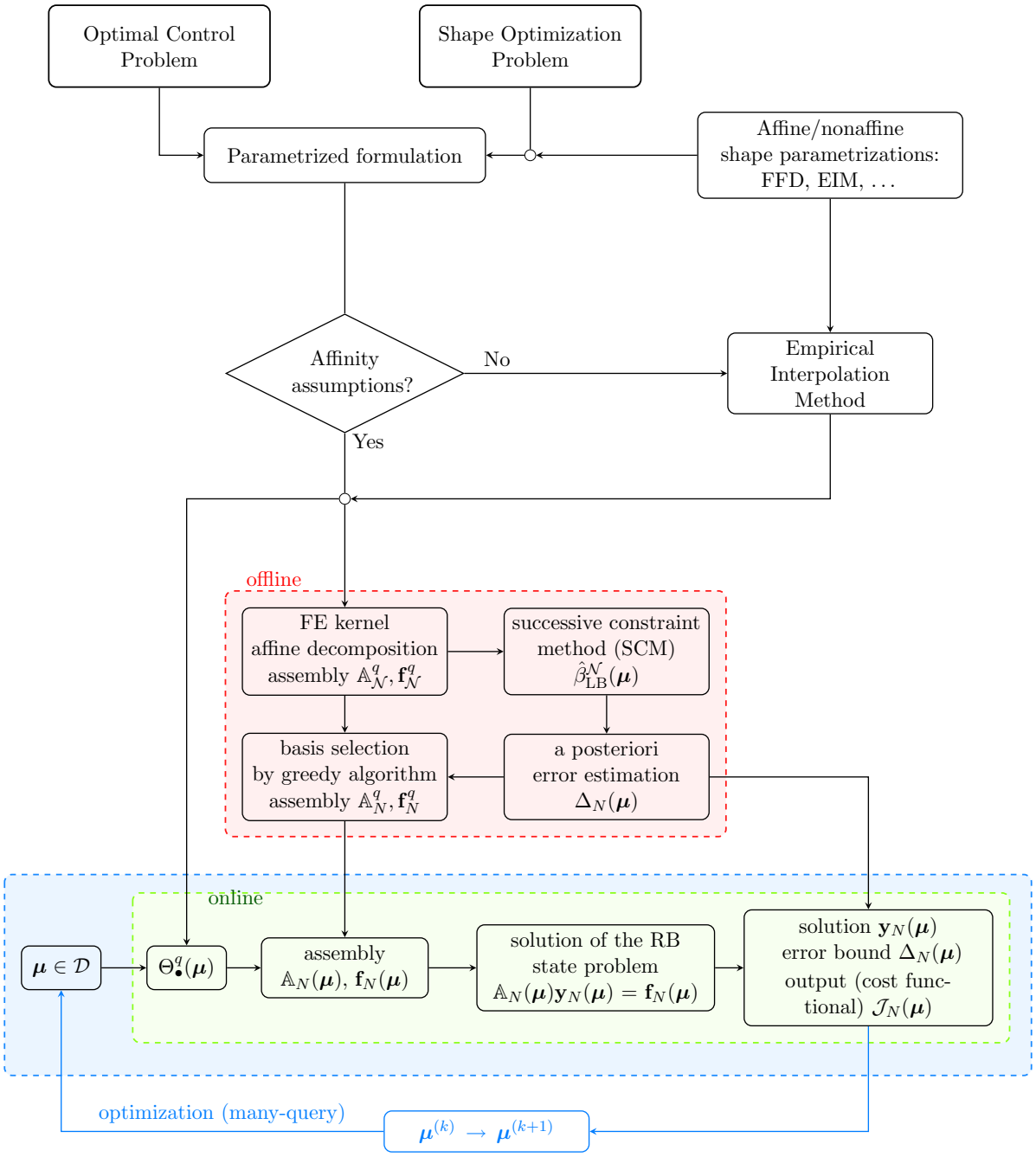



Figure 4.2: Numerical procedure for the solution of reduced parametrized optimization problems.

## 4.8 Design parameter space reduction

We close this chapter by showing some possible techniques for reducing the dimension of the design parameter space. The evaluation of input-output relationships requiring PDE solutions suffers from the *curse of dimensionality* if the input parameter space has a too large dimension, so that methods based on sampling and parameter space exploration become too much expensive. As a result, we are interested to keep the dimension of the parameter space as small as possible also when dealing with RB methods for parametrized PDEs, and parametric shape design problems are no exception. With this respect, we aim at reducing the (geometric) design parameter space in the case of FFD parametrization, by setting some automatic procedures for selecting the active control points – ultimately, the design parameters $\mu_1, \dots, \mu_p$ within the RB context. The problem of restricting the design parameters to a small but effective subset is sometimes called *variable screening* in the surrogate model optimization community (see e.g. [300, 314]). We might consider two different paradigms to reach this goal: *restriction* vs. *adaptivity*. In the former case, we start off with a parametrization with a large number of free parameters, and attempt to find their most effective subset for representing shapes of interest while fixing all the other parameters. For this purpose we take advantage of some methods originating from the statistical Design of Experiments (DoE). In the latter case, instead, we start from very few parameters and add one parameter at a time in an adaptive fashion, by minimizing at each step some defect functional.

### 4.8.1 Selection techniques based on restriction procedures

We have implemented four techniques within the *restriction* paradigm for reducing the geometric design parameter space and selecting a suitable set of control points in the FFD parametrization case. A *restriction* paradigm is better suited in this case, where a starting set of *candidate* parameters is given by the lattice of control points. Instead, an adaptive paradigm should be more feasible with RBF parametrizations, where the control points can be freely positioned.

Let us assume to be able to construct a *trial* parametrization of the reference domain $\Omega$ with a priori large number of parameters, $p' \gg 10$, and a large parameter space $\mathcal{D}' \subset \mathbb{R}^{p'}$, which can effectively explore all the parametric variability of the PDE system we are interested in. In our case a trial parametrization is obtained by FFD, by considering sufficiently many control points. The goal is to select a reduced parameter set $\mathcal{D} \subset \mathcal{D}'$ of all the possible parameters $\boldsymbol{\mu} \in \mathcal{D} \subset \mathbb{R}^p$, with $p \ll p'$, such that the parametric variability of a given quantity of interest $s(\boldsymbol{\mu})$ – e.g., the cost functional $s(\boldsymbol{\mu}) = S(y(\boldsymbol{\mu}))$ to be minimized – is still well explored in the lower-dimensional parameter space. Since the selection procedure has to be run before the construction of a reduced-order approximation, the exploration of the larger parameter space $\mathcal{D}'$ relies on the full-order approximation – i.e. on the input/output evaluation $\boldsymbol{\mu} \to s^{\mathcal{N}}(\boldsymbol{\mu}) = S(y^{\mathcal{N}}(\boldsymbol{\mu}))$. The number of these expensive input/output evaluations has to be as small as possible.
Without loss of generality, we can assume that the output functional $S : X \to \mathbb{R}$ (and its discrete version) are under the form

$$S(y) = \frac{1}{2} Q(y, y), \qquad S(y^{\mathcal{N}}) = \frac{1}{2} (\mathbb{Q}\mathbf{y}, \mathbf{y}),$$

where $Q : X \times X \to \mathbb{R}$ is a symmetric and coercive bilinear form and $\mathbb{Q} \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$ a symmetric positive definite matrix, respectively, and $\mathbf{y} \in \mathbb{R}^{\mathcal{N}}$ the vector of degrees of freedom representing $y^{\mathcal{N}} \in X^{\mathcal{N}}$ over a basis of the discrete space $X^{\mathcal{N}}$. In particular, let us denote $(\mathbf{y}, \mathbf{w})_Q = (\mathbb{Q}\mathbf{y}, \mathbf{w})$ the inner product defined over $X^{\mathcal{N}}$ by $\mathbb{Q}$ and $\mathbb{Q} = \mathbb{H}\mathbb{H}^T$ its Cholesky factorization.

In the following we sketch the four selection recipes:

1. **User's experience-based (UEB) selection**
   In this case, the reduced set of retained parameters is selected according to *empirical* rules, based e.g. on user's experience or problem-dependent. In fact, the closer the control point to the shape, the more influent its effect on shape deformation when it is displaced.

2. **One-at-a-time (OAT) selection** by experimental orthogonal design
   Here we exploit the simplest technique for *variable screening*, by adding to the retained set of parameters the ones with the largest observed *sensitivities*, with respect to varying one parameter at a time while keeping the others fixed. Let us define the parametric variations $\boldsymbol{\mu}^{i,\min}$ and $\boldsymbol{\mu}^{i,\max}$ for $i = 1, \dots, p'$, by varying one parameter at a time while keeping the other parameters fixed at a reference value $\bar{\boldsymbol{\mu}}$:

   $$\boldsymbol{\mu}^{i,\min} = (\,\bar{\mu}_1, \dots, \mu_i^{\min}, \dots, \bar{\mu}_{p'}\,), \quad \boldsymbol{\mu}^{i,\max} = (\,\bar{\mu}_1, \dots, \mu_i^{\max}, \dots, \bar{\mu}_{p'}\,), \quad \forall i = 1, \dots, p'.$$

   Then, let us evaluate the parametric range w.r.t to each parameter by solving the FE approximations for $y^{\mathcal{N}}(\boldsymbol{\mu}^{i,\min})$ and $y^{\mathcal{N}}(\boldsymbol{\mu}^{i,\max})$ for each $i = 1, \dots, p'$, and retain the $p$ parameters with the largest *observed* parametric sensitivities, such that

   $$\left|\frac{\partial s}{\partial \mu_i}(\bar{\boldsymbol{\mu}})\right| \approx \frac{|S(y^{\mathcal{N}}(\boldsymbol{\mu}^{i,\max})) - S(y^{\mathcal{N}}(\boldsymbol{\mu}^{i,\min}))|}{|\boldsymbol{\mu}^{i,\max} - \boldsymbol{\mu}^{i,\min}|} > \varepsilon_{tol}^{OAT}, \tag{4.49}$$

   where $\varepsilon_{tol}^{OAT} > 0$ is a prescribed tolerance (otherwise, we fix a priori the number $p$ of parameters to be selected). We underline that the calculation of the *exact* parametric sensitivities would need to approximate the solution $\partial y(\boldsymbol{\mu})/\partial \boldsymbol{\mu}$ of the sensitivity equations, and thus the parametric derivatives $\partial \Theta_\bullet / \partial \boldsymbol{\mu}$ for a very large number $p'$ of parameter components – rather unfeasible for such a preparatory stage. In any case, such an approach may fail when some design parameters are strongly correlated.

3. **Morris screening procedure (MR-OAT)**
   Morris' randomized *one-at-a-time* design is a screening procedure based on individual randomized OAT designs, accounting also for interactions among the factors [218, 57]. The basic idea is related to a sample of independently observed *elementary effects*, which are exploited to measure the output sensitivity for a particular input. Let us suppose that each component $\mu_i$, $i = 1, \dots, p'$ is scaled to have a region of interest equal to $[0, 1]$ and can take $k$ discrete values in the set $\{0, 1/(k-1), \dots, 1\}$. The experimental domain is then a $p'$-dimensional $k$-level grid in $\mathcal{D}'$. The elementary effect of the $i$-th factor at a given point in the design space is defined as

   $$\delta_i(\boldsymbol{\mu}) = \frac{s(\mu_1, \dots, \mu_{i-1}, \mu_i + \Delta, \mu_{i+1}, \dots, \mu_{p'})}{\Delta}$$

   where $\Delta$ is a fixed multiple of $1/(k-1)$. The ultimate goal of the Morris procedure is to estimate the mean and the standard deviation of the distribution of elementary effects associated with each input[11], which can be obtained by producing a random sample of $R$ elementary effects for each $i = 1 \dots, p'$. Here we do not provide the construction of the procedure (see [218] for details). We just mention that, by the MR-OAT procedure, $R$ elementary effects are produced for each input parameter $\mu_i$ at a total cost of $R(p'+1)$ input/output evaluations (which is a linear function of the numbers of factors involved), corresponding to a sample $\mathcal{B} = \{\boldsymbol{\mu}^{l,r}\} \subset \mathcal{D}$, $l = 1, \dots, p'+1$, $r = 1, \dots, R$.

[11] A large absolute value for the mean of $\delta_i(\boldsymbol{\mu})$ implies that the corresponding parameter component $\mu_i$ has an important *overall* effect on the output, whereas a high standard deviation indicates that the effect of $\mu_i$ is not constant, which may be implied by a parameters interacting with other parameters or whose effect is nonlinear.

In this way, at the $r$-th step, $(p'+1)$ parameters values $\{\boldsymbol{\mu}^{l,r}\}_{l=1,\dots,p'+1}$ originate a *design* matrix accounting for possible interactions among the parameters. Once the $R(p'+1)$ input/output evaluations have been performed, we rank the parameters $\mu_1,\dots,\mu_{p'}$ according to its *absolute mean effect*

$$\bar{\delta}(i) = \frac{1}{R}\sum_{r=1}^{R} |\delta_i(\boldsymbol{\mu}^{i,r})| \;\approx\; \frac{1}{|\mathcal{D}'|}\int_{\mathcal{D}} \left|\frac{\partial s(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right| d\boldsymbol{\mu} \tag{4.50}$$

and retain as active parameter only the $p \ll p'$ most effective ones.

4. **Approximate POD (APOD)**
In this case we consider a *greedy-like* procedure over the large parameter space $\mathcal{D}'$, based on a set of computed *snapshots*. Let us define the parametric variations $\boldsymbol{\mu}^{i,\min}$ and $\boldsymbol{\mu}^{i,\max}$ as in (OAT) case, by varying one parameter at a time while keeping the other parameters fixed at a reference value $\bar{\boldsymbol{\mu}}$. Then:

- evaluate the parametric range w.r.t to each parameter by computing $y^{\mathcal{N}}(\boldsymbol{\mu}^{i,\min})$ and $y^{\mathcal{N}}(\boldsymbol{\mu}^{i,\min})$, for each $i=1,\dots,p'$, and set the initial *parametric directions* as the vectors corresponding to $y^{\mathcal{N}}(\boldsymbol{\mu}^{i,\min})$ and $y^{\mathcal{N}}(\boldsymbol{\mu}^{i,\min})$:

$$\boldsymbol{\zeta}_i^{\min} = \mathbf{y}(\boldsymbol{\mu}^{i,\min}), \qquad \boldsymbol{\zeta}_i^{\max} = \mathbf{y}(\boldsymbol{\mu}^{i,\max}), \qquad i=1,\dots,p';$$

- sample a suitable training set $\Xi_{\text{train}} \subset \mathcal{D}'$ and form the matrix of *snapshots* $\mathbb{Y} = [\boldsymbol{\zeta}_1|\dots|\boldsymbol{\zeta}_S]^T$ by computing

$$\boldsymbol{\zeta}_s := \mathbf{y}(\boldsymbol{\mu}_s) - \mathbf{y}(\bar{\boldsymbol{\mu}}), \qquad \{\boldsymbol{\mu}_s\}_{s=1}^{S} \subset \Xi_{\text{train}}.$$

Once these quantities have been stored, we can set the following iterative procedure:

`set` $\mathcal{P}' = \{1,\dots,p'\}\, \mathcal{P} = \emptyset$;
`repeat`
  `for` $i \in \mathcal{P}'$ `compute` the $\mathbb{Q}-$correlation factors
    $\Delta_i^{\min} := \dfrac{\|\mathbb{Y}\mathbb{H}\boldsymbol{\zeta}_i^{\min}\|_2}{\|\boldsymbol{\zeta}_i^{\min}\|_2}, \qquad \Delta_i^{\max} := \dfrac{\|\mathbb{Y}\mathbb{H}\boldsymbol{\zeta}_i^{\max}\|_2}{\|\boldsymbol{\zeta}_i^{\max}\|_2}$
  `select` $\hat{p} = \arg\max_i\{\Delta_i^{\min}, \Delta_i^{\max}\}$ (largest total correlation)
  `set` $\mathcal{P} = \mathcal{P} \cup \hat{p},\ \mathcal{P}' = \mathcal{P}' \setminus \hat{p}$;
  `deflate` all the other parametric directions: for $q \neq \hat{p}$
    $\boldsymbol{\zeta}_q^{\min\,/\,\max} \longmapsto \boldsymbol{\zeta}_q^{\min\,/\,\max} - \dfrac{(\boldsymbol{\zeta}_{\hat{p}}^{\min\,/\,\max}, \boldsymbol{\zeta}_q^{\min\,/\,\max})_K}{(\boldsymbol{\zeta}_{\hat{p}}^{\min\,/\,\max}, \boldsymbol{\zeta}_{\hat{p}}^{\min\,/\,\max})_K}\boldsymbol{\zeta}_{\hat{p}}^{\min\,/\,\max}$;
`until` $\max_i\{\Delta_i^{\min}, \Delta_i^{\max}\} < \varepsilon_{tol}^{APOD}$.

being $\varepsilon_{tol}^{APOD} > 0$ a prescribed tolerance. We denoted this procedure as *approximate POD*. In fact, the ($K$-weighted) POD modes are given by the right singular vectors $\boldsymbol{\xi}_i$ of $\mathbb{Y}\mathbb{H} = \mathbb{U}\Sigma\mathbb{V}^T$, whereas their significance is measured by the singular values $\sigma_i$. We can consider the quantity $\boldsymbol{\zeta}_i^{\min\,/\,\max}$ corresponding to the largest $\mathbb{Q}$-correlation factor $\Delta_i^{\min\,/\,\max}$ at each step as an approximation to the largest singular vector of $\mathbb{Y}\mathbb{H}$ *in the subspace spanned by the current set of* $\boldsymbol{\zeta}_i^{\min\,/\,\max}$.
We point out that if all parameters are independent and the snapshots are taken exactly as the parametric directions $y^{\mathcal{N}}(\boldsymbol{\mu}^{i,\min})$, $y^{\mathcal{N}}(\boldsymbol{\mu}^{i,\max})$, then APOD and OAT strategies coincide.

We compare these techniques applied to a simple test case in the forthcoming section. We remark that, concerning instead RBFs parametrizations, the best choice of control points (or *centers*) to approximate a given function has been considered as a problem of pure approximation theory. A possible *greedy* algorithm has been proposed in [292]; see e.g. [269] for a possible application of efficient RBF parametrizations in the framework of a shape parameterization context. We are currently investigating selection procedures for RBF centers [182].

### 4.8.2 A practical example

We now consider the reduction of parameter space through the selection of a suitable set of control points in a FFD case defined over a starting lattice of $6 \times 4$ control points. Our goal is the selection of $p = 6$ design parameters $\mu_1, \ldots, \mu_6$ for constructing a FFD map, among all the possible $p' = 2 \times 6 \times 4$ horizontal/vertical displacements of the control points in the lattice (we refer to Fig. 5.25 for the geometrical configuration). The application related to this geometrical configuration deals with the shape optimization of a bypass graft and is discussed in Sect. 5.4.3. The quantity of interest chosen for the parameter screening is the *energy* functional

$$S(\mathbf{v}) = \frac{\nu}{2} \int_\Omega |\nabla \mathbf{v}|^2 \, d\Omega,$$

in order to take into account the interplay between global shape deformations and the flow field across the domain. The results obtained by applying the OAT and the MR-OAT procedures are shown in Fig. 4.3-4.4. To get the reduced set of parameters, $2p'$ – respectively $R(p'+1)$ – FE approximations have been required by OAT and MR-OAT procedures, being $R = 10$ in the MR-OAT case. In both cases, horizontal deformations effects were found to be negligible compared to the ones induced by vertical displacements of control points, so that only $p'/2 = 24$ effects corresponding to vertical displacements of control points are reported.
More effective parameters were found to be rather the same in both cases, even if the MR-OAT procedure provides also information on whether any significant interaction among the parameters exists. Concerning the APOD procedure, largest total correlations $\arg\max_i \{\Delta_i^{\min}, \Delta_i^{\max}\}$ and singular values of the matrix $\mathbb{YH}$ are represented in Fig. 4.5 and decay at a similar rate.

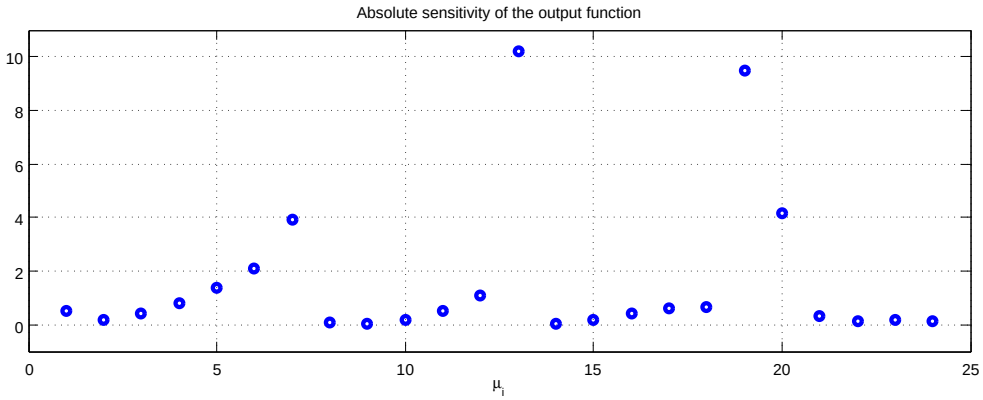



Figure 4.3: *Observed* parametric sensitivities (OAT procedure). Here only $p'/2 = 24$ effects corresponding to the (more relevant) vertical displacements of control points are represented.

FFD parametrizations originating from these selection procedures have been implemented in Sect. 5.4.3, taking into account some additional constraints related to the application. The subsets of active control points are represented in Fig. 5.25.

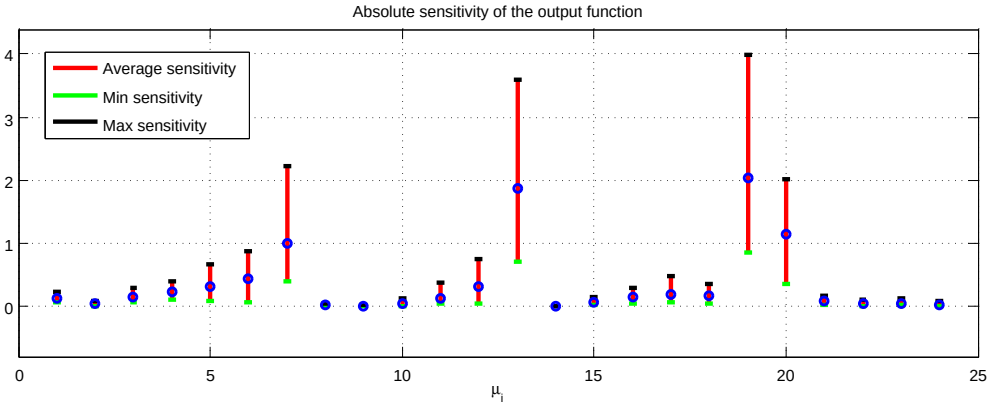



Figure 4.4: *Absolute mean effects*, minimum/maximum effects (MR-OAT procedure). Here only $p'/2 = 24$ effects corresponding to vertical displacements of control points are represented.

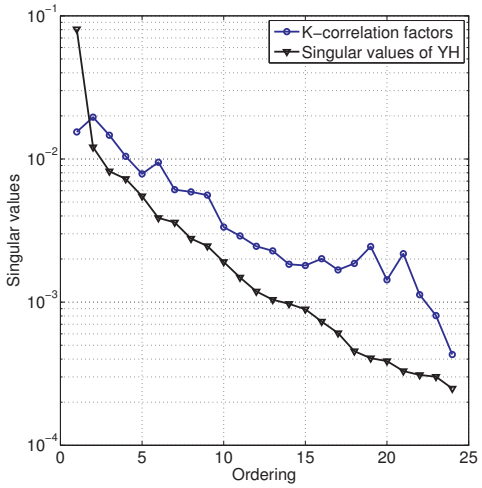



Figure 4.5: Convergence of the 24 $\mathbb{Q}$-correlation factors $\arg\max_i\{\Delta_i^{\min}, \Delta_i^{\max}\}$ for a random sample of 100 snapshots (APOD procedure).

We close this section by pointing out a further issue related with reduction of parameter space dimensions. Dealing with shape parametrizations, often the effect of a single parameter can be quite small (e.g. because of local shape deformation), but their combined effect is large. In this case a reduced $\tilde{\boldsymbol{\mu}}$-parametrization can be obtained from the starting $\boldsymbol{\mu}'$-parametrization, being $\boldsymbol{\mu}' \in \mathcal{D}' \subset \mathbb{R}^{p'}$, by means of a suitable transformation – finally, by a change of variables $\tilde{\boldsymbol{\mu}} = \mathbb{U}(\boldsymbol{\mu}')$ in the parameter space, so that $\tilde{\boldsymbol{\mu}} \in \tilde{\mathcal{D}} \subset \mathbb{R}^{\tilde{p}}$, where $\tilde{p} \ll p'$ and $\mathbb{U} \in \mathbb{R}^{\tilde{p} \times p'}$ is an orthogonal matrix.
A method recently proposed by Constantine and Wang [66] for *input subspace detection* is based on the introduction of the following correlation matrix $\mathbb{C}^{p' \times p'}$, whose components are given by

$$\mathbb{C}_{ij} = \int_{\mathcal{D}'} \frac{\partial s(\boldsymbol{\mu}')}{\partial \boldsymbol{\mu}'_j} \frac{\partial s(\boldsymbol{\mu}')}{\partial \boldsymbol{\mu}'_i} \, d\boldsymbol{\mu} \approx \frac{|\mathcal{D}'|}{R} \sum_{r=1}^{R} \delta_j(\boldsymbol{\mu}^{j,r}) \delta_i(\boldsymbol{\mu}^{i,r}), \tag{4.51}$$

where the approximation follows by (4.50). The quality of this approximation is crucial in order to represent the variations of the sensitivities over $\mathcal{D}$, and it can be controlled by the number of samples $R$. Further investigations are ongoing, in order to make this procedure more robust.

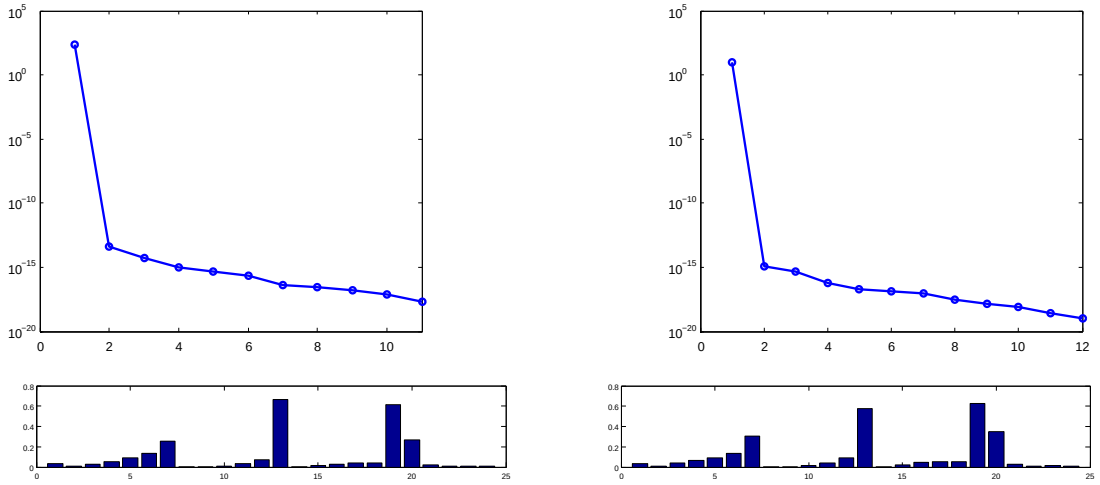


Figure 4.6: Top: eigenvalues decay for approximated correlation matrices (4.51) in the OAT and MR-OAT cases. Bottom: components of the first eigenvector in both cases.

By computing the eigenvalue decomposition $\mathbb{C} = \mathbb{Z}\Lambda\mathbb{Z}^T$, it is possible to show [66] that the $\tilde{p}$ (dominant) eigenvectors related to the $\tilde{p}$ largest eigenvalues indicate the dominant directions of output variation. Thus, a good choice is $\mathbb{U} = [\mathbf{z}_1|\dots|\mathbf{z}_{\tilde{p}}]$.

Moreover, if the decay of the eigenvalues $\lambda_1 \geq \dots \geq \lambda_{p'}$ is rapid (typically exponential) a very low-dimensional subspace may be sufficient to capture most of the variation of the output $s(\boldsymbol{\mu}')$. In Fig. 4.6 we represent the eigenvalues decay obtained by the approximation (4.51) for the OAT and MR-OAT cases (in the former case $R = 1$ and the absolute mean effects are replaced by the parametric sensitivities (4.49)). In this case retaining $\tilde{p} = 1$ eigenvectors is enough; moreover, from Fig. 4.6 it is possible to see that in this (lucky) case the components of the first eigenvector return the most effective parameters already selected (see Fig. 4.3-4.4).

# Part III Applications

# 5 Optimal design of cardiovascular prostheses

In this chapter we present the main application driving the development of the reduced framework introduced and analyzed in the previous parts. Our focus is on the reduced solution of optimal design problems related with cardiovascular prostheses, such as bypass anastomoses. We introduce the main features related to the optimal design of cardiovascular prostheses, then we give a brief description of the physiology of the cardiovascular system and the most important arterial diseases such as atherosclerosis, highlighting the strong correlation between blood flows and vessel geometrical configurations. Then, after providing a general framework for modeling the optimal design of end-to-side bypass anastomoses we recall the main results achieved in the last decades and we show how the reduced framework discussed in the previous parts can be applied in order to face this problem. We present some results concerning *(i)* a simplified version of an optimal flow control problem, *(ii)* two shape optimization problems dealing with Stokes and Navier-Stokes flows and *(iii)* a robust shape optimization problem of particular interest in this framework.

## 5.1 Cardiovascular physiology and arterial pathologies

Cardiovascular diseases are the first cause of death in developed countries and nowadays several methods and tools provided by mathematical modeling and scientific computing prove to be very helpful in improving our understanding of the physiology of the cardiovascular system, as well as of the development of pathological processes. Thanks to the results achieved by computational fluid dynamics in several engineering contexts, numerical simulation of blood flows have become widespread within the bioengineering and medical research community. This is mainly due to *(i)* the availability of increasing computational power, *(ii)* the progress in imaging and geometry extraction/reconstruction techniques [213] as well as *(iii)* the availability of more and more efficient numerical algorithms. For a general introduction to modeling and simulation of the circulatory system we refer to the work by Quarteroni, Tuveri and Veneziani [259] and to some recent books [9, 100] which collect many contributes spanning over a wide range of topics.

The driving factor behind this development is the awareness that numerical models can provide quantitative descriptions of blood behavior in important vascular districts or in vessel networks, and to explain and assess the relationships between vessels shape, haemodynamics, and a family of clinical indicators. Among the latters we mention wall shear stresses, vorticity, viscous energy dissipations: they can be correlated at various extents to the risk of failure in bypass grafting [193], or artery occlusion in presence of stenosis [241], or the one of aneurysm rupture [222]. Such kind of analyses can also help in understanding how different surgical solutions may affect blood circulation.

### 5.1.1 Arterial blood flows

The human cardiovascular system can be regarded as a transport system through which the blood delivers nutrients and oxygen to organ and muscles and removes carbon dioxide. It is composed of the heart, which pumps the blood, and the network of blood vessels that convey blood to the body and drain it from the body tissues back to the heart. This network basically consists of two separate loops, connected through the heart: the pulmonary circulation, a first loop through the lungs where blood is oxygenated, and the systemic circulation, a second loop through the rest of the body to provide oxygenated blood) (see Fig. 5.1, left). For reader's convenience, we summarize the basic elements of blood circulation and provide its mathematical description, as well as the main assumptions which lie at the bottom of the models we have considered. In this description of the human circulatory system we mainly refer to a review by Thiriet and Parker [308] and to the book by Ottesen, Olufsen and Larsen [233].

Blood is transported throughout the body within blood vessels. In the systemic circulation, it is ejected from the heart into the main arteries, it flows through a network of branching arteries of decreasing size to the arterioles and then the capillaries where it carries oxygen and nutrients to the tissues and removes carbon dioxide. Then, it is collected from the capillaries through merging venules and returns to the heart through a network of veins. In the pulmonary circulation, instead, de-oxygenated blood is pumped away from the heart to the lungs through the pulmonary artery, while oxygenated blood is carried back to the heart through the pulmonary vein. Each blood circuit, systemic and pulmonary, is thus composed of three main compartments: arteries, capillaries and veins. The primary purpose of the arterial and venous vessels is to carry blood to and from the various tissues, while the microcirculatory compartment carries out the various exchange processes. A fundamental subnetwork of the systemic circulation is the coronary circulation (see Fig. 5.1, right), which provides the blood supply to the heart muscle (the myocardium): the coronary arteries (left coronary (LCA), left anterior descending (LAD) and right coronary (RCA) arteries) deliver oxygenated blood to the myocardium, while cardiac veins remove deoxygenated blood.

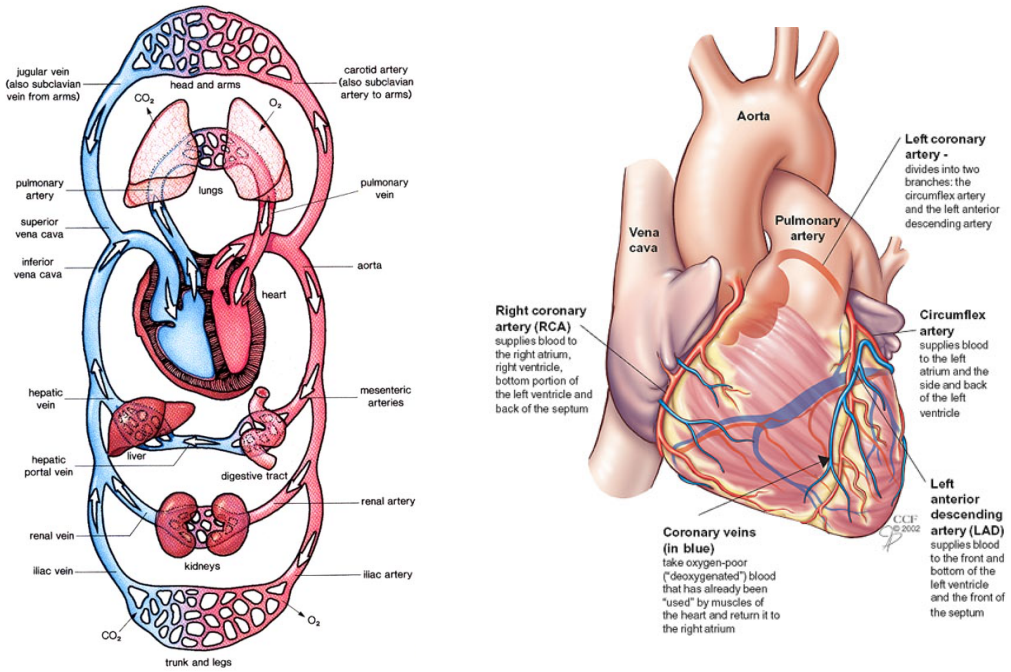



Figure 5.1: Left: systemic and the pulmonary circulation. Right: coronary circulation.

### 5.1.2 Arterial diseases

Among cardiovascular diseases, arterial diseases are meant to be pathological states for large and medium sized arteries, which can have fatal consequences on the whole cardiovascular system. For instance, heart failure – a progressive disease in which the myocardium loses the ability to adequately perfuse the body – is usually caused by chronic arterial disease, as well as by chronic valve disease. Blockages in the heart (heart attacks) or in the brain (strokes), which comprise approximately half of all deaths in developed countries, are often caused by the two most severe arterial diseases, which care atherosclerosis and aneurysms:

- Atherosclerosis is defined by the accumulation of fatty materials (such as *cholestherol*), fibrous elements and calcium between the innermost layer of an artery – called *intima* – and the *lumen*, which is the space where the blood flows. The development of the disease is very complex and there are conflicting theories about how it is initiated and progresses. In particular, the four main stages are: *(i)* the appearance of fatty streaks, *(ii)* the appearance of macrophages (foam cells), *(iii)* the development of fibrous plaque and *(iv)* calcification and the development of complicated plaques. As a result, the arterial wall first thickens as the plaque grows, and in a subsequent stage narrows, leading to partial or total occlusion. Lesions are commonly found on the inside of bends in tortuous arteries and on the outer walls of bifurcations. Atherosclerosis is a progressive disease, where the early stages are completely benign; as the lesion develops, it can intrude into the lumen of the vessel and create a stenosis which obstructs blood flow leading to clinical complications. More dangerous, clinically, is the development of thrombi or the rupture of the plaque leading to emboli which are lodged in downstream vessels causing infarctions.

- Aneurysms are gradual dilations of an arterial segment over a period of years: the aneurysm wall stretches and becomes thinner and weaker than normal arterial walls, so that untreated aneurysms can rupture causing massive haemorrhage, except in the brain where rupture leads to possibly lethal strokes. The plastic deformation of the arterial wall is associated with structural changes in the connective tissue. There are two types of aneurysms: fusiform aneurysms, cylindrical dilations where the entire circumference of the artery is weakened, and saccular aneurysms, resulting from a weakening of one side of the artery wall – for instance, as a complication of atherosclerosis.

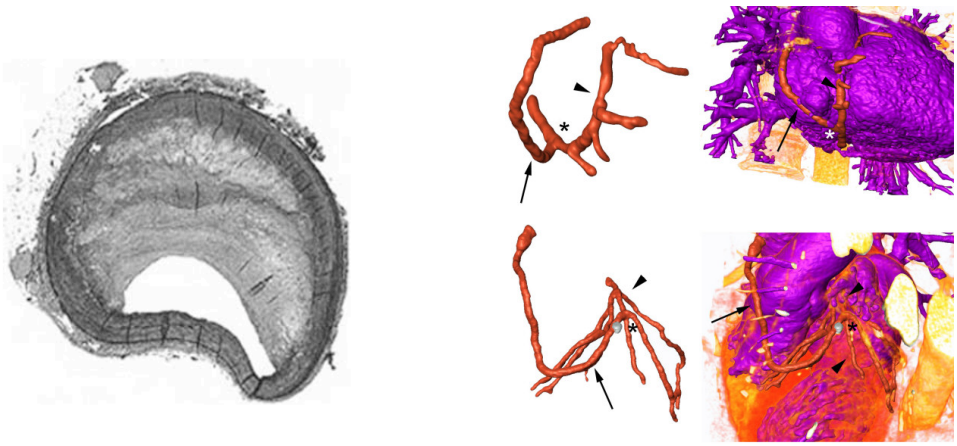


Figure 5.2: Left: cross-sectional view of a very advanced atherosclerotic plaque in a human coronary artery (picture taken from [100]). Right: possible sites of atherosclerosis development in the coronary arteries (arrowheads), aorto-coronary bypasses (arrows), anastomosis (asterisks) and their relation to the heart and aorta (picture taken from [102]).

### 5.1.3 Surgical procedures

Two main surgical procedures can be applied to severely stenosed arteries: endarterectomy, which removes the plaque and preserves the artery, and grafting[1]. There are two types of grafts: *(i)* vessel replacement and *(ii)* bypass. Replacement involves removing the diseased artery and replacing it with another vessel sutured end-to-end with the remaining arteries, while a bypass provides an alternative bridging path in order to overcome the stenoses, which are left in place. Different kinds and shapes for aorto-coronary bypass anastomoses are available, such as Miller cuffed models [98], Taylor patches, etc. [193, 136]. Consequently, different surgery procedures are used to set up a bypass; in general, the connection of the graft to the coronary artery can be done using an end-to-side or a side-to-side anastomosis. For a discussion and a comparative approach see, for example, Cole et al. [65]. Bypass can be made by organic material (e.g. the saphena vein taken from the leg or the mammary artery, either from a patient or a donor, or vessels from other species, e.g. bovine or porcine) or by prosthetic material.
The graft patency affects the long-term efficacy of coronary bypass procedures, and several post-surgical complications have to be taken into consideration. Among them, we mention intimal thickening hyperplasia (near suture lines), which is a narrowing of coronary arteries due to the excess proliferation of cells, restenosis, surgical injury, long term degeneration of the graft material or graft failure. In particular, during the first year after bypass surgery up to 15% of venous graft occlude, while by ten years after only 60% of vein grafts are patent and only 50% of patent vein grafts are free of significant stenosis [102, 235]. Moreover, every year 8% of all patients risk bypass occlusion, after ten years 80% bypasses must be replaced. In particular, atherosclerotic obstruction occurring on a foundation of neointimal hyperplasia, resulting from an accumulation of muscle cells and extracellular matrix in the intimal compartment, is the principal disease process in venous bypass grafting during the first post-surgical year and represents the foundation for later development of graft *atheroma* and *atheromatous* plaques.

### 5.1.4 Aorto-coronary and femoro-popliteal bypass grafts

In view of solving some optimal design problems related with arterial grafts, we consider two examples of geometrical configurations, related with an aorto-coronary bypass graft and a femoro-popliteal bypass graft.
Coronary arteries supply the oxygen-rich blood perfusion to the heart muscle: the lack of an adequate blood supply may cause tissue ischemia[2] and myocardial infarctions. *Coronary artery bypass graft* (CABG) surgery is a standard procedure to restore blood perfusion to the cardiac muscle by redirecting blood from the aorta through a graft vessel (either artificial or biological) to the downstream of the occluded coronary artery (see Fig. 5.3). The coronary-graft connection can be done using an end-to-side or a side-to-side anastomosis; a detailed survey of the predominant flow features of end-to-side anastomoses is provided for instance in [193].
*Femoro-popliteal bypass graft* (FPBG) surgery is used instead to bypass diseased blood vessels above or below the knee, and represents one of the most common surgical treatment of chronic lower-extremity ischemia (see Fig. 5.4). Chronic lower-extremity ischemia affects approximately 10% of patients over the age of 70. In addition to the functional limitations caused by pain with walking, several patients fear the loss of the limb as ultimate complication of ischemic conditions. Traditional infrainguinal surgical therapy for femoro-popliteal lesions involves FPBG surgery

[1] The number of vascular graft operations has declined during the past decade because of the development of stenting. Stenting is a minimally invasive procedure where a catheter is advanced from a peripheral artery to the site of stenosis. The stenosis is generally disrupted by dilating it with a balloon (angioplasty) used to expand a wire mesh stent which supports the vessel walls. The stent is left in place permanently, very often eluting drugs to limit cell migration and proliferation (intimal hyperplasia in the region of the stent) or to restore mechanical properties of the arterial tissue. Also in stenting procedures restenosis occurs at a significant rate.

[2] Ischemia is a restriction in blood supply to tissues, causing a shortage of oxygen and glucose. Ischemia is generally caused by arterial diseases, such as atherosclerosis, hypotension, thromboembolism.

(either above-knee or below-knee) using either a vein or a prosthetic bypass as alternative bridge. Revascularization surgical treatments of critically ischemic lower limbs include also *aorto-iliac* or *aorto-femoral* bypass (delivering blood from the aorta directly to the iliac arteries or to the femoral arteries, depending on the level of the blockage), *femoro-distal* bypass (in case the portion to be bypassed is longer than in the FPBG case), *aorto-iliac* angioplasty with stenting.

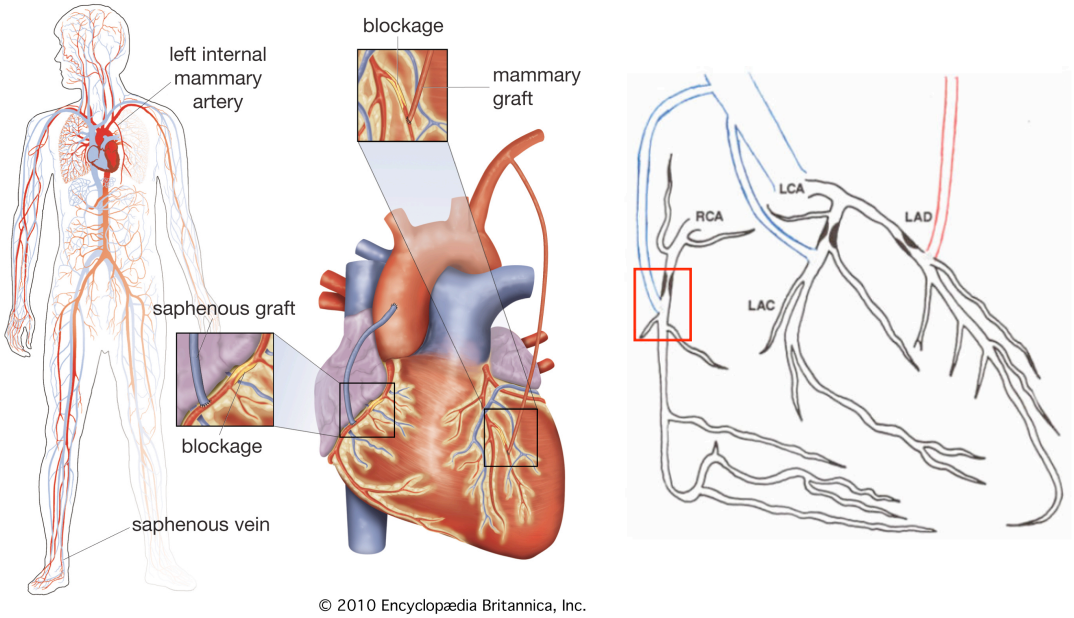



Figure 5.3: Left: double coronary artery bypass surgery, showing the grafting of a section of saphenous vein from the leg to bypass a blockage on the right side of the heart and the diversion of an internal mammary artery to bypass a blockage on the left side of the heart (picture taken from Encyclopædia Britannica, Inc.). Right: coronary tree and possible end-to-side CABG locations.

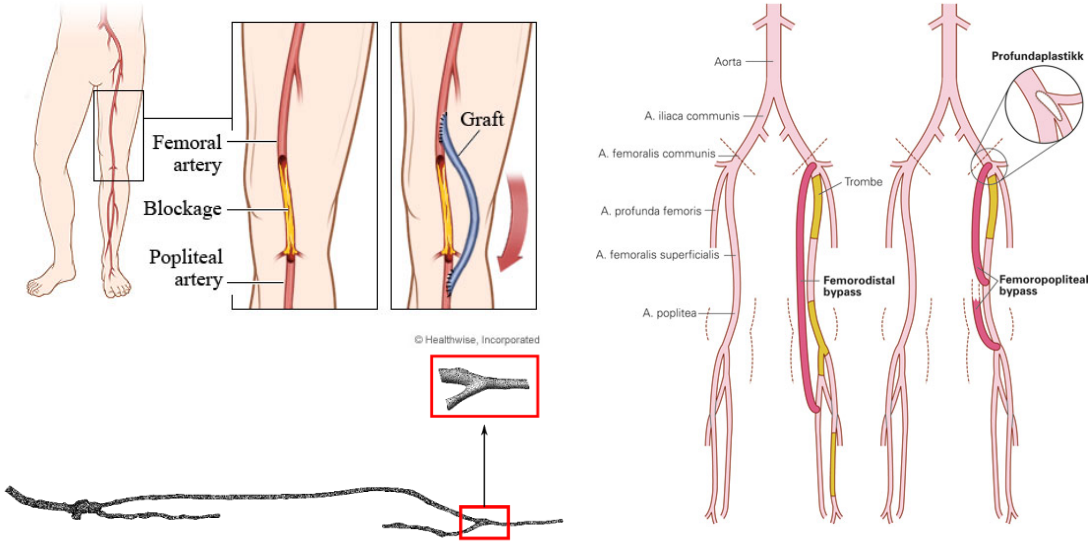



Figure 5.4: Left: femoro-popliteal bypass surgery to bypass a blockage at the level of the knee. Right: possible femoro-popliteal and femoro-distal bypass grafts.

## 5.2 Optimal design of end-to-side bypass anastomoses

The design of the end-to-side anastomosis that connects graft and host vessels is a critical factor in avoiding post-operative recurrence of the stenosis. As for natural arterial vessels, haemodynamics plays a role in the development of lesions in grafts. Factors like flow recirculation or stagnation, as well as high vorticity or dissipation regions, low and oscillatory wall shear stresses (WSS), play a driving role in the development of vascular diseases, such as atherosclerosis or intimal hyperplasia, which may lead to the growth of another stenosis downstream from the anastomosis. Hence, a better understanding of the flow patterns in bypass anastomoses may be very helpful in surgical planning of grafting and may improve the lifetime of grafts. A typical attempt to design a bypass graft is apt at minimizing some cost functionals related to these physical indices.
In this section we review the main features of mathematical modeling for haemodynamics and the most recent contributions to optimal design of bypass grafts. Then, we discuss the assumptions underlying our models and some possible choices of objectives for optimization purposes.

### 5.2.1 Main features of mathematical modelling of haemodynamics

The mathematical modelling of the cardiovascular system is a very challenging problem, due to the presence of different space and time scales, so that a hierarchy of models possibly coupled can be devised. The mathematical equations for fluid dynamics are the key component: even if blood is not a fluid but a suspension of particles in the *plasma*, in larger vessels viscoelastic and shear thinning effects can be neglected, so that a Newtonian model can be used. In smaller vessels viscoelastic effects can be very important and blood viscosity becomes dependent on the vessel radius, so that a more adequate non-Newtonian model has to be considered. In large and medium size vessels, flow is governed by Navier-Stokes equations

$$\begin{cases} \dfrac{\partial \mathbf{y}}{\partial t} - \nu \Delta \mathbf{y} + (\mathbf{y} \cdot \nabla)\mathbf{y} + \nabla p = \mathbf{f} & \mathbf{x} \in \Omega, \\ \nabla \cdot \mathbf{y} = 0 & \mathbf{x} \in \Omega, \end{cases} \tag{5.1}$$

in a domain $\Omega$ representing the lumen of the vessel under investigation, where $\mathbf{y}$ and $p$ are the velocity and the pressure of the blood, respectively. To account for the compliance of the vessel walls, we need to introduce an additional (*elastodynamics*) equation modeling the behavior of the wall displacement $\boldsymbol{\eta}$. The mechanical interaction between the blood flowing in the vessel and the vessel structure[3] is rather complex and still an open problem. Several details concerning applications in haemodynamics can be found for instance in [199, 96] and in some recent works [68, 67].

Another available methodology for *model reduction* of blood flow simulation in vessel networks can be devised by exploiting some specific features, such as the approximately cylindrical morphology of the vessels, and a multiscale paradigm. In this way, three-dimensional models for blood flows in (local) portions of large vessels have been coupled with one-dimensional (global) models for the network of arteries and veins, or even to zero-dimensional models for the capillary network. In this case, the geometrical downscaling involves time-dependent Navier-Stokes equations (3D), an Euler hyperbolic system (1D) in which the space dependence is reduced only to the axial coordinate (see e.g. [203]), and lumped parameter models based on ordinary differential equations (0D), often represented in terms of hydraulic or electric networks (see e.g. [261, 238]).

[3]We point out that the interaction between the blood flow and the arterial wall deformation has to be taken into account to correctly predict arterial flows in case of large deformations induced by pressure pulses, such as in large vessels. Nevertheless, even in the larger vessels – at least in physiological situations – the main features of the flow can be already captured by a model involving a fixed geometry. Only if more details are needed, such as a precise computation of shear stresses, then compliant models are better suited.

### 5.2.2 Recent contributions on optimal design of bypass grafts

Modelling and simulations of blood flows through arterial grafts have been carried out by many research groups in the past two decades, in order to describe flow and stress behavior, as well as to optimize surgical procedures or the design of medical devices. On the other hand, theoretical methods of optimal control and shape optimization have proved to be essential for a suitable formulation of these optimal design problems.
Many works have focused in the last decade on the optimal shape design of end-to-side anastomoses (see Fig. 5.5), typically by acting on the wall shape near the anastomosis by local shape variations. A very detailed survey on blood flows in end-to-side anastomoses is provided by Loth, Fischer and Bassiouny [193]. Early investigations on flow patterns in end-to-side anastomoses date back to the late 90s and are due to Ethier and Steinman [94, 93], Cole [65, 64], Sherwin and Peiró [114]. Parametric numerical investigation on haemodynamics in distal coronary anastomoses were presented in many works (see e.g. the more recent works [102, 323, 84]). Also concerning shape optimization techniques applied to bypass grafts design, first works date back to the late 90s [190, 98]. A non-exaustive list of contributions in the last decade include the works by Rozza, Quarteroni et al. [274, 6, 5, 254, 273], Behr [252] and Heinkenschloss [3, 4], Marsden [210, 288].

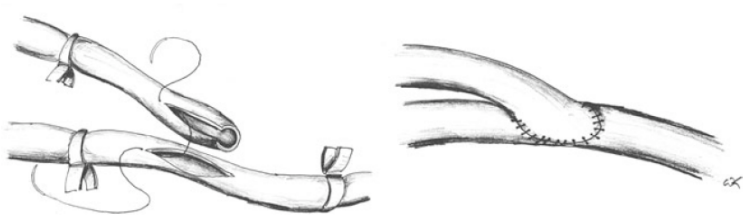


Figure 5.5: Sketch of end-to-side anastomosis construction. Figure taken from [193].

The three most significant design variables in end-to-side anastomoses are [193]: the anastomosis angle, the graft-to-host diameter ratio [164], and the toe shape (see Fig. 5.6). Also the flow split between the proximal outflow segment and the distal outflow segment affects greatly the distribution of WSS [114], as do the viscosity and the Reynolds number (different e.g. for aorto-coronary bypasses as compared with femoro-popliteal bypasses). The effect of the flow profile at or near the inlets must also be taken into account. The near-complete occlusion of stenotic arteries produces the largest (often turbulent) disturbances in the flow, and has been linked to triggering biochemical processes such as thrombosis, hemolysis, etc. While it is known that the physical unsteady and pulsatile flow can be replaced with a steady mean flow with the same Reynolds number for purposes of evaluating the mean WSS distribution [113, 94], the correct flow profile must be taken into account if accurate WSS predictions are desired[4]. Even if most of the studies have been performed in idealized geometries, there is a growing interest in computations carried out in realistic geometries determined from medical imaging [173].

From several analyses, it seems clear that in order to design a bypass graft in a robust way, we must take into account all the various sources of *uncertainty* that can effect the final optimized design. Only recently the effect of uncertainty in the design of bypass grafts has been taken into account. In [288] the bypass configuration was optimized under unsteady flow with an uncertain flow split between the occluded artery and the graft. The robust design was sought by minimizing a cost functional that measured the area of flow WSS in the downstream region of the anastomosis. To make the design robust, the authors added a penalty term for the standard deviation of the output due to input uncertainties. The cost of such an optimization method was reported as quite high, 11 days in the fully 3D unsteady case on a $18 \times 4$ cores parallel cluster.

[4]For example, in [326] the effect of cardiac motion on the flow in a coronary artery was studied: it was shown that the motion-induced change in the velocity profile could impact the WSS values by up to 150%.

In Sect. 5.3 we consider a similar problem (presented in [180]) for steady two-dimensional flows, but optimizing the whole shape rather than just the angle. In this respect, the computational cost has been sensibly diminished by introducing a reduced order model (ROM) for the fluid equation based on RB methods, making the robust design problem computationally feasible. In [252] the bypass shape was obtained by minimizing the total shear rate, and the sensitivity of the optimal shape with respect to the uncertain viscosity in a non-Newtonian rheology was considered. There was no attempt made to find a robust optimal shape over a range of viscosity values, likely due to prohibitive cost of running the full-fidelity three-dimensional FE simulations.

### 5.2.3 New contributions on optimal design of bypass grafts

Recent works already indicate that in presence of uncertainty effects the bypass design problem is not yet satisfactorily solved by existing classical computational approaches, and, furthermore, that some type of ROM is needed to reduce the computational cost. This Thesis aims at exploiting computational and geometrical reduction techniques presented in Parts I and II in order to improve the computational performances in the solution of bypass design problems. In particular, we aim at extending the framework illustrated in the first works by Rozza and Quarteroni [254, 274, 273], where the RB framework was applied for the first time to this problem. We thus aim at addressing the following features:

- more complex geometrical description of geometrical configurations through global nonaffine mappings based on FFD techniques, which allow the optimization of the whole shape rather than the control of basic geometrical parameters, such as angles or diameters;
- uncertainty elements featuring the nature of the residual flow in the partially occluded arterial branch, and the solution of corresponding robust optimization problems;
- a first temptative comparison between two-dimensional and three-dimensional flows.

We present two different paradigms based on optimal control and shape optimization, highlighting key points and difficulties. A first boundary control formulation has been considered (Sect. 5.3), in order to evaluate the impact of basic geometrical features on the flow pattern developed within a simplified configuration. Then, different shape optimization problems are presented (Sect. 5.4), dealing with geometrical configurations representing both aorto-coronary bypass grafts and femoro-popliteal bypass grafts. In the former case, both a simplified Stokes model and a more realistic Navier-Stokes model are considered, while in the latter we only deal with a Navier-Stokes model. In all these cases, fluid flows are simulated by using the RB approximation of Stokes or Navier-Stokes equations. Another new contribution of this Thesis work is aimed at inserting some uncertainty elements (featuring the nature of the residual flow in the partially occluded arterial branch) in both optimal control and shape optimization problems, solved within a suitable reduced framework, in view of simultaneous computational and geometrical reduction. We also test whether our simplifications affect the robust design obtained with or without the reduction to a pure boundary control problem, as well as comparing results between two-dimensional and three-dimensional flows. Results of the following sections are based on three works carried out in collaboration with Quarteroni, Rozza [206] and also with Lassila [180, 181].

### 5.2.4 Our models for optimal design of bypass grafts

A realistic model for blood circulation should take into account *(i)* the flow unsteadiness, *(ii)* the arterial wall deformability, described by suitable structural models [145] and possibly *(iii)* complex rheological model to characterize the aggregate nature of the blood [271] especially in small vessels and capillaries. In view of studying optimal control and shape optimization problems, which entail the repeated simulation of these flow equations (and the evaluation of

the cost functional to be minimized), we cannot afford the solution of PDE models involving such complex features – computational costs would be too prohibitive. For these reasons, in the applications discussed throughout this Part, we model blood flows adopting steady incompressible Navier-Stokes ($\delta = 1$) or Stokes ($\delta = 0$) equations in case of smaller vessels such as coronary arteries and lower blood velocities, with the velocity $\mathbf{y}$ and the pressure $p$ satisfying

$$\begin{cases} -\nu\Delta\mathbf{v} + \delta(\mathbf{v}\cdot\nabla)\mathbf{v} + \nabla p = \mathbf{0} & \text{in } \Omega \\ \nabla\cdot\mathbf{v} = 0 & \text{in } \Omega \\ \mathbf{v} = \mathbf{u}_{\mathrm{c}} & \text{on } \Gamma_{\mathrm{c}} \\ \mathbf{v} = \mathbf{v}_{\mathrm{in}} & \text{on } \Gamma_{\mathrm{in}} \\ \mathbf{v} = \mathbf{0} & \text{on } \Gamma_{\mathrm{w}} \\ -p\mathbf{n} + \nu\dfrac{\partial\mathbf{v}}{\partial\mathbf{n}} = \mathbf{0} & \text{on } \Gamma_{\mathrm{out}}. \end{cases} \tag{5.2}$$

Here $\Omega \subset \mathbb{R}^d$ for $d = 2, 3$ represents an end-to-side anastomosis (see Fig. 5.6), $\Gamma_{\mathrm{in}}$ the inlet from the stenosed section of the artery, $\Gamma_{\mathrm{c}}$ the bypass inlet and $\Gamma_{\mathrm{w}}$ the artery wall, where we prescribe two inflow profiles $\mathbf{u}_{\mathrm{in}}$ and $\mathbf{u}_{\mathrm{c}}$ and a no-slip boundary condition, respectively. We assume a free-stress boundary condition at the outlet $\Gamma_{\mathrm{out}} \neq \emptyset$ and no volumetric forces along the domain. Concerning physical parameters, we denote by $\mu = 0.04\ \mathrm{g\,cm^{-1}\,s^{-1}}$ the blood dynamic viscosity, by $\rho = 1\ \mathrm{g\,cm^{-3}}$ the blood density, thus yielding a kinematic viscosity $\nu = \mu/\rho = 0.04\ \mathrm{cm^2\,s^{-1}}$ and a Reynolds number $Re = \tilde{v}D/\nu$ of order 100.

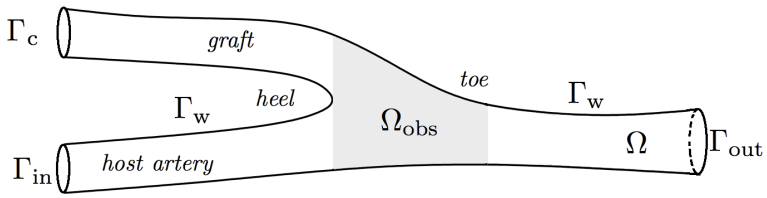



Figure 5.6: A schematic view of a bypass graft anastomosis.

Moreover, in order to have a physically meaningful problem, we enforce the total conservation of fluxes between the (partially or totally) occluded branch $\Gamma_{in}$ and the graft inlet $\Gamma_{\mathrm{c}}$, according to

$$Q_{\mathrm{in}} + Q_{\mathrm{c}} = \int_{\Gamma_{\mathrm{in}}} \mathbf{v}_{\mathrm{in}}\cdot\mathbf{n}\,d\Gamma + \int_{\Gamma_{\mathrm{c}}} \mathbf{v}_{\mathrm{c}}\cdot\mathbf{n}\,d\Gamma = \int_{\Gamma_{\mathrm{in}}} \mathbf{v}_{\mathrm{in}}\cdot\mathbf{n}\,d\Gamma =: Q_{\mathrm{tot}}\ \text{(constant)}. \tag{5.3}$$

To pose the optimal design problem, several cost functionals depending on the state solution $(\mathbf{v}, p)$ have been proposed in literature for the optimization of arterial bypass grafts or otherwise regularization of flows where recirculation and vortex generation are to be minimized. By denoting $\Omega_{\mathrm{obs}}$ the subdomain where physical indices of interest are observed, we list some typical choices together with references to previous works where such functionals have been employed:

1. *viscous energy dissipation* [180, 252]:

$$J_1(\mathbf{v}) := \frac{\nu}{2}\int_{\Omega_{\mathrm{obs}}} |\nabla\mathbf{v}|^2\,d\Omega \quad \text{or} \quad J_1(\mathbf{v}) := \frac{\nu}{2}\int_{\Omega_{\mathrm{obs}}} \varepsilon(\mathbf{v}) : \varepsilon(\mathbf{v})\,d\Omega,$$

   where $\varepsilon(\mathbf{v}) = (\nabla\mathbf{v} + \nabla\mathbf{v}^T)/2$ is the Cauchy strain tensor;

2. *Stokes-tracking type functional* [163, 142, 180]:

$$J_2(\mathbf{v}) := \int_{\Omega_{\mathrm{obs}}} |\mathbf{v} - \mathbf{v}_{\mathrm{Stokes}}|^2\,d\Omega,$$

   where $\mathbf{v}_{\mathrm{Stokes}}$ is the solution of (1) obtained after neglecting the term $c(\mathbf{v}, \mathbf{v}, \mathbf{z})$;

3. *vorticity* [6, 30, 163, 180, 273]:

$$J_3(\mathbf{v}) := \frac{\nu}{2} \int_{\Omega_{\text{obs}}} |\nabla \times \mathbf{v}|^2 \, d\Omega,$$

4. *Galilean invariant* vortex measure for two-dimensional flows [142, 163, 180]:

$$J_4(\mathbf{v}) := \int_{\Omega_{\text{obs}}} \max\{\det(\nabla \mathbf{v}), 0\} \, d\Omega \quad \text{or} \quad J_4(\mathbf{v}) := \int_{\Omega_{\text{obs}}} g(\det(\nabla \mathbf{v})) \, d\Omega,$$

where $g(z)$ is a smooth nonnegative function satisfying $g(z) = 0$ for $z \leq 0$ and $g(z) = \mathcal{O}(z)$ as $z \to \infty$ [163]. In fact, vortex cores are related to regions where the eigenvalues of $\nabla \mathbf{v}$ are complex, and in the two-dimensional case this is equivalent to $\det(\nabla \mathbf{v}) > 0$.

5. *wall shear stress gradient* [190]:

$$J_5(\mathbf{v}) := \int_{\Gamma_{\text{obs}}} \nabla \left( \nu \frac{\partial \mathbf{v}}{\partial \mathbf{n}} \cdot \mathbf{t} \right) \cdot \mathbf{t} \, d\Omega, \tag{5.4}$$

where $\mathbf{n}$ and $\mathbf{t}$ are the unit normal and tangent vectors respectively.

All of the above functionals are bounded in $[H^2(\Omega)]^d$, and the functionals $J_1$-$J_4$ are bounded in $[H^1(\Omega)]^d$. The energy functionals $J_1$ are analytically the simplest to handle. They are coercive and weakly coercive owing to the Poincaré and Korn inequalities, respectively[5]. The tracking functional $J_2$ is suitable only for low-Reynolds flows with negligible convective effects. The vorticity functional $J_3$ is the most common choice, but it has the problem that strong shear boundary layers can have a disproportionate weight compared to the vortices. The functional $J_4$ is not differentiable and needs to be regularized to make it regular enough to use the standard optimal control framework. The functional $J_5$ contains second-order derivatives of velocity evaluated on the boundary, which makes its computation from finite elements approximations difficult. However, it would be possible – through a bounded residual functional, exploiting the weak formulation of the state problem – to obtain good approximations (as well as error bounds) for the wall shear stress integrated over some segments of the boundary (or alternatively, (say) wall shear stress Legendre coefficients). These functionals might serve the ultimate cardiovascular intentions better than some of the other functionals listed above. For this reasons, further investigations are ongoing at the moment, in view of this goal. Based on these considerations, we concentrate in the numerical examples on three cost functionals: the viscous energy dissipation $J_1$, the vorticity $J_3$, and the vortex measure $J_4$.

## 5.3 Optimal control for end-to-side bypass grafts design

### 5.3.1 Formulation and numerical approximation

A first possible approach for the optimal design of bypass grafts is based on the solution of the *deterministic* optimal control problem (DD-OC) analyzed in Sect. 4.1.2, in the vein of [30, 71, 130, 142, 174], for which the control function is the Dirichlet boundary condition $\mathbf{u}_c$ representing the flow entering into the artery from the graft on the boundary[6] $\Gamma_c$, and the residual flow through the occluded artery $\mathbf{v}_{\text{in}}$ is known. Thus the geometrical properties of the bypass graft are only represented by the velocity profile $\mathbf{u}_c \in \mathcal{U}_c$ imposed at the anastomosis. This entails

[5] Coercivity holds at least if $\Omega_{\text{obs}} = \Omega$. If $\Omega_{\text{obs}} \subsetneq \Omega$, as is usually the in practice as we want to focus the reduction to a subregion near the anastomosis, we do not have analytical results but typically the convexity and coercivity of the cost functional is preserved as we shall see in the results section.

[6] With respect to Sect. 5.2.4, here $\Gamma_c$ is a (fictitious) boundary representing the anastomosis (see Fig. 5.7).

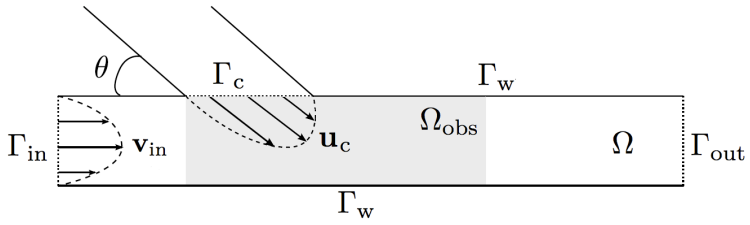



Figure 5.7: Domain and boundary segments for the optimal boundary control formulation

the solution of an OC problem on a *frozen*, fixed domain – representing the occluded artery – on which the state equations have to be solved repeatedly. For simplicity we refer also to problems following this formulation as design problems, even if they only involve boundary control.

On the other hand, assuming that the patency of the occluded artery (as well as the residual flow $\mathbf{v}_{\text{in}}$) is uncertain, we consider the following *worst-case* design problem: *find the bypass control function* $\mathbf{u}_c$ *in such a way that it minimizes the worst-case value of* $\mathcal{J}$ *over all admissible values of the residual flow function* $\mathbf{v}_{in}$, leading to the solution of the *robust design problem* (RD-OC). In order to work it out and find a suitable representation of the residual flow $\mathbf{v}_{\text{in}}$, we have introduced the *complementary uncertainty problem* (CU) – which provides somehow a dual perspective with respect to the deterministic OC problem (DD-OC): given a known flow profile across the anastomosis $\mathbf{u}_{\text{c}}$, find the residual flow function $\hat{\mathbf{v}}_{\text{in}}$ which maximizes $\mathcal{J}$. A suitable, parametrized version of this flow function is then considered as family of residual flows in the robust design problem (RD-OC), whose aim is the identification of the worst case among the family.

We now present some numerical results about the optimal design of aorto-coronary bypass grafts based on the solution of the optimal control problems recalled above, and analyzed in Sect. 4.1.2. We consider throughout this section a simplified 2D bypass configuration $\Omega = (0,5)\times(0,1)$, where $\Gamma_{\text{in}} = \{(x_1,x_2) : x_1 = 0,\, x_2 \in (0,1)\}$ and $\Gamma_{\text{c}} = \{(x_1,x_2) : x_1 \in (1,3/2),\, x_2 = 1\}$, respectively, thus considering the graft-to-host diameter ratio to be fixed at its (near-)optimal value 1.5 as discussed in [164]. In order to exploit the reduced framework discussed in the previous section, we make the simplifying assumption that the control functions are parametrized with respect to the anastomosis angle $\theta$ and are given by the following parabolic *Poiseuille* profiles for simplicity:

$$\mathbf{u}_{\text{c}}(x;\theta,\omega) := \frac{16}{9}\left(\frac{7}{6}-\frac{\omega}{6}\right)\Big(x-1\Big)\left(\frac{3}{2}-x\right)\begin{bmatrix}(\tan\theta)^{-1}\\ -1\end{bmatrix}, \qquad \theta\in(\theta_{\min},\theta_{\max}), \tag{5.5}$$

where $0 < \theta_{\min} < \theta_{\max} \le \pi/2$ . In this way, the set of admissible[7] boundary controls is defined by $\mathcal{U}_{\text{c}} := \{\mathbf{u}_{\text{c}}(x;\theta,\omega) :\ \theta\in[\theta_{\min},\theta_{\max}]\}$, being in our case $\theta_{\min} = 15^\circ$ and $\theta_{\max} = 85^\circ$. On the other hand, $\omega \in (0,4)$ is the variable controlling the flux split between the graft and the host vessel: a complete occlusion corresponds to $\omega = 0$ while a 50/50 split of total flux between graft and host artery is obtained for $\omega = 4$. The rescaling factor $\tan(\theta)$ in $\mathbf{u}_{\text{c}}$ is considered in order to fulfill (5.3). For the sake of the parametrized PDE to solve, here $\boldsymbol{\mu} = (\theta,\omega) \in (15^\circ, 85^\circ)\times(0,3)$. Thus, the control variable in the simplified deterministic design problem (DD-OC) is reduced to the angle $\theta$ of the bypass graft. Also the radius of the bypass could be taken as an optimization variable, in the case that this is under control of the surgeon performing the operation, but in general more complex geometrical properties such as cuff shapes cannot be incorporated into our model problem. We are interested in the minimization of the cost functionals $J_1 - J_4$ in the downfield subregion $\Omega_{\text{obs}}$ where a vortex may occur, leading to possible occlusions after grafting and plaque formation; here $\Omega_{\text{obs}} = \{(x_1,x_2)\in\Omega : (x_1,x_2)\in(1,4)\times(0,1)\}$ (see Figs. 5.6-5.7).

[7]Concerning the existence results provided in Sect. 4.1.2, it is clear that the admissible set $\mathcal{U}_{\text{c}}$ given by the control functions (5.5) is closed in $H^{1/2}(\partial\Omega)$. To see that it is also convex, note that $\varphi(\theta) = 1/\tan\theta$ is a continuous function in $\theta\in[\theta_{\min},\theta_{\max}]$ and its image is a closed interval, therefore convex.

The resulting problem is discretized with $\mathcal{N}_V = 14,260$ and $\mathcal{N}_Q = 1,827$ dofs for velocity and pressure, respectively, using $\mathbb{P}_2/\mathbb{P}_1$ finite elements; the dimension of the computed reduced basis space is $3N = 54$, thus yielding the possibility to solve a Navier-Stokes problem in a *real time* way (averaged time over 1,000 evaluations is 0.19 seconds).

### 5.3.2 Results: deterministic design optimal control problem

As a first approximation we assume that the residual flow profile is parabolic, in particular it is defined as $\mathbf{v}_{\text{in}}(y;\omega) := \omega y(1-y)\mathbf{e}_1$. This is a typical choice in numerical simulations of arteries yet its justification or effect on the outcomes seems to be rarely considered. We have solved the (DD-OC) problem for 7 different values of $\omega \in [0,3]$ and for 6 different values of the Reynolds number $Re \in (15,90)$ considered as a further input parameter in the formulation (5.2); nevertheless, for the sake of simplicity we report here the results for the maximum value experimented, Re = 90. Within the reduced framework illustrated, the solution of 42 optimal control problems takes about 3 hours of CPU time,[8] each of them implying about $10 \div 15$ iterations of the optimization procedure.
The optimal angle $\theta^*$ obtained by solving the problem (DD-OC) decreases as the magnitude of the residual flow increases. The specific behavior and values of the four cost functionals varies, however, leading to different ranges for $\theta^* = \theta^*(\omega)$. In particular, the vorticity functional $J_3$ and the Stokes tracking functional $J_2$ (see Fig. 5.8) exhibit a stronger convexity and lead to smaller values of the optimal angles: $\theta^* \in (29°, 43°)$ for $J_3$ and $\theta^* \in (27°, 30°)$ for $J_2$, respectively, which are very close to values usually treated as optimal for a graft anastomosis [193, 288]. On the other hand, minimization of functionals $J_1$ and $J_4$ yields larger values of the optimal angles, perhaps due to their weaker convexity. In Fig. 5.9 the flows corresponding to the optimal angles for the functional $J_3$ and $\omega = 0$, $\omega = 1$ are represented. We point out that in the case of total occlusion the main vortex core in the heel region can never be totally eliminated.

### 5.3.3 Results: complementary uncertainty optimal control problem

Since the residual flow profile in the (partially) occluded artery might play an important role in the fluid dynamics of a bypass model, instead of using the parabolic profile $\mathbf{v}_{\text{in}}$ we are interested in finding the profile of the worst residual flow so that the optimized graft is *robust* not just to the magnitude of this flow, but also to its profile. To this aim, we solve a relaxed version of the (CU) problem,[9] by considering parametrized control functions $\mathbf{v}_{\text{in}}$ under the form

$$\mathbf{v}_{\text{in}}(y;\boldsymbol{\pi}) = \sum_{i=1}^{6} \pi_i \boldsymbol{\phi}_i(y)\mathbf{e}_1$$

where $\boldsymbol{\phi}_1$ is the parabolic profile already introduced, $\boldsymbol{\phi}_2 = \exp(-100(y-0.5)^2)$, $\boldsymbol{\phi}_3 = \exp(-100(y-0.25)^2)$, $\boldsymbol{\phi}_4 = \exp(-100(y-0.75)^2)$ are three gaussian profiles centered at the points 0.25, 0.5 and 0.75, and $\boldsymbol{\phi}_5 = y(1-y)(y-0.25)$, $\boldsymbol{\phi}_6 = y(1-y)(y-0.75)$ are two cubic profiles; the control parameters $\{\boldsymbol{\pi}_i\}_{i=1}^6$ are such that the flux of $\mathbf{v}_{\text{in}}$ is constant. By solving the relaxed (CU) problem, we find that the worst case corresponds to the gaussian profile centered at the midpoint of the occluded section $y = 0.5$, thus corresponding to a severe occlusion in the host artery. The function $\mathbf{v}_{\text{in}} = \omega \exp(-100y^2)\mathbf{e}_1$ will be the boundary condition on $\Gamma_{\text{in}}$ from now on.

[8] Computations involving RB approximations have been executed on a personal computer with $2 \times 2$GHz Dual Core AMD Opteron (tm) processors 2214 HE and 16 GB of RAM.

[9] In case a near-complete occlusion is not expected, the admissible set of residual flow functions can be regularized to rule out extreme singular cases. In this case we can use the knowledge about the solutions of the (CU) to design a reasonable set of admissible residual flow functions that still contain some level of uncertainty while being mathematically better behaved. See also remark 4.6.

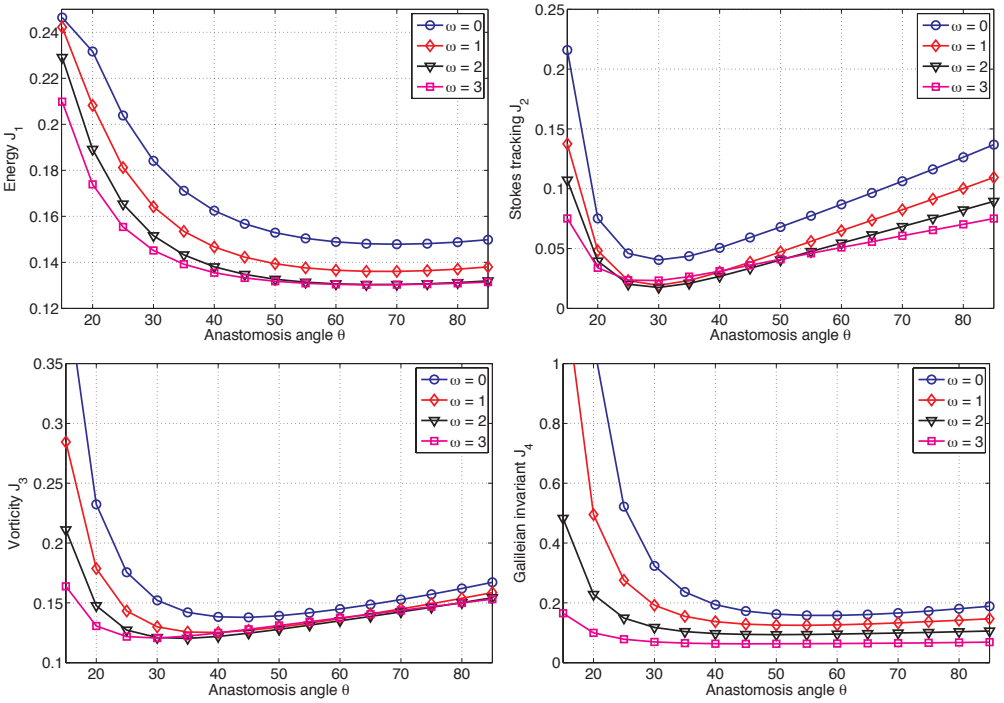



Figure 5.8: Functionals $J_1, J_2$, $J_3$, and $J_4$ in the subdomain as a function of the anastomosis angle $\theta$ for different values of the residual parameter $\omega$ with parabolic residual flow.

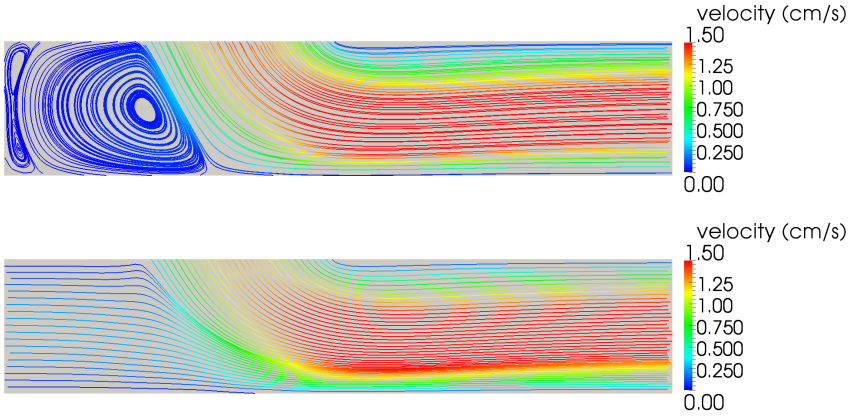



Figure 5.9: Flow in the optimal configuration for the cost functional $J_3$, $\omega = 0$ (top) and $\omega = 1$ (bottom), with parabolic residual flow.

### 5.3.4 Results: robust design optimal control problem

We can now solve the robust design (RD-OC) problem by considering the same setting as in the (DD-OC) case and the residual flow $\mathbf{v}_{\text{in}}$ given by the gaussian profile obtained by solving the complementary uncertainty (CU) problem. First of all, we consider the deterministic design (DD-OC) problem with a gaussian residual flow: the results, concerning the behavior of the vorticity functional $J_3$ w.r.t. $\theta$ and $\omega$, as well as the optimal angles obtained with the four functionals $J_1 - J_4$, are reported in Fig. 5.10 for both the parabolic and gaussian residual inflows. The results follow the same trends in both cases, even if flow patterns are remarkably different if the residual flow profile changes. In particular, the gaussian profile induces two secondary vortices near the occlusion, and a more complex vorticity pattern in the anastomosis region.

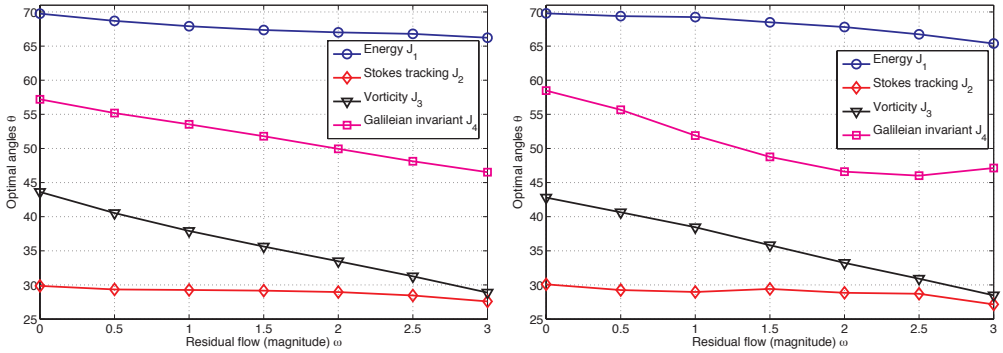



Figure 5.10: Optimal anastomosis angles as a function of the residual parameter $\omega$ for the functionals $J_1 - J_4$, for the parabolic residual flow (left) and the gaussian residual flow (right).

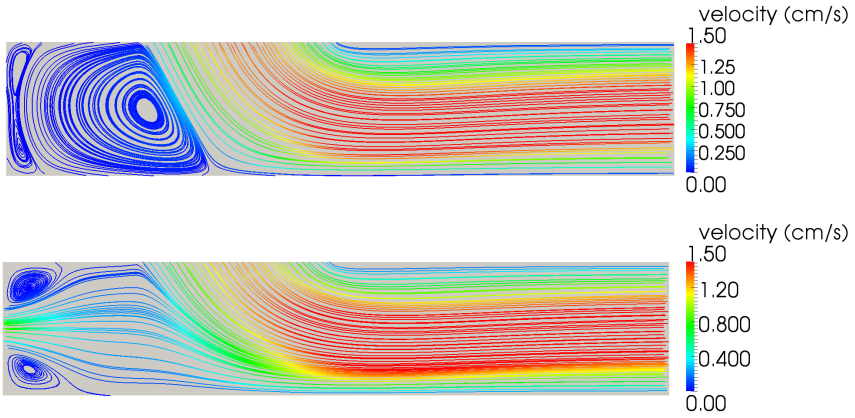



Figure 5.11: Flow in the optimal configuration for the cost functional $J_3$, $\omega = 0$ (top) and $\omega = 1$ (bottom), with gaussian residual flow.

Two major sources of vorticity can be observed from the streamlines. First, the primary vortex behind the incoming jet, which is generated by the interaction between the fast and slow flows coming into contact at the anastomosis exit. This vortex tends to disappear as we increase $\theta$ and/or $\omega$. Secondary vortices are generated by the peak residual flow on both sides of the entry from the occluded branch. However, due to the choice of the observation subdomain $\Omega_{\text{obs}}$ (considering only the flow downstream from the anastomosis) their effect is removed from the vorticity measure: this explains why the results, in terms of optimal angles, are very similar. In Fig. 5.11 the flows corresponding to the optimal angles for the functional $J_3$ and $\omega = 0, 1$ are represented.
Finally, the robust design problem (RD-OC) has been solved for the four cost functionals $J_1 - J_4$, providing the results listed in Table 5.1 (dealing with the most significant $J_3$ and $J_2$ cases). Each of these four problems takes approximately $500 \div 700$ seconds to be solved, requiring about $110 \div 150$ input/output evaluations, depending on each case. The robust angles are about the same as the ones obtained as solutions of the deterministic design (DD-OC) problem in the case $\omega = 0$. Hence, the most challenging situation for the minimization of vorticity appears to be the case of total or near-total occlusion of the stenosed branch.

### 5.3.5 Results: comparison with three-dimensional steady flows

The three-dimensional effects in a steady flow through a bypass anastomosis were considered in [93] and found to be highly significant when it comes to the WSS distribution, especially at

| functional | dimension | profile | $\omega = 0$ | $\omega = 1$ | $\omega = 2$ | $\omega = 3$ | robust solution |
|---|---|---|---|---|---|---|---|
| $J_3$ | 2D | $\mathbf{v}_{\text{in}}$ parabolic | 43.6 | 37.9 | 33.4 | 28.9 | - |
| | 2D | $\mathbf{v}_{\text{in}}$ gaussian | 42.8 | 38.5 | 33.2 | 28.5 | 42.6 |
| $J_2$ | 2D | $\mathbf{v}_{\text{in}}$ parabolic | 29.9 | 29.2 | 28.9 | 27.5 | - |
| | 2D | $\mathbf{v}_{\text{in}}$ gaussian | 30.1 | 28.9 | 28.8 | 27.1 | 30.0 |
| $J_3$ | 3D | $\mathbf{v}_{\text{in}}$ gaussian | 45.8 | 43.8 | 41.4 | 38.0 | - |

Table 5.1: Optimal angles and robust angles $\theta^*$ obtained through the (DD-OC) and (RD-OC) problems, vorticity functional $J_3$ and Stokes tracking functional $J_2$, both in 2D and 3D.

higher Reynolds numbers. To test the relevance of 3D effects on the optimal anastomosis angle in our simplified setup, we consider a 3D problem which is assumed plane symmetric along the centerline of the vessel – thus only the half-width of the configuration needs to be meshed.
The length and radius of the channel and the bypass are kept the same as in the 2D case, as well as the inflow profiles, which are chosen to be radially symmetric:

$$\mathbf{v}_{\text{in}}(y,z;\omega) := \omega \exp\left[-100\left(y^2+z^2\right)\right],$$
$$\mathbf{u}_{\text{c}}(x,z;\theta,\omega) := \left(\frac{7}{6}-\frac{\omega}{6}\right)\left[1-\frac{16}{9}\left(x-\frac{7}{4}\right)^2+4z^2\right]\begin{bmatrix}(\tan\theta)^{-1}\\ -1\end{bmatrix}.$$

We also choose the viscosity in such a way that the Reynolds number is comparable to the highest possible one used in the 2D case, i.e. $\text{Re} \approx 80$. This is obtained correspondingly to $\nu = 0.0125\, cm^2\, s^{-1}$. The resulting problem is discretized with $\mathcal{N}_V = 196,041$ and $\mathcal{N}_Q = 65,347$ dofs for velocity and pressure, respectively, using $\mathbb{P}_1/\mathbb{P}_1$ finite elements with an interior penalty stabilization scheme [55]. The nonlinear problem (5.2) is solved starting from the steady Stokes solution and performing pseudo-time stepping until convergence to a steady solution has been achieved. No model reduction was applied in this case and as a result each solution took roughly 20 minutes on 24 parallel 2.66 GHz cores of an Intel Xeon Nehalem cluster.
In Fig. 5.12 we display the obtained value of the vorticity functional $J_3$ for different values of the parameter $\omega$ (other functionals are less meaningful in 3D and are omitted here). Optimal angle is located around 45°, but as $\omega$ increases the optimal angle $\theta^*$ decreases. The qualitative behavior of the vorticity functional $J_3$ resembles that of the 2D case even if the 3D flow exhibits much more complex flow phenomena, which we will attempt to explain in the following.

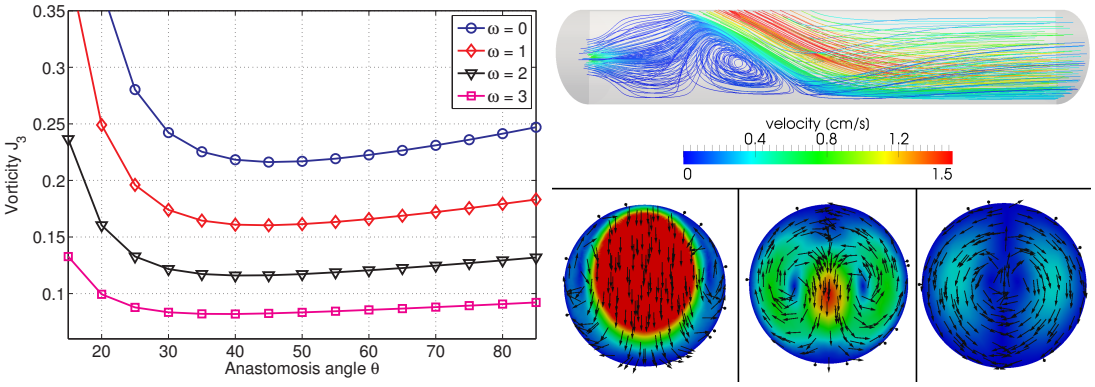



Figure 5.12: Vorticity $J_3$ in the subdomain $\Omega_{\text{obs}}$ as a function of the anastomosis angle $\theta$ for different values of the residual parameter $\omega$ in the 3D bypass case (left). Streamlines of steady flow and transversal velocities at $x = 2.5$, $x = 3.5$, and $x = 4.5$ for the case $\theta = 25°$, $\omega = 1$ (right).

Visualizations of the 3D flow field are reported in Figs. 5.12, 5.13 and 5.14. Three major sources of vortices can be observed – primary and secondary vortices already remarked for the 2D case, and a tertiary vortex structure.
As before, the primary vortex tends to disappear as we increase $\theta$ and/or $\omega$ (see Fig. 5.13); of course, in the case of total occlusion, $\omega = 0$, the primary vortex can never be totally eliminated. This fact is demonstrated in Fig. 5.14, where we display the flow at the maximum angle $\theta = 85°$ but with two different residual flows, $\omega = 0$ (total occlusion) an $\omega = 1$ (strong occlusion). Only in the first case the primary vortex is observed. Secondary vortices are generated as in the 2D case by the peak residual flow on both sides of the entry from the occluded branch. Tertiary transversal vortices, the so-called *Dean* vortices, appear downstream of the anastomosis at moderate Reynolds numbers [85, 80]. While these structures appear exclusively in the 3D flow, it seems their effect on the vorticity functional is an order of magnitude less when compared to the primary vortex, and thus they do not alter the conclusions we obtained earlier based on 2D simulations. The vorticity functional $J_3$ therefore measures and attempts to control mainly the primary vortex.

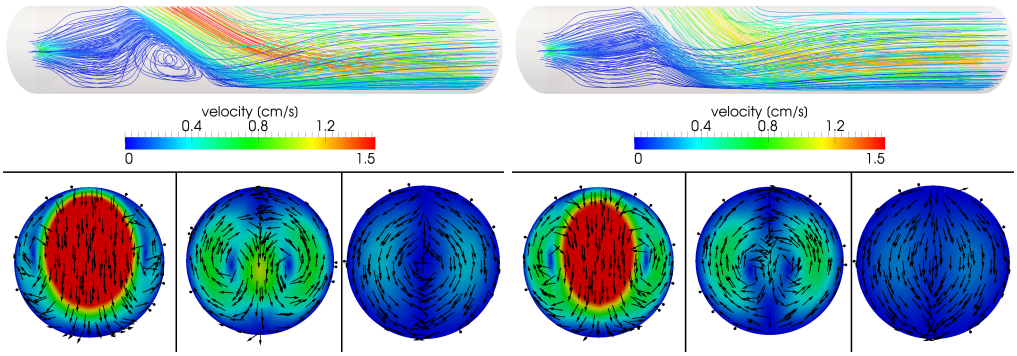



Figure 5.13: Streamlines of the steady flow and transversal velocities at $x = 2.5$, $x = 3.5$, and $x = 4.5$ for the case $\theta = 30°$ (left) and $\theta = 50°$ (right) with $\omega = 1$. For sufficiently large angles $\theta$ the primary vortex disappears, while the secondary and tertiary vortices remain.

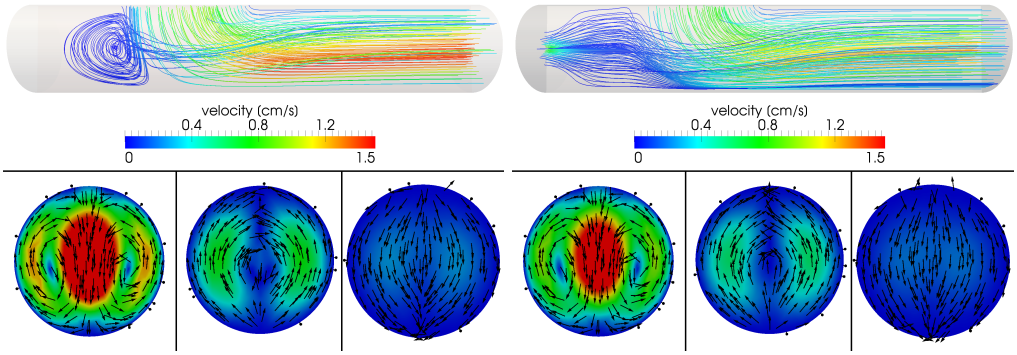



Figure 5.14: Streamlines of the steady flow and transversal velocities at $x = 2.5$, $x = 3.5$, and $x = 4.5$ for the case $\theta = 85°$ with $\omega = 0$ (left) and $\omega = 1$ (right). In the case of total occlusion, the primary vortex is present even with very large angles.

We can remark a strong similarity on the primary/secondary vortex structures between the 2D and the 3D case – see Fig. 5.15, obtained for the same values of $\theta$ and $\omega$ already considered in Fig. 5.12–5.14. For $\omega \leq 1$ a very strong primary vortex is generated for small angles $\theta$, causing a strongly convex behavior in the functional $J_3$ as a function of the angle, while for $\omega \geq 3$ the value of the vorticity functional becomes rather insensitive to the choice of the anastomosis angle. Thus we are able to conclude that – similarly to the 2D case – the most challenging situation for

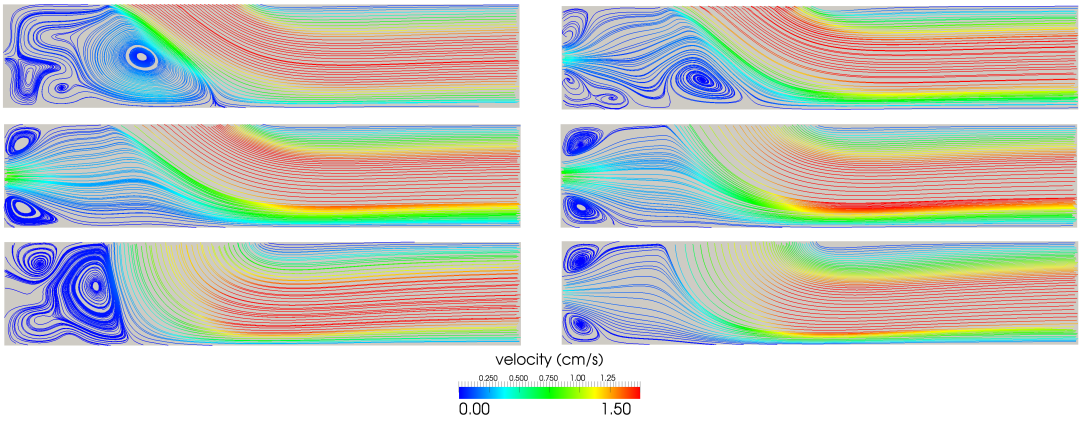



Figure 5.15: Streamlines of the 2D steady flow for the cases (from top to bottom, from left to right): $\theta = 25°$ with $\omega = 0$, $\omega = 1$; $\theta = 30°$, $\theta = 50°$ with $\omega = 0$; $\theta = 85°$ with $\omega = 0$, $\omega = 1$.

the minimization of vorticity is the case of total or near-total occlusion of the stenosed branch. We did not test the effect of the residual flow profile uncertainty on the 3D case as the results in the 2D case already highlighted the need to consider a "worst-case" flow profile in order to obtain robust results. In Table 5.1 we also include the estimated optimal angles in the 3D case for reference with the 2D results presented before. These were obtained by cubic spline interpolation of the curves in Fig. 5.12. For $\omega = 0$ the optimal angle is very close to the one obtained for the 2D problem, while a divergence of results occurs as $\omega$ is increased; the optimal angles in the 3D case tend to be somewhat larger. However, if the robust angle is assumed to correspond in both cases to the optimal angle for $\omega = 0$, we can state that the solution of the robust design problem in 2D gives a good indication to the choice of a robust angle in the more realistic 3D problem.

The remaining question to be answered is, whether the similarity of the 2D and 3D problems in the context of vorticity minimization extends also to the more difficult case of WSS-related functionals, such as $J_5$ given by (5.4). It is likely that the tertiary vortices have some effect on the downstream WSS, thus potentially changing the situation between the 2D and 3D cases. While some works on direct minimization of WSS-related quantities have been attempted [87, 288], a rigorous mathematical framework for the minimization of quantities depending on higher derivatives of velocity especially in the uncertainty quantification or robust design context seems beyond the reach of current methodology.

## 5.4 Shape optimization of end-to-side bypass grafts

An alternative approach to optimal control for grafts design relies on the solution of a shape optimization (SO) problem, for which the control variable is the shape of the domain $\Omega$ itself. This entails the minimization of a cost functional by finding the optimal shape of the domain where the PDE which models the flow is defined. In the following, we present and discuss several results, dealing with shape optimization of aorto-coronary bypass configurations (considering both a simplified Stokes and a more accurate Navier-Stokes flow model) and femoro-popliteal bypass configurations. We rely on the reduced framework addressed in Sect. 4.3, by describing the shape of these configurations through Free-Form Deformation (FFD) mappings.
First, we deal with some shape optimization problems under the form (DD-SO$\boldsymbol{\mu}$), by fixing the residual flow $\mathbf{v}_{\text{in}}$ through the occluded artery (Sect. 5.4.1). Then, we consider a robust optimization problem under the form (RD-SO$\boldsymbol{\mu}$), by parametrizing the residual flow profile as in Sect. 5.3 and seeking for the configuration corresponding to the *worst case* scenario (Sect. 5.4.2-5.4.3). In all these cases, we refer to the framework of Sect. 5.2.4 for state problems and notation.

The results of the following sections have been presented in three recent papers published in collaboration with Rozza, Quarteroni [206] and also with Lassila [180, 181].

### 5.4.1 Case I: Stokes flows for an aorto-coronary bypass

In this first case we are interested to the minimization of the *vorticity* cost functional $J_3$ in the downfield subregion $\Omega_{\text{obs}}$ where a vortex may occur, leading to possible occlusions after grafting and plaque formation. Concerning inflow profiles, we consider a Poiseuille profile on the inflow $\Gamma_{\text{c}}$ (graft) and two different different blockage conditions have been considered on the inflow $\Gamma_{\text{in}}$ (occluded artery). In Case I.A a completely occluded coronary is considered, corresponding to $\mathbf{v}_{\text{in}} = \mathbf{0}$, while in Case I.B a partially stenosed coronary is considered, choosing a residual flux given by a Gaussian profile $\mathbf{v}_{\text{in}} = \phi(y)\mathbf{e}_1$ like in Sect. 5.3.3, centered on the section $\Gamma_{\text{in}}$ with flow rate of about 15% w.r.t. the one in the bypass grafting.
In both cases, the same reference configuration $\Omega \subset \mathbb{R}^2$ of Fig. 5.16 – similar to some bypass configurations already considered in previous works [190, 254, 276] – has been chosen. The observation subregion is $\Omega_{\text{obs}} = \{(x_1, x_2) \in \Omega \, : \, x_1 > 1.5\}$. A set of admissible shapes under the form (4.14) is obtained through a FFD parametrization $T(\cdot; \boldsymbol{\pi})$ built over the reference domain $\Omega$, with $\boldsymbol{\pi} \in \mathcal{D}_{\boldsymbol{\pi}}$. Here we consider a $5 \times 6$ lattice of control points on the rectangle $D = [-1, 3] \times [-0.6, 0.4]$, for a total number of 60 degrees of freedom. Among them, $p = 8$ of these degrees of freedom (seven vertical and one horizontal deformations) have been selected as design variables on the basis of an *empirical* criterion: selected design variables aim at controlling both the anastomotic angle and the ratio between the diameters of the two branches.

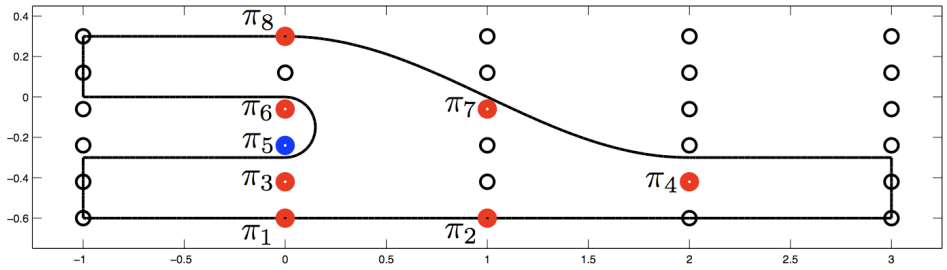



Figure 5.16: Reference domain $\Omega$ and FFD setting. Control points depicted in red/blue can be moved in vertical/horizontal direction.

In particular, since the closer the control point to the free-boundary, the larger its influence on local shape deformations, we select parameters $\pi_3$ and $\pi_5$ (the latter is the only one giving an horizontal deformation) to control the anastomotic angle, parameters $\pi_6$ and $\pi_8$ to control the diameters ratio (see Fig. 5.16). Moreover, parameters $\pi_4$ and $\pi_7$ control the upper shape of the anastomotic bifurcation, while $\pi_1$ and $\pi_2$ are responsible of the shape of the lower wall. The parametric domain $\mathcal{D}_{\boldsymbol{\pi}}$ is therefore given by

$$\mathcal{D}_{\boldsymbol{\pi}} = \{\boldsymbol{\pi} = (\pi_1, \dots, \pi_8) \in \mathbb{R}^8 : \; \pi_i \in [-\alpha, \alpha] \;\; \forall\, i \neq 5, \;\; \pi_5 \in [0, \beta]\}. \tag{5.6}$$

where the choice $\alpha = 0.2$ and $\beta = 0.8$ enables to guarantee that $\det(J_T) > 0$ (at least by direct numerical inspection). We point out that the number of parameters $P = 8$ has been selected to ensure a trade-off between *(i)* a significative shape versatility during the optimization process and *(ii)* a better computational performance at each iteration of the optimization process in terms of Online efficiency and certification by *a posteriori* error bounds. The points location (and their activation) have been emphasized in very sensible zone of the bypass [6, 276] where a Miller or Taylor patch shape is expected [254].

The ranges of variation of the control points positions with respect to the original locations ensure to keep small deformations and it is influenced by the lattice of points[10].
Before constructing the RB approximation of the flow problem, we need to recover the affinity assumption through the EIM procedure – only once for both cases I.A and I.B. By setting $\varepsilon_{tol}^{EIM} = 2.5 \times 10^{-4}$, we obtain an affine expansion of $\sum_{i,j} K_{ij}^a = 204$ terms for $\nu_{ij}$ components, of $\sum_{i,j} K_{ij}^b = 18$ terms for $\chi_{ij}$ components and of $\sum_i K_i^f = 5$ terms for the right hand-side components. The resulting problem is discretized with $\mathcal{N}_V = 33,330$ and $\mathcal{N}_Q = 4,269$ dofs for velocity and pressure, respectively, using $\mathbb{P}_2/\mathbb{P}_1$ finite elements; the dimension of the computed reduced basis space is $3N = 63$ (resp. $3N = 60$) for case I.A (resp. I.B), by choosing a tolerance $\varepsilon_{tol}^{RB} = 5 \cdot 10^{-4}$ for the greedy procedure.

For the test case I.A, a vorticity reduction of about 39% has been obtained through the described procedure; the cost functional decreases from an initial value $s_N(\boldsymbol{\pi}^{(0)}) = 459.40$ to an optimal value $s_N(\hat{\boldsymbol{\pi}}_{RB}) = 281.25$ after 36 optimization steps performed in $2968s$ until convergence.
In Fig. 5.17 both the reference and the optimal configurations are displayed, together with deformation induced by the displacement of the control points. Flow variables and corresponding vorticity layers for the initial and the optimal configurations are represented in Fig. 5.18.

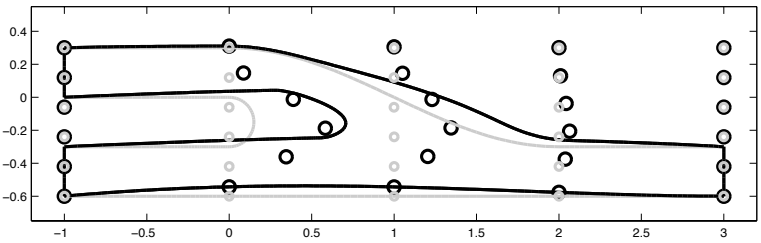


Figure 5.17: Case I.A: reference configuration (in grey), optimal shape (in black) and corresponding control points displacement.

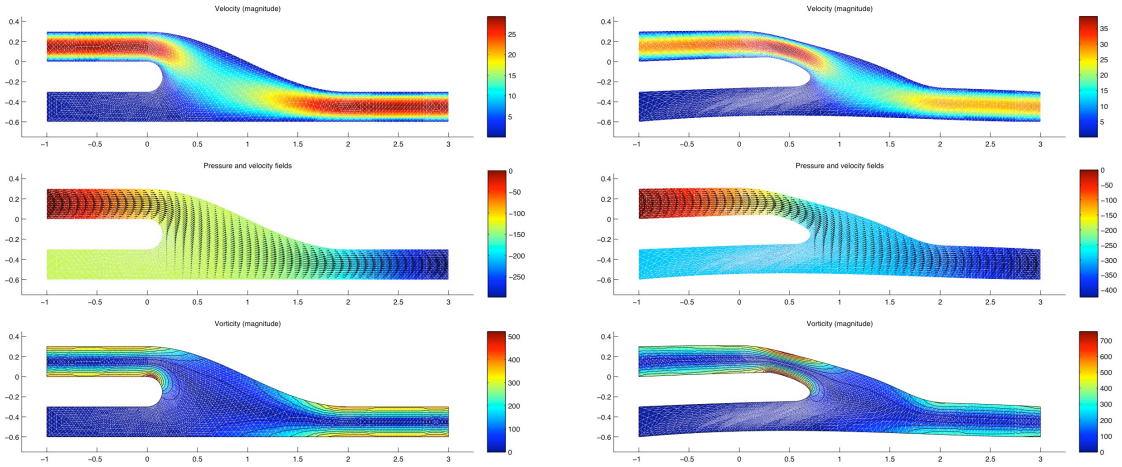



Figure 5.18: Case I.A: velocity field $[cm/s]$, pressure field $[dyn/cm^2]$ and vorticity magnitude for the reference (left) and the optimal (right) configuration.

[10] We remark that the density of control points in the lattice is instead responsible of the chosen degree of the polynomials describing the FFD map (quantities $L$ and $M$ in (2.56)).

For the test case I.B, a vorticity reduction of about 45% has been obtained, passing from an initial value of the cost functional $s_N(\boldsymbol{\pi}^{(0)}) = 839.45$ to an optimal value of $s_N(\hat{\boldsymbol{\pi}}_{RB}) = 466.05$ after 45 optimization steps, which have been performed in $3554s$. Optimal configuration obtained in this case is displayed in Fig. 5.19, while flow variables and the corresponding vorticity layers for the initial and the optimal configurations are represented in Fig. 5.20.

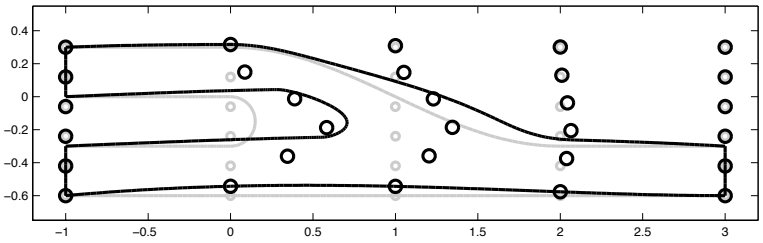


Figure 5.19: Case I.B: reference configuration and optimal shape with corresponding control points displacement.

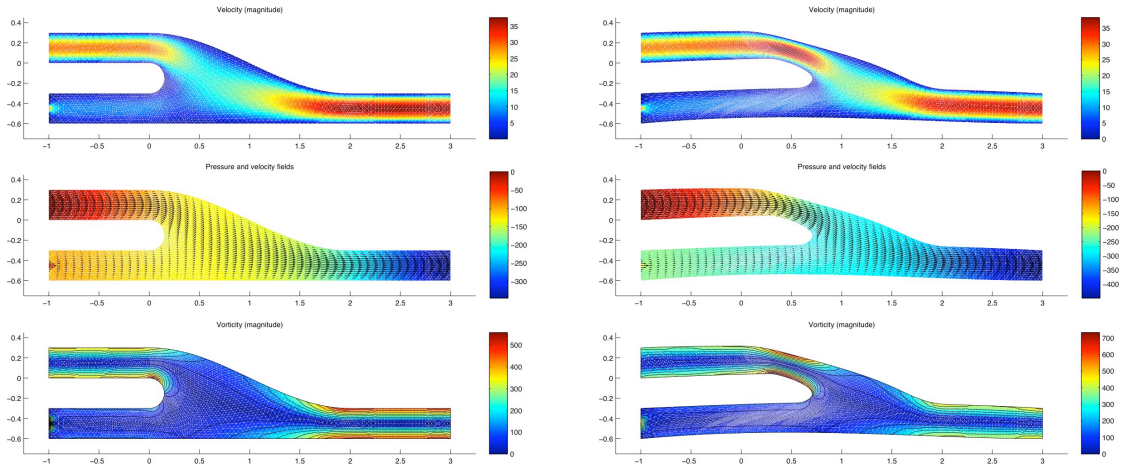



Figure 5.20: Case I.B: velocity field $[cm/s]$, pressure field $[dyn/cm^2]$ and vorticity magnitude for the reference (left) and the optimal (right) configuration.

Moreover, in both cases I.A (total occlusion) and I.B (partial occlusion) the optimized shapes are quite similar, showing a much smoother toe and heel than the initial shape; they also show smoother curvatures at the heel with a gradual transition in the toe region, as already pointed out in the work by Lei et al. [190]. Only the graft section before the anastomosis – say, at about $x_1 = 0.5$ – results a bit narrower than usual grafts, maybe due to the choice of the downfield observation region. The localization of maximum values of vorticity at the heel and toe is expected, because this is the region where disturbed flows occur, even with a Stokes model; the same conclusion can be drawn for wall shear stress gradient and Navier-Stokes flows [190]. Moreover, higher values in the heel region are not as clinically significant as the high vorticity values near the toe region and in the down-field region, which is a well-known location where restenosis might develop.
The shape optimization procedure however enables to reduce the vorticity in the down-field region of the anastomosis, even if increasing values of vorticity arise at heel and top segments. In the case where the observation region $\Omega_{obs}$ is the whole domain $\Omega$ instead of the subdomain $\Omega_{\text{obs}}$ – and thus vorticity is minimized all over the anastomosis – different results can be obtained, both in terms of optimal shapes and vorticity reduction. By using the same procedure as before, we

obtain (e.g. for test case B) the vorticity patterns in Fig. 5.21: the minimization of the vorticity all over the anastomosis leads to optimal shapes characterized by bigger sections and narrow bifurcations. Moreover, the parameter $\pi_5$ – which strongly affects the anastomotic angle – seems to play an important role in the behavior of the vorticity field (see also Fig. 5.22). For both test cases numerical results are summarized in Tabs. 5.2.

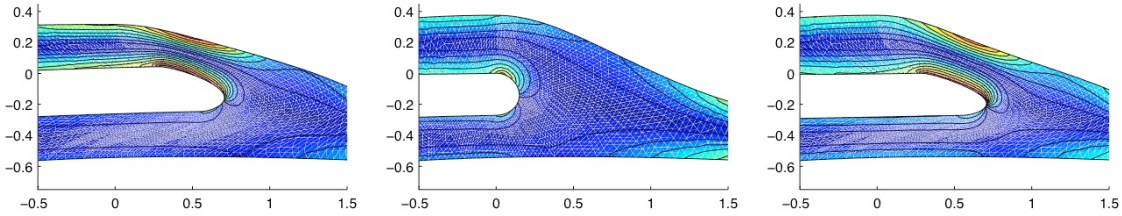


Figure 5.21: Case I.B: optimal shapes (and related vorticity patterns) obtained through vorticity minimization on the down-field region $\Omega_{\text{obs}}$ (left) and on the whole domain $\Omega$ (middle, right). In the third case the parameter $\pi_5$ has been fixed equal to the optimal value obtained in the first case by vorticity minimization on $\Omega_{\text{obs}}$.

| Case I.A | $\Omega_{obs}$ | $s_N(\boldsymbol{\pi}^{(0)})$ | $s_N(\hat{\boldsymbol{\pi}}_{RB})$ | $\Delta s_N$ | # iterations | cpu (s) |
|---|---|---|---|---|---|---|
| | $\Omega_{\text{obs}}$ | 459.40 | 281.25 | 38.8% | 36 | 2968 |
| | $\Omega$ | 844.93 | 576.75 | 31.7% | 45 | 3504 |

| Case I.B | $\Omega_{obs}$ | $s_N(\boldsymbol{\pi}^{(0)})$ | $s_N(\hat{\boldsymbol{\pi}}_{RB})$ | $\Delta s_N$ | # iterations | cpu (s) |
|---|---|---|---|---|---|---|
| | $\Omega_{\text{obs}}$ | 839.45 | 466.05 | 44.5% | 45 | 3554 |
| | $\Omega$ | 1492.5 | 934.25 | 37.4% | 158 | 10492 |

Table 5.2: Cases I.A (top) and I.B (bottom): initial and optimal value of cost functional, total reduction, number of iterations of the optimization procedure and cpu times for a vorticity observation on $\Omega_{\text{obs}}$ and $\Omega$.

We also compare the results of the downfield vorticity minimization problem based on the reduced (basis) procedure with the ones obtained using the truth FE approximation of the Stokes flow (see Tab. 5.3). In the considered cases, the solution of the reduced optimization problem $s_N(\hat{\boldsymbol{\pi}}_{RB})$ shows a slight difference (less than 0.1% in the output value) and results sub-optimal with respect to the solution $s^{\mathcal{N}}(\hat{\boldsymbol{\pi}}_{FE})$ computed by using the truth FE approximation; this fact has already been remarked both in an optimal control [72] and in a shape optimization [185] context.

| | Case I.A | Case I.B |
|---|---|---|
| $s_N(\hat{\boldsymbol{\pi}}_{RB})$ | 281.25 | 466.05 |
| $s^{\mathcal{N}}(\hat{\boldsymbol{\pi}}_{FE})$ | 281.17 | 465.89 |
| $\dfrac{\lvert s_N(\hat{\boldsymbol{\pi}}_{RB}) - s^{\mathcal{N}}(\hat{\boldsymbol{\pi}}_{FE})\rvert}{s^{\mathcal{N}}(\hat{\boldsymbol{\pi}}_{FE})}$ | $3 \cdot 10^{-4}$ | $3 \cdot 10^{-4}$ |

Table 5.3: Optimal values of cost functional (downfield vorticity, $\Omega_{obs} = \Omega_{\text{obs}}$) for the optimization based on RB and FE approximations, for test cases A and B.

Moreover, by means of the Morris' factor *screening* method introduced in Sect. 4.8, we can determine which are the most important parameters affecting the vorticity output in the two

cases I.A and I.B analyzed throughout this section. By performing $R = 20$ (randomly selected) experiments, each consisting of $(p + 1)$ random parameter configurations, we are able to get $R = 20$ estimates of the elementary effects of each parameter component on the vorticity output, by taking into considerations the possible interactions among the $p = 8$ parameter components $\pi_1, \ldots, \pi_8$. The outcome of this procedure (obtained through $R(p+1) = 180$ input/output Online evaluations) is shown in Fig. 5.22: we can point out how, at least in the case I.B, the parameter $\pi_5$ tuning the anastomotic angle is the one giving the largest effect on vorticity output variations.

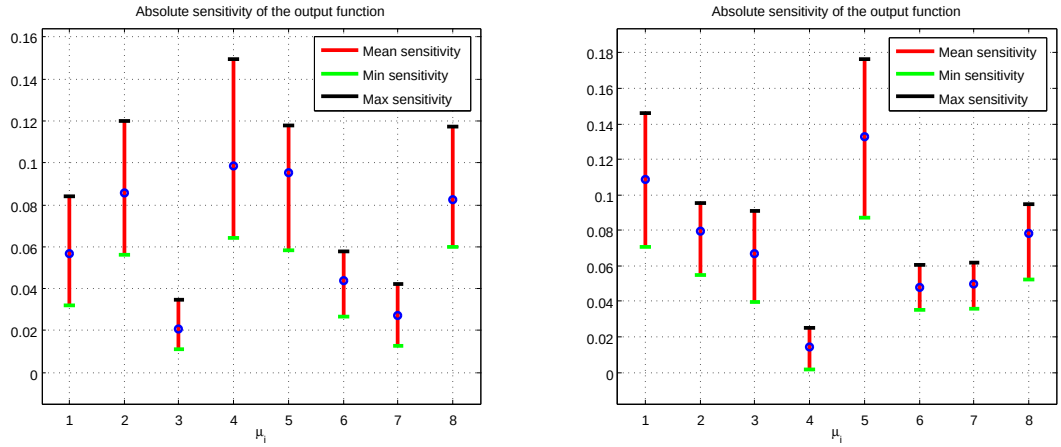



Figure 5.22: Cases I.A and I.B: sensitivity analysis by Morris' screening method. For each parameter component $i = 1, \ldots, 8$, we report the minimum, the maximum and the average sensitivities over a sample given by $R = 20$ estimates of elementary effects on the vorticity output.

In the end, we provide some computational details concerning the reduction pursued by the coupled RB+FFD framework. Compared to a FE simulation, in our 2D aorto-coronary bypass case, a RB Online evaluation of flow variables enables a computational speedup of about 100 times. Average time over 100 Online evaluations is of $t_{RB}^{online} = 2.204s$ and of $t_{RB}^{online} = 2.012s$ for the test cases I.A and I.B, while the corresponding FE computations take an average time of $t_{FE} = 209s$ and $t_{FE} = 216s$, respectively (see Tab. 5.4). This is basically due to the reduction of about 500 times in the dimension – $\mathcal{N} = 35,997$ vs. $N = 72$ (case I.A) and $N = 66$ (case I.B) – of the linear systems obtained by FE discretization and RB approximation. Consequently, computational times for the whole optimization process are reduced too, depending almost linearly by the time for a single Online evaluation.
We take into account also the time spent for the Offline construction and storage; this allows to determine the break-even point, given by $\mathcal{Q}_{BE} = t_{RB}^{offline}/t_{FE}$. In particular, we obtain a break-even point of $\mathcal{O}(10^2)$ in both cases. Note that Offline procedures for basis assembling and certification of the error are the same even if we change the cost functional or the optimization goal: in this way Offline efforts can be reused for different Online optimization procedures.

| | FE | RB (min) | RB (average) | RB (max) | Speedup |
|---|---|---|---|---|---|
| Case I.A | 209.06s | 1.939s | 2.012s | 3.078s | 94.85 |
| Case I.B | 215.76s | 1.945s | 2.204s | 3.314s | 107.18 |

Table 5.4: Computational times and speedup for reduced basis Online evaluations and finite element computations. RB Online evaluations have been executed considering a sample of 100 randomly chosen parameter vectors $\boldsymbol{\mu} \in \mathcal{D}$.

### 5.4.2 Case II: Navier-Stokes flows for an aorto-coronary bypass

We now consider a steady Navier-Stokes model to describe blood flows. Before considering the optimal design of aorto-coronary bypasses exploiting a RB approximation of Navier-Stokes equations, we provide a feedback validation of the optimized bypass obtained in the previous section (through a RB Stokes model) by a "high-fidelity" FE Navier-Stokes model. We thus compare the results obtained by using these two flow models on the same shape configurations. We consider the initial shape and the optimal shape obtained by minimizing the blood vorticity corresponding to the Stokes flow, in both cases A and B. We then compute the Navier-Stokes flows and vorticity distributions corresponding to the same configuration and boundary conditions. We obtain in this case a vorticity reduction in the downfield region $\Omega_{\text{obs}}$ of 39.9% in case I.A and of 44.5% in case I.B, respectively – which only slightly differs from that obtained by using a Stokes model. Moreover, in the four considered configurations (initial and optimal shapes for cases I.A and I.B) the difference between the computed vorticity outputs $s_N(\cdot)$ and $s_{NS}^{\mathcal{N}}(\cdot)$ is smaller than 5% (see Tab. 5.5).

| | Stokes case | Navier-Stokes case |
|---|---|---|
| case I.A | $s_N(\boldsymbol{\pi}^{(0)}) = 459.40$ | $s_{NS}^{\mathcal{N}}(\boldsymbol{\pi}^{(0)}) = 469.14$ |
| | $s_N(\hat{\boldsymbol{\pi}}_{RB}) = 281.25$ | $s_{NS}^{\mathcal{N}}(\hat{\boldsymbol{\pi}}_{RB}) = 281.81$ |
| case I.B | $s_N(\boldsymbol{\pi}^{(0)}) = 839.45$ | $s_{NS}^{\mathcal{N}}(\boldsymbol{\pi}^{(0)}) = 878.32$ |
| | $s_N(\hat{\boldsymbol{\pi}}_{RB}) = 466.05$ | $s_{NS}^{\mathcal{N}}(\hat{\boldsymbol{\pi}}_{RB}) = 488.50$ |

Table 5.5: Feedback validation (downfield vorticity minimization): cost functional values for initial and optimal configurations computed with a Stokes and a Navier-Stokes flow model.

This feedback is provided with the sole aim of testing with a higher-fidelity fluid model the optimal shape obtained with the low-fidelity model. Nevertheless, we can consider this shape *optimal* only for the Stokes model and just *improved* for the Navier-Stokes model (Tab. 5.5). Previous feedback results were provided in [254]. We report in Fig. 5.23 the velocity and the vorticity fields computed on the optimal configurations in cases I.A and I.B.

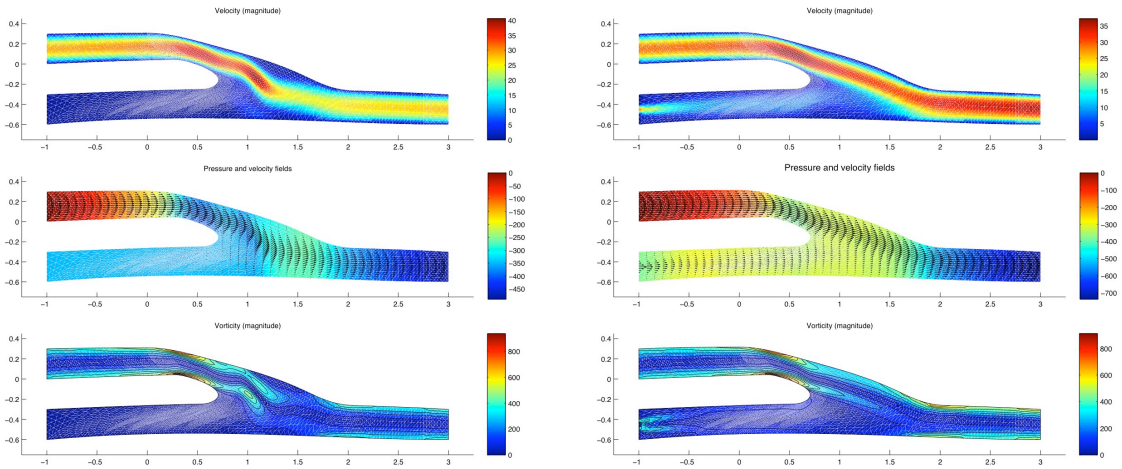



Figure 5.23: Feedback validation: velocity, pressure and vorticity fields computed by using a Navier-Stokes model in the optimal configuration of cases I.A (left) and I.B (right).

A second problem solved within this RB+FFD reduced framework is related to the design of a graft which is *robust* with respect to the magnitude of the residual flow through the occluded branch. In other words, we aim at finding the optimal shape of the graft in presence of the worst case scenario in terms of residual flow, by solving the problem (RD-SO$\boldsymbol{\mu}$). We exploit the same configuration and FFD parametrization as in the previous Stokes case, now being $\alpha = 0.15$ and $\beta = 0.6$ the ranges of vertical and horizontal displacements of the control points depicted in red/blue in Fig. 5.16. We now consider an observation subregion close to the heel, given here by $\Omega_{\text{obs}} = \{(x_1, x_2) \in \Omega : x_1 \in (1, 2)\}$.
Moreover, concerning inflow profiles, we consider a Poiseuille profile on the inflow $\Gamma_{\text{c}}$ (graft) and a parametrized gaussian profile $\mathbf{v}_{\text{in}} = \omega\phi(y)\mathbf{e}_1$ on the inflow $\Gamma_{\text{in}}$ (occluded artery), where $\omega \in \mathcal{D}_{\boldsymbol{\omega}} = [0, \omega_{\max}]$ is the uncertainty parameter tuning the degree of occlusion. In particular, the dependence of the two flows on $\omega$ is such that the downfield flowrate is constant, with a flow split ranging from 1/0 (complete occlusion of the host artery, for $\omega = 0$) to 2/1 (flowrate through the occluded artery equal to one half of the flowrate through the graft, for $\omega_{\max} = 20$).
The resulting RB approximation, relying on the same EIM procedure and FE approximation spaces (now related to discretization of Navier-Stokes equations) is defined on RB spaces of dimension $3N = 36$, thus yielding the possibility to solve a Navier-Stokes problem in a very rapid way ($1.84s$, averaged time over 1,000 evaluations).

For the shape optimization, we consider three different cost functionals, given by the vorticity functional $J_3$, the Stokes tracking functional $J_2$ and the functional based on Galilean invariant $J_4$, which have been introduced in Sect. 5.2.4. The choice of these three cost functionals is based on the considerations developed in Sect. 5.3 for the case of optimal flow control problems.
In Fig. 5.24 the velocity fields within the optimal shapes for the $J_3$, $J_2$ and $J_4$ functionals (from top to bottom) are represented, in the cases $\omega = 0$ (complete occlusion, left) and $\omega = \omega_{\max}$ (maximum residual flow, right), respectively. Similarly as in the OC case analyzed in Sect. 5.3, the condition leading to the strongest development of vorticity cores is the presence of a complete occlusion, for which the flow through the bypass starts creating a strong and complex vorticity pattern close to the heel. Comparing the optimal shapes obtained in this case with the results obtained through a Stokes model, we can remark how the different degrees of occlusion (here $\omega = 0$ vs. $\omega = \omega_{\max}$) lead in the Navier-Stokes case to different graft profiles, showing a less elongated anastomosis in the case $\omega = 0$. This effect might be also due to the different observation region, in this case closer to the heel.

We also find that the minimum values of the three cost functionals are decreasing functions with respect to $\omega$, an indication that the case $\omega = 0$ is the most difficult one concerning shape optimization (see Table 5.7); reduction in the cost functionals ranges from 24% to 70% for the different cases. The vorticity cores are clearly observable also in the optimal configuration in presence of a complete artery occlusion; moreover, although we obtain a clear reduction of vorticity also in this case, the vorticity cores never disappears completely.
In the end, as for the OC case, we remark that the anastomosis angle decreases as the residual flow increases, since optimal shapes obtained in the case $\omega = \omega_{\max}$ (see Fig. 5.24) show a more elongated heel. We come back to these considerations about robust shape optimization of a bypass graft and behavior of different cost functionals in the forthcoming section, devoted to the design of a femoro-popliteal bypass.

Concerning the solution of the robust design (RD-SO$\boldsymbol{\mu}$) problem, we find that the robust configurations correspond to the optimal shapes computed for $\omega = 0$ above. As in the case of robust OC problems, the condition leading to the strongest development of vorticity cores is the presence of a complete occlusion, for which the flow through the bypass starts creating a strong vortex in the proximity of the anastomosis. Thus, it is clear that in order to be robust over the entire range $\omega \in [0, \omega_{\max}]$, the bypass must be tuned mainly for the case of total occlusion.

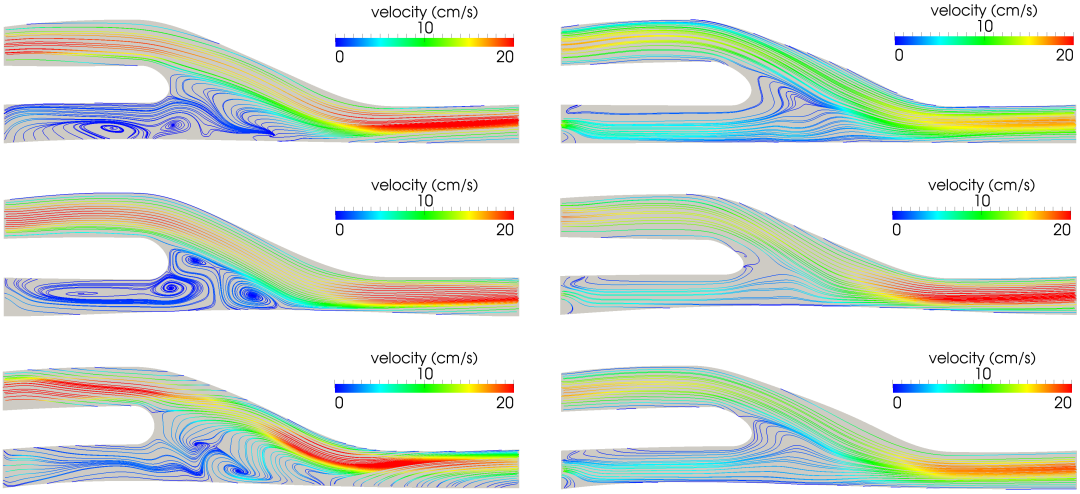



Figure 5.24: Case II. Optimal shapes for $\omega = 0$ (left) and $\omega = \omega_{\max}$ (right) for the vorticity $J_3$, the Stokes tracking $J_2$ and the Galilean invariant $J_4$ functionals (from top to bottom).

| | $\frac{J^*(\omega=0)}{J^*(\omega=\omega_{\max})}$ | $\Delta J$ $(\omega = 0\,,\ \omega = \omega_{\max})$ | # I/O evals $(\omega = 0\,,\ \omega = \omega_{\max})$ | # I/O evals (robust) |
|---|---|---|---|---|
| $J_3$ | 1.257 | 26.3% , 24.2% | 125 , 27 | 389 |
| $J_2$ | 1.924 | 63.2% , 55.3% | 99 , 64 | 416 |
| $J_4$ | 1.267 | 65.4% , 70.7% | 183 , 63 | 973 |

Table 5.6: Case II. Results obtained by solving the shape optimization problem (DD-SO$\boldsymbol{\mu}$) in the cases $\omega = 0$ and $\omega = \omega_{\max}$, as well as the robust shape optimization problem (RD-SO$\boldsymbol{\mu}$).

Regarding instead computational aspects, we remark that the solution of the robust (shape) optimization problem requires about $\mathcal{O}(10^3)$ input/output evaluations, thus entailing a CPU time which is at least one order of magnitude larger than a (shape) optimization problem (see Table 5.7), ranging from $1 \div 4$ hours for the latter case to $13 \div 35$ hours for the former case. This indicates that a design that is robust over the entire range $\omega \in [0, \omega_{\max}]$ must be tuned mainly for the case of total occlusion.

### 5.4.3 Case III: Navier-Stokes flows for a femoro-popliteal bypass

In this last case we consider the shape optimization of a femoro-popliteal bypass graft, as well as the robust optimization of this configuration with respect to residual flows through the occluded branch. Even if it is considered at the end of this chapter, the analysis of this problem – performed before the Navier-Stokes aorto-coronary case – has been the first attempt to the development of the current framework to robust optimization problems and to more general inverse problems.

In particular, we aim at solving the problems above by using a RB+FFD approach, comparing three cost functionals – the Stokes-tracking functional $J_2$, the vorticity $J_3$ and the one based on Galilean invariant $J_4$. Before setting the reduced framework, we performed a preliminary test, in order to select $p_{\boldsymbol{\pi}} = 6$ design parameters $\pi_1, \ldots, \pi_6$ for constructing the FFD map $T(\cdot; \boldsymbol{\pi})$, among the possible $2 \times 6 \times 4$ displacements of a $6 \times 4$ lattice of control points.

We have compared the strategies presented in Sect. 4.8: *(i)* a user's experience-based (UEB) selection; *(ii)* a one-at-a-time (OAT) selection; *(iii)* a selection based on the Morris screening procedure (MR-OAT) and *(iv)* an approximate POD (APOD) procedure. In the cases *(ii)-(iv)*, the *selection criteria* were based on the parametric sensitivities of the viscous energy output $J_1$ computed all over the domain. This choice has been made before selecting more accurate outputs and observation regions for sake of optimization, to *measure* the interplay between global shape deformations and the flow field across the domain. Moreover, the choice of 7 control points (denoted with a cross in Fig. 5.25) was excluded due to geometric constraints, since we experienced strong correlations between the displacement of these control points and the variation of the diameters of the inflow and outflow sections, which have to be considered as *prescribed* in this application.

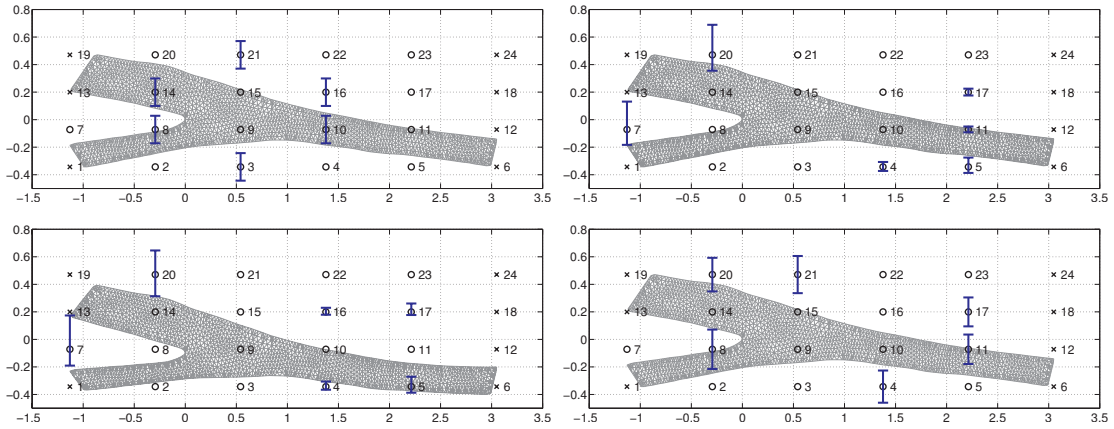


Figure 5.25: Reference domain $\Omega$, "truth" finite element mesh, and selected FFD control points used to model the displacements of the shape by means of the *(i)* UEB, *(ii)* OAT, *(iii)* MR-OAT and *(iv)* APOD procedures (from left to right, from top to bottom). Control points indicated with a cross are not allowed for selection. Vertical displacements $\pi_i \in [-0.2, 0.2]$ of the points denoted with a bar are allowed. Stems are referred to parametric sensitivities (see Sect. 4.8) of the viscous energy output $J_1$, scales are not uniform among the procedures *(ii)-(iv)*.

In any case, vertical displacements were found to be uniformly more influent than horizontal displacements. The vertical displacements of the control points selected by means of the procedures *(i)-(iv)* are different. By constructing in each of these cases the RB approximation of the flow problem and solving a shape optimization problem with the viscous energy cost functional $J_1$, we found a reduction $\Delta J_1$ which ranges between 18% (UEB case) and 44% in the APOD case; in the OAT and MR-OAT cases reduction was rather similar and about 40%, thus leading in this case to a remarkable improvement with respect to the UEB choice.
We thus consider the FFD mapping obtained through the APOD selection procedure *(iv)* and the computational reference domain represented in Fig. 5.25.
The admissible parameter range for each control point displacement is $\boldsymbol{\pi} \in \mathcal{D}_{\boldsymbol{\pi}} = [-0.2, 0.2]^6$ for $i = 1, 2, \ldots, 6$. A seventh parameter $\omega \in \mathcal{D}_{\boldsymbol{\omega}} = [0, \omega_{\max}]$ takes into account the magnitude of the residual flow – and thus the flow splitting. The inflow velocities on $\Gamma_{\mathrm{c}}$ (bypass) and $\Gamma_{\mathrm{in}}$ (blocked host artery) are given by a Poiseuille profile $\mathbf{v}_{\mathrm{c}}(\omega)$ and a Gaussian profile $\mathbf{v}_{in}(\omega)$, respectively. The dependence of the two flows from the uncertain parameter $\omega \in \mathcal{D}_{\omega} = [0, \omega_{\max}]$ is such that the downfield flowrate is constant, with a flow split ranging from 1/0 (complete occlusion of the host artery, for $\omega = 0$) to 2/1 (flowrate through the occluded artery equal to one half of the flowrate through the graft, for $\omega = \omega_{\max} = 15$). Moreover, by choosing physiological values for physical flow viscosity and density we end up with a steady incompressible Navier-Stokes flow characterized by $\mathrm{Re} = 150$ at the location of the anastomosis.

Before constructing the RB approximation of the flow problem, we need to recover the affinity assumption through the EIM procedure. By setting $\varepsilon_{tol}^{EIM} = 10^{-4}$, we obtain an affine expansion of $Q_a + 2Q_b + Q_c = 289$ terms. The resulting problem is discretized with $\mathcal{N}_V + \mathcal{N}_Q \approx 16{,}000$ degrees of freedom, using $\mathbb{P}_2/\mathbb{P}_1$ finite elements; the dimension of the computed reduced basis space is $3N = 60$ by choosing a tolerance $\varepsilon_{tol}^{RB} = 5 \cdot 10^{-2}$ for the greedy procedure.
In Fig. 5.26 the results of the shape optimization problem for each cost functional $J_i$, $i = 2, 3, 4$ are reported, in case of fixed degrees of occlusion; indeed, $n_{opt} = 16$ shape optimization problems have been solved, corresponding to the cases $\omega = 0, 1, \ldots, 15$. We can notice that the minimum values obtained with the vorticity functional $J_3$ and the functional based on Galilean invariant $J_4$ are decreasing function with respect to $\omega$, an indication that the case $\omega = 0$ is the most difficult one concerning vorticity minimization.

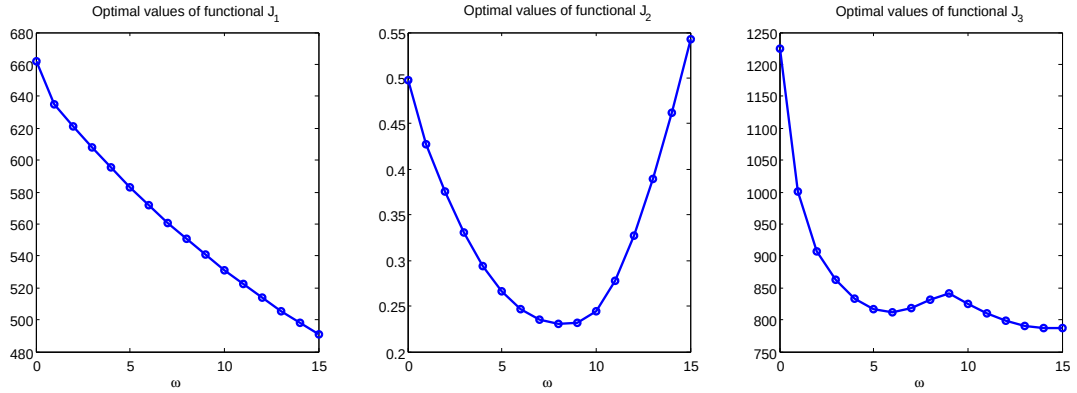



Figure 5.26: Minimum values for the three cost functionals (from left to right: vorticity $J_3$, Stokes-tracking $J_2$, Galilean invariant $J_4$) as a function of the unknown residual flow parameter $\omega$. Results obtained by shape optimization for $n_{opt} = 16$ different cases between $\omega = 0, 1, \ldots, 15$.

As expected, the condition leading to the strongest development of vorticity cores is that corresponding to a complete occlusion, for which the flow through the bypass starts creating a strong vortex in the proximity of the anastomosis. As soon as a small residual flow crosses the occluded branch, the vortex cores are removed. On the other hand, also the tracking-type functional detects the presence of the vortex core for small values of $\omega$, but for larger values of $\omega$ the misfit between the two flows increases considerably due to the presence of the strong nonlinear convective term. Moreover, the reduction of the cost functional for the three cases ranges from 53% to 73%, averaging the results on the $n_{opt} = 16$ problems for each case (see Table 5.7).
In Fig. 5.27 the velocity fields for the initial and the optimal shapes obtained with the three cost functionals are reported, considering the minimum ($\omega = 0$) and the maximum ($\omega = 15$) magnitude of the residual flow through the occluded branch. The vorticity cores are clearly observable in presence of a complete artery blockage; moreover, although we get a sensible reduction of the cost functional also in this case, the vorticity cores never disappear completely.

| | $\Delta J$ (average) | # optim. iters | # output evals | # total output evals | CPU time |
|---|---|---|---|---|---|
| $J_3$ | 0.5297 | 3 ÷ 7 | 28 ÷ 99 | 771 | 5.9 h |
| $J_2$ | 0.6618 | 3 ÷ 7 | 28 ÷ 60 | 641 | 9.4 h |
| $J_4$ | 0.7340 | 3 ÷ 27 | 28 ÷ 230 | 1 120 | 11.3 h |

Table 5.7: Results of the shape optimization problem ($n_{opt} = 16$ cases, $\omega = 0, 1, \ldots, 15$). Here $\Delta J = (J^{(0)} - \hat{J})/J^{(0)}$, and the average is computed over the 16 cases.

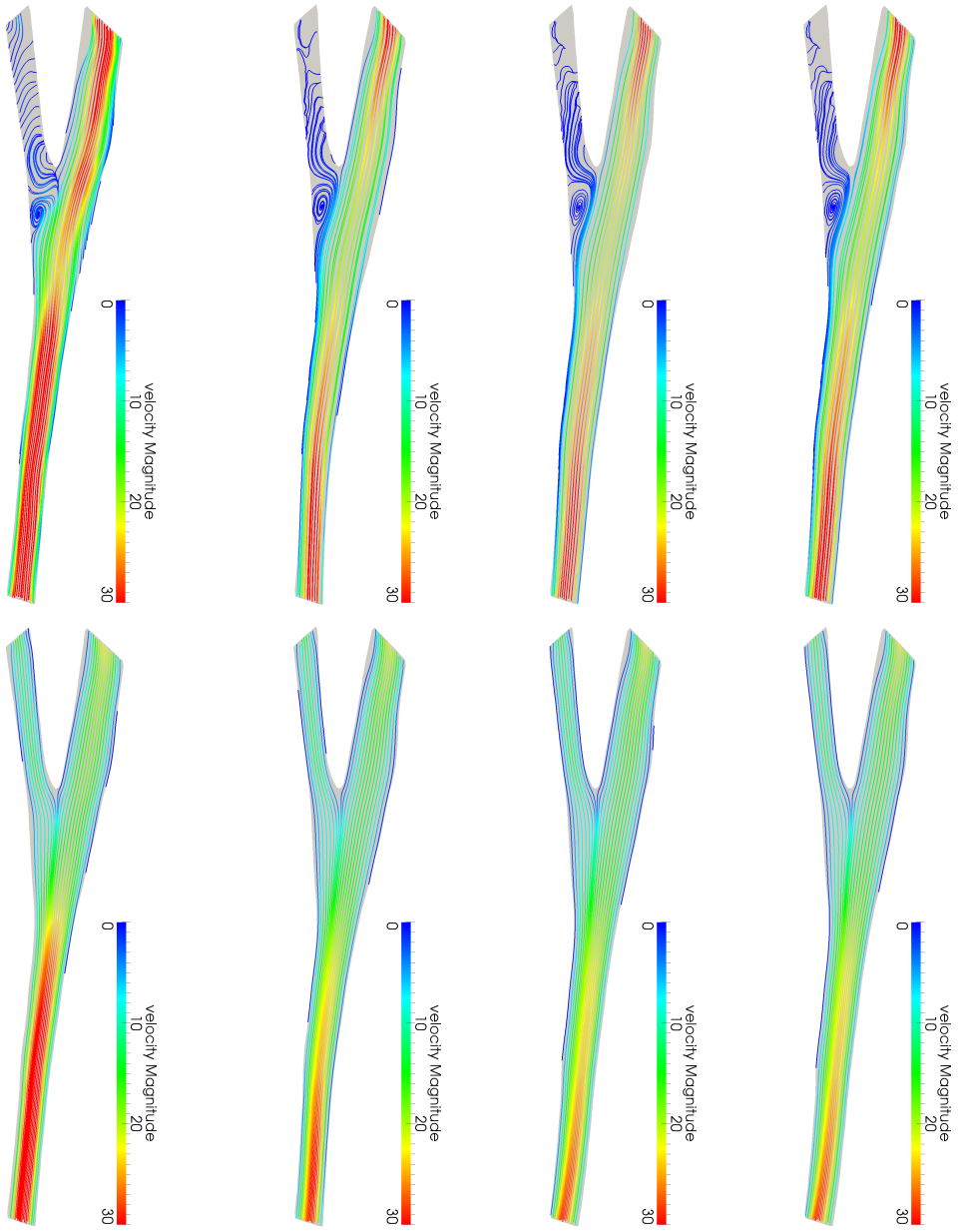



Figure 5.27: Velocity magnitude $[cm/s]$ and streamlines of the steady incompressible Navier-Stokes (RB simulation); from left to right: initial and optimal configurations obtained with shape optimization of $J_i$, $i = 3, 2, 4$, using the two extremal values of the residual flow parameter ($\omega = 0$ on top, $\omega = 15$ on bottom).

Concerning the solution of a robust SO problem under the form (RD-SO$\boldsymbol{\mu}$), we consider only the case of the functionals $J_3$ and $J_4$, since the behavior of the cost functional $J_2$ may be influenced more by the presence of the nonlinear convection than by the vorticity patterns for large values of $\omega$. As in the case of aorto-coronary bypass grafts, we find that the robust configurations correspond, for both cost functionals, to the optimal shapes computed for $\omega = 0$ in the previous case (see Fig. 5.27, columns 2 and 4). As in the case discussed in Sect. 5.4.2, the solution of the robust (shape) optimization problem requires a CPU time which is at least one order of magnitude larger than that used for a (shape) optimization problem (see Table 5.8).

| | # optim. iters | # output evals | # total PDE solves | CPU time |
|---|---|---|---|---|
| $J_3$ | 7 | 87 | 1,482 | 4.0 h |
| $J_4$ | 27 | 230 | 3,684 | 19.8 h |

Table 5.8: Numerical details for robust shape optimization. Here the number of output evaluations refers to the solution of the minimization problem (for the max $J_i(\boldsymbol{\mu})$ functional).

Here robust shape optimization problems took about $4 \div 20$ hours of computational time by using the RB method for the fluid simulation. Solving the same problem with the full FEM simulation would have taken considerably more (about 700 hours) if shape deformations had been handled by standard methods like local boundary variations and not taking advantage of the geometrical reduction afforded by the proposed shape parametrization.

We close this section by pointing out that in both cases we have considered, we found that a design that is robust over the entire range $\omega \in [0, \omega_{\max}]$ of (parametrized) residual flows through the occluded artery must be adjusted mainly for the case of total occlusion, i.e. $\omega = 0$.
In the case that total occlusion of the artery is not expected, we can exclude some region near the point of total occlusion and instead consider the range of uncertainty $\omega \in [\omega_{\min}, \omega_{\max}]$ for some $\omega_{\min} > 0$. In the typical case that the amount of occlusion changes over time, it is imperative that we should be able to monitor the amount of occlusion over time by simple noninvasive measurements to guarantee that, in fact, $\omega \geq \omega_{\min}$, since otherwise our bypass design will not be robust any more. This leads us naturally to consider the inverse problem of determining the magnitude of the residual flow, which is discussed in Sect. 6.5 of next chapter.

## 5.5 Remarks and future focuses

We have presented some optimal design problems aimed at improving the shape of cardiovascular prostheses. Two different worst-case optimization formulations have been proposed to solve the problem of bypass design under uncertainty, in order to reduce the downfield vorticity: (i) a boundary control formulation, which simplifies the geometry and treats only the angle of the anastomosis as a boundary control variable, and (ii) a shape optimization formulation relying on a FFD shape parametrization to represent the bypass anastomosis. Thanks to computational reduction allowed by the reduced basis method, the numerical solution of these problems can be greatly enhanced, above all in the *robust* optimization case.

In the optimal control case (Sect. 5.3) we have considered a simplified two-dimensional bypass configuration, parametrized with respect to the anastomosis angle and the residual flow through the occluded artery. Four different cost functionals taken from literature and suited for the reduction of downstream vorticity were studied. The optimal anastomosis angle was found to depend strongly on the total residual flow from the occluded branch, but not as strongly on the particular shape of the flow profile. Some numerical tests confirm the robustness of the obtained anastomosis angle with respect to the unknown residual incoming flows. We validated the 2D model by comparing the results obtained against a 3D boundary control problem. Three dimensional effects were found not to have a large impact on the total downstream vorticity at moderate Reynolds numbers. The largest vorticity was observed for the case of total occlusion in the host artery, so that a robust bypass shape should be tailored for that particular situation.

Similar conclusions were drawn, for two different bypass configurations, by considering a shape optimization problem (Sect. 5.4). In this case, a FFD shape parametrization provides a strong reduction in geometrical complexity – in term of the number of parameters, reduction is of about

100 times with respect to traditional shape parametrization based on local boundary variation – yet allowing to describe a wide family of shapes. Moreover, active control points can be selected according to suitable *ad hoc* procedures. Also in the shape optimization case, different cost functionals proposed for the reduction of downstream vorticity were studied. The condition leading to the strongest development of vorticity cores is the presence of a complete occlusion, for which the flow through the bypass starts creating a strong vortex in the proximity of the anastomosis. Regarding the robust shape optimization, we find in both the *aorto-coronary* and the *femoro-popliteal* case that a bypass should be designed mainly for the case of total occlusion, in order to be robust over a range of possible residual flows.

The models we have presented are very simplified, since they do not take into account important features like the pulsatility of the flow and the periodic detachment/separation of the flow layer from the wall. Extension of the proposed applications to *(i)* more realistic three-dimensional configurations and *(ii)* unsteady flows, already taken into consideration in several problems solved by traditional discretization schemes (also in presence of uncertainties), represent the focus of future research work, parallel to the development of the reduced basis methodology to more complex unsteady and/or three-dimensional fluid dynamics problems.

# 6 Inverse parametrized problems for blood flows

In this chapter we present a second group of applications dealing with inverse identification problems related to blood flows in parametrized geometrical configurations. Our goal is to investigate the dependence of simple blood flows by shape parameters using the reduced framework discussed in Parts I and II. After a brief survey on the interplay between blood flows and geometrical features of the vessels, we present some possible strategies to represent realistic vessel geometries based on suitable shape parametrization techniques. Thanks to shape parametrization, we can deal with both simple shape identification problems and more realistic parametric coupling problems for representing fluid-structure interactions. Then, three inverse identification problems are presented and discussed: *(i)* identification of arterial wall and flow properties aiming at atherosclerosis risk assessment in a deformable stenotic artery; *(ii)* identification of residual flows in a bypass model; *(ii)* identification of shape features related to pathologies. Both a deterministic and a statistical inverse framework are employed, fitting in both cases the many-query opportunity context which has driven the development of our reduced framework.

## 6.1 Parametrized problems of interest in haemodynamics

As already pointed out in chapter 6, a strong mutual interaction exists between haemodynamic factors and vessels geometry. On the one hand, flow adapts to the vessel shape; on the other hand, it also exerts forces affecting the vessel behavior and its morphology, resulting in a closed-loop mechanism. For instance, plaques occur preferentially in regions where arteries bend and in the vicinity of branches, but at the same time the changes in vessel shape due to this pathology clearly affect blood flows. Nevertheless, this interaction is far from being completely understood. Thus, improving the knowledge of the interplay between flows and geometries may be useful not only for the sake of design of better prosthetic devices, but also to characterize physical and geometrical properties of the flows which may be related for instance to pathological risks, such as in the case of narrowing or thickening of an arterial vessel.

It is therefore interesting to analyze the sensitivity of physical outputs related to viscous flows – such as wall shear stresses, wall shear stress gradients, vorticity, dissipated energy – with respect to shape morphology and possible variations. All these indices are influenced by lumen geometry, characterize local haemodynamics effects and may assess a risk of artery occlusion. Hence, rather than numerical simulations on a relatively small number of different configurations, it becomes crucial to explore a wide family of geometries – thus spanning a broad variety of shapes – in order to take into account their variability (for instance, among patients) and provide a more complete representation of blood flows and related outputs with respect to shape variation.

This may yield to the solution either of a *forward problem* – for any geometrical configuration, we wish to characterize the flow – or of a more complex *inverse problem* – where we wish to identify some properties, for instance related to the geometry, from data measurements. Thus, we need to face both with *real-time* problems (whenever, given a new geometry, we are interested to obtain a numerical approximation of the flow in a very small amount of time) and *many-query* contexts (because of the need to span a large set of configurations, or to solve an inverse problem), for which suitable model reduction techniques are requested, since full simulations may result very expensive if they have to be carried out from the beginning for each new geometry.

In general, whenever some parameters related to the flow can be isolated as representative features, parametrized frameworks and reduction techniques provide an excellent computational opportunity in view of sensitivity studies. Several works have focused on flow parameters such as Reynolds, Dean or Womersley numbers [58, 85], aiming at the description of three-dimensional characteristic flow patterns occurring where vessels bend, bifurcate or narrow. Other cases of significative interest can be the blood flow analysis with respect to inflow/outflow boundary conditions [20], or the sensitivity analysis of optimal shapes of artificial grafts with respect to flow parameters [323, 114, 252]. Since we basically deal with reduced basis simulations of two-dimensional flows, we do not take into account the effect of flow parameters such as Reynolds or Dean numbers. Nevertheless, the reduced framework can be easily extended in view of dealing with this kind of features [79, 168], provided that the truth approximation beneath can provide reliable numerical simulations of this kind of flows.
Besides parametric studies, inverse identification problems have been playing an ever increasing role in haemodynamics, mainly because of the huge amount of data (provided for instance through measures or images) that can be merged into a numerical model in order to get improved simulations. In this framework – the so-called *data assimilation* procedures [76] – several problems can be faced. However, our focus consists in exploiting the reduced framework analyzed in chapter 4 for the solution of inverse parametrized problems, where the goal is the identification of parameter values or combinations leading to some observed or measured flow conditions – for the time being, through the assimilation of some (surrogate) output of interest. Although restrictive (we will not deal neither with distributed observations of field variables, nor with the identification of physical or geometrical fields), this class of problems can be seen as a first step towards the integration of reduction strategies into more general inverse problems.

As motivated above, in the examples presented in this chapter we focus on parametrized flows with respect to inflow boundary conditions, to geometrical configuration of blood vessels and to physical properties related to the arterial wall. This choice is dictated by the capabilities of our current reduced approximation framework and motivated by the investigation of inverse problems related with shape variation fitting our parametrized context. In particular, we aim at:

1. providing a real-time simulator for blood flows in parametrized realistic geometries, under some modelling simplifications, possibly supported by some shape identification preprocessing stage;
2. performing parametric investigations of outputs of clinical interest over a range of possible configurations;
3. identifying flow and/or shape parametric features from surrogate measures of physical outputs through a suitable inverse identification process.

Since observations and measures are usually affected by experimental noise, we also want to deal with some basic issues arising in uncertainty quantification, for instance in order to quantify how these noise sources affect the predicted quantities of the simulations.
Results of the following sections are based on two works carried out in collaboration with Quarteroni, Rozza [205] and also with Lassila [180].

## 6.2 Reduced representation of realistic vessel geometries

The vascular networks are closed, tortuous systems of branching or merging junctions of vessels with widely different diameters and lengths. On the one hand, the complexity of the network arises from the biological functioning, the complicated structure and the large number of blood vessels; on the other hand, there is a huge variability between individuals in vessel shape, path and branching. These features make the numerical simulation of blood flows in realistic geometries a very challenging task. If mathematical modelling and numerical simulation of large portions of (or in the whole) the arterial network is almost impossible without resorting to a geometrical multiscale paradigm [203], in order to perform exploration studies related to the geometrical variability of a single arterial portion repeated full numerical simulations over a set of different configurations are in order. This may be helpful for instance in view of investigating the relationship between local geometrical variations and the occurrence of lesions or pathologies. Typical portions of cardiovascular network where they may develop are made up by curved vessels and bifurcations; an important segment where vessel diseases are often clinically observed is the human carotid artery, which supplies blood to the head and neck. Of the two common carotid arteries, which extend headward on each side of the neck, the left originates in the aortic arch, while the right originates in the brachiocephalic trunk (the largest branch departing from the aortic arch). Each common carotid artery (CCA) bifurcates in the lower neck into two branches (see Fig. 6.1), the internal and the external carotid arteries (ICA and ECA, respectively). Stenoses, that is the narrowing of the inner portion of an artery, manifest quite often in the ICA. It is well known that *(i)* the carotid shape yields a complex flow dynamics, *(ii)* its anatomy changes considerably among individuals, and *(iii)* the risk of strokes is directly related to carotid lesions severity. All what is reported above motivates the effort to characterize blood flows in this arterial tract.

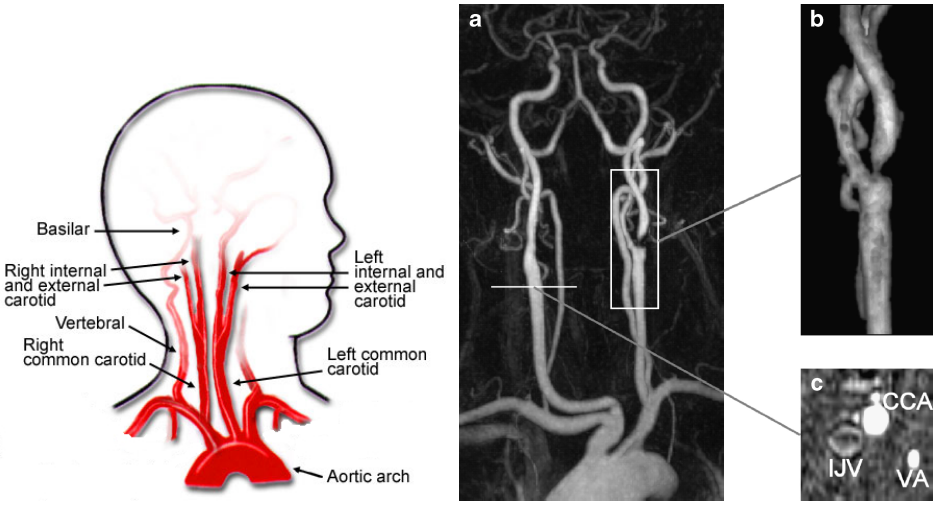



Figure 6.1: Left: schematic representation of the carotid artery bifurcation. Right: illustrative carotid magnetic resonance angiography (MRA), showing an 80% stenosis of the proximal left internal carotid artery, the right common carotid artery (CCA), vertebral artery (VA) and the internal jugular vein (IJV). Source: Eur. Radiol. (2009) 19: 2654–2662.

Several results have been obtained in the past two decades. We cite the first contributions by Ku and Giddens [171, 172] in the analysis of pulsatile flows in a first model carotid bifurcation, the works by Perktold [240, 241] on the geometrical factor in atherogenesis, and (among a very long list) the contributions by Kleinstreuer, Steinman [213, 204] and Lee [187], related to effects of the carotid artery geometry on blood velocity profiles and distribution of physical indices of interest.

Figure 6.2: Different carotid bifurcation specimens obtained by autopsy (adults aged 30-75). Source: J. Biomech. (2001) 34: 1555–1562.

Furthermore, we can observe a strong variability in geometry and morphology of carotid artery bifurcations (see Fig. 6.2) making unavoidable the setting of a numerical framework able to manage with local and global shape characteristics. Whether several works have been focused on the numerical simulation of blood flows in one or few carotid artery bifurcations – obtained directly by medical images, by using computational meshes reconstructed from patient data in a very precise way – only few works [45, 169, 187] have tried to investigate the relationship between infra-individuals variations of some risk factors and *geometrical indices*.
With this aim, some simplified shape parametrizations of the carotid bifurcation have been introduced, mainly based on sections areas or diameters, curvature of the branches, bifurcation angle and so on. In particular, we can mention (see Fig. 6.3):

- *Y model*: introduced in [35, 172] and derived from biplanar angiograms, it defines a carotid bifurcation as a function of the diameter $D$ of the CCA, the diameters $d_i$ and $d_e$ of the ICA and ECA outflows, the diameters at the proximal section $d_r$, mid-sinus section $d_s$ of the ICA, the distance $l_s$ between proximal and midsinus sections, the angles $\alpha_i$, $\alpha_e$ between the ICA (resp. ECA) and the CCA. A second version, using inwardly curved branches instead of straight branches, has been discussed in [241]; here a further parameter $d_m$ representing the mid-sinus cross section is introduced, whereas $d_s$ represents the diameter of the section where the maximum sinus width occurs. Because of the curvature, two angles between CCA and ICA are considered, at the origin of the branching ($\alpha_{i,o}$) and at its end ($\alpha_{i,e}$).
- *Tuning fork model*: introduced in [83] and derived from carotid specimens obtained from autopsy, it represents the carotid bifurcation like a tuning fork. The geometrical parameters are in this case the five diameters $D$, $d_i$, $d_e$, $d_r$ and $d_s$, the bifurcation angle $\alpha = \alpha_i + \alpha_e$, the distance $l_s$ (as in the previous case), the distance between the two branches $L$, and two more lengths representing the sectional diameter $d_u$ of the ICA at the end of the sinus and the distance $G$ of this section from the divider tip. A parametric CAD model based on the tuning fork model has been introduced in [46, 45] and is based on the following parameters: the ICA and ECA angles $\theta_i$, $\theta_e$, the ICA and ECA diameters $d_i$, $d_e$, the position $d$ of a Bézier curve control point $P$, which enables to control the maximum width of the sinus bulb. Also the inner curve of the bifurcation is described by means of a *Bézier* curve.

With respect to these models, our approach to define a shape parametrization is based on a free choice of the *input* parameters, which are given by the displacements of some control points located on the boundary of a reference carotid configuration; hence, deformed shapes will be simply obtained by moving the control points. In this way, we get a versatile shape parametrization, since control points may be located either *a)* on peripheral positions on the branches for describing global deformations or *b)* on selected positions for modeling local modifications such as stenoses or restrictions.

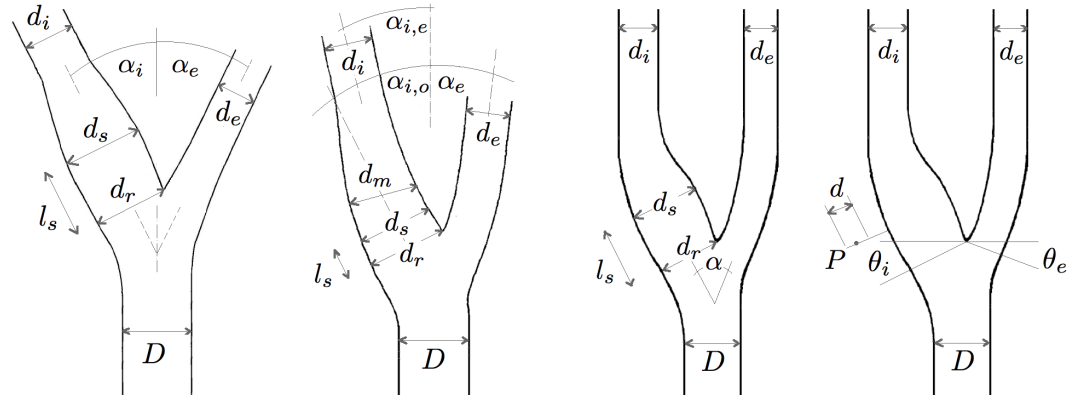



Figure 6.3: Parametrized models of the carotid artery bifurcation. From left to right: *Y model* with straight branches, *Y model* with curved branches, *tuning fork model* and parametric tuning fork *CAD model*.

Moreover, since we are interested in performing real-time blood flow simulations (see Sect. 6.3), the shape parametrization has to be low-dimensional. For these reasons, a parametrization based on radial basis functions (RBF) is a good compromise: despite its low-dimensionality, its interpolatory nature enables to track some selected points on the shape boundary, and to account for both local and global deformations depending on the chosen basis function (see Sect. 2.7).

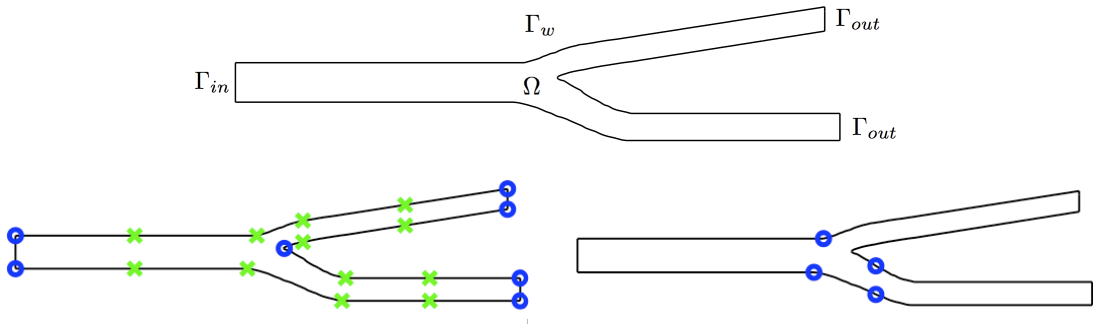



Figure 6.4: Radial basis function shape parametrization models of the carotid artery bifurcation. Top: reference configuration. Bottom: control points (represented as blue dots) for the definition of global (left) and local (right) shape deformations. Green crosses represent registration points needed for shape reconstruction, see Sect. 6.3.1.

In particular, we deal either with *global* and *local* deformations of the two-dimensional reference configuration represented in Fig. 6.4, obtained by displacing a set of control points in the cross direction with respect to the vessel (i.e. in the vertical direction in Fig. 6.4). The reference shape has been obtained as a longitudinal two-dimensional projection of a carotid configuration extracted from medical images and post-processed. For instance, to get global shape deformations, $p = 7$ control points have been located at the extremities of the three branches and cubic RBFs have been used in order to define the global parametric map; instead, to get local shape deformations, $p = 4$ control points have been located in one of the branches and close to the bifurcation, and Gaussian RBFs have been used in order to describe local but moderate deformations representing possible stenoses. Some representative deformations obtained within this RBF setting by are shown in Fig. 6.5. We remark that, instead of a fixed, *a priori* chosen location of the control points, we may select a good set of control points apt to represent a given *atlas* of shapes (e.g. obtained from MRA or Doppler images), e.g. through a *greedy-like* procedure, starting from a suitable reference *averaged* configuration. This possibility is currently under investigation [182].

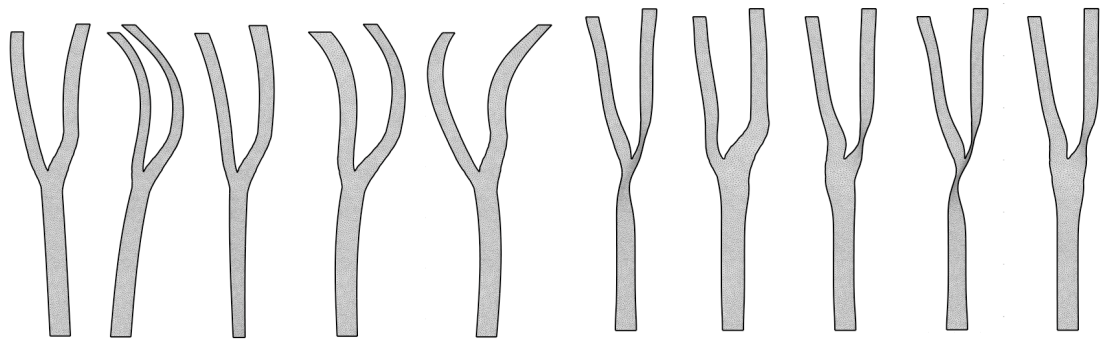


Figure 6.5: Different carotid bifurcations obtained through the RBF parametrizations by displacing the control points (left: global deformations, right: local deformations).

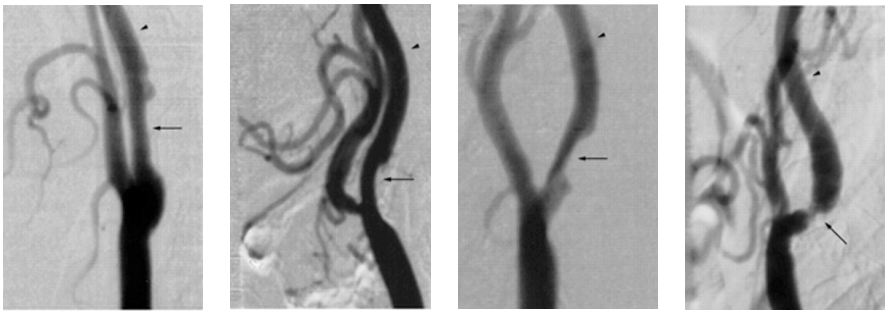


Figure 6.6: Doppler sonography of carotid stenoses. Arrow indicates point of maximum stenosis, and arrowhead indicates normal distal internal carotid artery. Source: Americ. J. Neuroradiol. (2000) 21: 639–642.

In the end, we point out that, even if RBF-based shape parametrizations are low dimensional and defined through a general interpolation tool, they allow to represent complex shapes, approaching some features shown by true vessel configurations, in case both of global large deformations (see Fig. 6.2) and of smaller, localized deformations (see Fig. 6.6). The capability to represent complex and realistic shapes by few control points, and corresponding deformations through a small set of parameters, is a strong key point in view of rapid flow simulations.

## 6.3 Rapid blood flow simulations in parametrized vessels

Ideally, numerical simulations of blood flows ought to be carried out in a very fast way. This is relevant to provide a quantitative output/response for each new (possibly patient-specific) geometrical configuration and derive some metamodel able to explain the dependence of some output of clinical interest with respect to simple shape or flow features.
We show how to take advantage of the reduced framework to address these goals. We present two *snapshot* applications related to a set of carotid artery bifurcations, dealing with *(i)* simulation in simplified reconstructed geometries, where each configuration can be easily represented as a deformation of a reference configuration, and *(ii)* a sensitivity analysis with respect to shape parameters. A recent application to fast patient-specific simulations close to the former case can be found in [212], while some results related to the latter are presented in [45, 169].
We remark that shape sensitivity analyses may be relevant in this context also for stenting procedures, where carotid endarterectomy may require the possibility to modify or reconstruct vessel geometry in order to optimize the flow characteristics therein.

### 6.3.1 Shape parameter identification

We address in this section a simplified instance of geometrical reconstruction from data, ideally suited for a possible application of the proposed methodology in a patient-specific framework. Shape reconstruction consists of recovering a transformation that establishes some desired correspondences between two geometrical configurations, according to some similarity/distance measures. This process, and more generally, representation of shapes and their deformations, is central in many applications of image processing and medical image analysis, such as segmentation (extraction of geometrical structures from images), tracking (recovering the temporal evolution from consecutive frames), or registration.

Our (limited) ambition is to set a (much simpler) preliminary procedure in order to reconstruct a parametrized version of a given shape configuration, by means of the RBF setting presented in the previous section, on which we aim to perform a flow simulation. In order to be compatible with the reduced real-time framework, the reconstruction procedure has to be *(i)* based on a small number of parameters and *(ii)* performed in a very small amount of time. Let us denote by $\mathcal{S}$ the initial shape which has to be deformed in order to reconstruct the target shape $\mathcal{T}_d$ by means of the parametric map $T(\cdot;\boldsymbol{\pi})$. Moreover, let us introduce a shape representation $\mathcal{R}(\cdot)$, a distance (or similarity/dissimilarity metric) $d(\cdot)$ defined in the space of representations and a loss function $\rho(\cdot)$. The reconstruction process can be in general expressed as a minimization problem:

$$\hat{\boldsymbol{\pi}} = \arg\min_{\boldsymbol{\pi}\in\mathcal{D}_{ad}} \rho(d(\mathcal{R}(T(\mathcal{S};\boldsymbol{\pi})),\mathcal{R}(\mathcal{T}_d))) + \beta\omega(\boldsymbol{\pi}) \tag{6.1}$$

where $\mathcal{D}_{ad}$ is the space of admissible parameters, $\omega(\boldsymbol{\pi})$ is a suitable regularization term and $\beta > 0$ is a weighting (regularization) parameter. A usual choice is $\rho(s) = s^2$ for the loss function and the Euclidean distance for $d(\mathbf{x})$; more complex transformations or weighting can be introduced in order to improve the robustness of the matching.
The most difficult task is the choice of the shape representation $\mathcal{R}(\cdot)$. Many strategies have been developed (see e.g. [216, 38]). In view of model reduction, we represent a shape configuration by means of a set of $N_R$ *registration points* (or landmarks) $\{\mathbf{R}_j\}_{j=1}^{N_R} \in \mathcal{S}$, such that

$$\mathcal{R}(T(\mathcal{S};\boldsymbol{\pi})) = T(\mathbf{R}_j;\boldsymbol{\pi}), \quad j = 1,\dots N_R;$$

in the same way, we suppose to know the target position

$$\mathcal{R}(\mathcal{T}_d) = \mathbf{R}_j^o, \quad j = 1,\dots,N_R$$

of the registration points in the target configuration $\mathcal{T}_d$. Moreover, we introduce a set of control points $\{\mathbf{P}_i\}_{i=1}^{N_C}$ on the initial shape in order to define a parametric RBF mapping $T(\cdot;\boldsymbol{\pi})$, whose image is given by $\{\mathbf{P}_i^o\}_{i=1}^{N_C}$ (see Fig. 6.7).

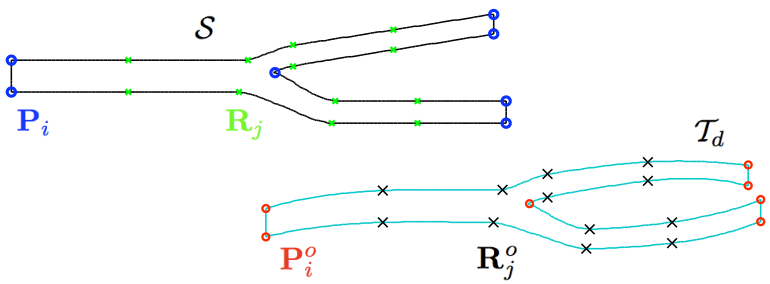



Figure 6.7: Schematic diagram of the shape reconstruction: above, the reference (or initial) configuration $\mathcal{S}$, the unperturbed control points $\mathbf{P}_i^o$, $i = 1,\dots,N_C$ (in black), and the unperturbed registration points $\mathbf{R}_j$, $i = 1,\dots,N_R$ (in green); below, the target configuration $\mathcal{T}_d$, the target control points $\mathbf{P}_i^o$ (in red), and the target registration points $\mathbf{R}_j^o$ (in black).

The idea is that, since the RBF mapping may fail in matching the two configurations, the registration points can be used to drive the mapping in order to enhance its fitting capabilities. In this way, the parameter identification problem (6.1) becomes a matching problem between two point-sets, that can be written as a *least squares* minimization problem

$$\hat{\boldsymbol{\pi}} = \arg \min_{\boldsymbol{\pi} \in \mathcal{D}_{ad}} \left( \sum_{j=1}^{N_R} \| T(\mathbf{R}_j; \boldsymbol{\pi}) - \mathbf{R}_j^o \|^2 + \beta \sum_{i=1}^{N_C} \| T(\mathbf{P}_i; \boldsymbol{\pi}) - \mathbf{P}_i^o \|^2 \right), \tag{6.2}$$

where the regularization term (a quadratic function of $\boldsymbol{\pi}$) is given by the distance between the images $\{T(\mathbf{P}_i; \boldsymbol{\pi})\}$ of the control points and their target positions $\{\mathbf{P}_i^o\}_{i=1}^{N_C}$. Note that the introduction of the registration points does not increase the problem dimension, as the number of input parameters of shape representation (related to the control points of the map) remains unchanged. Moreover, we require that control points are scattered all over the domain in order to describe a wide family of shapes and global deformations.

In particular, we build a RBF parametrization by exploiting the first setting presented in Sect. 6.2 (cubic RBF, $N_C = 7$ control points on the branches and $p = 7$ input parameters $\pi_1, \ldots, \pi_7$, given by the vertical displacements of the control points, being $\pi_i \in [-0.2, 0.2]\ \forall\ i = 1, \ldots, 7$). Concerning the shape reconstruction, the parameter identification procedure (6.2) has been tested on 20 different configurations[1], by considering $N_R = 12$ registration points, divided in 6 couples, and placing two couples of points on each carotid branch. The minimization of the cost functional (6.2) takes (average over the 20 run cases) $5.3\,s$ by SQP optimization. We remark that the choice of the weighting parameter $\beta > 0$ in (6.2) plays an important role (see Fig. 6.8): too small values do not allow a good control of the shape at the extremities, whereas penalizing too much the matching constraint on the control points results in large deformations on the central sections. A possible way to select a good weighting parameter $\beta > 0$ might be based on the minimization of a second cost functional, i.e. we may select

$$\hat{\beta} = \arg \min_{[\beta_-, \beta_+]} L(\beta),$$

where $L : [\beta_-, \beta_+] \to \mathbb{R}$ is a simple objective to evaluate. Two examples we considered are: *(i)* $L_1(\beta) = J(\hat{\boldsymbol{\pi}}, \beta)$, where $J(\boldsymbol{\pi}, \beta)$ and $\hat{\boldsymbol{\pi}} = \hat{\boldsymbol{\pi}}(\beta)$ are the cost functional and the solution of problem (6.2); the following objective:

$$L_2(\beta) = \left( 1 - \frac{E(\mathbf{d}(\hat{\boldsymbol{\pi}}(\beta)))}{E(\mathbf{d}_T)} \right)^2,$$

where $E(\cdot) > 0$ is the elastic energy of the deformation, $d(\boldsymbol{\pi}) = T(\mathbf{x}; \boldsymbol{\pi}) - \mathbf{x}$ the deformation produced by the mapping $T(\cdot; \boldsymbol{\pi})$ and $\mathbf{d}_T$ the deformation of the target shape with respect to the reference configuration (if available). We experienced better reconstruction results by using the functional $L_2(\cdot)$, often resulting in a convex function of $\beta$. The two configurations reported in Fig. 6.8 refer to the $\hat{\beta}$ values computed by using these two criteria on the same test configuration.

Concerning flow simulation, we use a steady Navier-Stokes model by imposing (see Fig. 6.4) no-slip conditions on the wall boundary $\Gamma_{\mathrm{w}}$, homogeneous Neumann conditions (free-stress) on the outflow sections $\Gamma_{out}$ and a Poiseuille profile $\mathbf{v}_{\mathrm{in}}$ on the inflow boundary $\Gamma_{in}$, being $\max |\mathbf{v}_{\mathrm{in}}| = 30 cm/s$. Resulting Reynolds number is of order 300.
Before constructing the RB approximation of the flow problem, we recover the affinity assumption through the EIM procedure, by obtaining an affine expansion of $Q_a + Q_b + Q_c = 78$ terms. The problem is discretized with $\mathcal{N}_V + \mathcal{N}_Q \approx 26{,}000$ degrees of freedom, using $\mathbb{P}_2/\mathbb{P}_1$ finite elements;

[1]The (surrogate) target configuration to be reconstructed has been generated as a random FFD of the reference geometry (built on a $8 \times 6$ lattice of control points) in order to avoid a scaling/translation preprocessing.

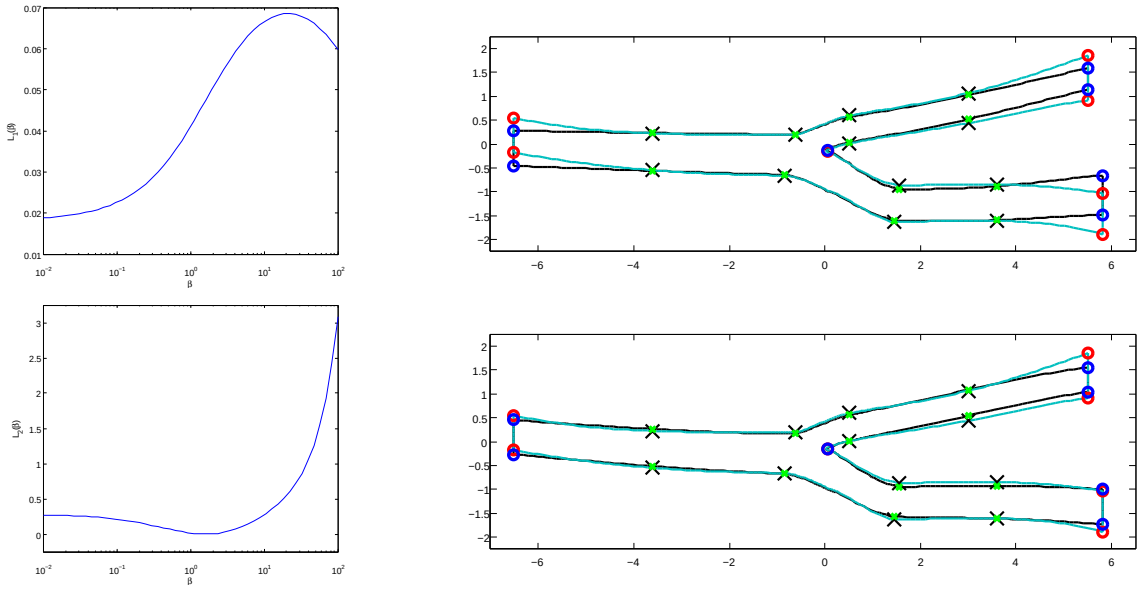


Figure 6.8: Shape parameter identification with $\beta = 0.01$ (top) and $\beta = 1$ (bottom); these values of $\beta$ result from the minimization of $L_1(\beta)$ and $L_2(\beta)$ objectives, respectively.

the dimension of the computed RB space is $3N = 30$ by choosing a tolerance $\varepsilon_{tol}^{RB} = 2.5 \cdot 10^{-3}$ for the greedy procedure. This yields the possibility to perform almost real-time (steady) blood flow simulations, each query $\boldsymbol{\pi} \to Y_N(\boldsymbol{\pi})$ requiring $\mathcal{O}(1\,s)$. In the case at hand, solving the nonlinear RB Navier-Stokes problem requires about $2.3\,s$, thus resulting in overall Online costs (for shape reconstruction and flow simulation) of $\mathcal{O}(10\,s)$.

This example may be seen as an extremely simplified version of a possible application of reduced order modelling in this context. Accounting for the storage of a large dataset of patient-dependent shape configurations (from medical images) and corresponding flow fields obtained by numerical simulation (*Offline* or pre-deployment stage), the evaluation of the flow field (and related outputs of clinical interest) for a new configuration can be afforded during the *Online* (or deployed) stage thanks to the reduced-order model. In this way, the Offline stage requires huge computational power or storage capacity, but is run before the Online stage. The Online task is thus affordable and capable to provide the outcome of interest in a very rapid time.

### 6.3.2 Sensitivity analysis of blood flows and metamodels

A second natural application of our reduced framework deals with sensitivity analysis of blood flows with respect to physical/geometrical input parameters and the possibility to draw simple metamodels to describe input/output relationships.
Here we present two possible examples, dealing with parametrized flows through *(i)* an idealized bypass configuration and *(ii)* a carotid bifurcation in presence of growing stenoses.

#### Case I. Parametrized flows through an idealized bypass configuration

In the first case, we exploit the geometrical setting already introduced in Sect. 5.3 for the analysis of an optimal control approach to the bypass design problem (see Fig. 5.7). Steady Navier-Stokes flows are considered through an idealized bypass configuration, parametrized with respect to the Reynolds number $\mu_1 = Re \in [10, 80]$ and the magnitude of the residual flow $\mu_2 = \omega \in [0, 2.5]$.

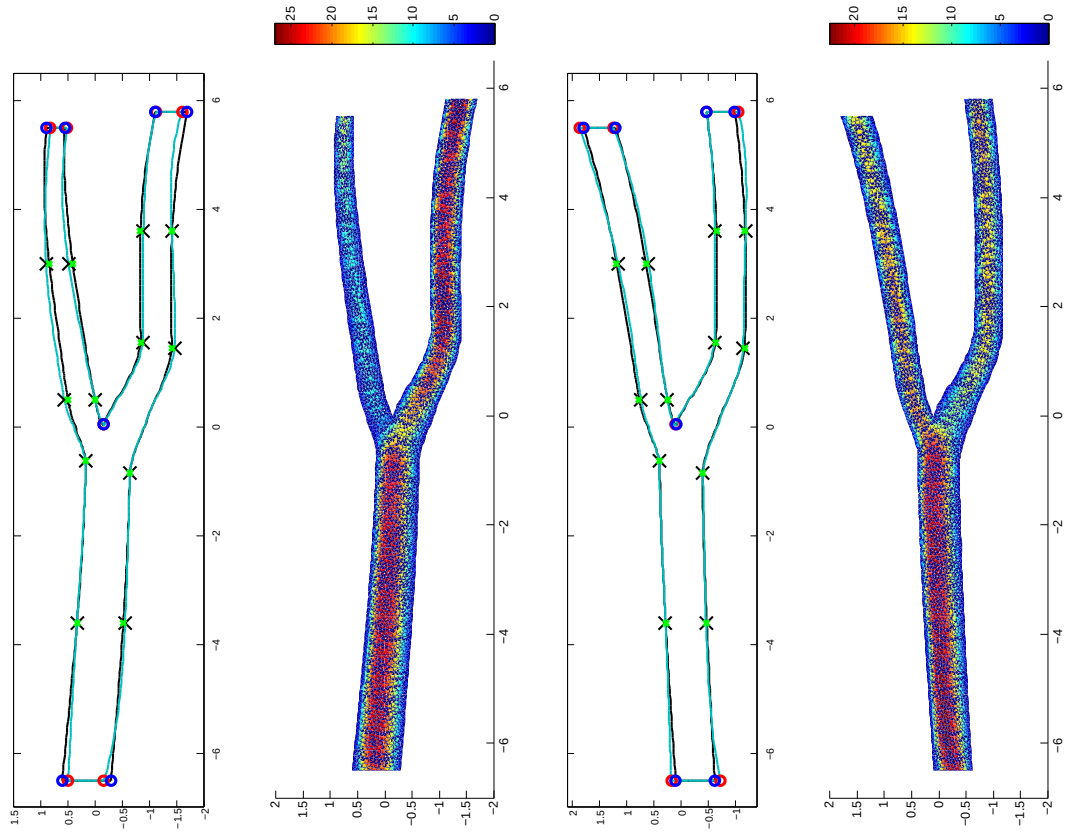


Figure 6.9: RB approximation of blood flows in two different reconstructed configurations (target and reconstructed shapes, velocity and pressure fields).

We are interested in evaluating the minimum downstream shear rate:

$$\mathcal{J}(\boldsymbol{\mu}) = \nu \min_{x_1 \in \Gamma_{\text{wall}}} \left\{ \frac{\partial v_1}{\partial x_2}\Big|_{x_2=0} \right\}, \tag{6.3}$$

where in this case $\Gamma_{\text{wall}} = \{(x_1, x_2) : x_1 \in (0, 5),\, x_2 = 0\}$. We exploit the RB approximation already analyzed in Sect. 5.3.1 to evaluate the flow field and the output $\mathcal{J}(\boldsymbol{\mu})$ over the parametric set $\mathcal{D} \in [10, 80] \times [0, 2.5]$. Two examples of computed RB solutions are reported in Fig. 6.10. We remark the different position of the minimum shear rate in the two cases.
Thanks to the rapid Online evaluation afforded by the RB approximation, we can evaluate the flow field and the output over a test sample $\Xi_{\text{test}}$ randomly selected over $\mathcal{D}$ of size $n_{\text{test}} = 500$ in 212 seconds, thus giving an average time of 0.42 seconds for each input/output evaluation $\boldsymbol{\mu} \to \mathcal{J}_N(\boldsymbol{\mu})$. We also compute the $x_1$ position of the minimum shear rate for each case. The results are reported in Fig. 6.11. We can observe a stronger sensitivity of the output with respect to the parameter $\mu_2 = \omega$; moreover, the $x_1$ location of the minima cluster in two groups, for $\mu_2 = \omega < 0.5$ and $\mu_2 = \omega > 0.5$. Recalling the results presented in Sect. 5.3.2, this behavior can be related to the vorticity pattern. In fact, the primary vortex behind the incoming jet tends to disappear as we increase $\omega$, as soon as secondary vortices generated by the peak residual flow from the occluded branch appear: this explains why for $\omega < 0.5$ the $x_1$ position of the minimum is moved back, while it is moved forward for $\omega > 0.5$.

In order to explain the behavior of this simulation model – in terms of a functional relationship between the output and the input parameters – we consider a *metamodel* computed through a response surface method (RSM), which is based on low-order *polynomial regression* and aims at

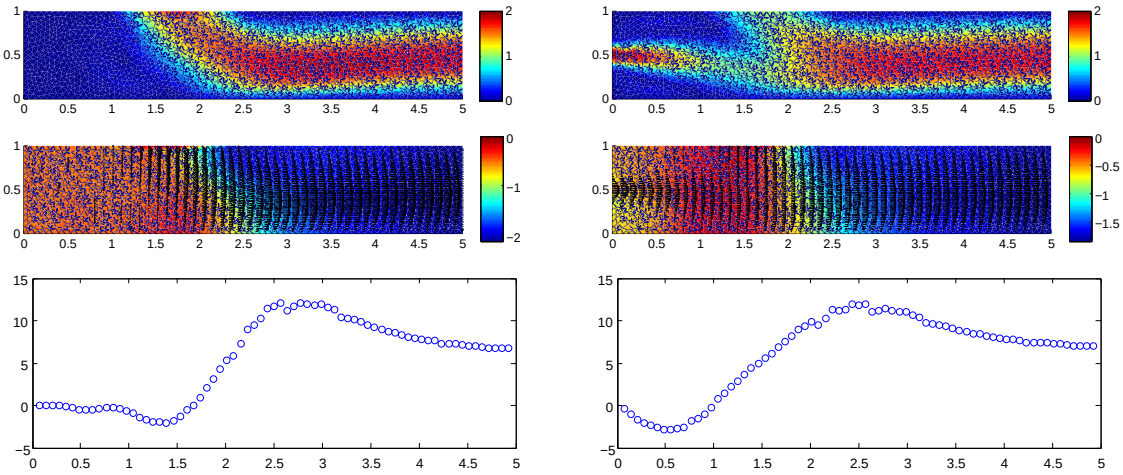


Figure 6.10: Velocity profiles $[cm/s]$, pressure fields $[dyn/cm^2]$ and shear rate on $\Gamma_{\text{wall}}$ for the idealized bypass configurations: $\boldsymbol{\mu} \in (60, 0)$ (left) and $\boldsymbol{\mu} = (80, 2.5)$ (right).

representing the output (or *response*) surface as a polynomial function of the input parameters. From Fig. 6.11 it is rather evident the presence of a *curvature* effect in the response. Thus, we approximate the relationship $\boldsymbol{\mu} \to \mathcal{J}_N(\boldsymbol{\mu})$ through a *second-order* polynomial regression model:

$$\mathcal{J}_{RSM}(\boldsymbol{\mu}) = \beta_0 + \sum_{i=1}^{p} \beta_i \mu_i + \sum_{i=1}^{p} \beta_{ii} \mu_i^2 + \sum_{\substack{i=1,j=1 \\ i<j}}^{p} \beta_{ij} \mu_i \mu_j + \varepsilon_{reg},$$

where $\beta_0, \ldots, \beta_{pp}$ are the coefficient of the metamodel and $\varepsilon_{reg}$ is the error term including the lack of fit of the RSM surface. The response surface obtained by fitting[2] a quadratic model over

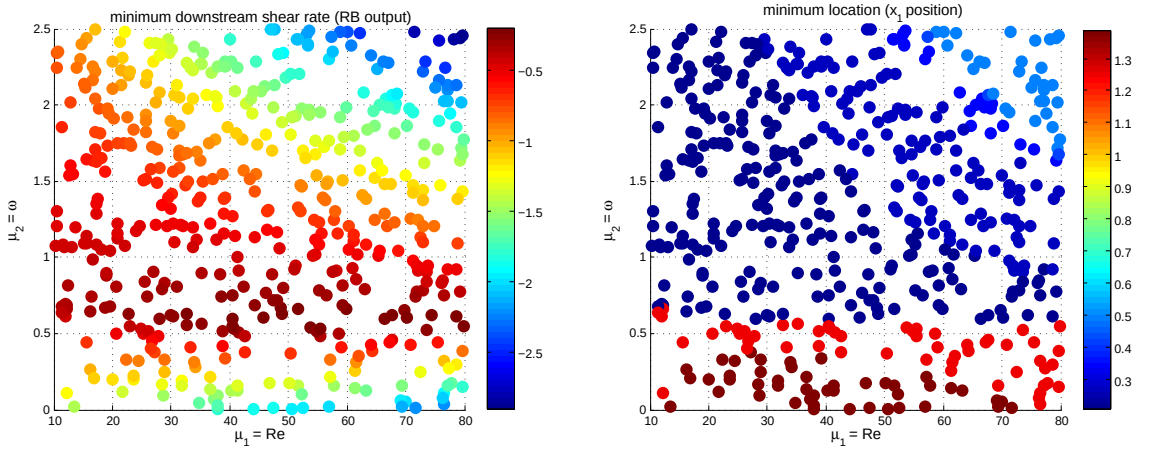



Figure 6.11: Computed values (left) and $x_1$ positions (right) of the minimum shear rate $\mathcal{J}_N(\boldsymbol{\mu})$ for a random sample $\Xi_{\text{test}} \subset \mathcal{D}$ of size $n_{\text{test}} = 500$ (RB approximation).

[2]Several design optimality criteria are available for choosing the input combinations where we intend to simulate the output values, where the response surface is fitted, in order to get an accurate estimation of the RSM coefficients. We do not take care of this issue in these illustrative examples.

a randomly selected train sample $\Xi_{\text{train}} \subset \Xi_{\text{test}}$ of size $n_{\text{train}} = 25$ is given by

$$\mathcal{J}_{RSM}(\boldsymbol{\mu}) = -1.328 - 0.019\mu_1 + 2.046\mu_2 - 0.0085\mu_1\mu_2 + 0.0001\mu_1^2 - 0.740\mu_2^2,$$

and is represented in Fig. 6.12, together with the relative error $(\mathcal{J}_N(\boldsymbol{\mu}) - \mathcal{J}_{RSM}(\boldsymbol{\mu}))/\mathcal{J}_N(\boldsymbol{\mu})$ between the RB approximation and the RSM metamodel. The RSM model captures the larger sensitivity of the response with respect to the parameter $\mu_2 = \omega$. We observe that a quadratic response surface provides a good approximation of the input/output relationship, except for a limited region in the parameter space – $\mu_2 \in (0.5, 1)$ – corresponding to the transition between the two clusters highlighted above.

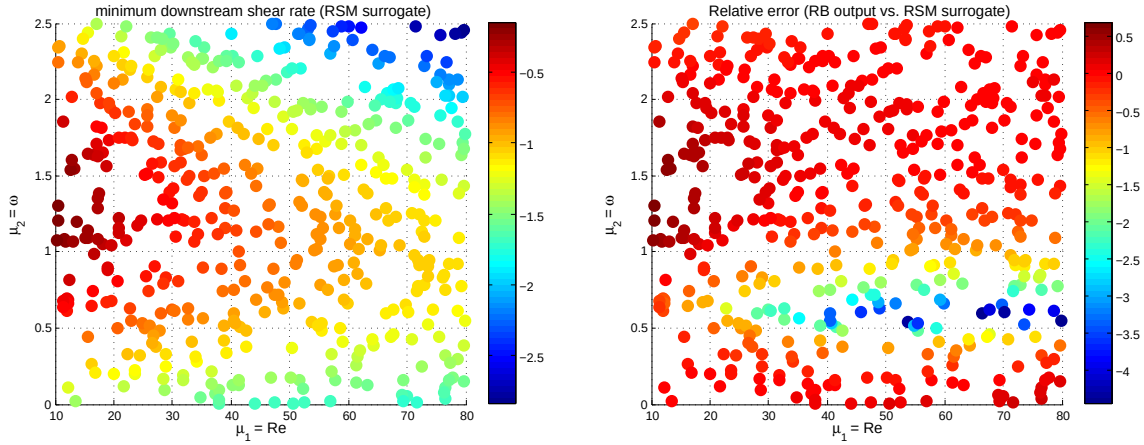



Figure 6.12: Minimum shear rate $\mathcal{J}_{RSM}(\boldsymbol{\mu})$ computed through a response surface method of degree 2 (left) and relative error between the RB approximation and the RSM metamodel (right) for a random sample $\Xi_{\text{test}} \subset \mathcal{D}$ of size $n_{\text{test}} = 500$.

### Case II. Parametrized flows through a set of carotid artery bifurcations

A second case we have considered deals with flows through a carotid artery bifurcation in presence of growing stenoses. We exploit the RBF parametrization introduced in Sect. 6.2 (see Fig. 6.5, right) for local deformations (Gaussian RBF with 4 control points close to the bifurcation and $p = 4$ input parameters $\pi_1, \dots, \pi_4$, given by the vertical displacements of the control points, being $\pi_i \in [-0.25, 0.25]\ \forall\ i = 1, \dots, 4$).

Before constructing the RB approximation of the flow problem, we recover the affinity assumption through the EIM procedure, by obtaining an affine expansion of $Q_a + 2Q_b + Q_c = 62$ terms. The problem is discretized with $\mathcal{N}_V + \mathcal{N}_Q \approx 26{,}000$ degrees of freedom, using $\mathbb{P}_2/\mathbb{P}_1$ finite elements; the dimension of the computed RB space is $3N = 45$ by choosing a tolerance $\varepsilon_{tol}^{RB} = 2.5 \cdot 10^{-3}$ for the greedy procedure. In the case at hand, Online evaluations of the Navier-Stokes solution takes about $2.5\,s$. Four examples of computed RB solutions are reported in Fig. 6.13. We remark the strong sensitivity of the flow with respect to varying diameters $d_c = d_c(\pi_1, \pi_2)$ of the CCA at the bifurcation and $d_b = d_b(\pi_3, \pi_4)$ of the mid-sinus level of the ICA, respectively.

Here we are interested in evaluating the viscous energy dissipation:

$$\mathcal{J}(\boldsymbol{\pi}) = \frac{\nu}{2} \int_{\Omega_o(\boldsymbol{\pi})} |\nabla \mathbf{v}(\boldsymbol{\pi})|^2 d\Omega_o \tag{6.4}$$

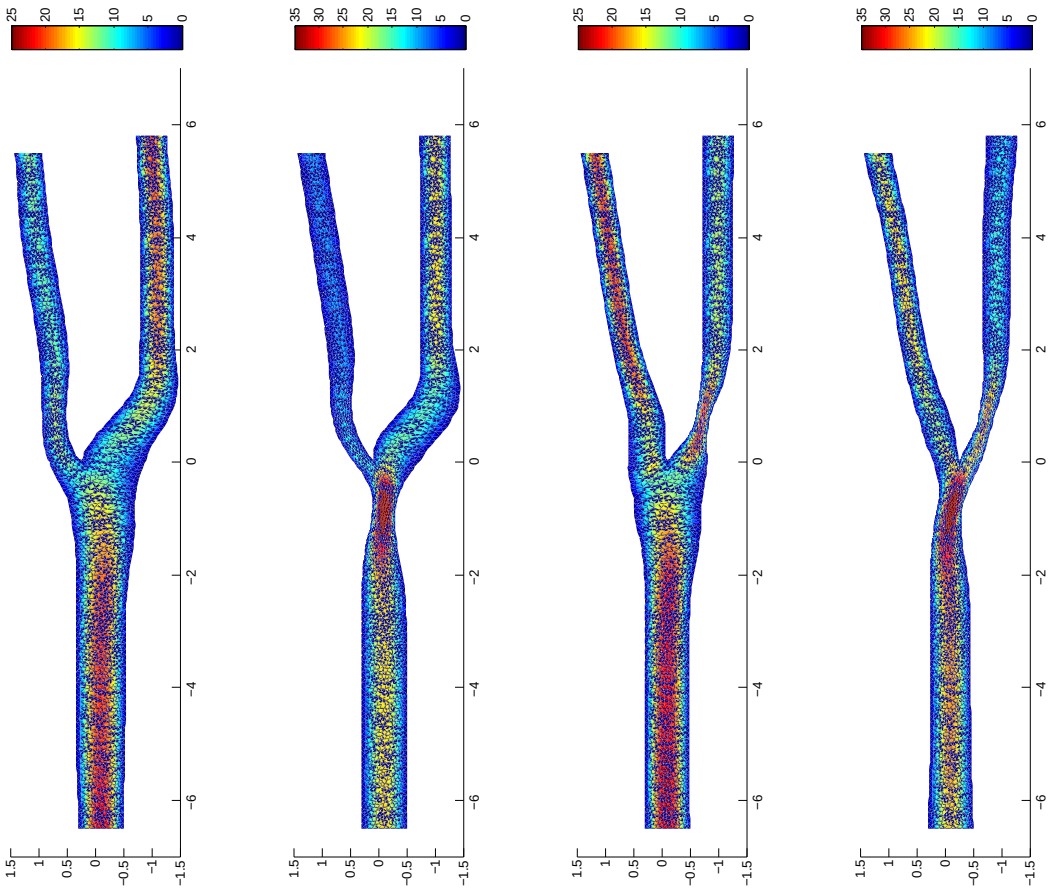


Figure 6.13: Velocity profiles $[cm/s]$ in four different carotid bifurcations parametrized with respect to the diameters $d_c = d_c(\pi_1, \pi_2)$ of the CCA at the bifurcation and $d_b = d_b(\pi_3, \pi_4)$ of the mid-sinus level of the ICA.

and to analyze its behavior with respect to the diameters $d_c = d_c(\pi_1, \pi_2)$ and $d_b = d_b(\pi_3, \pi_4)$ of the two branches in presence of different degrees of occlusion. We consider symmetric variations for each couple of parameters tuning a diameter, i.e. $\pi_2 = -\pi_1$ and $\pi_4 = -\pi_3$. In Fig. 6.14 we display the viscous energy dissipation as a function of the two diameters, obtained from the RB flow simulation over a test sample $\Xi_{\text{test}}$ of $n_{\text{test}} = 500$ parameter values. As expected, flow disturbances caused by stenoses lead to higher values of the dissipated energy, the maximum occurring for the smallest diameters on both sections.
In order to better understand the role of the two diameters on the resulting output, we consider a *metamodel* computed through a quadratic RSM as in case I. The response surface fitted over a randomly selected train sample $\Xi_{\text{train}} \subset \Xi_{\text{test}}$ of size $n_{\text{train}} = 20$ is given by

$$\mathcal{J}_{RSM}(\boldsymbol{\pi}) = 0.1637 - 0.1201 d_c - 0.0713 d_b - 0.003 d_c d_b + 0.0479 d_c^2 + 0.0353 d_b^2,$$

and is represented in Fig. 6.15, together with the relative error $(\mathcal{J}_N(\boldsymbol{\mu}) - \mathcal{J}_{RSM}(\boldsymbol{\mu}))/\mathcal{J}_N(\boldsymbol{\mu})$ between the RB approximation and the RSM surface.

The RSM model approximates very well the behavior of the input/output relationship (relative error is less than 0.05 over the whole diameters range). Output variations seem to be more sensitive to the variations of the diameter $d_c$ of the common carotid artery (closer to bifurcation) than to the variations of the diameter $d_b$ of the internal carotid artery.

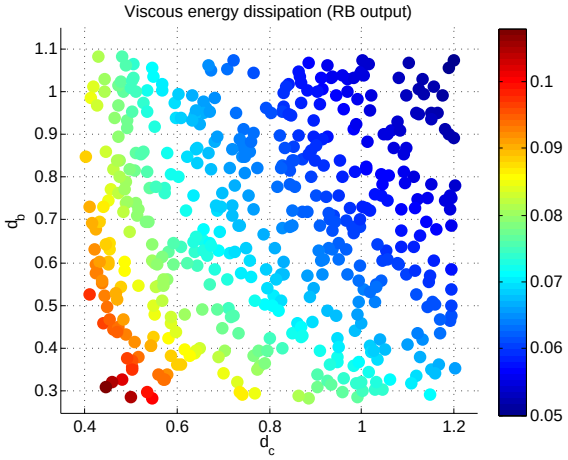



Figure 6.14: Computed values of the viscous energy dissipation $\mathcal{J}_N(\boldsymbol{\mu})$ for a random sample $\Xi_{\text{test}} \subset \mathcal{D}$ of size $n_{\text{test}} = 500$ (RB approximation).

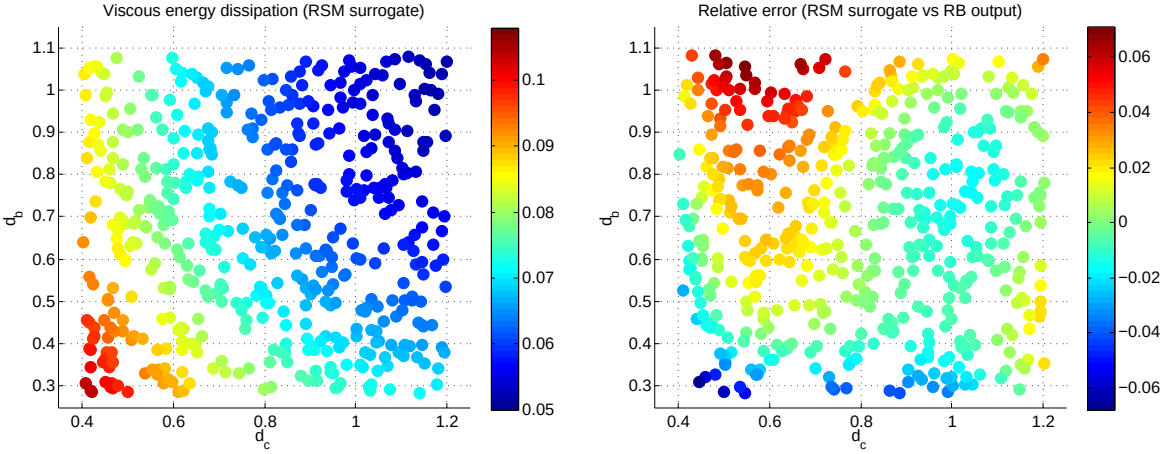



Figure 6.15: Viscous energy dissipation $\mathcal{J}_{RSM}(\boldsymbol{\mu})$ computed through a response surface method of degree 2 (left) and relative error between the RB approximation and the RSM metamodel (right) for a random sample $\Xi_{\text{test}} \subset \mathcal{D}$ of size $n_{\text{test}} = 500$.

The forthcoming sections are devoted to the solution of three inverse problems related to applications of haemodynamics interest. Based on the measure of a pressure output, we try to recover the structural coefficients in a steady FSI model (Sect. 6.4), the magnitude of the residual flow through a bypass graft (Sect. 6.5) and the diameters of the stenosed branches in a carotid bifurcation (Sect. 6.6).

## 6.4 Atherosclerosis risk assessment in elastic arteries

We present in this section a first inverse problem related to the identification of structural parameters in a steady FSI problem, based on the reduced model introduced in Sect. 4.6. Fluid-structure interaction problems are of great interest in haemodynamics: for instance, in pulsatile flows through compliant vessels, the wall deformability causes large displacement effects that cannot be neglected when attempting to model the fluid dynamics. Thus, estimating the structural parameters related to the arterial wall may greatly enhance the quality of the results [239, 34].

Other problems of interest related to estimation of physiological parameters may arise e.g. in electrocardiography [41], as well as in patient-specific design of biomedical devices, such as stents [306] or dental implants. RB techniques have been recently applied e.g. to estimate the elastic moduli of the interfacial tissue between dental implants and the surrounding bones in [325].

Here we consider the model reduction of a simplified fluid-structure interaction model predicting flow over a stenosis inside an artery. To estimate the structural parameters related to arterial wall, we need to solve two *inverse identification* problems: *(i)* the former to formulate the FSI coupling as a parametric coupling under the form (4.43), and *(ii)* the latter for the identification of the elastic moduli entering in the structural equations from some output measurements.
Following the approach presented in Sect. 4.6, we consider a steady Navier-Stokes flow and represent an ideal stenosed arterial segment by the domain $\Omega_o := (0, L) \times (w_s, \bar{R} + \eta)$, where $\bar{R}$ and $L$ are the width and length of the artery in the reference configuration, respectively. We denote by $\eta \in H_0^2(0, L)$ the (unknown) function representing displacement of the upper wall of the artery from the reference configuration, and by $w_s$ the bottom wall shape function

$$w_s(x_1) := \begin{cases} h \cos \dfrac{\pi(x_1 - x_s)}{2\delta} & \text{if } x_s - \delta < x_1 < x_s + \delta \\ 0 & \text{otherwise,} \end{cases} \tag{6.5}$$

where $\delta > 0$ is the width, $x_s$ the centerpoint, and $h$ the height of the stenosis (see Fig. 6.16).

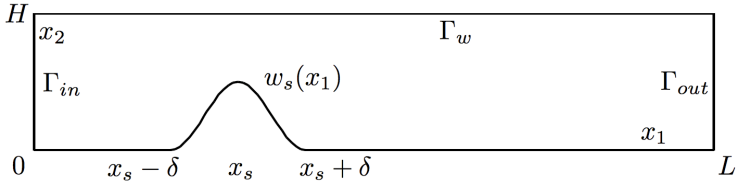



Figure 6.16: Definition of the geometrical domain for the stenotic artery case.

The reference fluid domain – with $\bar{R} = 1\,cm$ and $L = 5\,cm$ – is represented in Fig. 6.17. The parametrization of its displacement is performed with a $12 \times 2$ regular grid of control points, where only the 8 central points on the upper row were allowed to move freely in the $y$-direction. The admissible parameter range is $\pi_i \in [-0.25, 0.25]$ for $i = 1, 2, \ldots, p = 8$. In particular, the two left- and rightmost columns of control points are kept fixed in order to guarantee a parametric displacement that conforms to the boundary conditions set upon the structure. The parabolic inflow velocity fixed as $\mathbf{v}_{\text{in}} = (5(1 - (y - 0.5)^2/0.25), 0)\,cm/s$, yielding a Reynolds number of about $Re = 80$ at the location of the stenosis.
We recover an affine expansion of $Q_a + 2Q_b + Q_c = 106$ terms through the EIM procedure. The problem is discretized with $\mathcal{N}_V + \mathcal{N}_Q \approx 35{,}000$ degrees of freedom, using $\mathbb{P}_2/\mathbb{P}_1$ finite elements; the dimension of the computed RB space is $3N = 24$ by choosing a tolerance $\varepsilon_{tol}^{RB} = 5 \cdot 10^{-2}$ for the greedy procedure. Convergence of the parametric FSI coupling algorithm (4.43) requires about $10 - 15$ iterations, each step involving the solution of a RB problem for the (steady) flow field and a FE (1D) problem for the structural displacement.

In Fig. 6.18 we display a RB simulation of the flow field, namely the velocity streamlines in the case that the artery is rigid and does not deform. The stenosis induces a strong double vortex downstream, resulting in an area of low wall shear stress immediately after the stenosed part. The presence of the vortices in the reduced solution already with quite few basis functions highlights the fact that important physical features of the flow can be captured by the RB method as long as the features are part of the snapshots used to construct the basis (i.e. recirculation zones, viscous boundary layers, etc.). Thus a proper sampling procedure via the greedy algorithm is absolutely mandatory to guarantee a good "physical" selection of states for the reduced system.

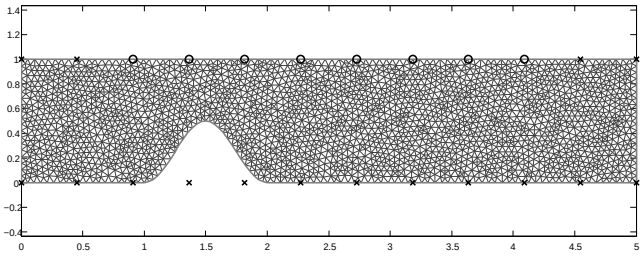


Figure 6.17: Reference domain $\Omega$, FE mesh and FFD control points used to model the displacements of the upper wall for the FSI problem (eight control points are allowed to move).

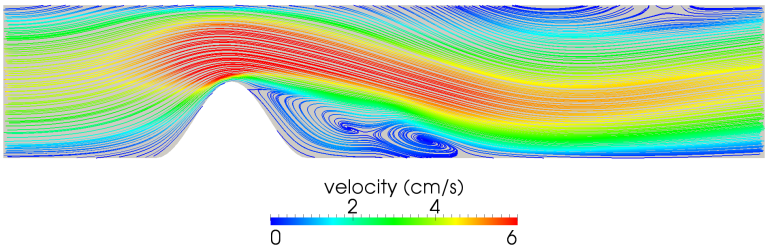



Figure 6.18: Velocity magnitude and streamlines of the flow in a rigid stenosed artery (RB simulation).

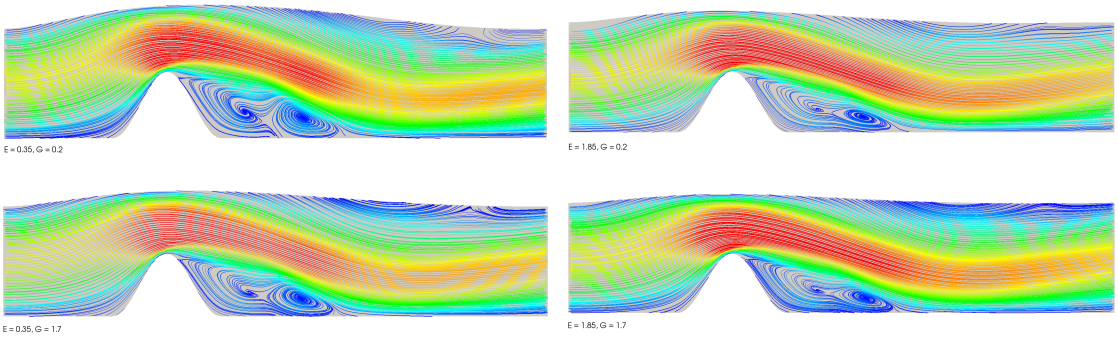



Figure 6.19: Velocity magnitude and streamlines (steady Navier-Stokes flow) in compliant stenosed artery (RB simulation) for different elastic moduli values, $E \in \{0.35, 1.85\}$ and $G \in \{0.2, 1.7\}$.

We also point out that the shape of the upper wall has a very strong effect on the type of vortices created and ultimately the potential growth of the stenosis. In Fig. 6.19 the RB simulation of the flow is given for each of the four corners of the parametric uncertainty domain. It can be observed that the more compliant the arterial wall is (small values of elastic moduli), the larger is the recirculation region created behind the stenosis. To explore the uncertainty related to the arterial wall properties, we define the uncertainty parameters $\boldsymbol{\omega} = (E, G)$ as the Young modulus and the shear modulus being[3] $E \in [0.35, 1.85] \cdot 10^6$ dyn/cm$^2$ and $G \in [0.2, 1.7] \cdot 10^6$ dyn/cm$^2$. The rest of the material constants in the 1D structural equation (4.38) were chosen as $\nu_P = 0.5$, $K = 0.9643$, and $h = 0.1$ cm. The fourth-order perturbation term was chosen as $\varepsilon = 0.1$ to obtain reasonable displacements and enhance the convergence for the partitioned algorithm.

To measure the effect of the uncertainty in the wall properties we look at four different output functionals, such as the total viscous energy dissipation:

[3]The unit dyn/cm$^2$ is omitted from hereon in for brevity and should be implicitly understood.

$$\mathcal{J}_1(\boldsymbol{\pi}, \boldsymbol{\omega}) = \frac{\nu}{2} \int_{\Omega_o(\boldsymbol{\pi})} |\nabla \mathbf{v}|^2 \, d\Omega_o, \tag{6.6}$$

the minimum downstream shear rate (where in this case $\Gamma_{\text{wall}} = (x_s + \delta, L)$):

$$\mathcal{J}_2(\boldsymbol{\pi}, \boldsymbol{\omega}) = \nu \min_{x_1 \in \Gamma_{\text{wall}}} \left\{ \frac{\partial v_1}{\partial x_2}\Big|_{x_2=0} \right\}, \tag{6.7}$$

the mean downstream shear rate:

$$\mathcal{J}_3(\boldsymbol{\pi}, \boldsymbol{\omega}) = \frac{\nu}{|\Gamma_{\text{wall}}|} \int_{\Gamma_{\text{wall}}} \left[ \frac{\partial v_1}{\partial x_2}\Big|_{x_2=0} \right] dx_1, \tag{6.8}$$

and the mean pressure drop in the stenosed section:

$$\mathcal{J}_4(\boldsymbol{\pi}, \boldsymbol{\omega}) = \int_{\Gamma_{\text{in}}} p \, dx_2 - \int_{\Gamma_{\text{out}}} p \, dx_2. \tag{6.9}$$

The first output measures the total energy loss in the flow due to the stenosed section, including recirculation effects, the second and third outputs are local indicators that have been linked to plaque buildup and onset of atherosclerosis (see e.g. [299]), whereas the fourth output is obtainable from noninvasive measurements and can be used as an indicator or surrogate for the first three outputs. Moreover, we denote by $J_i(\boldsymbol{\omega}) = \mathcal{J}_i(\boldsymbol{\pi}^*(\boldsymbol{\omega}), \boldsymbol{\omega})$ the output as a function of the structural parameters, where $\boldsymbol{\pi}^*(\boldsymbol{\omega})$ is the value of the shape parameters computed through the parametric coupling algorithm, given the structural parameters $\boldsymbol{\omega}$. In Fig. 6.20 we present the outputs $J_1, \ldots, J_4$ evaluated at 100 randomly selected sample points in the parametric uncertainty region.

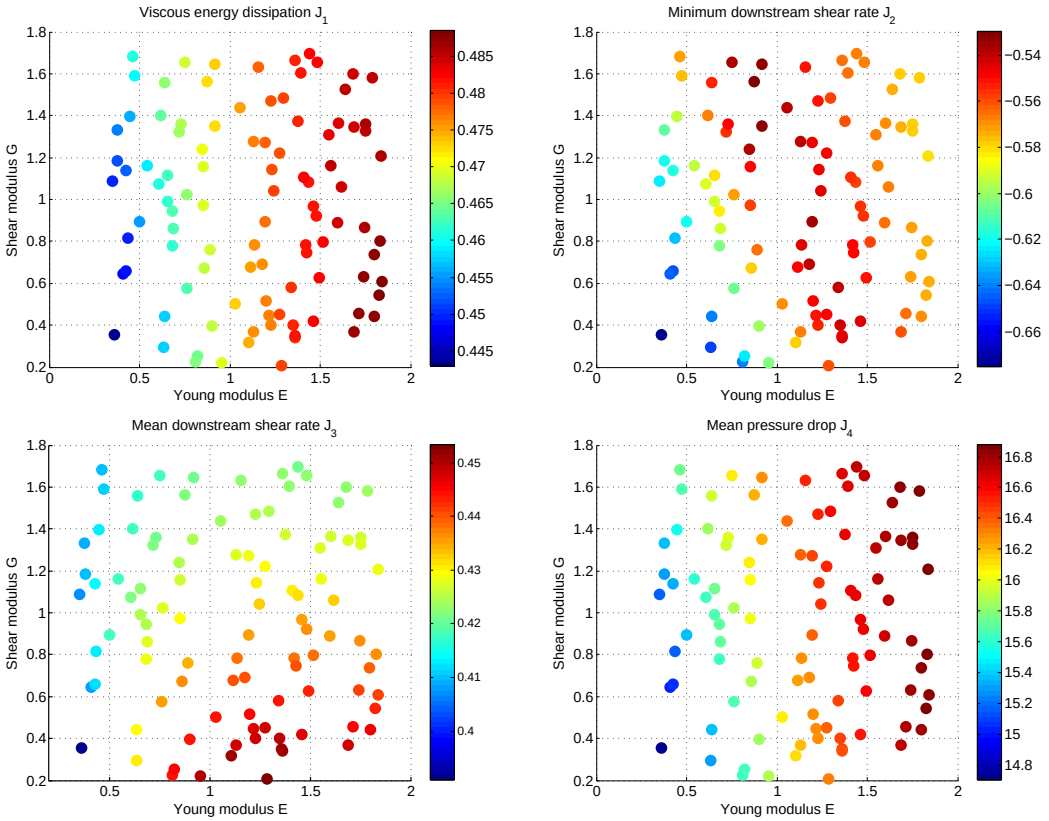



Figure 6.20: Flow observations in the compliant stenosed artery: values of the output $J_1$, $J_2$, $J_3$ and $J_4$ at randomly sampled points in the parametric uncertainty range (RB simulations).

Some conclusions can be drawn:

1. The effect of the Young modulus $E$ is considerably larger than the effect of the shear modulus $G$ on both functionals (6.6)-(6.7), at least for this simplified structural model.
2. The viscous dissipation is a monotonic function of the Young modulus, i.e. the stiffer the arterial wall the more dissipation is observed. This is consistent with the fact that atherosclerotic stiff arteries are at greater risk of stenosis occurrence.
3. The minimum shear rate depends primarily on the Young modulus and behaves in a nonlinear way. There seems to be an optimal range of the Young modulus ($E \geq 0.7$), after which the observed shear rate starts to go down again. This is likely due to the fact that further increasing of the compliance of the structure leads to a larger recirculation zone and a strong separation of the flow from the bottom wall.
4. The mean downstream shear rate has different behavior from the minimum downstream shear rate. This is likely due to the strong local shear rate induced by the jet impacting the downstream arterial wall.

So far, we have considered a *forward* problem with respect to the structural parameters $\boldsymbol{\omega} = (E, G)$, analyzing their role in the coupled FSI problem by measuring some outputs of interest under the form $J(\boldsymbol{\omega}) = \mathcal{J}(\boldsymbol{\pi}^*(\boldsymbol{\omega}), \boldsymbol{\omega})$ – yet, each input/output evaluation relies on an inverse identification problem for sake of geometrical coupling, in order to determine $\boldsymbol{\pi}^* = \boldsymbol{\pi}^*(\boldsymbol{\omega})$ for any given $\boldsymbol{\omega}$. We move on to the solution of an inverse problem related with the structural parameters.

### 6.4.1 Identification of elastic moduli based on measured pressure drop

An inverse problem of interest in haemodynamics is the determination of arterial system properties from pressures and flows measured at the entrance of this system [262]. Usually, some reduced arterial system models are fitted to data, yielding to the identification of parameters entering in the model as arterial properties. Assuming the stenosis geometry described above, here we consider a modified version of this problem, by asking whether it is possible to identify the material properties $\boldsymbol{\omega} = (E, G)$ by measuring the inflow and the (mean) pressure drop $s^* := J_4(\boldsymbol{\omega}^*)$ over the segment. It is clear from Fig. 6.20 that in general the shear modulus can not be reliably identified, so we concentrate first on identification of the Young modulus, assuming that the value of the shear modulus is known.

We show the solution of the deterministic inverse problem for two different observed values of the Young modulus: $E = 1.2334 \cdot 10^6$ dyn/cm$^2$ and $E = 0.9175 \cdot 10^6$ dyn/cm$^2$, and a fixed value $G = 0.5000$ dyn/cm$^2$ for the shear modulus. The corresponding values for the pressure drops were $J_4 = 16.3881$ dyn/cm$^2$ and $J_4 = 15.9390$ dyn/cm$^2$. In the first case we assume 1% additive noise in the measurements, in the second case 3% noise.
The results of the inverse identification problem without regularization are given in Table 6.1 for one particular noise realization. For the first case a very good reconstruction $\hat{E}$ of the value of the Young modulus $E$ is obtained, while in the second case the estimate is quite poor. This is due to the low sensitivity of the pressure drop $J_4$ to the material parameters, leading to increasingly poor reconstruction as the noise level is increased.

| Young modulus $E$ | Inverse estimate $\hat{E}$ | Relative error | # optim iters | # PDE solves |
|---|---|---|---|---|
| 1.2334 | 1.2253 | 0.66 % | 10 | 45 |
| 0.9175 | 0.6501 | 29.1 % | 9 | 41 |

Table 6.1: Results of the deterministic inverse problem of haemodynamics in two different cases.

Following instead the Bayesian approach, we provide a probability distribution function for the elastic moduli $\boldsymbol{\omega} = (E, G)$ encompassing the uncertainties related to measurements. If we assume that the prior distribution is $p(\boldsymbol{\omega}) \sim \mathcal{U}([0.35, 1.85], [0.2, 1.7])$, i.e. $E$ and $G$ are two independent uniformly distributed random variables, and that the measurements of the pressure drop are independent and normally distributed variables with expected value equal to the true mean pressure drop $J_4(\boldsymbol{\omega})$ and variance $\sigma^2$, i.e. $p(s \mid \boldsymbol{\omega}) = p(s - J_4(\boldsymbol{\omega})) \sim \mathcal{N}(0, \sigma^2)$, then we obtain the *posterior probability density* $p(\boldsymbol{\omega} \mid s)$ given by

$$p(\boldsymbol{\omega} \mid s^*) \propto \begin{cases} \exp\left(-\frac{1}{2\sigma^2}\left(s^* - J_4(\boldsymbol{\omega})\right)^2\right) & \text{if } \boldsymbol{\omega} \in [0.35, 1.85] \times [0.2, 1.7], \\ 0 & \text{otherwise.} \end{cases} \tag{6.10}$$

The posterior distribution $p(\boldsymbol{\omega} \mid s^*)$ for the two different test cases and their observations $s^*$ are shown in Fig. 6.21. In both cases we see that only the Young modulus $E$ is identifiable from the pressure drop alone, and the value of the shear modulus $G$ can have an effect on the maximum likelihood estimate of $E$.

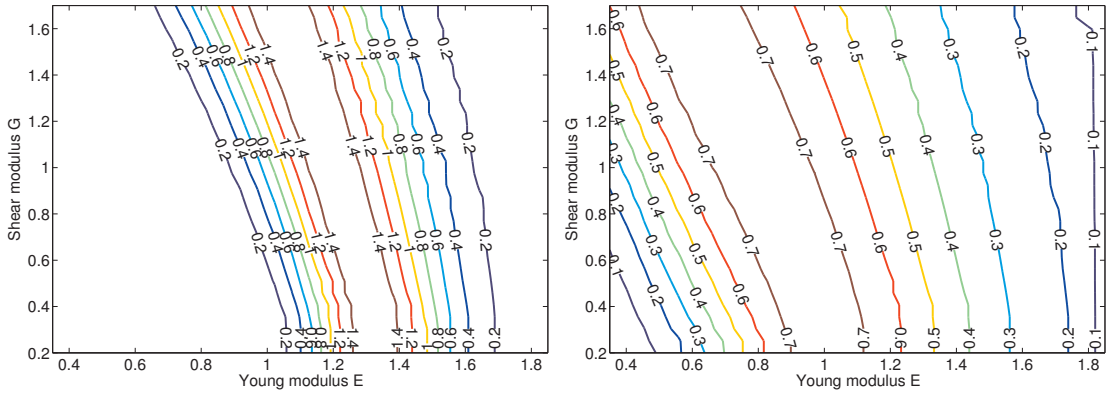



Figure 6.21: Contour plot of the posteriori distribution $p(\boldsymbol{\omega} \mid s^*)$ for two different cases: $(s^*, \sigma) = (16.3881, 0.17)$ and $(s^*, \sigma) = (15.9390, 0.51)$. Only the Young modulus $E$ can be reliably identified.

From the second case we already observed that increasing the noise very quickly smears out the probability distribution. Even when the topology of the artery is assumed to be known, the inverse problem of haemodynamics is rather ill-posed and susceptible to noise. A purely deterministic approach to the solution of this inverse problem can therefore mask the underlying uncertainty and lead us to erroneously optimistic conclusions about the quality of inverse estimates such as those dictated in Table 6.1. If no prior information is at hand concerning the indeterminable parameters, such as the shear modulus, the inverse problem of haemodynamics can not be reliably solved.

### 6.4.2 Identification of shear rate based on measured pressure drop

Alternatively to the inverse problem faced in Sect. 6.4.1 we can raise the question: if we have a measurement $s^*$ of the pressure drop, can we predict the shear rate on the arterial wall downstream from the stenosis? Even if we have shown that the elastic moduli of the arterial wall are not strictly identifiable from the pressure drop measurement alone, it might still be possible that the wall shear rate (which in the case of predicting atherosclerosis risk is the primary quantity of interest) can be reasonably accurately identified. In this way the pressure drop measurement would act as a surrogate for identifying the wall shear rate.

As the value of the mean shear rate $J_3(\boldsymbol{\omega})$ can be obtained by performing the forward simulation given the material parameters $\boldsymbol{\omega}$, we can explore the posterior p.d.f. $p(J_3(\boldsymbol{\mu}) \mid s^*)$ provided by means of the Metropolis-Hastings algorithm, introduced in Sect. 4.5.2. In Figs. 6.22-6.23 we display the observed histograms for the distribution of the minimum shear rate $J_2$ and the mean shear rate $J_3$ in two cases, given by $(s^*, \sigma) = (16.3881, 0.17)$ and $(s^*, \sigma) = (15.9390, 0.51)$, respectively. We can remark how the pressure drop can indeed be used to determine the minimum value of the shear rate downstream from the stenosis, while the mean shear rate distribution is more smeared out. This is mainly due to the fact that the unidentifiable shear modulus $G$ has a larger effect on the mean shear rate than the minimum shear rate.

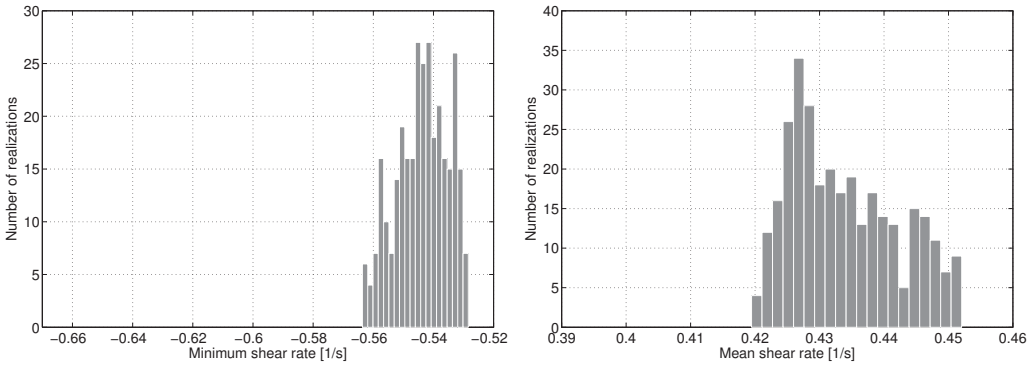



Figure 6.22: Histograms of the posterior distributions $p(J_2(\boldsymbol{\omega}) \mid s^*)$ (left) and $p(J_3(\boldsymbol{\omega}) \mid s^*)$ (right) for the case $(s^*, \sigma) = (16.3881, 0.17)$. The range of the plot is scaled to correspond with the entire range of observed variation in each output.

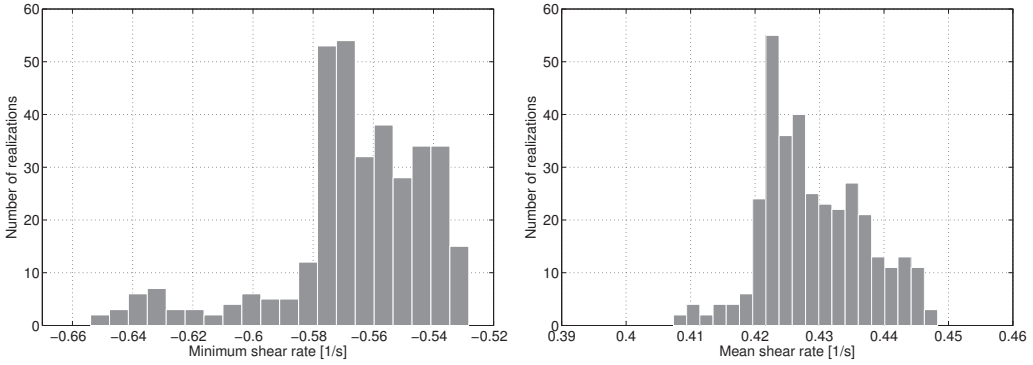



Figure 6.23: Histograms of the posterior distributions $p(J_2(\boldsymbol{\omega}) \mid s^*)$ (left) and $p(J_3(\boldsymbol{\omega}) \mid s^*)$ (right) for the case $(s^*, \sigma) = (15.9390, 0.51)$. The range of the plot is scaled to correspond with the entire range of observed variation in each output.

The model we have presented is of course very much an abstraction: it does not take into account the pulsatility of the flow, the transition to turbulence, the oscillation of the wall shear stress in time and the periodic detachment/separation of the flow layer from the wall. The exploration of the parameter space with 100 samples to produce the Fig. 6.20 took around 70 minutes of CPU time[4] when all the fluid simulations were performed with the use of the RB method. We estimate that exploring the same parameter space with the full FE simulation would have taken around 33 hours of CPU time – and considerably more if the FSI problem had been solved using standard methods and not taking advantage, as done here, of the geometrical reduction afforded by the parametrization of the structural deformation.

[4]Computations have been executed on a personal computer with 2 × 2GHz Dual Core AMD Opteron (tm) processors 2214 HE and 16 GB of RAM.

## 6.5 Identification of residual blood flows in a bypass model

The results of the shape optimization problems shown in the previous chapter (see Sect. 5.4.2-5.4.3) highlight the criticality occurring in presence of a complete blockage of the occluded artery. For this reason, we might be interested in identifying the uncertain residual flow magnitude $\omega \in \mathcal{D}_{\boldsymbol{\omega}}$ from some possible, noninvasive measurements of quantities related with the blood flow.
In particular, the chance of identifying whether the bypass will be operating under low residual flows or high residual flows might guide the optimal graft design. As in Sect. 6.4, here inverse identification is driven by the numerical observation of the mean pressure drop in the bypass:

$$J_4(\omega) = \int_{\Gamma_{\text{bc}}} p\, dx_2 - \int_{\Gamma_{\text{out}}} p\, dx_2. \tag{6.11}$$

where $\Gamma_{\text{in}}$ is the inlet from the stenosed section of the artery, $\Gamma_{\text{bc}}$ the bypass inlet and $\mathbf{v}_{\text{in}} = \mathbf{v}_{\text{in}}(\omega)$ the residual flow across $\Gamma_{\text{in}}$, of magnitude $\omega \in \mathcal{D}_{\boldsymbol{\omega}} =\in [0, \omega_{\max}]$.

Here we refer to the optimal design of a a femoro-popliteal bypass analyzed in Sect. 5.4.3. The behavior of the mean pressure drop with respect to the parameter $\omega$ in the two robust optimal bypass shapes (see Fig. 5.27) obtained in the previous analysis is reported in Fig. 6.24. In both cases a local maximum is observed at $\omega = 0$, and the value of the functional decreases as the magnitude of the residual flow increases, with a smaller and smaller decrease rate. This fact yields some interesting consequences in the inverse identification problem. We show the solution of the deterministic inverse problem for two given values of the magnitude of the residual flow $\omega$:

1. for the robust optimal shape (1), obtained with the vorticity functional, $\omega = 2.6097$ and $\omega = 10.9817$, and corresponding observed pressure drops $s^* = -440$ and $s^* = -490$;
2. for the robust optimal shape (2), obtained with the Galilean-invariant functional, $\omega = 3.6274$ and $\omega = 10.0130$, and corresponding observed pressure drops $s^* = -735$ and $s^* = -765$.

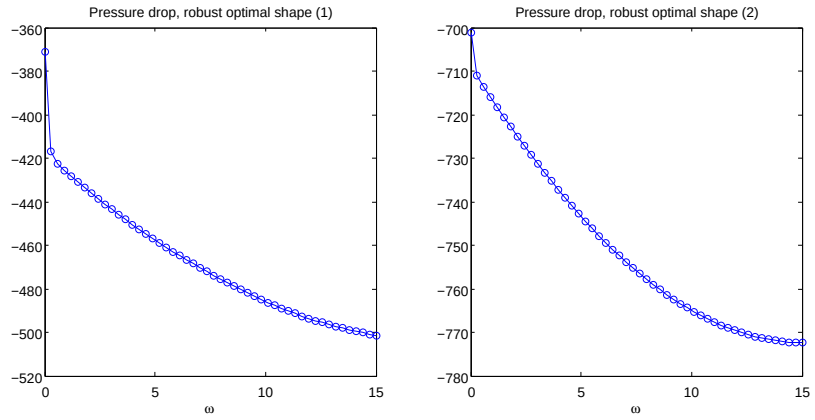



Figure 6.24: Computed mean pressure drop $J_4(\boldsymbol{\mu})$ for different values of $\omega$, for the robust optimal shapes obtained with a vorticity functional (case 1) and a Galilean-invariant functional (case 2).

The results of the inverse identification problem without regularization are given in Tables 6.2-6.3 for three particular noise realizations ($\sigma = 0.01, 0.03, 0.05$), for the robust shapes obtained with the vorticity and the Galilean-invariant functionals, respectively. Clearly, the lower the noise the better the reconstruction. The estimate becomes increasingly poor for larger amounts of noise, especially for small values of $\omega$, which represent more critical cases not only for the fluid dynamics, but also concerning the decrease rate of the pressure drop. Thus, also this example shows that the total occlusion represents a critical condition to deal with.

| $\omega$ | $\sigma$ | Inverse estimate | Relative error | # optim iters | # PDE solves |
|---|---|---|---|---|---|
| 2.6097 | 0.01 | 2.6295 | 0.76% | 5 | 13 |
| | 0.03 | 3.8004 | 45.63% | 5 | 13 |
| | 0.05 | 6.3890 | 144.82% | 4 | 11 |
| 10.9817 | 0.01 | 11.2151 | 2.13 % | 5 | 13 |
| | 0.03 | 8.2964 | 24.45 % | 4 | 11 |
| | 0.05 | 7.3632 | 32.95 % | 4 | 11 |

Table 6.2: Results of the deterministic inverse problem, robust optimal shape (1).

| $\omega$ | $\sigma$ | Inverse estimate | Relative error | # optim iters | # PDE solves |
|---|---|---|---|---|---|
| 3.6274 | 0.01 | 3.5594 | 1.87 % | 4 | 11 |
| | 0.03 | 4.1799 | 15.23 % | 6 | 15 |
| | 0.05 | 9.8734 | 172.19 % | 5 | 13 |
| 10.0130 | 0.01 | 9.2899 | 1.83 % | 5 | 13 |
| | 0.03 | 6.6272 | 33.81 % | 4 | 11 |
| | 0.05 | 5.4044 | 46.03 % | 4 | 11 |

Table 6.3: Results of the deterministic inverse problem, robust optimal shape (2).

Following the Bayesian approach we provide a probability distribution function for the (uncertain) magnitude of the residual flow $\omega$ encapsulating the noise related to measurements. In particular, if we assume that the prior distribution is normal, $p(\omega) \sim \mathcal{N}(\omega_M, \tau^2)$ (we choose a plausible value $\omega_M$ for the mean, corresponding to a partial occlusion case, e.g. $\omega_M = 10$), and that the measurements of the pressure drop are $n = 250$ independent and normally distributed variables with expectation equal to the true mean pressure drop and variance $\sigma^2$, i.e. $p(s \mid \omega) = p(s - J_4(\omega)) \sim \mathcal{N}(0, \sigma^2)$, we obtain that also the posterior probability density $p(\omega \mid s)$ is normally distributed, with mean and variance given by

$$\mathbb{E}[\omega \mid s] = \frac{1/\tau^2}{1/\tau^2 + n/\sigma^2}\omega_M + \frac{n/\sigma^2}{1/\tau^2 + n/\sigma^2}\bar{J}_4, \qquad \mathrm{Var}[\omega \mid s] = \frac{\sigma^2/n\tau^2}{\tau^2 + \sigma^2/n},$$

being

$$\bar{J}_4 = \frac{1}{n}\sum_{i=1}^{n} J_4(\omega_i), \qquad i = 1, \ldots, n.$$

The posterior probability density $p(\omega \mid s^*)$ for the two different test cases and their observations $s^*$ are shown in Figs. 6.25-6.26. We can remark that, in presence of small noises, the difference in the means of the prior and the posterior distribution may be significantly different, because of the information provided by the measurements. A Bayesian statistic approach can provide a better indication of the identified parameter, including the effect of the measurement noise. Moreover, in the case of a *Gaussian/Gaussian* model, like the one applied in this case, it is simple to evaluate the *updating* effect provided by the knowledge based on measured quantities.

The exploration of the parameter space with 250 samples to produce the posterior distributions took around 90 minutes of CPU time. As for the previous scenario, we estimate that exploring the same parameter space with the full FE simulation would have taken a computational time eight times larger. If shape deformations had been handled using standard methods instead of the reduced FFD representation, this cost would have increased even more.

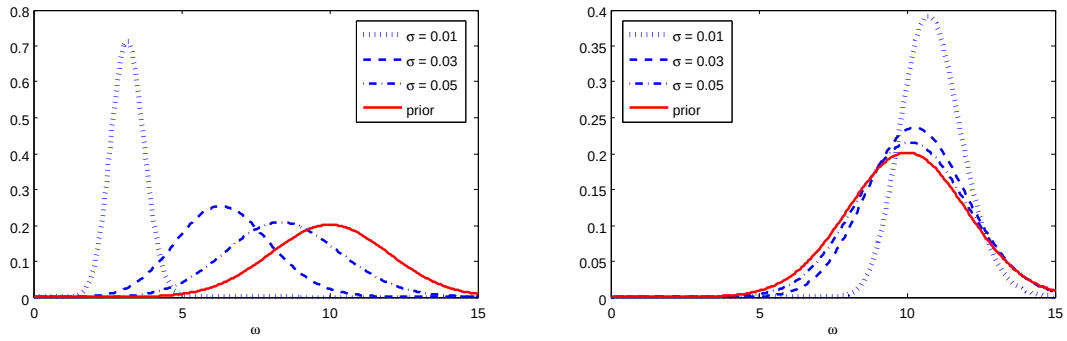



Figure 6.25: Robust shape (1): posterior p.d.f. $p(\omega \mid s^*)$ for the cases $s^* = -440$ and $s^* = -490$, with three different noise levels: $\sigma^2 = 0.01, 0.03, 0.05$. The prior distribution is $p(\omega) \sim \mathcal{N}(10, 2)$.

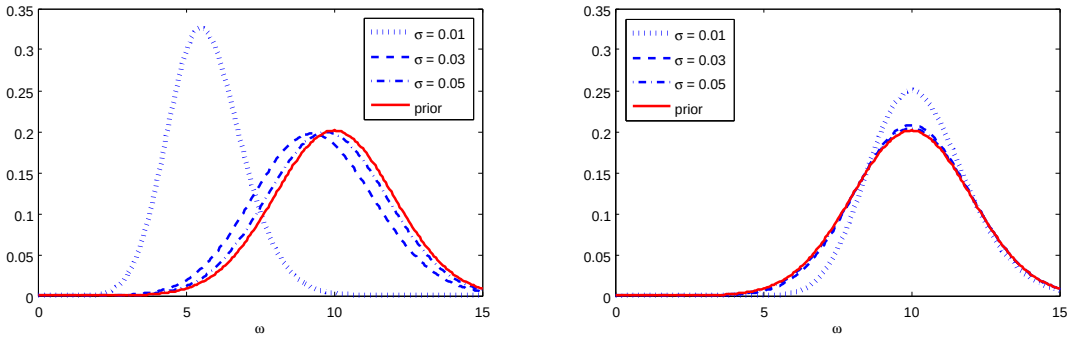



Figure 6.26: Robust shape (2): posterior p.d.f. $p(\omega \mid s^*)$ for the cases $s^* = -735$ and $s^* = -765$, with three different noise levels: $\sigma^2 = 0.01, 0.03, 0.05$. The prior distribution is $p(\omega) \sim \mathcal{N}(10, 2)$.

## 6.6 Inverse shape identification from output measurements

A third possible scenario fitting the reduced inverse framework exploited throughout the chapter deals with the inverse identification of shape features from observations of physical indices related with flows. The solution of such an inverse problem is meaningful, since a strong correlation exists between haemodynamic forces and the location, development and morphology of plaque deposits within the lumen, which are responsible of the narrowing of blood vessels.
As shown for a set of carotid artery bifurcations (Sect. 6.3.2), we may establish some input/output relationships between geometrical parameters and physical indices related with blood flows in case one or more stenoses occur. Here we consider the same setting, dealing with local deformations of a carotid bifurcation, in order to recover the diameters $d_c$ of the CCA at the bifurcation and $d_b$ of the mid-sinus level of the ICA, respectively, from the observation of the mean pressure drop between the internal carotid outflow $\Gamma_{\text{out}}$ and the inflow $\Gamma_{\text{in}}$. In particular, we take advantage of both the deterministic and the Bayesian framework to solve this inverse identification problem, by considering surrogate measurements of the mean pressure drop.

In the first case, we demonstrate the solution of the deterministic inverse problem for two different observed values of the pressure drop, $s^* = -1400$ and $s^* = -2200$, by assuming 5% relative additive noise in the measurements. The results of the inverse identification problem are given in Fig. 6.27 for 100 realization of random noise in both cases: each point in the graph corresponds to the recovered diameters $(d_c, d_b)$ given a noisy observation. We observe that in the case $s^* = -1400$ recovered values of the diameters are more smeared out, due to the flatter pattern of the pressure drop, but result is closed in values to the considered observation.

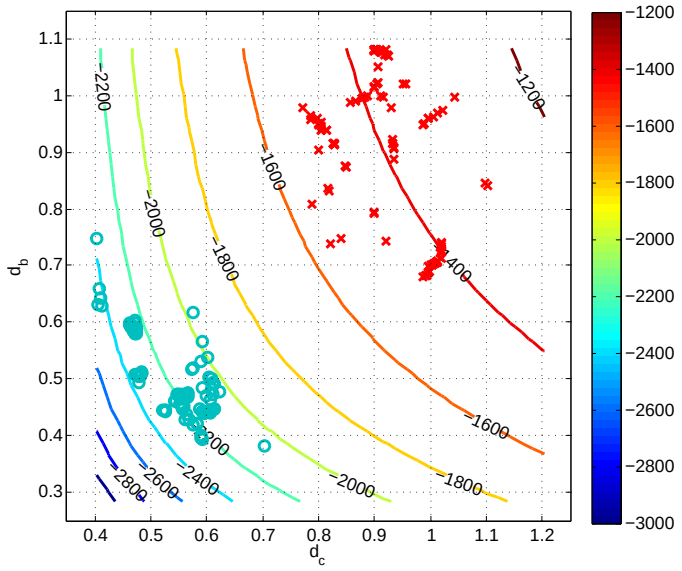



Figure 6.27: Results of the deterministic inverse problems for $s^* = -1400$ (in red) and $s^* = -2200$ (in green). Isocontours of the pressure drop (RB approximation of the Navier-Stokes problem).

Thus, in the former case $s^* = -1400$ the inverse problem is worse conditioned than in the latter $s^* = -2200$, where the recovered values $(d_c, d_b)$ lie in a smaller region of the space. However, the solution of a single optimization problem is more feasible in the former case compared to the latter: solving 100 optimization problems took about 14 hours in the former and about 25.6 hours in the latter case, respectively. We remark that solving 100 inverse problems of this type through a full-order discretization technique would have been unfeasible on a standard workstation.

Even in presence of small noises, the result of a deterministic inverse problem may be very sensitive – just when one diameter is known, the second one can be recovered. This is due to the fact that several geometrical configurations – in terms of diameters $(d_c, d_b)$ – may correspond to the same output observation. Following instead the Bayesian approach, we are able to characterize a *set* of configurations, rather than a single configuration: this is done by providing the *joint* probability distribution function for the (uncertain) diameters $(d_c, d_b)$ encapsulating the noise related to measurements, as we discuss hereon.

Let us denote by $\mathbf{d} = (d_c, d_b)^T \in \mathbb{R}^2$ the vector of the two diameters and assume that the prior distribution $p(\mathbf{d})$ has a (bivariate) *Gaussian* shape, $p(\mathbf{d}) \sim \mathcal{N}(\mathbf{d}_M, \Sigma_M)$, where $\mathbf{d}_M \in \mathbb{R}^2$ is the (prior) mean and $\Sigma_M \in \mathbb{R}^{2\times 2}$ is the (prior) covariance matrix, encapsulating a possible prior knowledge on the diameters distribution (e.g. from observations of previous shape configurations). By supposing that also the measurements of the pressure drop are expressed by $n$ independent and normally distributed variables with expectation equal to the true pressure drop and variance $\sigma^2$, i.e. $p(s \mid \mathbf{d}) = p(s - J_4(\mathbf{d})) \sim \mathcal{N}(0, \sigma^2)$, we can compute the explicit form of the posterior probability density $p(\mathbf{d} \mid s)$. Thus, provided some preliminary information on plausible values of the diameters, the observation of a (large) sample of outputs allows to characterize a set of configurations as the ones maximizing the *posterior* distribution relying on a Bayesian inversion procedure.
In particular, we consider two different realizations of prior normal distributions, obtained by choosing the mean $\mathbf{d}_M = (0.803, 0.684)^T$ as given by the diameters corresponding to the reference carotid configuration, and two different covariance matrices:

$$\Sigma_{M_1} = \begin{bmatrix} 0.025 & 0 \\ 0 & 0.0125 \end{bmatrix}, \qquad \Sigma_{M_2} = \begin{bmatrix} 0.025 & -0.0125 \\ -0.0125 & 0.0125 \end{bmatrix},$$

i.e., by assuming that the two diameters are *a priori* independent ($\Sigma_{M_1}$ case) or correlated ($\Sigma_{M_2}$ case), respectively. The two prior distributions, as well as the resulting posterior distribution obtained for two different observed values $s^* = -1400$ and $s^* = -2200$ of the pressure drop are reported in Fig. 6.28.

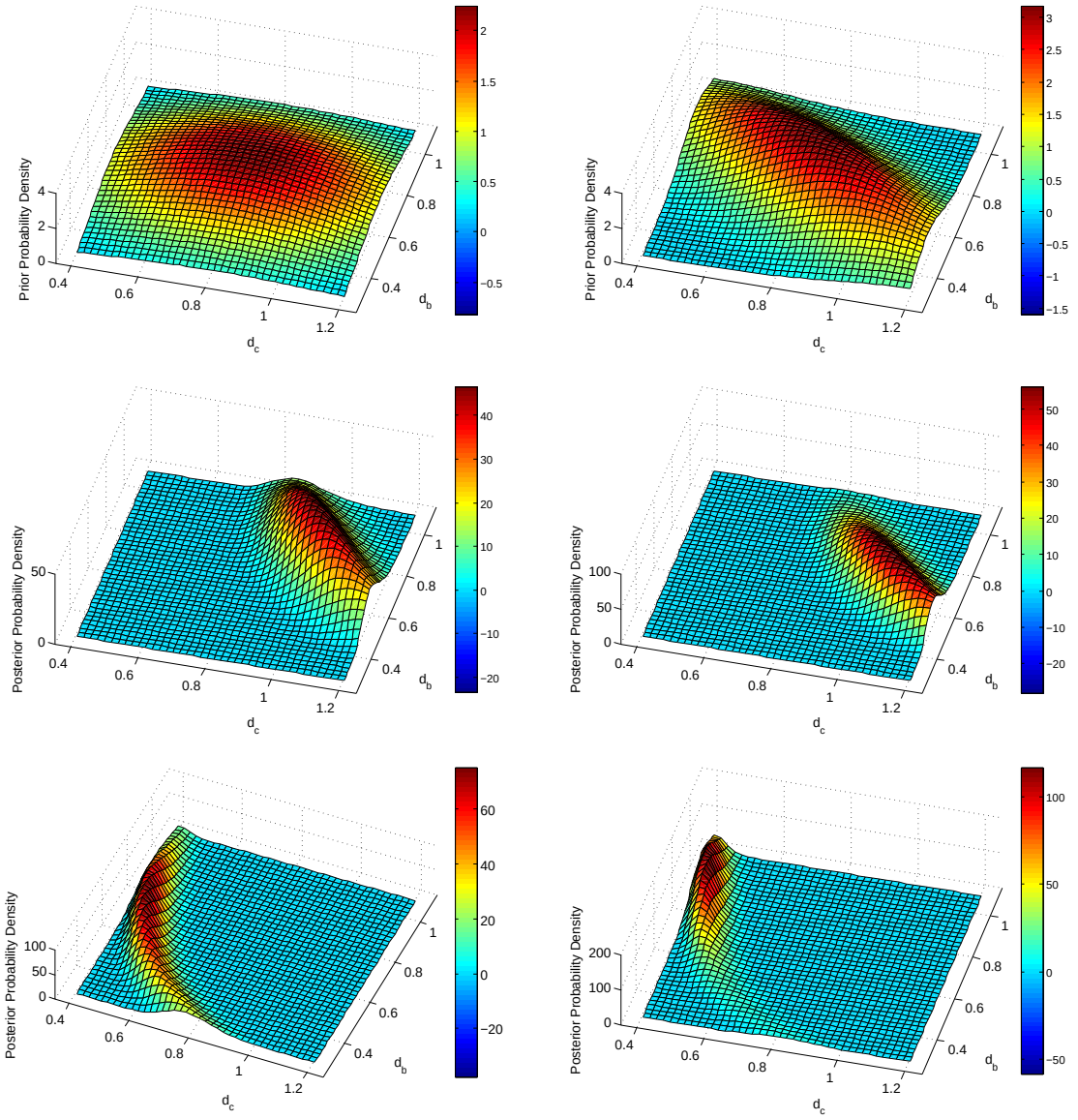



Figure 6.28: Top: two different choices of the prior distribution on diameters $\mathbf{d} = (d_c, d_b)^T$; left: $\mathbf{d} \sim \mathcal{N}(\mathbf{d}_M, \Sigma_{M_1})$, right: $\mathbf{d} \sim \mathcal{N}(\mathbf{d}_M, \Sigma_{M_2})$. Bottom: results of the Bayesian inverse problems for two different choices of the prior distribution and observed pressure drop $s^* = -1400$ (second row) and $s^* = -2200$ (third row).

In this case, we do not rely on the Metropolis-Hastings algorithm for the evaluation of the posterior distribution, since the expression of this distribution can be computed explicitly. We evaluated a sample of $n = 1600$ values of pressure drops on a uniform $40 \times 40$ grid on the $(d_c, d_b)$ space – by the Metropolis-Hastings algorithm we would have determined instead a random trajectory on the space $(d_c, d_b)$ and still many queries under the form $\mathbf{d} \to J_4(\mathbf{d})$. To get efficiently the observation

sample, we exploited the RB approximation of the Navier-Stokes flow problem, which has been solved repeatedly, on $n$ different geometrical configurations. In this way, a reduced order model such as the one we have developed allows to recast a Bayesian inverse problem – which turns to be much more powerful than the deterministic paradigm in this case – in a many query context.

## 6.7 Remarks and future focuses

In this chapter we presented a reduced framework for treating some haemodynamic inverse problems considered as test cases. In the first example we formulated a steady FSI problem in a portion of a stenosed artery; by solving a suitable inverse problem, we predicted downstream shear rates linked to stenosis growth given a noninvasive pressure drop measurement. This inversion could be performed even in the likely case that the pressure measurements are too noisy to accurately identify the vessel wall material parameters themselves. In the second example we considered the identification of the residual flow magnitude in a bypass graft. The critical case of total occlusion in the artery could be either identified or ruled out based on similar noninvasive pressure drop measurements. Finally, we showed how to identify some geometrical features within a family of parametrized carotid bifurcations. This problem can be more easily afforded exploiting a Bayesian inverse framework than using a deterministic paradigm.

The three scenarios illustrated represent an idealized more than realistic case, due to the very simple nature of the observed outputs (scalar quantities) and to the assumptions considered on both fluid and geometrical models. Nevertheless, they provide a good proof of the flexibility of the reduced framework developed and of the possibility to recast in a many-query scenario a large variety of problems.

A possible perspective stemming from these test cases is the development of an integrated framework where:

1. geometrical information can be extracted or *reconstructed* from an *atlas* of collected shape configurations and synthetized through a reduced parametrization. Possibly, some reduction in the number of parameters should be performed. By this stage we can build some *prior* knowledge on geometry;
2. the Bayesian inverse problem allows to obtain a *posterior* distribution over the shapes family, thus resulting from the merging of physical models, output measurements and prior information on geometrical setting. Within this stage, the numerical simulation of a fluid flow problem should to be considered, in order to take into account the interplay between flow and geometry.

However, once the *updated* atlas is built *Offline*, the information concerning a new *Online* query could be provided through a *data-fitting* procedure, by means of a *black-box* metamodel. Unfortunately, in this way the physics of the problem would not be considered anymore at the *Online* stage, and no information could be provided on outputs not involved in the Offline stage. A reduced basis approximation of the fluid equations allows instead to model the physics of the problem also at the *Online* stage, thus accounting for additional output and information, with reasonable accuracy.

# Conclusion and perspectives

**Concluding remarks**

In this work we have set, analyzed and applied some reduced order modelling techniques for the efficient numerical solution of optimization and inverse problems governed by PDEs. The general framework we have developed relies on the Reduced Basis (RB) method for parametrized PDEs to reduce the computational complexity hindered by optimization problem, e.g. in optimal flow control problems. The extension of this framework to shape-dependent problems, e.g. in shape optimization problems, is made possible by coupling the reduced basis method with flexible but low-dimensional shape parametrizations, in order to represent a general set of admissible shapes, such as the Free-Form Deformations (FFD) or the Radial Basis Functions (RBF) techniques.

We have extended the reduced basis framework to general parametrized Stokes and Navier-Stokes problems, focusing on nonaffine geometrical parametrizations. In both cases, *a posteriori* error estimation methods have been developed for velocity and pressure fields jointly, moreover in the Stokes case error bounds for linear and quadratic outputs have been computed. Several numerical tests show that a strong improvement in efficiency can be achieved with respect to full-order approximation techniques, like the Finite Element (FE) method: the reduction in linear systems dimensions between RB and FE discretization is of order $\mathcal{O}(10^2 \div 10^3)$. This entails a computational speedup of about the same order, and the possibility to solve any new instance of a parametrized Stokes or Navier-Stokes problem in $\mathcal{O}(10^{-2} \div 10^0)$ seconds. In particular, computational times are almost constant also when dealing with problems parametrized w.r.t. Reynolds number. These remarkable performances are very well suited for the many-query contexts of our interest, as well as for real-time simulations. Moreover, error bounds ensuring the reliability of the reduced approximation can be evaluated in about the same amount of time.
We have analyzed both FFD and RBF techniques, proving some theoretical results of stability with respect to parametric variations. We have also set some algorithms for the selection of control points in the FFD case. Both FFDs and RBFs turned to be very flexible in spanning a set of realistic shapes, keeping the parametric space dimension small ($\mathcal{O}(10)$). In particular, RBFs allow a better control of local deformations thanks to their interpolation properties.

This parametrized reduced framework has been successfully applied to optimal control and shape optimization problems. Several other inverse problems related to both fluid flows and shape variations have also been faced, such as robust optimization under uncertain conditions, parametric coupling for multi-physics problems, inverse identification. To address this last case, both a deterministic and a Bayesian framework have been developed.
Several applications have been presented, dealing with meaningful problems arising in haemodynamics, like the optimal design of cardiovascular prostheses (e.g. bypass grafts), or the inverse identification of potentially risk factors related to shape or flow features in simplified and more realistic configurations (e.g. carotid artery bifurcations).

Although derived under simplifying assumptions (no flow pulsatility, steady flows, ...), the models we considered allow to characterize in (almost) real-time blood flows in 2D geometries, by capturing several features related e.g. to vessel shape, fluid behavior and also structural parameters. Interesting remarks about optimal sizing of bypass grafts, also in presence of uncertain conditions, have been obtained.
We point out that reduced order modelling proves to be necessary for optimal design and inverse problems related with shape variation. For instance, a robust shape optimization problem took about $4 \div 20$ hours of CPU time on a workstation by using the RB method for fluid simulation. Solving the same problem with a full-order FE discretization would have taken about 700 hours of CPU time – and even more if shape deformations had been handled without FFD techniques.

### Further developments of the current methodology

The proposed reduction paradigm can be extended in many different directions. Beyond some evident but not straightforward developments, such as the application to three-dimensional configurations or the validation of some results with e.g. *in vitro* experiments, we point out three main paths for further methodological developments:

1. *Reduced basis method.* The reduced basis methodology has already been extended to unsteady Boussinesq equations [168]. Even if exponential instability seriously compromises *a posteriori* error estimates, by carefully treating the stability growth rate (closely related to hydrodynamic stability theory) error bounds remain practicable for parameter domains and finite final times of physical interest. Nevertheless, error bounds for velocity and pressure fields jointly are still missing, also in the simpler linear Stokes case. In this respect, we are currently developing *a posteriori* error bounds of this type, in the framework of the Babuška stability theory. Other techniques, based e.g. on fractional step methods, may be considered in RB approximation and error bounds for Navier-Stokes equations. Other issues, such as estimation of stability factors and affine expansion for evaluating dual norms of residuals, become true bottlenecks in case of nonaffine problems or large parameter spaces, so that alternative strategies may be devised.

2. *Optimization framework.* The current reduced framework, based on a *reduce-then-optimize* paradigm, exploits the reduced basis methodology as a rapid and reliable *input/output* evaluator. Thanks to advanced optimization subroutines, we rely on numerical approximation of the gradient (with respect to parameter components) of the cost functional to be minimized, rather than solving the adjoint problem and evaluate the gradient. Extending the RB methodology to approximate the reduced adjoint problem, as well as exploiting the EIM procedure to approximate the derivatives of parameter-dependent functions appearing in affine expansions, are further steps which can be taken into consideration, still subject to the same bottlenecks previously highlighted.

3. *Inverse problems.* We have considered only simple inverse identification problems, since the parameter values were recovered by inverting (surrogate) output measures given by scalar functions of the parameters (even if depending on the solution of a parametrized PDE). Extending the current framework in order to deal with (either surrogate or true) measures of parametric fields – instead of functions – and to quantify the propagation of uncertainty coming from noisy data represents a twofold issue to be further developed.

### Longer-term outlook

The great potential made available by reduced order modelling, in terms both of real-time input/output evaluation and possibility to face large many-query problems, indicate two possible long-term goals.
From the theoretical point of view, the interplay between numerical and statistical methodologies could be further strengthened. In this respect, an accurate but reduced-order model can be seen as a possible *emulator* of the original (physical/geometrical) model, capable to generate in a cheap way large sets of data – corresponding to different input configurations – whose analysis can take advantage of accurate statistical tools, such as advanced methods for analysis of complex geometries, images and surfaces (object-oriented data analysis, Bayesian functional data analysis, response surface methods).
Depending on the context, this approach could lead for instance to *(i)* a better detection of parameters of interest within a PDE model, *(ii)* a better interpretation of numerical results coming from large computer experiments, and *(iii)* a simpler availability of datasets for statistical analysis within a rapid time and acceptable effort.

Concerning applications, a reduced framework such as the one developed in this work could be ideally exported to many engineering contexts requiring a rapid *in the field* (or *in situ*) evaluation process, such as parameter estimations, identification procedures, optimization problems, visualization environments. All these situations require high-order mathematical models, but rapid queries, such as the ones provided by a reduced-order method. Moreover, computational power or storage capacity available in the field are usually limited. With this respect, the *Offline-Online* paradigm exploited by the reduced framework developed in this work could fit for the purpose: it is based on *deployed* platforms for rapid and reliable reduced approximations, yet taking advantage of expensive computations stored in a former pre-deployment stage.

# A Insights on RB approximation and error estimation

We provide in this Appendix further insights on the reduced basis approximation of Stokes and Navier-Stokes equations, and *a posteriori* error estimations. We deal with the algebraic and the approximation stability of the Stokes and Navier-Stokes problems, guaranteed by the Gram-Schmidt orthonormalization procedure and the enrichment of the velocity space through the solutions of the supremizer equation. Then, we show how to compute the quantities entering in the error bounds introduced in Sect. 3.4: *(i)* the dual norm of the residuals, *(ii)* the lower bound of stability factors, and *(iii)* Sobolev embedding constants. In particular, we show how to exploit the Offline/Online procedure for computing the dual norms of the residuals in Stokes/Navier-Stokes problems, we characterize the *natural norm* SCM algorithm for the Stokes problems and we propose an extension to Navier-Stokes problems, and introduce a *fixed-point* algorithm for the approximation of Sobolev embedding constants entering in the error bounds for Navier-Stokes equations. In the end, we briefly recall the construction of affine geometrical mappings, used in the simpler parametrized problems we have presented.

## A.1 Algebraic and approximation stability

We present in this section the rigorous formulation of the Stokes problem (3.65) introduced in Sect. 3.3, taking into account *(i)* the $\boldsymbol{\mu}$-dependence in the supremizers enriching the velocity space and *(ii)* the orthonormalization procedure ensuring the algebraic stability. The same procedure is used to construct also the reduced spaces in the Navier-Stokes case, since the approximation stability relies on the *inner pressure* supremizers in that case as well. We recall that the *RB pressure space* $Q_N^{\mathcal{N}} \subset Q^{\mathcal{N}}$ and the *RB velocity space* $V_N^{\mathcal{N},\boldsymbol{\mu}} \subset V^{\mathcal{N}}$, defined by (3.63)-(3.64), are given by

$$Q_N^{\mathcal{N}} = \text{span}\{\tilde{\zeta}_n^p := p^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n = 1, \dots, N\}, \qquad N = 1, \dots N_{\max},$$

$$V_N^{\mathcal{N},\boldsymbol{\mu}} = \text{span}\{\tilde{\boldsymbol{\zeta}}_n^{\mathbf{v}} := \mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}^n),\ T_p^{\boldsymbol{\mu}}\tilde{\zeta}_n^p,\ n = 1, \dots, N\}, \qquad N = 1, \dots N_{\max},$$

respectively. For computational convenience, $V_N^{\mathcal{N},\boldsymbol{\mu}}$can be rewritten using the affine dependence of $b(\cdot, \cdot; \boldsymbol{\mu})$ on the parameter and the relation (3.51), yielding to

$$V_N^{\mathcal{N},\boldsymbol{\mu}} = \text{span}\ \{\sum_{k=1}^{\overline{Q}_b} \Theta_b^k(\boldsymbol{\mu})\boldsymbol{\sigma}_{kn},\ n = 1, \dots, 2N\}, \tag{A.1}$$

where $\overline{Q}_b = Q_b + 1, \Theta_b^{\overline{Q}_b} = 1$ and, for $n = 1, \dots, N$,

$$\boldsymbol{\sigma}_{kn} = 0, \quad k = 1, \dots, Q_b; \qquad \boldsymbol{\sigma}_{\overline{Q}_b n} = \tilde{\boldsymbol{\zeta}}_n^{\mathbf{v}} = \mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}^n), \tag{A.2}$$

while, for $n = N+1, \dots, 2N$ (in order to take account of the supremizer operator),

$$(\boldsymbol{\sigma}_{kn}, \mathbf{w})_V = b^k(\tilde{\zeta}_{n-N}^p, \mathbf{w}), \ \forall\ \mathbf{w} \in V^{\mathcal{N}}, \quad k = 1, \dots, Q_b; \qquad \boldsymbol{\sigma}_{\overline{Q}_b n} = 0. \tag{A.3}$$

For a new parameter value $\boldsymbol{\mu}$, the RB solution can be written as a combination of (previously computed and stored) basis functions as:

$$\mathbf{v}_N(\boldsymbol{\mu}) = \sum_{j=1}^{2N} u_{Nj}(\boldsymbol{\mu}) \Big( \sum_{k=1}^{\overline{Q}_b} \Theta_b^k(\boldsymbol{\mu}) \boldsymbol{\sigma}_{kj} \Big), \qquad p_N(\boldsymbol{\mu}) = \sum_{l=1}^{N} p_{Nl}(\boldsymbol{\mu}) \tilde{\zeta}_l^p,$$

whose weights $v_{Nj}$ and $p_{Nl}$ are given by the following RB linear system:

$$\begin{cases} \displaystyle\sum_{j=1}^{2N} \sum_{q=1}^{Q_a} \Theta_a^q(\boldsymbol{\mu})\ A_{ij}^q(\boldsymbol{\mu}) v_{Nj}(\boldsymbol{\mu}) + \sum_{l=1}^{N} \sum_{q=1}^{Q_b} \Theta_b^q(\boldsymbol{\mu})\ B_{il}^q(\boldsymbol{\mu}) p_{Nl}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_a} \Theta_F^q(\boldsymbol{\mu}) f_i^q(\boldsymbol{\mu}), & 1 \le i \le 2N, \\ \displaystyle\sum_{j=1}^{2N} \sum_{q=1}^{Q_b} \Theta_b^q(\boldsymbol{\mu})\ B_{jl}^q(\boldsymbol{\mu}) v_{Nj}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_G} \Theta_G^q(\boldsymbol{\mu}) g_l^q(\boldsymbol{\mu}) & 1 \le l \le N, \end{cases}$$

where for $1 \le i, j \le 2N$ and $1 \le l \le N$:

$$A_{ij}^q(\boldsymbol{\mu}) = \sum_{k'=1}^{\overline{Q}_b} \sum_{k''=1}^{\overline{Q}_b} \Theta_b^{k'}(\boldsymbol{\mu}) \Theta_b^{k''}(\boldsymbol{\mu}) a^q(\boldsymbol{\sigma}_{k'i}, \boldsymbol{\sigma}_{k''j}), \quad B_{il}^q(\boldsymbol{\mu}) = \sum_{k'=1}^{\overline{Q}_b} \Theta_b^{k'}(\boldsymbol{\mu}) b^q(\tilde{\zeta}_l^p, \boldsymbol{\sigma}_{k'i}),$$

$$f_i^q(\boldsymbol{\mu}) = \sum_{k'=1}^{\overline{Q}_b} \Theta_b^{k'}(\boldsymbol{\mu}) F^q(\boldsymbol{\sigma}_{k'i}); \qquad g_l^q(\boldsymbol{\mu}) = G^q(\zeta_l^p).$$

### Orthonormalization: Gram-Schmidt (GS) algorithm

Since the direct choice of the retained snapshots as basis functions would lead to a very poorly conditioned RB system, we apply the Gram-Schmidt (GS) orthogonalization procedure to create a well-conditioned set of basis functions [237]. In this way, given a set of vectors $\mathbf{v}_j, j = 1, ..., N$, we obtain an orthonormal basis $\{\mathbf{z}_1, \dots, \mathbf{z}_N\}$ of $< \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N >$ defined as

$$\mathbf{z}_j = \frac{\mathbb{P}_j \mathbf{v}_j}{||\mathbb{P}_j \mathbf{v}_j||}, \qquad \text{where} \quad \mathbb{P}_j = \mathbb{I}_j - \mathbb{L}_{j-1} \mathbb{L}_{j-1}^T \quad \text{and} \quad \mathbb{L}_{j-1} = \{\mathbf{v}_1, ..., \mathbf{v}_{j-1}\},$$

where $\mathbb{I}_{j-1} \in \mathbb{R}^{j-1 \times j-1}$ is the identity matrix. The GS algorithm reads as follows:

$\mathbf{z}_1 = \mathbf{v}_1 / \|\mathbf{v}_1\|;$
`for` $j = 2 : N$
$\mathbb{P}_j \mathbf{v}_j = \mathbf{v}_j - \sum_{m=1}^{j-1} (\mathbf{z}_m^T \mathbf{v}_j) \mathbf{z}_m;$
$\mathbf{z}_j = \mathbb{P}_j \mathbf{v}_j / \|\mathbb{P}_j \mathbf{v}_j\|;$
`end.`

The norm $\|.\|$ used here is the $V = (H^1(\Omega))^2$ for velocity (and supremizers) and $L^2(\Omega)$ for pressure, respectively; the scalar product is the one induced by the functional space and the norm we use. The orthonormalization procedure has been applied, separately, to our set of velocity snapshots, of supremizer snapshots and to our set of pressure snapshots. For velocity and pressure snapshots the procedure is standard. For the supremizer (computed) snapshots we have to introduce some considerations. Referring to previous supremizer formulation of (A.3) and to the compact notation already introduced in the reference domain, thanks to the linearity of supremizer operator and to the affine composition property we have for $n = N+1, \dots, 2N$:

$$(\boldsymbol{\sigma}_n, \mathbf{w})_V = \sum_{q=1}^{Q^b} \Theta_b^q(\boldsymbol{\mu})(\boldsymbol{\sigma}_{qn}, \mathbf{w})_V = \sum_{q=1}^{Q^b} \Theta_b^q(\boldsymbol{\mu}) b^q(\tilde{\zeta}^p_{n-N}, \mathbf{w}) \qquad \forall \mathbf{w} \in (H^1_{0,\Gamma_D}(\Omega))^2; \quad \text{(A.4)}$$

in fact $(\boldsymbol{\sigma}_n, \mathbf{w})_V = b(\tilde{\zeta}^p_{n-N}, \mathbf{w}; \boldsymbol{\mu})$, but recalling the affine decomposition of the bilinear form $b(\cdot, \cdot; \boldsymbol{\mu})$, for each component $q = 1, \dots, Q^b$ we have:

$$(\boldsymbol{\sigma}_{qn}, \mathbf{w})_V = b^q(\tilde{\zeta}^p_{n-N}, \mathbf{w}).$$

At this point we have two possibilities (referring to $n-th$ supremizer $\boldsymbol{\sigma}_n$, $n = N+1, \dots, 2N$) in applying orthonormalization:

a) an orthonormalization (GS) directly on $\boldsymbol{\sigma}_n$ done Online (since $\boldsymbol{\sigma}_n$ is dependent on $\boldsymbol{\mu}$) to obtain $\boldsymbol{\sigma}_n^\perp$ as new element (basis function) to enrich the reduced basis velocity space:

$$\boldsymbol{\sigma}_n^\perp = \frac{\mathbb{P}_n^\perp \boldsymbol{\sigma}_n}{||\mathbb{P}_n^\perp \boldsymbol{\sigma}_n||} = \frac{\mathbb{P}_n^\perp (\sum_{q=1}^{Q_b} \Theta_b^q(\boldsymbol{\mu}) \boldsymbol{\sigma}_{qn})}{||\mathbb{P}_n^\perp (\sum_{q=1}^{Q_b} \Theta_b^q(\boldsymbol{\mu}) \boldsymbol{\sigma}_{qn})||},$$

where $\mathbb{P}_n^\perp = \mathbb{I} - \mathbb{L}_{n-1}\mathbb{L}_{n-1}^T$ and $\mathbb{L}_i = \{\boldsymbol{\sigma}_1^\perp, ..., \boldsymbol{\sigma}_i^\perp\}$;

b) an orthonormalization (GS) on components $\boldsymbol{\sigma}_{qn}$ made Offline once and for all, since $\boldsymbol{\sigma}_{qn}$ are not depending on $\boldsymbol{\mu}$, to get $\boldsymbol{\sigma}_{qn}^{\perp *}$:

$$\boldsymbol{\sigma}_{qn}^{\perp *} = \frac{\mathbb{P}_{qn}^\perp \boldsymbol{\sigma}_{qn}}{||\mathbb{P}_{qn}^\perp \boldsymbol{\sigma}_{qn}||},$$

where $\mathbb{P}_{qn}^\perp = \mathbb{I} - \mathbb{L}_{q(n-1)}\mathbb{L}_{q(n-1)}^T$ and $\mathbb{L}_{qi} = \{\boldsymbol{\sigma}_{q1}^{\perp *}, ..., \boldsymbol{\sigma}_{qi}^{\perp *}\}$.

Nevertheless, if we apply the orthonormalization algorithm to the RB spaces assembled as proposed above, we may lose the guarantee of the approximation stability – heuristically we do not have any guarantee to fulfill (3.68). For this reason, other options in building the RB velocity space when applying orthonormalization procedures have been proposed.

**Approximation stability: other supremizer options**

In particular, by orthonormalizing the supremizer solutions according to the approach *(a)* we could lose approximation stability in the attempt of preserving algebraic stability by reducing the condition number. We can orthonormalize just using method *(a)* the pressure $\tilde{\zeta}^p$ and the velocity $\tilde{\boldsymbol{\zeta}}^{\mathbf{v}}$ basis functions and not the supremizer $\sigma_n$ and use the approach *(b)* to orthogonalize the supremizer on its component $\sigma_{kn}$ (before summation) to preserve approximation stability. To achieve this goal we may introduce two further different options in assembling the supremizer solutions for stabilization procedure by building in a different way the RB velocity approximation space so that we may guarantee both approximation and algebraic stability.

i) *First option ("splitted supremizer").* We have considered the following reduced basis spaces:

$$\tilde{V}_N^{\mathcal{N}} = \text{span}\,\{\boldsymbol{\sigma}_n, n = 1, \dots, N\overline{Q}_b\},$$

where

$$\boldsymbol{\sigma}_n = \tilde{\boldsymbol{\zeta}}_n^{\mathbf{v}} = \mathbf{v}(\boldsymbol{\mu}^n), \qquad \text{for } n = 1, \dots, N$$

and

$$(\boldsymbol{\sigma}_n, \mathbf{z})_V = (\tilde{\boldsymbol{\sigma}}_{mk}, \mathbf{z})_V, \qquad \text{for } n = N+1, \dots, N\overline{Q}_b,$$

condensing index $m$ and $k$ in $n$, where for all $\mathbf{z} \in X$

$$(\tilde{\boldsymbol{\sigma}}_{mk}, \mathbf{z})_V = b^k(\tilde{\zeta}_m^p, \mathbf{z}), \qquad k = 1, \dots, Q^b,\ m = 1, \dots, N.$$

In this way, the RB solution can be expressed as

$$\mathbf{v}_N(\boldsymbol{\mu}) = \sum_{j=1}^{N\overline{Q}_b} v_{Nj}(\boldsymbol{\mu})\boldsymbol{\sigma}_j, \qquad p_N(\boldsymbol{\mu}) = \sum_{l=1}^{N} p_{Nl}(\boldsymbol{\mu})\tilde{\zeta}_l^p;$$

and the reduced basis system (3.69) becomes:

$$\begin{cases} \displaystyle\sum_{j=1}^{N\overline{Q}_b}\sum_{q=1}^{Q_a} \Theta_a^q(\boldsymbol{\mu})\, A_{ij}^q\, \mathbf{v}_{Nj}(\boldsymbol{\mu}) + \sum_{l=1}^{N}\sum_{q=1}^{Q_b} \Theta_b^q(\boldsymbol{\mu})\, B_{il}^q\, p_{Nl}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_F} \Theta_F^q(\boldsymbol{\mu}) f_i^q, & 1 \le i \le N\overline{Q}_b, \\ \displaystyle\sum_{j=1}^{N\overline{Q}_b}\sum_{q=1}^{Q_b} \Theta_b^q(\boldsymbol{\mu})\, B_{jl}^q\, \mathbf{v}_{Nj}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_G} \Theta_G^q(\boldsymbol{\mu}) g_l^q, & 1 \le l \le N, \end{cases} \tag{A.5}$$

where:

$$(A^q)_{ij} = a^q(\boldsymbol{\sigma}_i, \boldsymbol{\sigma}_j), \qquad 1 \le i, j \le N\overline{Q}_b,$$

$$(B^q)_{il} = b^q(\tilde{\zeta}_l^p, \boldsymbol{\sigma}_i), \qquad 1 \le i \le N\overline{Q}_b,\ 1 \le l \le N,$$

$$f_i^q = F^q(\boldsymbol{\sigma}_i), \quad 1 \le i \le N\overline{Q}_b, \qquad g_l^q = G^q(\tilde{\zeta}_l^p), \quad 1 \le l \le N.$$

In this case the basis is no longer $\boldsymbol{\mu}$ (Online) dependent, and the RB velocity space, enriched by supremizers, has a bigger dimension ($N\overline{Q}_b > 2N$) than previously. The computational costs are as follows: $O(Q_a(\overline{Q}_b)^2N^2)$ for sub-matrix $\mathbb{A}_N$, $O(\overline{Q}_bQ_bN^2)$ for $\mathbb{B}_N$, $O(\overline{Q}_bQ_FN)$ for $\mathbf{f}$, $O(Q_GN)$ for $\mathbf{g}$ but the cost for inversion of the full RB matrix increases now to $O((\overline{Q}_b + 1)^3N^3)$. This approach has the big advantage to preserve the approximation stability, to let us apply orthonormalization (method (ii)) and to preserve stability also after orthonormalization. Nevertheless, the cost for assembling and inverting the global RB matrix is still dependent on the parametric complexity $Q_a, Q_b$ of the problem, which can be rather high, above all in nonaffinely parametrized problems.

ii) *Second option ("global supremizer").* Another approach is based on the idea that the supremizers are built upon summation using the same $\boldsymbol{\mu}^j$ values used to store velocity $\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}^j)$ and pressure solutions $p^{\mathcal{N}}(\boldsymbol{\mu}^j)$ (also in this case the basis for velocity is not dependent on the Online value of $\boldsymbol{\mu}$ and it is completely assembled Offline):

$$V_N^{\mathcal{N},\overline{\boldsymbol{\mu}}} = \text{span}\,\{\boldsymbol{\sigma}_n = \sum_{k=1}^{\overline{Q}_b} \Theta_b^k(\boldsymbol{\mu}^n)\boldsymbol{\sigma}_{kn},\ n = 1,\dots,2N\}, \tag{A.6}$$

where $\overline{Q}_b = Q_b + 1$, $\Theta_b^{\overline{Q}_b} = 1$ and, for $n = 1,\dots,N$,

$$\boldsymbol{\sigma}_{kn} = 0, \quad \text{for } k = 1,\dots,Q_b, \qquad\qquad \boldsymbol{\sigma}_{\overline{Q}_b n} = \tilde{\boldsymbol{\zeta}}_n^{\mathbf{v}} = \mathbf{v}(\boldsymbol{\mu}^n),$$

while for $n = N+1,\dots,2N$

$$(\boldsymbol{\sigma}_{kn},\mathbf{z})_V = b^k(\tilde{\zeta}_{n-N}^p,\mathbf{z}) \ \ \forall\ \mathbf{z}\in Y, \quad \text{for } k = 1,\dots,Q_b, \qquad\qquad \boldsymbol{\sigma}_{\overline{Q}_b n} = 0.$$

The reduced basis solution is thus given by

$$\mathbf{v}_N(\boldsymbol{\mu}) = \sum_{j=1}^{2N} v_{Nj}(\boldsymbol{\mu})\Big(\sum_{k=1}^{\overline{Q}_b} \Theta_b^k(\boldsymbol{\mu}^j)\boldsymbol{\sigma}_{kj}\Big), \qquad\qquad p_N(\boldsymbol{\mu}) = \sum_{l=1}^{N} p_{Nl}(\boldsymbol{\mu})\tilde{\zeta}_l^p,$$

and is obtained by solving the following system:

$$\begin{cases} \displaystyle\sum_{j=1}^{2N}\sum_{q=1}^{Q_a} \Theta_a^q(\boldsymbol{\mu})\, A_{ij}^q\, \mathbf{v}_{Nj}(\boldsymbol{\mu}) + \sum_{l=1}^{N}\sum_{q=1}^{Q_b} \Theta_b^q(\boldsymbol{\mu})\, B_{il}^q\, p_{Nl}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_F} \Theta_F^q(\boldsymbol{\mu}) f_i^q, & 1 \le i \le 2N, \\ \displaystyle\sum_{j=1}^{2N} B_{jl}^q\, \mathbf{v}_{Nj}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_G} \Theta_G^q(\boldsymbol{\mu}) g_l^q, & 1 \le l \le N, \end{cases} \tag{A.7}$$

where, for $1 \le i,j \le 2N$, $1 \le l \le N$:

$$(A^q)_{ij} = a^q(\boldsymbol{\sigma}_i,\boldsymbol{\sigma}_j) = \sum_{k'=1}^{\overline{Q}_b}\sum_{k''=1}^{\overline{Q}_b} \Theta_b^{k'}(\boldsymbol{\mu}^i)\Theta_b^{k''}(\boldsymbol{\mu}^j) a^q(\boldsymbol{\sigma}_{k'i},\boldsymbol{\sigma}_{k''j});$$

$$B_{il}^q = b^q(\tilde{\zeta}_l^p,\boldsymbol{\sigma}_i) = \sum_{k'=1}^{\overline{Q}_b} \Theta_b^{k'}(\boldsymbol{\mu}^i) b^q(\tilde{\zeta}_l^p,\boldsymbol{\sigma}_{k'i});$$

$$f_i^q = F^q(\sigma_i) = \sum_{k'=1}^{\overline{Q}_b} \Theta_b^{k'}(\boldsymbol{\mu}^i) F^q(\boldsymbol{\sigma}_{k'i}), \qquad\qquad g_l^q = G^q(\tilde{\zeta}_l^p).$$

This option is also competitive as regards the computational costs dealing with $3N \times 3N$ RB matrix instead of $(\overline{Q}_b+1)N \times (\overline{Q}_b+1)N$ matrix (usually $(\overline{Q}_b+1) \gg 3$). We thus have the following computational costs to build the RB matrices, given also the supremizer components in the velocity space: $O(Q_a 4N^2)$ for sub-matrix $\mathbb{A}_N$, $O(Q^b 2N^2)$ for $\mathbb{B}_N$, $O(Q_F N)$ for $\mathbf{f}$, $O(Q_G N)$ for $\mathbf{g}$ and $O(27N^3)$ for the inversion of the full RB matrix, which are usually smaller than computational costs of the previous option.
Using this option we cannot rigorously demonstrate that the approximation stability is preserved (even without orthonormalization); nevertheless, after several numerical tests, we can safely argue that this option is very efficient – by allowing us to reduce the matrix dimension of the RB system – and reasonably stable. Certified *a posteriori* error bounds are another proof of guaranteed stability using this approach combined with orthonormalization.

In the RB approximation procedure used throughout this thesis work, we have used this second *"global supremizer"* option, denoting by $V_N$ the RB velocity space $V_N^{\mathcal{N},\overline{\boldsymbol{\mu}}}$ for the sake of simplicity. Numerical tests and comparisons about the different supremizer options have been reported in previous works [278, 282] for Stokes and [255, 277] for Navier-Stokes equations.

## A.2 Offline/Online procedure for dual norms of residuals

In order to compute the error estimates in a very efficient way, the same Offline/Online procedure introduced for the RB approximation is used. In this section we describe the evaluation of the dual norm of the residuals in both Stokes and Navier-Stokes cases.

### A.2.1 Stokes equations

Recalling Sect. 3.4.1, residuals $r_{\mathbf{v}}(\cdot\,;\boldsymbol{\mu})$ and $r_p(\cdot\,;\boldsymbol{\mu})$ are defined as:

$$\begin{aligned} &r_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu}) := F(\mathbf{w};\boldsymbol{\mu}) - a(\mathbf{v}_N(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}) - b(p_N(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}), \quad \forall\ \mathbf{w}\in V^{\mathcal{N}},\\ &r_p(q;\boldsymbol{\mu}) := G(q;\boldsymbol{\mu}) - b(q,\mathbf{v}_N(\boldsymbol{\mu});\boldsymbol{\mu}), \qquad\qquad\qquad \forall\ q\in Q^{\mathcal{N}}. \end{aligned} \tag{A.8}$$

In order to evaluate these quantities exploiting an Offline/Online procedure, we introduce the Riesz representations $\hat{\mathbf{e}}_{\mathbf{v}}(\boldsymbol{\mu})\in V^{\mathcal{N}}$ and $\hat{e}_p(\boldsymbol{\mu})\in Q^{\mathcal{N}}$, defined as follows:

$$(\hat{\mathbf{e}}_{\mathbf{v}}(\boldsymbol{\mu}),\mathbf{w})_X = r_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu}),\ \ \forall \mathbf{w}\in V^{\mathcal{N}}, \quad (\hat{e}_p(\boldsymbol{\mu}),q)_Q = r_p(q;\boldsymbol{\mu}),\ \ \forall q\in Q^{\mathcal{N}}. \tag{A.9}$$

This allows us to write (A.8) as

$$\begin{aligned} a(\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}) + b(e_p(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}) &= (\hat{\mathbf{e}}_{\mathbf{v}}(\boldsymbol{\mu}),\mathbf{w})_V \qquad \forall\ \mathbf{w}\in V^{\mathcal{N}},\\ b(q,\mathbf{e}_{\mathbf{v}}(\boldsymbol{\mu});\boldsymbol{\mu}) &= (\hat{e}_p(\boldsymbol{\mu}),q)_Q \qquad \forall\ q\in Q^{\mathcal{N}}. \end{aligned}$$

and to evaluate the dual norm of residuals as

$$\|r_{\mathbf{v}}(\cdot\,;\boldsymbol{\mu})\|_{V'} = \sup_{\mathbf{w}\in V^{\mathcal{N}}}\frac{r_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu})}{\|\mathbf{w}\|_V} = \|\hat{\mathbf{e}}_{\mathbf{v}}(\boldsymbol{\mu})\|_V, \qquad \|r_p(\cdot\,;\boldsymbol{\mu})\|_{Q'} = \sup_{q\in Q^{\mathcal{N}}}\frac{r_p(q;\boldsymbol{\mu})}{\|q\|_Q} = \|\hat{e}_p(\boldsymbol{\mu})\|_Q. \tag{A.10}$$

Let us consider the global Stokes operator $A: X\times X\to\mathbb{R}$ defined by (3.12), in the case $\delta=0$. From the affine decomposition (3.16) we can write, equivalently,

$$A(Y,W;\boldsymbol{\mu}) = \sum_{q=1}^{Q_A}\Theta_A^q(\boldsymbol{\mu})A^q(Y,W), \qquad \tilde{F}(W;\boldsymbol{\mu}) = \sum_{q=1}^{Q_{\tilde{F}}}\Theta_{\tilde{F}}^q(\boldsymbol{\mu})\tilde{F}^q(W) \tag{A.11}$$

where $Q_A = Q_a + 2Q_b$, $Q_{\tilde{F}} = Q_F + Q_G$,

$$\Theta_A^q(\boldsymbol{\mu}) = \Theta_a^q(\boldsymbol{\mu}), \quad \Theta_A^{q'+Q_a}(\boldsymbol{\mu}) = \Theta_A^{q'+Q_a+Q_b}(\boldsymbol{\mu}) = \Theta_b^{q'}(\boldsymbol{\mu}), \qquad q=1,\dots,Q_a,\ q'=1,\dots,Q_b,$$

$$\Theta_{\tilde{F}}^q(\boldsymbol{\mu}) = \Theta_F^q(\boldsymbol{\mu}), \quad \Theta_{\tilde{F}}^{q'+Q_F}(\boldsymbol{\mu}) = \Theta_G^{q'}(\boldsymbol{\mu}), \qquad q=1,\dots,Q_F,\ q'=1,\dots,Q_G.$$

and, denoting $Y=(\mathbf{v},p)$, $W=(\mathbf{w},q)$,

$$\begin{aligned} &A^q(Y,W) = a^q(\mathbf{v},\mathbf{w}) && q=1,\dots,Q_a,\\ &A^q(Y,W) = b^q(p,\mathbf{w}) && q=Q_a+1,\dots,Q_a+Q_b,\\ &A^q(Y,W) = b^q(q,\mathbf{v}) && q=Q_a+Q_b+1,\dots,Q_a+2Q_b; \end{aligned}$$

$$\begin{aligned} &\tilde{F}^q(W) = F^q(\mathbf{w}) && q=1,\dots,Q_F,\\ &\tilde{F}^q(W) = G^q(q) && q=Q_F+1,\dots,Q_F+Q_F. \end{aligned}$$

The global residual $r(W;\boldsymbol{\mu}) := r_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu}) + r_p(q;\boldsymbol{\mu})$ can be expressed, considering the "global

supremizer" option of Sec. 4.2, as

$$r(W;\boldsymbol{\mu}) = \tilde{F}(W) - A(Y_N(\boldsymbol{\mu}), W; \boldsymbol{\mu}) = \sum_{q=1}^{Q_{\tilde{F}}} \Theta_{\tilde{F}}^q(\boldsymbol{\mu}) \tilde{F}^q(W) - \sum_{n=1}^{3N} Y_{N\,n}(\boldsymbol{\mu}) \sum_{q=1}^{Q_A} \Theta_A^q(\boldsymbol{\mu}) A^q(\boldsymbol{\Phi}_n, W),$$

where $Y_N(\boldsymbol{\mu}) = (\mathbf{v}_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu}))$ is the global RB solution and $\boldsymbol{\Phi}_n = (\boldsymbol{\sigma}_n, 0)$ for $n = 1, \ldots, 2N$, $\boldsymbol{\Phi}_n = (\mathbf{0}, \tilde{\zeta}_n^p)$ for $n = 2N+1, \ldots, 3N$. Together with (A.9) and linear superposition, this gives us

$$(\hat{\mathbf{e}}(\boldsymbol{\mu}), W)_X = (\hat{\mathbf{e}}_{\mathbf{v}}(\boldsymbol{\mu}), \mathbf{w})_V + (\hat{e}_p(\boldsymbol{\mu}), q)_Q = \sum_{q=1}^{Q_{\tilde{F}}} \Theta_{\tilde{F}}^q(\boldsymbol{\mu}) \tilde{F}^q(W) - \sum_{n=1}^{3N} Y_{N\,n}(\boldsymbol{\mu}) \sum_{q=1}^{Q_A} \Theta_A^q(\boldsymbol{\mu}) A^q(\boldsymbol{\Phi}_n, W);$$

the element $\hat{\mathbf{e}}(\boldsymbol{\mu}) = (\hat{\mathbf{e}}_{\mathbf{v}}(\boldsymbol{\mu}), \hat{e}_p(\boldsymbol{\mu})) \in X^{\mathcal{N}}$ can thus be expressed as

$$\hat{\mathbf{e}}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{\tilde{F}}} \Theta_{\tilde{F}}^q(\boldsymbol{\mu}) \tilde{\mathcal{F}}^q + \sum_{q=1}^{Q_A} \sum_{n=1}^{3N} \Theta_A^q(\boldsymbol{\mu}) Y_{N\,n}(\boldsymbol{\mu}) \mathcal{A}_n^q, \tag{A.12}$$

where $\tilde{\mathcal{F}}^q \in X^{\mathcal{N}}$ and $\mathcal{A}_n^q \in X^{\mathcal{N}}$ (called FE "pseudo"-solutions) satisfy

$$(\tilde{\mathcal{F}}^q, W)_X = \tilde{F}^q(W), \quad \forall\, W \in X^{\mathcal{N}}, \qquad 1 \le q \le Q_{\tilde{F}} \tag{A.13}$$

$$(\mathcal{A}_n^q, W)_X = -A^q(\boldsymbol{\Phi}_n, W), \quad \forall\, W \in X^{\mathcal{N}}, \qquad 1 \le n \le 3N, \quad 1 \le q \le Q_A. \tag{A.14}$$

We note that (A.13) and (A.14) are simple parameter-independent problems and thus can be solved once in the Offline stage. By using (A.12), it then follows that:

$$\|\hat{\mathbf{e}}(\boldsymbol{\mu})\|_X^2 = \sum_{q''=1}^{Q_{\tilde{F}}} \sum_{q'''=1}^{Q_{\tilde{F}}} \Theta_{\tilde{F}}^{q''}(\boldsymbol{\mu}) \Theta_{\tilde{F}}^{q'''}(\boldsymbol{\mu}) (\tilde{\mathcal{F}}^{q''}, \tilde{\mathcal{F}}^{q'''})_X + \sum_{q=1}^{Q_A} \sum_{n=1}^{3N} \Theta_A^q(\boldsymbol{\mu}) Y_{N\,n}(\boldsymbol{\mu})$$
$$\left\{ 2 \sum_{q''=1}^{Q_{\tilde{F}}} \Theta_{\tilde{F}}^{q''}(\boldsymbol{\mu}) (\tilde{\mathcal{F}}^{q''}, \mathcal{A}_n^q)_X + \sum_{q'=1}^{Q_A} \sum_{n'=1}^{3N} \tilde{\Theta}^{q'}(\boldsymbol{\mu}) Y_{N\,n'}(\boldsymbol{\mu}) (A_n^q, A_{n'}^{q'})_X \right\}. \tag{A.15}$$

This expression can be related to the dual norm of the residual through (A.10). It is the sum of products of $\boldsymbol{\mu}$-dependent known functions and $\boldsymbol{\mu}$-independent inner products, formed of more complicated but precomputable quantities. The Offline/Online decomposition is thus clear:

(i) in the Offline stage we first solve (A.13), (A.14) for the $\boldsymbol{\mu}$-independent FE "pseudo"-solutions $\tilde{\mathcal{F}}^{q''}$ and $\mathcal{A}_n^q$, $1 \le n \le 3N$, $1 \le q \le Q_A$, $1 \le q'' \le Q_{\tilde{F}}$ and form/store the $\boldsymbol{\mu}$-independent inner products $(\tilde{\mathcal{F}}^{q''}, \tilde{\mathcal{F}}^{q'''})_X$, $(\tilde{\mathcal{F}}^{q''}, \mathcal{A}_n^q)_X$ and $(\tilde{\mathcal{A}}_n^q, \mathcal{A}_{n'}^{q'})_X$, being $1 \le n, n' \le 3N$, $1 \le q, q' \le Q_A$, $1 \le q'', q''' \le Q_{\tilde{F}}$. The Offline operation count depends on $N$, $Q_A + Q_{\tilde{F}}$ and $\mathcal{N}$;

(ii) in the Online stage - performed for each new value of $\boldsymbol{\mu}$ - we simply evaluate the sum (A.15) in terms of the $\Theta_A^q(\boldsymbol{\mu})$, $1 \le q \le Q_A$, $\Theta_{\tilde{F}}^{q''}(\boldsymbol{\mu})$, $1 \le q'' \le Q_{\tilde{F}}$, and $Y_{N\,n}(\boldsymbol{\mu})$, $1 \le n \le 3N$ (already computed for the output evaluation) and the precomputed and stored ($\boldsymbol{\mu}$-independent) $(\cdot,\cdot)_X$ inner products. The Online operation count is only $\mathcal{O}((Q_A + Q_{\tilde{F}})^2 9N^2)$, and thus the crucial point - the independence of $\mathcal{N}$ - is achieved again.

### A.2.2 Navier-Stokes equations

The same Offline/Online evaluation procedure for the residuals can be applied to the Navier-Stokes case in the same way. Clearly, nonlinear terms make much more complicated the Offline assembling stage, due to the increased parametric complexity. Recalling Sect. 3.4.4, residuals

$r_{\mathbf{v}}(\cdot\,;\boldsymbol{\mu})$ and $r_p(\cdot\,;\boldsymbol{\mu})$ for velocity and pressure equations, respectively, are defined in this case by:

$$
\begin{aligned}
r_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu}) &:= F(\mathbf{w};\boldsymbol{\mu}) - \tilde{a}(\mathbf{v}_N(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}) - b(p_N(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}) - c(\mathbf{v}_N(\boldsymbol{\mu}),\mathbf{v}_N(\boldsymbol{\mu}),\mathbf{w};\boldsymbol{\mu}),\\
r_p(q;\boldsymbol{\mu}) &:= G(q;\boldsymbol{\mu}) - b(q,\mathbf{v}_N(\boldsymbol{\mu});\boldsymbol{\mu}),
\end{aligned}
$$

for any $\mathbf{w} \in V^{\mathcal{N}}$, $q \in Q^{\mathcal{N}}$, respectively, being $\tilde{a}(\cdot,\cdot;\boldsymbol{\mu}) = a(\cdot,\cdot;\boldsymbol{\mu}) + d(\cdot,\cdot;\boldsymbol{\mu})$. In the more compact form, the global residual $r(W;\boldsymbol{\mu}) := r_{\mathbf{v}}(\mathbf{w};\boldsymbol{\mu}) + r_p(q;\boldsymbol{\mu})$ is given by

$$
r(W;\boldsymbol{\mu}) = \tilde{F}(W;\boldsymbol{\mu}) - \tilde{A}(Y_N(\boldsymbol{\mu}),W;\boldsymbol{\mu}) \qquad \forall\, W \in X^{\mathcal{N}}, \tag{A.16}
$$

where $\tilde{A}(V,W;\boldsymbol{\mu}) = A(V,W;\boldsymbol{\mu}) + C(V,V,W;\boldsymbol{\mu})$, where $C(Y,Y,W;\boldsymbol{\mu}) = c(\mathbf{v},\mathbf{v},\mathbf{w};\boldsymbol{\mu})$ is the term relative to the nonlinear operator. Following the same construction and notation introduced in the Stokes case, we can replace $A(\cdot,\cdot;\boldsymbol{\mu})$ by $\tilde{A}(\cdot,\cdot;\boldsymbol{\mu})$ and $Q_A$ by $Q_{\tilde{A}} = Q_A + Q_C$ in (A.11), exploiting the affine decomposition of the nonlinear operator (3.19). By keeping the contribution of the nonlinear term separate from the linear Stokes operator, we can express (A.16) as

$$
\begin{aligned}
r(W;\boldsymbol{\mu}) \equiv (\hat{\mathbf{e}}(\boldsymbol{\mu}),W)_X &= \tilde{F}(W) - A(Y_N(\boldsymbol{\mu}),W;\boldsymbol{\mu}) - C(Y_N(\boldsymbol{\mu}),Y_N(\boldsymbol{\mu}),W;\boldsymbol{\mu})\\
&= \sum_{q=1}^{Q_{\tilde{F}}} \Theta_{\tilde{F}}^q(\boldsymbol{\mu})\tilde{F}^q(W) - \sum_{n=1}^{3N} Y_{N\,n}(\boldsymbol{\mu}) \sum_{q=1}^{Q_A} \Theta_A^q(\boldsymbol{\mu}) A^q(\boldsymbol{\Phi}_n,W)\\
&\quad - \sum_{n=1}^{2N}\sum_{n'=1}^{2N} Y_{N\,n}(\boldsymbol{\mu}) Y_{N\,n'}(\boldsymbol{\mu}) \sum_{q=1}^{Q_C} \Theta_C^q(\boldsymbol{\mu}) C^q(\boldsymbol{\Phi}_n,\boldsymbol{\Phi}_{n'},W),
\end{aligned}
$$

so that $\hat{\mathbf{e}}(\boldsymbol{\mu})$ can be expressed as

$$
\hat{\mathbf{e}}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{\tilde{F}}} \Theta_{\tilde{F}}^q(\boldsymbol{\mu})\tilde{\mathcal{F}}^q + \sum_{q=1}^{Q_A}\sum_{n=1}^{3N} \Theta_A^q(\boldsymbol{\mu}) Y_{N\,n}(\boldsymbol{\mu}) \mathcal{A}_n^q + \sum_{q=1}^{Q_C}\sum_{n=1}^{2N}\sum_{n'=1}^{2N} \Theta_C^q(\boldsymbol{\mu}) Y_{N\,n}(\boldsymbol{\mu}) Y_{N\,n'}(\boldsymbol{\mu}) \mathcal{C}_{nn'}^q
$$

by introducing the FE "pseudo"-solutions $\tilde{\mathcal{F}}^q \in X^{\mathcal{N}}$, $\mathcal{A}_n^q \in X^{\mathcal{N}}$, $\mathcal{C}_{nn'}^q \in X^{\mathcal{N}}$, which satisfies:

$$
\begin{aligned}
(\tilde{\mathcal{F}}^q,W)_X &= \tilde{F}^q(W), \quad \forall\, W \in X^{\mathcal{N}},\ 1 \le q \le Q_{\tilde{F}}\\
(\mathcal{A}_n^q,W)_X &= -A^q(\boldsymbol{\Phi}_n,W), \quad \forall\, W \in X^{\mathcal{N}},\ 1 \le n \le 3N,\ 1 \le q \le Q_A\\
(\mathcal{C}_{nn'}^q,W)_X &= -C^q(\boldsymbol{\Phi}_n,\boldsymbol{\Phi}_{n'},W), \quad \forall\, W \in X^{\mathcal{N}},\ 1 \le n,n' \le 2N,\ 1 \le q \le Q_C.
\end{aligned}
$$

It then follows that

$$
\begin{aligned}
\|\hat{\mathbf{e}}(\boldsymbol{\mu})\|_X^2 = \|\hat{\mathbf{e}}_{\mathrm{S}}(\boldsymbol{\mu})\|_X^2 + \sum_{q=1}^{Q_C}\sum_{n,n'=1}^{2N} \Theta_C^q(\boldsymbol{\mu}) Y_{N\,n}(\boldsymbol{\mu}) Y_{N\,n'}(\boldsymbol{\mu}) \Bigg\{ &\sum_{q''=1}^{Q_{\tilde{F}}} \Theta_{\tilde{F}}^{q''}(\boldsymbol{\mu})(\mathcal{C}_{nn'}^q,\tilde{\mathcal{F}}^{q'''})_X +\\
+ \sum_{q'=1}^{Q_A}\sum_{n''=1}^{3N} \Theta_A^{q'}(\boldsymbol{\mu}) Y_{N\,n''}(\boldsymbol{\mu})(\mathcal{C}_{nn'}^q,\mathcal{A}_{n''}^{q'})_X + \sum_{q'=1}^{Q_C}\sum_{n'',n'''=1}^{2N} &\Theta_C^{q'}(\boldsymbol{\mu}) Y_{N\,n''}(\boldsymbol{\mu}) Y_{N\,n'''}(\boldsymbol{\mu})(\mathcal{C}_{nn'}^q,\mathcal{C}_{n''n'''}^q)_X \Bigg\},
\end{aligned}
$$

where $\|\hat{\mathbf{e}}_{\mathrm{S}}(\boldsymbol{\mu})\|_X^2$ is the expression computed as in (A.15) for the Stokes case, but accounting also for the lifting terms arising from the nonlinear operator.
The Offline/Online decomposition thus features the same structure of the Stokes case, but entails a stronger complexity, due to the presence of nonlinear terms, which requires the storage of the following, additional $\boldsymbol{\mu}$-independent inner products $(\mathcal{C}_{nn'}^q,\tilde{\mathcal{F}}^{q''})_X$, $(\mathcal{C}_{nn'}^q,\mathcal{A}_{n''''}^{q'''})_X$, $(\mathcal{C}_{nn'}^q,\mathcal{C}_{n''n''''}^{q'})_X$, being $1 \le n,n',n'',n''' \le 2N$, $1 \le n'''' \le 3N$, and $1 \le q,q' \le Q_C$, $1 \le 1'' \le Q_{\tilde{F}}$, $1 \le q''' \le Q_A$, with respect to the Stokes case.

## A.3 Successive Constraint Method for stability factors

In this section we review the Successive Constraint Method (SCM) used for the estimation of lower and upper bounds of stability factors. Based on the successive solution of suitable linear optimization problems, this algorithm has been developed for the special requirements of the RB method; it thus features an efficient Offline-Online strategy, making the Online computation of the lower bound independent of $\mathcal{N}$ – a fundamental requisite in view of a rapid error bound evaluation. This algorithm has been first introduced in [150] for both coercive and noncoercive problems, analyzed in [280] in the coercive case and afterwards improved in [61]. A general version using the so-called "natural norm" [298] has been analyzed in [146], where it has been applied to noncoercive problems such as Helmholtz equations – the simpler coercive case can be seen as a particular instance where the stability factor is just the coercivity constant. We apply for the first time this algorithm to saddle point Stokes problems – as shown in the work with Rozza and Huynh [279] – and provide a first extension to Navier-Stokes problems.

### A.3.1 Stokes equations

Our goal is to build a lower bound for the (Babuška) discrete inf-sup stability factor

$$\beta(\boldsymbol{\mu}) := \inf_{V^{\mathcal{N}}\in X^{\mathcal{N}}} \sup_{W^{\mathcal{N}}\in X^{\mathcal{N}}} \frac{A(V^{\mathcal{N}},W^{\mathcal{N}};\boldsymbol{\mu})}{\|V^{\mathcal{N}}\|_X\|W^{\mathcal{N}}\|_X} = \inf_{V^{\mathcal{N}}\in X^{\mathcal{N}}} \frac{A(V^{\mathcal{N}},T^{\boldsymbol{\mu}}V^{\mathcal{N}};\boldsymbol{\mu})}{\|V^{\mathcal{N}}\|_X\|T^{\boldsymbol{\mu}}V^{\mathcal{N}}\|_X} \tag{A.17}$$

(we omit subscript $_{Ba,\mathcal{N}}$ from now on) where $T^{\boldsymbol{\mu}} : X^{\mathcal{N}} \to X^{\mathcal{N}}$ is the (global) supremizer operator defined as

$$(T^{\boldsymbol{\mu}}V^{\mathcal{N}},W^{\mathcal{N}})_X = A(V^{\mathcal{N}},W^{\mathcal{N}};\boldsymbol{\mu}), \quad \forall\, W^{\mathcal{N}} \in X^{\mathcal{N}}, \tag{A.18}$$

such that

$$T^{\boldsymbol{\mu}}V^{\mathcal{N}} = \arg \sup_{W^{\mathcal{N}}\in X^{\mathcal{N}}} \frac{A(V^{\mathcal{N}},W^{\mathcal{N}};\boldsymbol{\mu})}{\|W^{\mathcal{N}}\|_X} \tag{A.19}$$

In particular, we shall observe that the computation of the stability factor (A.17) can be formulated as a suitable eigenproblem. In fact, thanks to (A.19), we have that

$$(\beta(\boldsymbol{\mu}))^2 = \left(\inf_{V^{\mathcal{N}}\in X^{\mathcal{N}}} \frac{\|T^{\boldsymbol{\mu}}V^{\mathcal{N}}\|_X}{\|V^{\mathcal{N}}\|_X}\right)^2 = \inf_{V^{\mathcal{N}}\in X^{\mathcal{N}}} \frac{\|T^{\boldsymbol{\mu}}V^{\mathcal{N}}\|_X^2}{\|V^{\mathcal{N}}\|_X^2}. \tag{A.20}$$

Let us introduce the matrix norm $\mathbb{X} = \mathrm{diag}(\mathbb{X}_V, \mathbb{X}_Q)$ of the space $X^{\mathcal{N}}$, where $\mathbb{X}_V$ and $\mathbb{X}_Q$ are defined in (3.57), its Cholesky decomposition $\mathbb{X} = \mathbb{H}^T\mathbb{H}$, and denote by $\underline{\mathbf{V}} = (\underline{\mathbf{v}}, \underline{\mathbf{p}})^T$ the vector representation of $V^{\mathcal{N}} = (\mathbf{v}^{\mathcal{N}}, p^{\mathcal{N}})^T \in \mathbb{R}^{\mathcal{N}}$, being $\mathcal{N} = \mathcal{N}_X + \mathcal{N}_Q$, so that $\|V^{\mathcal{N}}\|_X^2 = \underline{\mathbf{V}}^T\mathbb{X}\underline{\mathbf{V}}$. Moreover, denoting by $\underline{\mathbf{T}}$ the vector whose components are the coefficients in the expansion of $T^{\boldsymbol{\mu}}V^{\mathcal{N}} \in X^{\mathcal{N}}$, from (A.18) we have that $\underline{\mathbf{T}}^T\mathbb{X}\underline{\mathbf{W}} = \underline{\mathbf{V}}^T\mathbb{A}(\boldsymbol{\mu})\underline{\mathbf{W}}$. We thus obtain that $(\beta(\boldsymbol{\mu}))^2$ is the smallest (generalized) eigenvalue of the following eigenvalue problem[1]:

$$\left(\mathbb{H}^{-T}\mathbb{A}^T(\boldsymbol{\mu})\mathbb{X}^{-1}\ \mathbb{A}(\boldsymbol{\mu})\mathbb{H}^{-1}\right)\underline{\mathbf{V}} = \beta^2(\boldsymbol{\mu})\,\underline{\mathbf{V}}, \qquad \forall\ \underline{\mathbf{V}} \neq \underline{\mathbf{0}}. \tag{A.21}$$

[1] Inserting $\underline{\mathbf{T}} = \mathbb{X}^{-1}\mathbb{A}(\boldsymbol{\mu})\underline{\mathbf{V}}$ in (A.20) $(\beta^{\mathcal{N}}(\mu))^2$ can be expressed as a Rayleigh quotient, i.e.

$$(\beta(\boldsymbol{\mu}))^2 = \inf_{\underline{\mathbf{V}}\in\mathbb{R}^{\mathcal{N}}} \frac{\underline{\mathbf{T}}\mathbb{X}\underline{\mathbf{T}}}{\underline{\mathbf{V}}\mathbb{X}\underline{\mathbf{V}}} = \inf_{\underline{\mathbf{V}}\in\mathbb{R}^{\mathcal{N}}} \frac{\underline{\mathbf{V}}^T\mathbb{A}^T(\boldsymbol{\mu})\mathbb{X}^{-T}\mathbb{X}\mathbb{X}^{-1}\mathbb{A}(\boldsymbol{\mu})\underline{\mathbf{V}}}{\underline{\mathbf{V}}\mathbb{X}\underline{\mathbf{V}}} = \inf_{\underline{\mathbf{V}}\in\mathbb{R}^{\mathcal{N}}} \frac{\underline{\mathbf{V}}^T\mathbb{A}^T(\boldsymbol{\mu})\mathbb{X}^{-1}\mathbb{A}(\boldsymbol{\mu})\underline{\mathbf{V}}}{\underline{\mathbf{V}}\mathbb{X}\underline{\mathbf{V}}},$$

so that $(\beta(\mu))^2 = \sqrt{\boldsymbol{\lambda}_{\min}}$, where $\boldsymbol{\lambda}_{\min}$ is the smallest (generalized) eigenvalue, solving $\mathbb{A}^T(\boldsymbol{\mu})\mathbb{X}^{-1}\ \mathbb{A}(\boldsymbol{\mu})\underline{\mathbf{V}} = \boldsymbol{\lambda}\mathbb{X}\underline{\mathbf{V}}$.

## Appendix A. Insights on RB approximation and error estimation

The original version of SCM proposed in [150] deals with affine coercive operators and thus features a complexity of order $\mathcal{O}(Q_A)$; its extension to affine noncoercive operators – and then to the solution of problem (A.21) – is straightforward, even if this implies a complexity which is of order $\mathcal{O}(Q_A^2)$ – too cumbersome for problems with larger $Q_A$ like the ones coming from EIM. The *natural norm* approach overcomes this limitation by means of a different strategy, based on the computation of a lower bound for a *surrogate* inf-sup stability factor $\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$, given by

$$\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) = \inf_{V^{\mathcal{N}} \in X^{\mathcal{N}}} \sup_{W^{\mathcal{N}} \in X^{\mathcal{N}}} \frac{A(V^{\mathcal{N}}, W^{\mathcal{N}}; \boldsymbol{\mu})}{\|T^{\mu^*} V^{\mathcal{N}}\|_X \|W^{\mathcal{N}}\|_X} \tag{A.22}$$

for some selected parameter value $\boldsymbol{\mu}^*$. Since we assume that $\beta(\boldsymbol{\mu}) > 0$ for all $\boldsymbol{\mu} \in \mathcal{D}$, $\|T^{\mu^*} \cdot\|_X$ is a well-defined norm (which is equivalent to $\|\cdot\|_X$ in a neighborhood $\mathcal{P}_{\boldsymbol{\mu}^*} \ni \boldsymbol{\mu}^*$), called *natural norm*. A lower bound for the surrogate $\tilde{\beta}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ is thus given by

$$\tilde{\beta}^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) = \inf_{V^{\mathcal{N}} \in X^{\mathcal{N}}} \frac{A(V^{\mathcal{N}}, T^{\mu^*} V^{\mathcal{N}}; \boldsymbol{\mu})}{\|T^{\mu^*} V^{\mathcal{N}}\|_X^2}, \tag{A.23}$$

thanks to (A.18). Following the same analogy introduced before, $\tilde{\beta}^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ can be seen as the solution of the following eigenproblem: find the smallest $\beta_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ such that

$$\left(\mathbb{H}\, \mathbb{A}^{-T}(\boldsymbol{\mu}^*) \mathbb{A}(\boldsymbol{\mu}) \mathbb{H}^{-1}\right) \underline{\mathbf{V}} = \beta_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})\, \underline{\mathbf{V}} \quad \text{for each } \underline{\mathbf{V}} \neq \underline{\mathbf{0}}. \tag{A.24}$$

The point is that, unlike the version (A.21), for fixed $\boldsymbol{\mu}^*$ the operator on the right-hand-side of (A.24) contains only $Q_A$ terms. Moreover, since it can be shown [298] that $\beta(\boldsymbol{\mu}^*)\tilde{\beta}^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) \leq \beta(\boldsymbol{\mu})$, it is sufficient to compute a lower bound $\beta^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) \leq \tilde{\beta}^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ for the surrogate (A.23) and the *true* value $\beta(\boldsymbol{\mu}^*)$ on the selected $\boldsymbol{\mu}^*$, and then translate it into a lower bound for $\beta(\boldsymbol{\mu})$.
The *natural norm* SCM procedure we adopt is based on the patching of some local (or *surrogate*) inf-sup stability factors properly computed for a set of $J$ parameter values $\mathcal{S} = \{\boldsymbol{\mu}^{1*}, \ldots, \boldsymbol{\mu}^{J*}\}$. In this way, a *global* lower bound, i.e. valid for each $\boldsymbol{\mu} \in \mathcal{D}$, can be seen as the combination between two main ingredients, discussed in the following subsections: *(i)* the construction of a local lower bound $\beta^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ upon a given parameter value $\boldsymbol{\mu}^* \in \mathcal{S}$, and *(ii)* the combination of the local lower bounds computed upon each $\boldsymbol{\mu}^* \in \mathcal{S}$.

### Construction of a local lower bound $\beta^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$

Let us analyze the construction of a local lower bound (A.23) for the surrogate inf-sup stability factor (A.22), considering a chosen $\boldsymbol{\mu}^* \in \mathcal{D}$ and a very rich training sample $\Xi_{train} \subset \mathcal{D}$; since this surrogate problem is coercive, the standard SCM [150] can be used. First of all, we rewrite (A.23) as the following minimization (*linear programming*) problem:

$$\tilde{\beta}^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) = \inf_{\mathbf{y} \in \mathcal{Y}_*} J^{obj}(\mathbf{y}; \boldsymbol{\mu}), \tag{A.25}$$

where $J^{obj}(\mathbf{y}; \mu)$ is the following linear objective functional:

$$J^{obj}(\mathbf{y}; \boldsymbol{\mu}) = \sum_{q=1}^{Q_A} \Theta^q(\mu) y_q, \qquad \text{with } \mathbf{y} = (y_1, \ldots, y_{Q_A}),$$

and $\mathcal{Y}_* \subset \mathbb{R}^{Q_A}$ the following constraint set (exploiting the affine decomposition of $A(\cdot, \cdot; \boldsymbol{\mu})$):

$$\mathcal{Y}_* = \left\{ \mathbf{y} \in \mathbb{R}^{Q_A} \ : \ \exists\, W^{\mathcal{N}}_{\mathbf{y}} \in X^{\mathcal{N}} \text{ s.t. } y_q = \frac{A^q(W^{\mathcal{N}}_{\mathbf{y}}, T^{\mu^*} W^{\mathcal{N}}_{\mathbf{y}})}{\|T^{\mu^*} W^{\mathcal{N}}_{\mathbf{y}}\|_X^2}, \quad 1 \leq q \leq Q_A \right\}.$$

The goal is to build a sequence of suitable relaxed problems of the original LP problem (A.25) by seeking the minimum of the objective on a descending sequence of larger sets, built by adding successively linear constraints. In order to define this sequence, let us consider the following steps:

1. *Bounding box construction.* In order to guarantee *a priori* that all relaxations which will be considered are well-posed, we construct once for all a (continuity) bounding box given by

$$B_{\boldsymbol{\mu}^*} = \prod_{q=1}^{Q_A} \left[ -\frac{\gamma_q}{\beta(\boldsymbol{\mu}^*)}, \frac{\gamma_q}{\beta(\boldsymbol{\mu}^*)} \right],$$

where $\beta(\boldsymbol{\mu}^*)$ is the solution of the eigenproblem (A.21) computed for $\boldsymbol{\mu} = \boldsymbol{\mu}^*$ (equivalently given by (A.22)) and $\gamma_q$ the ($\boldsymbol{\mu}^*$-independent) continuity factor of the bilinear form $A^q(\cdot,\cdot)$.

2. *Relaxed LP problem.* Given a properly selected constraints sample (or SCM sample) $\mathcal{C}_k^* = \{\boldsymbol{\mu}_1^*, \ldots, \boldsymbol{\mu}_k^*\}$ associated to $\boldsymbol{\mu}^*$, compute the (surrogate) lower bounds $\tilde{\beta}^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}')$ defined by (A.23), for each $\boldsymbol{\mu}' \in \mathcal{C}_k^*$; then, define the relaxation set

$$\mathcal{Y}_*^{\mathrm{LB}}(\mathcal{C}_k^*) = \left\{ \mathbf{y} \in \mathbb{R}^{Q_A} \ : \mathbf{y} \in B_{\boldsymbol{\mu}^*} \ \middle| \ \sum_{q=1}^{Q_A} \Theta^q(\boldsymbol{\mu}') y_q \geq \tilde{\beta}^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}'), \quad \forall \boldsymbol{\mu}' \in \mathcal{C}_k^* \right\}$$

by selecting a set of additional linear constraints associated to $\mathcal{C}_k^*$. Let us remark that the desired *local lower bound* is given by the solution $\beta^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$ of the following relaxed problem:

$$\beta^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) \equiv \beta^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}; \mathcal{C}_k^*) = \inf_{\mathbf{y} \in \mathcal{Y}_*^{\mathrm{LB}}(\mathcal{C}_k^*)} J^{obj}(\mathbf{y}; \boldsymbol{\mu}), \qquad \forall \boldsymbol{\mu} \in \mathcal{D}_{\boldsymbol{\mu}^*}; \tag{A.26}$$

the definition of the subset $\mathcal{D}_{\boldsymbol{\mu}^*} \subset \mathcal{D}$ will be subsequently made precise. In fact, we have that $\tilde{\beta}^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) \geq \beta^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$, where $\mathcal{Y}_* \subset \mathcal{Y}_*^{\mathrm{LB}}(\mathcal{C}_k^*)$ and the minimum is taken over a larger set. We underline that problem (A.26) has to be solved for each $\boldsymbol{\mu} \in \Xi_{train}$.

3. *Selection of the successive constraint.* The last step deals with the selection of the set $\mathcal{C}_k^*$, which is performed by means of a *greedy* procedure. In order to measure the quality of the lower bounds, we need to introduce an upper bound, defined as follows:

$$\beta^{\mathrm{UB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}) \equiv \beta^{\mathrm{UB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}; \mathcal{C}_k^*) = \inf_{\mathbf{y} \in \mathcal{Y}_*^{\mathrm{UB}}(\mathcal{C}_k^*)} J^{obj}(\mathbf{y}; \boldsymbol{\mu}), \qquad \forall \boldsymbol{\mu} \in \mathcal{D}_{\boldsymbol{\mu}^*}, \tag{A.27}$$

where $\mathcal{Y}_*^{\mathrm{UB}}(\mathcal{C}_k^*)$ is the set given by

$$\mathcal{Y}_*^{\mathrm{UB}}(\mathcal{C}_k^*) = \left\{ \tilde{\mathbf{y}} \in \mathbb{R}^Q \ : \ \tilde{\mathbf{y}} = \arg\min_{\mathbf{y} \in \mathcal{Y}_*} J^{obj}(\mathbf{y}; \boldsymbol{\mu}'), \quad \forall \boldsymbol{\mu}' \in \mathcal{C}_k^* \right\}.$$

Since $\mathcal{Y}_*^{\mathrm{UB}}(\mathcal{C}_k^*) \subset \mathcal{Y}_*$ – see [150] for the proof – (A.27) is in fact an upper bound for $\tilde{\beta}^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$, i.e. $\tilde{\beta}^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\mu) \leq \beta^{\mathrm{UB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$; observe that (A.27) is just an enumeration problem. Finally, we can show how to add the *successive constraint*, by means of a *(local) greedy procedure*. Starting from an arbitrarily chosen $\mathcal{C}_1^* = \{\boldsymbol{\mu}_1^*\}$, at step $k$ we enrich the set $\mathcal{C}_k^* = \{\boldsymbol{\mu}_1^*, \ldots, \boldsymbol{\mu}_k^*\}$, by means of the value $\boldsymbol{\mu}_{k+1}^*$ given by

$$\boldsymbol{\mu}_{k+1}^* = \arg\max_{\boldsymbol{\mu} \in \Xi_{train}} \rho(\boldsymbol{\mu}; \mathcal{C}_k^*), \qquad \rho(\boldsymbol{\mu}; \mathcal{C}_k^*) = \frac{\beta^{\mathrm{UB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}; \mathcal{C}_k^*) - \beta^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu}; \mathcal{C}_k^*)}{\beta^{\mathrm{UB}}_{\boldsymbol{\mu}^*}(\mu; \mathcal{C}_k^*)};$$

i.e. choosing the element corresponding to the largest ratio $\rho(\boldsymbol{\mu}; \mathcal{C}_k^*)$ over $\Xi_{train}$. The stopping criterium for this *successive* enrichment is given by

$$\rho(\boldsymbol{\mu}; \mathcal{C}_k^*) \leq \varepsilon_*, \tag{A.28}$$

i.e. the procedure for the local lower bound finishes when the largest ratio is under a chosen SCM (local) tolerance $\varepsilon_* \in (0,1)$. At the end of this procedure, we end up with $K$ constraints, corresponding to the set $\mathcal{C}_K^* = \{\boldsymbol{\mu}_1^*, \dots, \boldsymbol{\mu}_K^*\}$.

**Computation of a global lower bound**

We now translate the local lower bound $\beta^{\mathrm{LB}}_{\boldsymbol{\mu}^*}(\boldsymbol{\mu})$, computed upon a selected value $\boldsymbol{\mu}^*$, to a global lower bound. We shall make a distinction between the iterative procedure by which we "cover" the parameter space $\mathcal{D}$ and the relationship between the local and the global lower bounds.
Let us start from this second point; the output of the coverage procedure are the set $\mathcal{S} = \{\boldsymbol{\mu}^{1*}, \dots, \boldsymbol{\mu}^{J*}\}$, $J \leq J_{\max}$ and the associated SCM samples $\mathcal{C}^{j*}_{K(j)} = \{\boldsymbol{\mu}_1^{j*}, \dots, \boldsymbol{\mu}^{j*}_{K(j)}\}$, for any $j = 1, \dots, J$, where $K(j) < K_{\max}$ is the number of constraints points related to each $\boldsymbol{\mu}^{j*} \in \mathcal{S}$. The global lower bound for $\beta(\boldsymbol{\mu})$ can be defined (see [146] for the proof) as

$$\beta^{\mathrm{LB}}(\boldsymbol{\mu}) = \beta(\boldsymbol{\mu}^{\sigma*})\beta^{\mathrm{LB}}_{\boldsymbol{\mu}^{\sigma*}}(\boldsymbol{\mu}; \mathcal{C}^{\sigma*}_{K(\sigma)}), \quad \sigma \equiv \sigma(\boldsymbol{\mu}) = \arg \max_{j \in \{1,\dots,J\}} \beta(\boldsymbol{\mu}^{j*})\beta^{\mathrm{LB}}_{\boldsymbol{\mu}^{j*}}(\boldsymbol{\mu}; \mathcal{C}^{j*}_{K(j)}). \quad \text{(A.29)}$$

In practice, for each $\boldsymbol{\mu}$ the global lower bound is given by the maximum among the products between the stability factors $\beta(\boldsymbol{\mu}^{j*})$ and the local lower bounds $\beta^{\mathrm{LB}}_{\boldsymbol{\mu}^{j*}}(\boldsymbol{\mu}; \mathcal{C}^{j*}_{K(j)})$, corresponding to the selected $\{\boldsymbol{\mu}^{1*}, \dots, \boldsymbol{\mu}^{J*}\}$. Previous equation also implicitly defines the subdomains $\mathcal{D}_{\boldsymbol{\mu}^{j*}}$:

$$\mathcal{D}_{\boldsymbol{\mu}^{*j}} = \{\boldsymbol{\mu} \in \mathcal{D} \ : \ \beta(\boldsymbol{\mu}^{j*})\beta^{\mathrm{LB}}_{\boldsymbol{\mu}^{j*}}(\boldsymbol{\mu}; \mathcal{C}^{j*}_{K(j)}) \geq \beta(\boldsymbol{\mu}^{j'})\beta^{\mathrm{LB}}_{\boldsymbol{\mu}^{j'}}(\boldsymbol{\mu}; \mathcal{C}^{j'}_{K(j')}).,\ \ \forall j' \in \{1, \dots, J\}\}.$$

We remark that the global lower bound $\beta^{\mathrm{LB}}(\boldsymbol{\mu})$ given by this method interpolates $\beta(\boldsymbol{\mu})$ at each $\boldsymbol{\mu}^* \in \mathcal{S}$, being $\beta^{\mathrm{LB}}(\boldsymbol{\mu}^*) = \beta(\boldsymbol{\mu}^*)$ in these cases.

We now discuss the procedure by which we select the set $\mathcal{S} = \{\boldsymbol{\mu}^{1*}, \dots, \boldsymbol{\mu}^{J*}\}$ and the associated SCM samples; also in this case, we use a *(global) greedy procedure*, which encapsulates the local ones used for the construction of each SCM sample. Starting from a chosen $\mu^{1*}$, we set $\mathcal{S} = \{\boldsymbol{\mu}^{1*}\}$ and initialize the corresponding SCM sample $\mathcal{C}_1^{1*} = \{\boldsymbol{\mu}_1^{1*}\}$, being $\boldsymbol{\mu}_1^{1*} = \boldsymbol{\mu}^{1*}$. At step $j$, we have

$$\mathcal{S}^{(j-1)} = \{\boldsymbol{\mu}^{1*}, \dots, \boldsymbol{\mu}^{(j-1)*}\}, \qquad \mathcal{C}^{s*}_{K(s)} = \{\boldsymbol{\mu}_1^{s*}, \dots, \boldsymbol{\mu}^{s*}_{K(s)}\}, \qquad s = 1, \dots, j-1,$$

(through the construction of the local lower bounds around $\boldsymbol{\mu}^{1*}, \dots, \boldsymbol{\mu}^{(j-1)*}$) and

$$\boldsymbol{\mu}^{j*} = \arg \min_{\boldsymbol{\mu} \in \Xi_{train}} \beta^{\mathrm{LB}}_{\boldsymbol{\mu}^{(j-1)*}}(\boldsymbol{\mu}; \mathcal{C}^{(j-1)*}_{K(j-1)})$$

i.e. the new $\boldsymbol{\mu}^{j*}$ is selected by taking the minimum over $\Xi_{train}$ of the local lower bound computed w.r.t. the previous $\boldsymbol{\mu}^{(j-1)*}$. Then, we build the *covered set*

$$\mathcal{R}_j = \left\{ \boldsymbol{\mu} \in \Xi_{train} \ \middle| \ \beta^{\mathrm{LB}}_{\boldsymbol{\mu}^{j*}}(\boldsymbol{\mu}; \mathcal{C}_1^{j*}) > 0 \right\}$$

and start the procedure for the construction of the local lower bound (upon $\boldsymbol{\mu}^{j*}$): for $k = 1, \dots, K(j)$, we build iteratively the set $\mathcal{C}_k^{j*}$ and compute the *actual covered set*

$$\mathcal{R}^{act}_{j,k} = \left\{ \boldsymbol{\mu} \in \Xi_{train} \ \middle| \ \beta^{\mathrm{LB}}_{\boldsymbol{\mu}^{j*}}(\boldsymbol{\mu}; \mathcal{C}_k^{j*}) > 0 \right\},$$

checking at each step $k$ if the current $\boldsymbol{\mu}^{j*}$ does not give the possibility to increase the coverage, i.e if $\mathcal{R}^{act}_{j,k} \setminus \mathcal{R}_j = \emptyset$, and the stopping criterium (A.28) is fulfilled. If these conditions are not fulfilled ($k < K(j)$), we keep on adding linear constraints, and setting $\mathcal{R}_j = \mathcal{R}^{act}_{j,k}$; instead, if they are verified ($k = K(j)$), we stock the (local) covered set, by putting $\Xi_{train} := \Xi_{train} \setminus \mathcal{R}_j$, and

seek for the subsequent $\boldsymbol{\mu}^{(j+1)*}$. The global procedure ends up when all the train sample has been covered, i.e. when $\Xi_{train} = \emptyset$. For the reader's convenience, we sum up the local/global procedure in the following schematic algorithm:

$\mathcal{S}^{(1)} = \{\boldsymbol{\mu}^{1*}\}, \;\; \mathcal{C}_1^{1*} = \{\boldsymbol{\mu}_1^{1*}\}, \;\; \boldsymbol{\mu}_1^{1*} = \boldsymbol{\mu}^{1*}$
`for` $j = 1 : J_{\max}$
  $\mathcal{C}_1^{j*} = \{\boldsymbol{\mu}_1^{j*}\}, \;\; \boldsymbol{\mu}_1^{j*} = \boldsymbol{\mu}^{j*}$
  $\mathcal{R}_j = \left\{ \boldsymbol{\mu} \in \Xi_{train} \;\middle|\; \beta^{\mathrm{LB}}_{\boldsymbol{\mu}^{j*}}(\boldsymbol{\mu}; \mathcal{C}_1^{j*}) > 0 \right\}$
  `for` $k = 1 : K_{\max}$
    `compute the lower bound` (A.26) : $\beta^{\mathrm{LB}}_{\boldsymbol{\mu}^{j*}}(\boldsymbol{\mu}; \mathcal{C}_k^{j*})$
    `compute the upper bound` (A.27) : $\beta^{\mathrm{UB}}_{\boldsymbol{\mu}^{j*}}(\boldsymbol{\mu}; \mathcal{C}_k^{j*})$
    `add the successive constraint` : $\boldsymbol{\mu}^*_{k+1} = \arg\max_{\boldsymbol{\mu} \in \Xi_{train}} \rho(\boldsymbol{\mu}; \mathcal{C}_k^*)$
    `set` $\mathcal{C}_{k+1}^{j*} = \mathcal{C}_k^{j*} \cup \boldsymbol{\mu}^*_{k+1}; \quad \mathcal{R}^{act}_{j,k+1} = \left\{ \boldsymbol{\mu} \in \Xi_{train} \;\middle|\; \beta^{\mathrm{LB}}_{\boldsymbol{\mu}^{j*}}(\boldsymbol{\mu}; \mathcal{C}_{k+1}^{j*}) > 0 \right\};$
    `if` $\mathcal{R}^{act}_{j,k+1} \setminus \mathcal{R}_j = \emptyset$ `and` $\rho(\boldsymbol{\mu}; \mathcal{C}_k^{j*}) \leq \varepsilon_*$
      `set` $K(j) = k; \quad \mathcal{C}^{j*}_{K(j)} = \mathcal{C}^{j*}_{k+1}; \quad \Xi_{train} := \Xi_{train} \setminus \mathcal{R}^{act}_{j,k+1};$
      `exit for`
    `else`
      $\mathcal{R}_j = \mathcal{R}^{act}_{j,k+1}; \quad$ `set` $k = k + 1;$
    `end`
  `end`
  `if` $\Xi_{train} \neq \emptyset$
    $\boldsymbol{\mu}^{(j+1)*} = \arg\min_{\boldsymbol{\mu} \in \Xi_{train}} \beta^{\mathrm{LB}}_{\boldsymbol{\mu}^{j*}}(\boldsymbol{\mu}; \mathcal{C}^{j*}_{K(j)})$
    `set` $\mathcal{S}^{(j+1)} = \mathcal{S}^j \cup \boldsymbol{\mu}^{(j+1)*}; \quad j = j + 1;$
  `else`
    `set` $J = j; \mathcal{S} = \mathcal{S}^{(j)};$
    `return`
  `end`
`end`

### A.3.2 Navier-Stokes equations

In the Navier-Stokes case, error bounds require the evaluation of the (parametric) lower bound $\beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$ of the stability factor $\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}) = \beta_{\tilde{A},\mathcal{N}}(Y_N(\boldsymbol{\mu}); \boldsymbol{\mu})$, recalling that in this case

$$\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}) \equiv \beta_{\tilde{A},\mathcal{N}}(Y_N(\boldsymbol{\mu}); \boldsymbol{\mu}) = \inf_{V \in X^{\mathcal{N}}} \sup_{W \in X^{\mathcal{N}}} \frac{d\tilde{A}(Y_N(\boldsymbol{\mu})(\boldsymbol{\mu}); \boldsymbol{\mu})(V, W)}{\|V\|_X \|W\|_X},$$

where $d\tilde{A}(Y_N(\boldsymbol{\mu}); \boldsymbol{\mu})(V, W)$ is the Fréchet derivative of the global Navier-Stokes operator $\tilde{A}(\cdot, \cdot; \boldsymbol{\mu}) = A(V, W; \boldsymbol{\mu}) + C(V, V, W; \boldsymbol{\mu})$ with respect to the first variable, evaluated for $Y_N(\boldsymbol{\mu})$.

In order to compute $\beta^{\mathrm{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$ we have exploited the *natural norm* SCM algorithm presented above, by considering the Fréchet derivative of the Navier-Stokes operator instead of the Stokes operator. In this way, the SCM algorithm constructs a set of local lower bounds for the selected values $\boldsymbol{\mu}^{j*}$, $j = 1, \ldots$, where we assume that $d\tilde{A}(Y_N(\boldsymbol{\mu})(\boldsymbol{\mu}); \boldsymbol{\mu})(V, W)$ is replaced by $d\tilde{A}(Y^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu})(V, W)$, i.e. the global operator is linearized around the FE approximation $Y_N(\boldsymbol{\mu}^{j*})$. This assumption does not seem so inconsistent, since *(i)* when performing the SCM algorithm, we still do not have any RB approximation available – SCM is in fact performed before the construction of the RB spaces through the greedy algorithm – and *(ii)* at the end, the error $\|Y^{\mathcal{N}}(\boldsymbol{\mu}) - Y_N(\boldsymbol{\mu})\|_X$ is very small.

With respect to the Stokes case, this version of the SCM algorithm requires additional calculations, such as the update of the *bounding box* regarding the terms arising from the (linearization of) nonlinear operator and the solution of a FE Navier-Stokes problem for each selected $\boldsymbol{\mu}^{j*}$ and for each $\boldsymbol{\mu} \in \mathcal{C}^{j*}_{K(j)}$, giving the local lower bounds for each added constraints. This features a large additional cost, since the SCM algorithm may take several iterations to converge.

Thus, we have been pushed to seek for alternative, possibly cheaper procedures. We have considered the calculation of a surrogate lower bound through interpolation by using the Radial Basis Function (RBF) technique (see Sect. 2.7). At the simplest level, we only need to select a *trial* set $\Xi_{\text{trial}} \subset \mathcal{D}$ of size $|\Xi_{\text{trial}}| = n_{\text{trial}}$ and to compute the stability factors

$$\beta_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}) := \beta_{\tilde{A},\mathcal{N}}(Y^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu}), \qquad \boldsymbol{\mu} \in \Xi_{\text{trial}}, \tag{A.30}$$

by solving an eigenproblem for each $\boldsymbol{\mu} \in \Xi_{\text{trial}}$. Even though this computation is expensive, RBF interpolation is well-suited for scattered data, so that small trial samples can be considered. Concerning the choice of the radial basis, *thin-plate spline* have been employed. Results presented in Sect. 3.6.2 are very promising about the possibility to apply a surrogate model in order to compute a (Offline) cheaper but effective lower bound for $\beta^{\text{LB}}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu})$, e.g. for nonlinear problems. Positive lower bounds have been obtained by starting from positive *trial* stability factors (A.30). However, compactly supported RBFs (or explicit imposition of positivity constraints) can also be considered [50, 322] in case of loss of positiveness in large regions of the parameter space.
Moreover, an integrated SCM-RBF procedure is under investigation, in order to combine some local, *trial* lower bounds $\beta^{LB}_{\tilde{A},\mathcal{N}}(\boldsymbol{\mu}^{j*})$, $j = 1, \ldots, n_{\text{trial}}$ calculated by running $n_{\text{trial}}$ iterations of the SCM procedure, and a subsequent RBF interpolation built over these *trial* lower bounds instead than over a set of *a priori* computed stability factors (A.30). We might also exploit a *greedy* algorithm for selecting the *trial* set $\Xi_{\text{trial}} \subset \mathcal{D}$, by retaining at each step the *worst candidate* point, maximizing a suitable distance between the current RBF interpolation and, e.g., an upper bound computed during earlier stages of the SCM algorithm. In case of larger parameter dimensions, we might also consider some *screening* procedures (see e.g. Sect. 4.8) before RBF interpolation, in order to fix some parameter components and allow only the most relevant ones to vary.

A modified version of the *natural norm* SCM algorithm has been set to improve computational performances in case of large (recovered) affine expansions, as mentioned in Sect. 3.5. Our proposed approach – submitted in a recent paper [183] with Lassila and Rozza – is to limit the number of affine terms used in the SCM and to exploit the current SCM procedure at a *coarser* level, by recovering the lower bounds at the *finer* level, by introducing suitable correction factors. Some tests on elliptic linear scalar problems, both in the coercive (Poisson) and in the noncoercive (Helmholtz) case, underline that some of the proposed corrections in the stability factor are reliable and accurate and represents a good compromise in terms of effectivity. Important improvements are expected in the application of this methodology in nonlinear problems with nonaffine complex geometric parametrizations governed by Stokes or Navier-Stokes equations.

## A.4 Sobolev embedding constants

We present here a fixed-point algorithm [236] for the computation of the (discrete) Sobolev constant $\rho$ defined in (3.46), given by

$$\rho^2_{\mathcal{N}} = \sup_{v \in V^{\mathcal{N}}} \frac{\|v\|^2_{L^4(\Omega)}}{(v,v)_H}. \tag{A.31}$$

where $H^1_0(\Omega) \subset V \subset H^1(\Omega)$, $V^{\mathcal{N}} \subset V$ is the truth approximation space, and $\|w\|_p = \left(\int_\Omega |w|^p\right)^{1/p}$.

This algorithm has been firstly addressed in [78]; here we provide a complete proof of the theoretical results it is built over. To set this procedure, we need to reformulate problem (A.31) as a fixed point iteration for a suitable operator and to introduce a suitable eigenproblem. First of all, let us denote with $u_*$ the element of $X^{\mathcal{N}}$ (not necessarily unique) which satisfies

$$u_* = \arg\max_{v \in X^{\mathcal{N}}} \frac{\|v\|_4^2}{(v,v)_{X^{\mathcal{N}}}}, \qquad (u_*, u_*)_{X^{\mathcal{N}}} = 1; \tag{A.32}$$

Then, let us define the operator $\sigma : X^{\mathcal{N}} \to X^{\mathcal{N}}$ as

$$\sigma(w) = \frac{w^2}{\|w\|_4^2}; \tag{A.33}$$

note that $\|\sigma(w)\|_2 = 1$, for all $w \in X^{\mathcal{N}}$. Thus, given a nonnegative function $z \in L^2(\Omega)$, let us introduce the following eigenproblems: for each $1 \le i \le \mathcal{N}$, $u_i(z) \in X^{\mathcal{N}}$ and $\lambda_i(z) \in \mathbb{R}_+$ are solution of

$$\int_\Omega z u_i(z) v = \lambda_i(z)(u_i(z), v)_{X^{\mathcal{N}}}, \quad \forall v \in X^{\mathcal{N}}, \quad \text{with} \quad (u_i(z), u_i(z))_{X^{\mathcal{N}}} = 1. \tag{A.34}$$

Here we consider the eigenvalues $\{\lambda_i(z)\}_{i=1}^{\mathcal{N}}$ in increasing order, with $0 \le \lambda_1(z) \le \ldots \le \lambda_{\mathcal{N}}(z)$, and denote with $u_{max}(z) = u_{\mathcal{N}}(z)$. We remark that

$$\lambda_{max}(z) = \max_{v \in X^{\mathcal{N}}} \left( \frac{1}{(v,v)_{X^{\mathcal{N}}}} \int_\Omega z v^2 d\Omega \right) \tag{A.35}$$

by definition of Rayleigh quotient. The following Lemma, providing the basis of the fixed-point iteration for the approximation of $\rho^2_{\mathcal{N}}$, holds:

**Lemma A.1.** *The discrete Sobolev embedding constant $\rho_{\mathcal{N}}$ and the element $u_*$ defined by* (A.31) *and* (A.32)*, respectively, satisfy*

$$\lambda_{max}(\sigma(u_*)) = \rho^2_{\mathcal{N}}, \qquad u_{max}(\sigma(u_*)) = u_*.$$

*Proof.* Let us recall that

$$\max_{\substack{v \in X^{\mathcal{N}} \\ (v,v)_{X^{\mathcal{N}}} = 1}} (u_*^2, v^2) = \frac{\|u_*^2\|_2^2}{(u_*, u_*)^2_{X^{\mathcal{N}}}}$$

by Cauchy-Schwarz inequality, since $(u_*^2, v^2)_2 \le \|u_*^2\|_2 \|v^2\|_2$ and the maximum is reached for an element $v$ such that $v^2 = \alpha^2 u_*^2$ and $(v,v)_{X^{\mathcal{N}}} = 1$, i.e. $v = \alpha u_*$ and $\alpha = 1/(u_*, u_*)_{X^{\mathcal{N}}}$. Then, by using definitions (A.35) and (A.33), we find

$$\lambda_{max}(\sigma(u_*)) = \max_{v \in X^{\mathcal{N}}} \frac{\int_\Omega u_*^2 v^2 d\Omega}{\|u_*\|_4^2 (v,v)_{X^{\mathcal{N}}}} = \max_{\substack{v \in X^{\mathcal{N}} \\ (v,v)_{X^{\mathcal{N}}} = 1}} \frac{\int_\Omega u_*^2 v^2 d\Omega}{\|u_*\|_4^2} = \frac{\int_\Omega u_*^4 d\Omega}{\|u_*\|_4^2 (u_*, u_*)^2_{X^{\mathcal{N}}}} \tag{A.36}$$

$$= \frac{\|u_*\|_4^4}{\|u_*\|_4^2 (u_*, u_*)^2_{X^{\mathcal{N}}}} = \frac{\|u_*\|_4^2}{(u_*, u_*)_{X^{\mathcal{N}}}} = \frac{\|u_*\|_4^2}{(u_*, u_*)^2_{X^{\mathcal{N}}}} = \rho^2, \tag{A.37}$$

since $(u_*, u_*)_{X^{\mathcal{N}}} = 1$. Consequently, since $u_{max}(z)$ is the maximizer of (A.35), i.e. it is such that

$$u_{max}(z) = \arg\max_{v \in X^{\mathcal{N}}} \left( \frac{1}{(v,v)_{X^{\mathcal{N}}}} \int_\Omega z v^2 d\Omega \right),$$

from (A.36) and $(u_*, u_*)_{X^\mathcal{N}} = 1$ we can conclude that $u_{max}(\sigma(u_*)) = u_*/((u_*, u_*)_{X^\mathcal{N}}) = u_*$. □

In order to develop the fixed-point procedure, we point out that

$$\begin{aligned}
&\lambda_{max}(z_2) - \lambda_{max}(z_1) = \int_\Omega z_2 u_{max}^2(z_2) d\Omega - \int_\Omega z_1 u_{max}^2(z_1) d\Omega \\
&= \int_\Omega (z_2 - z_1) u_{max}^2(z_1) d\Omega + \int_\Omega z_2 u_{max}(z_2)(u_{max}(z_2) - u_{max}(z_1)) d\Omega \\
&+ \int_\Omega z_1 u_{max}(z_1)(u_{max}(z_2) - u_{max}(z_1)) d\Omega + \int_\Omega (z_2 - z_1) u_{max}(z_1)(u_{max}(z_2) - u_{max}(z_1)) d\Omega,
\end{aligned} \tag{A.38}$$

and we show the following Lemma:

**Lemma A.2.** *Given two non-negative functions $z_1, z_2 \in L^2(\Omega)$, the corresponding maximum eigenvalues $\lambda_{max}(z_1)$, $\lambda_{max}(z_2)$ defined by* (A.35) *are such that*

$$\lambda_{max}(z_2) - \lambda_{max}(z_1) = \int_\Omega (z_2 - z_1) u_{max}^2(z_1) d\Omega + \mathcal{O}(\|z_2 - z_1\|_2^2).$$

*Proof.* It is clear that $u_{max}(z_2) - u_{max}(z_1) = \mathcal{O}(\|z_2 - z_1\|_2)$ and $\lambda_{max}(z_2 - z_1) = \mathcal{O}(\|z_2 - z_1\|_2)$. Then, the last term of (A.38) is at least $\mathcal{O}(\|z_2 - z_1\|_{X^\mathcal{N}})$. Then, by definition of the eigenproblem (A.34), we have:

$$\int_\Omega z_2 u_{max}(z_2)(\partial u) d\Omega = \lambda_{max}(z_2)(u_{max}(z_2), \partial u)_{X^\mathcal{N}}, \tag{A.39}$$

$$\int_\Omega z_1 u_{max}(z_1)(\partial u) d\Omega = \lambda_{max}(z_1)(u_{max}(z_1), \partial u)_{X^\mathcal{N}} \tag{A.40}$$

where $\partial u = u_{max}(z_2) - u_{max}(z_1)$. To simplify these expressions, we now exploit the two relationships (valid for any $A, B$ and any scalar product)

$$\begin{aligned}
(A - B, A) &= \frac{1}{2}(A, A) - \frac{1}{2}(B, B) + \frac{1}{2}(A - B, A - B), \\
(B - A, A) &= \frac{1}{2}(A, A) - \frac{1}{2}(B, B) - \frac{1}{2}(A - B, A - B),
\end{aligned} \tag{A.41}$$

by applying the former to the right-hand side of (A.39) with $A = u_{max}(z_2)$, $B = u_{max}(z_1)$ and the latter to the right-hand side(A.39) with $A = u_{max}(z_1)$, $B = u_{max}(z_2)$, respectively. Since $\|u_{max}(z_i)\|_{X^\mathcal{N}} = 1$ for $i = 1, 2$, we end up with

$$\begin{aligned}
&\int_\Omega z_2 u_{max}(z_2)(u_{max}(z_2) - u_{max}(z_1)) d\Omega = \frac{1}{2}\lambda_{max}(z_2)\|u_{max}(z_2) - u_{max}(z_1)\|_{X^\mathcal{N}}^2, \\
&\int_\Omega z_1 u_{max}(z_2)((u_{max}(z_2) - u_{max}(z_1)) d\Omega = -\frac{1}{2}\lambda_{max}(z_1)\|u_{max}(z_1) - u_{max}(z_2)\|_{X^\mathcal{N}}^2,
\end{aligned}$$

which proves the thesis. □

Based on the relationship above, we can compute the Sobolev embedding constant by means of

the following fixed point algorithm:

$$\begin{cases} z^{(0)} = 1; \\ \text{for } k = 1, 2, \dots \\ \qquad \phi^{(k)} = \lambda_{max}(z^{(k-1)}); \\ \qquad z^{(k)} = \sigma(u_{max}(z^{(k-1)})); \end{cases}$$

In order to prove that $\phi^{(k)} \to \rho^2_{\mathcal{N}}$ as $k \to \infty$, we observe that, owing to Lemma A.2, the following relationship holds:

$$\begin{aligned} \phi^{(k+1)} - \phi^{(k)} &= \lambda_{max}(z^{(k)}) - \lambda_{max}(z^{(k-1)}) \\ &= \int_\Omega \Big(\sigma(u_{max}(z^{(k-1)})) - \sigma(u_{max}(z^{(k-2)}))\Big)\, u^2_{max}(z^{(k-1)}) d\Omega + \mathcal{O}(\|z^{(k-1)} - z^{(k-2)}\|_2^2). \end{aligned}$$

At this point, it just remains to show that the first term in the last expression is non-negative, so that the fixed-point iteration at least heads in the right direction – we remark that a fixed point of the algorithm is not necessary the supremizer of (A.35), but it is at least a *local* supremizer. Thanks to definition (A.33) of the operator $\sigma : X^{\mathcal{N}} \to X^{\mathcal{N}}$, the last factor appearing in the integral can be rewritten as $u^2_{max}(z^{(k-1)}) = \sigma(u_{max}(z^{(k-1)}))\|u_{max}(z^{(k-1)})\|^2_4$, so that the integral

$$\begin{aligned} \int_\Omega &(\sigma(u_{max}(\ z^{(k-1)})) - \sigma(u_{max}(z^{(k-2)}))\Big)\, u^2_{max}(z^{(k-1)}) d\Omega \\ &= \|u_{max}(z^{(k-1)})\|^2_4 (\sigma(u_{max}(z^{(k-1)})), \sigma(u_{max}(z^{(k-1)})) - \sigma(u_{max}(z^{(k-2)})))_2 \end{aligned}$$

is positive thanks to the relationship (A.41) – by choosing $A = \sigma(u_{max}(z^{(k-1)}))$ and $B = \sigma(u_{max}(z^{(k-2)}))$ – and to the fact that $\|\sigma(u_{max}(z^{(k-1)}))\|_2 = \|\sigma(u_{max}(z^{(k-2)}))\|_2 = 1$.

## A.5 Automatic generation of piecewise affine maps

We provide in this section some details related to the construction of basic RB triangulations built on (standard) triangles, elliptical triangles and general "curvy" triangles; a more extended overview can be found in [280, 208]. This methodology has been used for the examples of Sect. 3.6.1-3.6.2. As illustrated in Sect. 2.5.1, considering the domain decomposition (2.14), in the affine case for the $k^{th}$ subdomain ($1 \le k \le K_{\text{dom}}$) the concrete affine transformation is given, for any $\boldsymbol{\mu} \in \mathcal{D}$ and $\mathbf{x} \in \Omega^k$, by

$$T^k_i(\mathbf{x}, \boldsymbol{\mu}) = C^k_i(\boldsymbol{\mu}) + \sum_{j=1}^{d} G^k_{ij}(\boldsymbol{\mu}) x_j, \ 1 \le i \le d \tag{A.42}$$

for given translation vectors $\mathbf{C}^k : \mathcal{D} \to \mathbb{R}^d$ and linear transformation matrices $\mathbf{G}^k : \mathcal{D} \to \mathbb{R}^{d \times d}$. The associated Jacobians are given by $J^k_T(\boldsymbol{\mu}) = |\det(\mathbf{G}^k(\boldsymbol{\mu}))|, \quad 1 \le k \le K_{\text{dom}}$ and for invertible mappings they are strictly positive. Thus, in the two-dimensional case $d = 2$ affine mappings contains $d(d+1) = 6$ degrees of freedom (the mapping coefficients), and it is therefore sufficient, for any given $\boldsymbol{\mu} \in \mathcal{D}$, to consider the relationship between three non-collinear pre-image points in $\Omega$, $(\boldsymbol{z}^1, \boldsymbol{z}^2, \boldsymbol{z}^3)$ and three parametrized image nodes in $\Omega_o(\boldsymbol{\mu})$, $(\boldsymbol{z}^1_o(\boldsymbol{\mu}), \boldsymbol{z}^2_o(\boldsymbol{\mu}), \boldsymbol{z}^3_o(\boldsymbol{\mu}))$. Note that every point consists of two components $(z^i_1, z^i_2)$, $1 \le i \le 3$, resp. $(z^i_{o\,1}, z^i_{o\,2})$, $1 \le i \le 3$, and therefore the application of (A.42) to these points gives a system of six independent equations to

determine[2] the six mapping coefficients:

$$z_{o\,i}^{m}(\boldsymbol{\mu}) = C_i(\boldsymbol{\mu}) + \sum_{j=1}^{2} G_{ij}(\boldsymbol{\mu}) z_j^m, \quad 1 \leq i \leq 2, \quad 1 \leq m \leq 3. \tag{A.43}$$

The RB triangulation may be built on (standard) triangles, elliptical triangles and general "curvy" triangles, which admit symbolic and numerical automation and are the building blocks in the `rbMIT` software [149]. Extension to 3D tethraedra is quite straightforward [109]. More complex geometrical parametrizations are represented by FFD and RBF techniques (see Sect. 2.6-2.7).

### Standard Triangles

In the case of a standard triangle subdomain the three vertices of the triangle in the reference domain shall serve as pre-image nodes while the three vertices of the triangle in the actual ($\boldsymbol{\mu}$-dependent) domain shall serve as image nodes. In this case, our three points uniquely define not only the transformation but also the reference domain and parametrized domains [280]. We can then readily establish the system of six linear equations to determine the six unknown mapping coefficients. In this way, we can construct an affine transformation from any reference triangle in $\mathbb{R}^2$ onto any desired triangle in $\mathbb{R}^2$. It is not mandatory to choose the vertices of the triangles as our nodes defining the transformation, and other characteristic points e.g. the barycentric coordinates of the FE context are also possible.

### Elliptical and Curvy Triangles

The class of elliptical triangles covers a much greater range of possible geometries and their formulation is also necessary for the more general case dealing with curvy triangles. We can distinguish two different kinds of elliptic triangles: "inwards" and "outwards" triangles. Both types are depicted in Figure A.1. In both cases, the elliptical triangle $\Omega_o(\boldsymbol{\mu})$ is defined by the three vertices $\boldsymbol{z}_o^1(\boldsymbol{\mu})$, $\boldsymbol{z}_o^2(\boldsymbol{\mu})$, $\boldsymbol{z}_o^3(\boldsymbol{\mu})$, the two straight lines $\overline{\boldsymbol{z}_o^1(\boldsymbol{\mu})\boldsymbol{z}_o^2(\boldsymbol{\mu})}$ and $\overline{\boldsymbol{z}_o^1(\boldsymbol{\mu})\boldsymbol{z}_o^3(\boldsymbol{\mu})}$ as well as the elliptical arc $\overline{\boldsymbol{z}_o^2(\boldsymbol{\mu})\boldsymbol{z}_o^3(\boldsymbol{\mu})}^{\text{arc}}$.

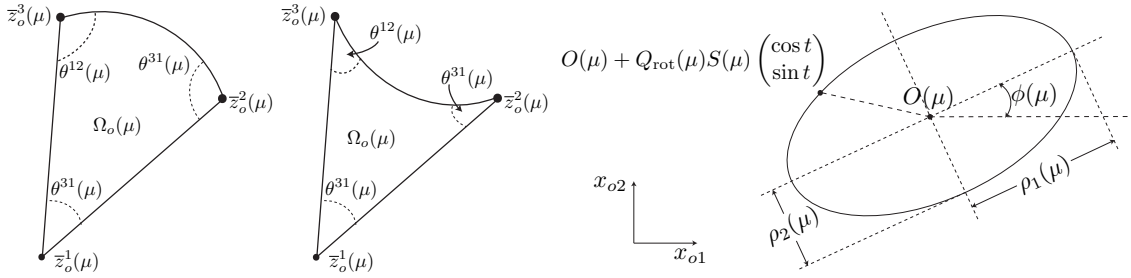



Figure A.1: From left to right: "inwards" elliptical triangle, "outwards" elliptical triangle and definition of a point on a prescribed parametrized ellipse.

We shall now precise the definition and description of the elliptical arc and explain the constraints that must be met by the location of the third point $\boldsymbol{z}_o^1(\boldsymbol{\mu})$ to ensure "proper" triangles and a continuous and well-defined global mapping in the multidomain context.
First, the description of the elliptical arc shall be derived from the definition of a parametrized ellipse (see Figure A.1 (c)).

[2]The assumption that the affine transformation is bijective thereby ensures that image nodes are perforce also non-collinear (if pre-image nodes are non-collinear) and hence equations are linear independent.

The ellipse is described implicitly by

$$(\mathbf{x}_o - O(\boldsymbol{\mu}))^T Q_{\mathrm{rot}}(\boldsymbol{\mu}) S^{-2}(\boldsymbol{\mu}) Q_{\mathrm{rot}}(\boldsymbol{\mu})^T (\mathbf{x}_o - O(\boldsymbol{\mu})) = 1, \tag{A.44}$$

while a particular point on this ellipse is given by

$$\mathbf{x}_o \equiv \begin{pmatrix} x_{o1} \\ x_{o2} \end{pmatrix} = O(\boldsymbol{\mu}) + Q_{\mathrm{rot}}(\boldsymbol{\mu}) S(\boldsymbol{\mu}) \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \tag{A.45}$$

for given $t \in \mathbb{R}$. We denote (see Fig. A.1) by $O(\boldsymbol{\mu}) : \mathcal{D} \to \mathbb{R}^2$ the center of the ellipse, by $\rho_1(\boldsymbol{\mu}) : \mathcal{D} \to \mathbb{R}_+$ and $\rho_2 : D \to \mathbb{R}_+$ the length of its semi-axes and by $\phi(\boldsymbol{\mu}) : \mathcal{D} \to \mathbb{R}$ the angle of inclination. With these quantities, the scaling matrix $S(\boldsymbol{\mu})$ and the rotation matrix $Q_{\mathrm{rot}}(\boldsymbol{\mu})$ can be defined:

$$S(\boldsymbol{\mu}) \equiv \begin{pmatrix} \rho_1(\boldsymbol{\mu}) & 0 \\ 0 & \rho_2(\boldsymbol{\mu}) \end{pmatrix}, \qquad Q_{\mathrm{rot}}(\boldsymbol{\mu}) = \begin{pmatrix} \cos\phi(\boldsymbol{\mu}) & -\sin\phi(\boldsymbol{\mu}) \\ \sin\phi(\boldsymbol{\mu}) & \cos\phi(\boldsymbol{\mu}) \end{pmatrix}.$$

The description of the elliptical arc with these means is then as follows:

$$\overline{\boldsymbol{z}_o^2(\boldsymbol{\mu})\boldsymbol{z}_o^3(\boldsymbol{\mu})}^{\mathrm{arc}} = \left\{ O(\boldsymbol{\mu}) + Q_{\mathrm{rot}}(\boldsymbol{\mu}) S(\boldsymbol{\mu}) \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \,\middle|\, t_2 \le t \le t_3 \right\}. \tag{A.46}$$

with $t_2 \in \mathbb{R}$ and $t_3 \in \mathbb{R}$ chosen such that the points $\boldsymbol{z}_o^2(\boldsymbol{\mu})$ and $\boldsymbol{z}_o^3(\boldsymbol{\mu})$ are given as the endpoints of the elliptical arc for $t = t_2$ and $t = t_3$:

$$\boldsymbol{z}_o^m(\boldsymbol{\mu}) = O(\boldsymbol{\mu}) + Q_{\mathrm{rot}}(\boldsymbol{\mu}) S(\boldsymbol{\mu}) \begin{pmatrix} \cos t_m \\ \sin t_m \end{pmatrix}, \quad m = 2, 3. \tag{A.47}$$

In addition, we have to make sure that $0 \le t_3 - t_2 < \pi$. It remains to specify the location of the third point $\boldsymbol{z}_o^1(\boldsymbol{\mu})$. For elliptical triangles, this location is not arbitrary but has to be chosen in a way that ensures that the affine transformation generates the desired elliptical arc (A.46). First, this ensures a continuous global mapping; second, to obtain well-defined elliptical triangles (and thus a well defined domain), several internal angle conditions have to be met by the choice for $\boldsymbol{z}_o^1(\boldsymbol{\mu})$: $0 < \Theta^* < \pi$, $\forall \Theta^* \in \{\Theta^{12}, \Theta^{23}, \Theta^{31}\}$. The first requirement can be fulfilled by the expression of the three corner points as

$$\boldsymbol{z}_o^m(\boldsymbol{\mu}) = O(\boldsymbol{\mu}) + \omega_m Q_{\mathrm{rot}}(\boldsymbol{\mu}) S(\boldsymbol{\mu}) \begin{pmatrix} \cos t_m \\ \sin t_m \end{pmatrix}, \quad 1 \le m \le 3, \tag{A.48}$$

for given $\omega_1 = \omega \in \mathbb{R}$, $\omega_2 = \omega_3 = 1$ and $t_1 \in [t_2, t_3]$. Pre-image points are thus given as

$$\boldsymbol{z}_o^m(\boldsymbol{\mu}_{\mathrm{ref}}) = O(\boldsymbol{\mu}_{\mathrm{ref}}) + \omega_m Q_{\mathrm{rot}}(\boldsymbol{\mu}_{\mathrm{ref}}) S(\boldsymbol{\mu}_{\mathrm{ref}}) \begin{pmatrix} \cos t_m \\ \sin t_m \end{pmatrix}, \quad 1 \le m \le 3. \tag{A.49}$$

From these representations we can identify our affine mapping as

$$\begin{aligned} \boldsymbol{z}_o^m(\boldsymbol{\mu}) = \mathbf{C}(\boldsymbol{\mu}) + \mathbf{G}(\boldsymbol{\mu}) \boldsymbol{z}^m = (O(\boldsymbol{\mu}) - Q_{\mathrm{rot}}(\boldsymbol{\mu}) S(\boldsymbol{\mu}) S(\boldsymbol{\mu}_{\mathrm{ref}})^{-1} Q_{\mathrm{rot}}(\boldsymbol{\mu}_{\mathrm{ref}})^T O(\boldsymbol{\mu}_{\mathrm{ref}})) \\ + (Q_{\mathrm{rot}}(\boldsymbol{\mu}) S(\boldsymbol{\mu}) S(\boldsymbol{\mu}_{\mathrm{ref}})^{-1} Q_{\mathrm{rot}}(\boldsymbol{\mu}_{\mathrm{ref}})^T) \boldsymbol{z}^m. \end{aligned}$$

The second requirement - the internal angle conditions - is illustrated in Figure A.2. In the inwards case, a necessary and sufficient condition to ensure the conditions $0 < \Theta^* < \pi$, for any $\Theta^* \in \{\Theta^{12}, \Theta^{23}, \Theta^{31}\}$ is given for an inwards elliptical triangle by $\boldsymbol{z}_o^1(\boldsymbol{\mu}) \in R_{\mathrm{in}}(\boldsymbol{\mu})$, where

$$\begin{aligned} R_{in}(\boldsymbol{\mu}) = \{ \boldsymbol{z}_o^1(\boldsymbol{\mu}) \in \mathbb{R}^2 | (\boldsymbol{z}_o^1(\boldsymbol{\mu}) - \boldsymbol{z}_o^2(\boldsymbol{\mu}))^T \boldsymbol{n}^2(\boldsymbol{\mu}) < 0, \\ (\boldsymbol{z}_o^1(\boldsymbol{\mu}) - \boldsymbol{z}_o^3(\boldsymbol{\mu})))^T \boldsymbol{n}^3(\boldsymbol{\mu}) < 0, (\boldsymbol{z}_o^1(\boldsymbol{\mu}) - \boldsymbol{z}_o^{2,3}(\boldsymbol{\mu}))^T \boldsymbol{n}^{2,3}(\boldsymbol{\mu}) < 0 \}, \end{aligned} \tag{A.50}$$

and for the outwards elliptical triangle by $\boldsymbol{z}_o^1(\boldsymbol{\mu}) \in R_{\text{out}}(\boldsymbol{\mu})$, where

$$R_{out}(\boldsymbol{\mu}) = \{\boldsymbol{z}_o^1(\boldsymbol{\mu}) \in \mathbb{R}^2 | (\boldsymbol{z}_o^1(\boldsymbol{\mu}) - \boldsymbol{z}_o^2(\boldsymbol{\mu}))^T \boldsymbol{n}^2(\boldsymbol{\mu}) > 0, (\boldsymbol{z}_o^1(\boldsymbol{\mu}) - \boldsymbol{z}_o^3(\boldsymbol{\mu}))^T \boldsymbol{n}^3(\boldsymbol{\mu}) > 0\}. \quad \text{(A.51)}$$

Here $\boldsymbol{n}^2(\boldsymbol{\mu})$ and $\boldsymbol{n}^3(\boldsymbol{\mu})$ are the outwards-facing normals to the ellipse at $\boldsymbol{z}^2(\boldsymbol{\mu})$ and $\boldsymbol{z}^3(\boldsymbol{\mu})$ respectively, $\boldsymbol{z}_o^{2,3}(\boldsymbol{\mu}) = \frac{1}{2}(\boldsymbol{z}_o^2(\boldsymbol{\mu}) + \boldsymbol{z}_o^3(\boldsymbol{\mu}))$ and $\boldsymbol{n}^{2,3}(\boldsymbol{\mu})$ is the "outwards-facing" normal to the line segment $\overline{\boldsymbol{z}_o^2(\boldsymbol{\mu})\boldsymbol{z}_o^3(\boldsymbol{\mu})}$ at $\boldsymbol{z}_o^{2,3}(\boldsymbol{\mu})$.

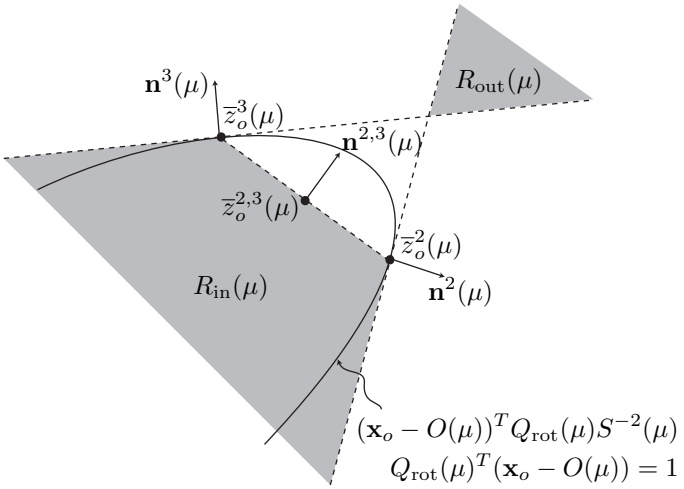



Figure A.2: Regions in which $z_o^1(\boldsymbol{\mu})$ must reside in the inwards ($R_{\text{in}}(\boldsymbol{\mu})$) and the outwards case ($R_{\text{out}}(\boldsymbol{\mu})$).

An important feature of the elliptical triangles is that they are consistent under refinement: so, if we split an elliptical triangle for which the internal angle conditions (A.50) and (A.51) are fulfilled, the resulting two elliptical triangles also satisfy the internal angle conditions.
To enlarge the possible range of geometries even more, the elliptical triangles are extended to "curvy" triangles. This is done by replacing $(\cos t, \sin t)^T$ in (A.45) with a general parametrization $(g_1(t), g_2(t))^T$, where the curvy arcs have to be either strictly convex ("inwards") or strictly concave ("outwards") for all $\boldsymbol{\mu} \in \mathcal{D}$. In this case, it is not possible to derive a simple closed form as it was the case for the elliptical triangles.

### Piecewise affine mappings for multiple subdomains

To treat more complex geometries, it is necessary to allow that the domain is built of several (standard, elliptical or curvy) triangles. We are then not restricted to a single affine mapping, but we deal with a piecewise affine mapping based on this domain decomposition. We can thus consider geometrical domains for which the boundary can be represented either by straight edges or by elliptical triangles as presented above (see e.g. Sect. 3.6.1, case 3). The multi-domain mapping process is then performed in three steps. First, the RB triangulation is generated on the reference domain $\Omega$ together with the associated reference regions by means of an automatic procedure[3]. In the second step, the necessary parameter-dependent affine mappings for each subdomain are constructed, as described in the previous section. In the last step we have to translate the parametric mappings obtained for each subdomain into PDE coefficients.

[3]The RB triangulation has to be compatible with the mapping continuity condition (2.51) and all elliptical and curvy subtriangles have to be well-defined and to fulfill the internal angle conditions (A.50) and (A.51).

# Bibliography


[1] F. Abergel and E. Casas. Some control problems of multistage equations appearing in fluid mechanics. *R.A.I.R.O. Math. Modelling Numer. Anal.*, 27(2):223–247, 1993.

[2] F. Abergel and R. Temam. On some control problems in fluid mechanics. *Theoret. Comput. Fluid Dynamics*, 1:303–325, 1990.

[3] F. Abraham, M. Behr, and M. Heinkenschloss. Shape optimization in stationary blood flow: a numerical study of non-newtonian effects. *Comput. Meth. Biomech. Biomed. Engrg.*, 8:127–137, 2005.

[4] F. Abraham, M. Behr, and M. Heinkenschloss. Shape optimization in unsteady blood flow: a numerical study of non-newtonian effects. *Comput. Meth. Biomech. Biomed. Engrg.*, 8:201–212, 2005.

[5] V. Agoshkov, A. Quarteroni, and G. Rozza. A mathematical approach in the design of arterial bypass using unsteady Stokes equations. *J. Sci. Comput.*, 28(2–3):139–165, 2006.

[6] V. Agoshkov, A. Quarteroni, and G. Rozza. Shape design in aorto-coronaric bypass anastomoses using perturbation theory. *SIAM J. Numer. Anal.*, 44(1):367–384, 2006.

[7] G. Allaire. *Conception optimale de structures*. Springer Verlag, 2007. Vol. 58 Mathématiques et Applications.

[8] B. O. Almroth, P. Stern, and F. A. Brogan. Automatic choice of global shape functions in structural analysis. *AIAA Journal*, 16:525–528, 1978.

[9] D. Ambrosi, A. Quarteroni, and G. Rozza (Eds.). *Modelling of Physiological Flows*, volume 5 of *Modeling, Simulation and Applications (MS&A)*. Springer-Verlag Italia, Milano, 2011.

[10] E.I. Amoiralis and I.K. Nikolos. Freeform deformation versus B-spline representation in inverse airfoil design. *J. Comput. Inf. Sci. Engrg.*, 8(2), 2008.

[11] M. Andreoli, A. Janka, and J.A. Désidéri. Free-form-deformation parametrization for multilevel 3D shape optimization in aerodynamics. Technical Report 5019, INRIA Sophia Antipolis, 2003.

[12] H. Antil, M. Heinkenschloss, and R.H.W. Hoppe. Domain decomposition and balanced truncation model reduction for shape optimization of the Stokes system. *Optim. Methods Softw.*, 26(4–5):643–669, 2011.

[13] H. Antil, M. Heinkenschloss, R.H.W. Hoppe, C. Linsenmann, and A. Wixforth. Reduced order modeling based shape optimization of surface acoustic wave driven microfluidic biochips. *Math. Comput. Simul.*, 2010. In press.

[14] A.C. Antoulas. *Approximation of Large-Scale Dynamical Systems*. Society for Industrial and Applied Mathematics, Philadephia, PA, 2005.

[15] N. Aubry. On the hidden beauty of the proper orthogonal decomposition. *Theor. Comp. Fluid. Dyn.*, 2:339–352, 1991.

[16] I. Babuška. The finite element method with lagrangian multipliers. *Numer. Math.*, 20:179–192, 1973.

[17] I. Babuška. Error-bounds for finite element method. *Numerische Mathematik*, 16:322–333, 1971.

[18] F. Ballarin, A. Manzoni, G. Rozza, and S. Salsa. Free form deformation as a perturbation of identity: analysis and applications to shape optimization problems. *In preparation*, 2012.

[19] E. Balmes. Parametric families of reduced finite element models: Theory and applications. *Mechanical Systems and Signal Processing*, 10(4):381–394, 1996.

[20] R. Balossino, G. Pennati, F. Migliavacca, L. Formaggia, A. Veneziani, M. Tuveri, and G. Dubini. Computational models to predict stenosis growth in carotid arteries: Which is the role of boundary conditions? *Comput. Methods Biomech. Biomed. Engrg.*, 12(1):113–123, 2009.

[21] M. Barrault, Y. Maday, N.C. Nguyen, and A.T. Patera. An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. *C. R. Math. Acad. Sci. Paris*, 339(9):667–672, 2004.

[22] A. Barrett and G. Reddien. On the reduced basis method. *Z. Angew. Math. Mech.*, 75(7):543–549, 1995.

[23] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. *Nonlinear programming - theory and algorithms (2nd Edition)*. Wiley-Interscience, 1993.

[24] R. Becker and R. Rannacker. An optimal control approach to a posteriori error estimation in finite element methods. *In: Acta Numerica 2001, A. Iserles, ed., Cambridge University Press, Cambridge, UK*, pages 1–102, 2001.

[25] A. Beckert and H. Wendland. Multivariate interpolation for fluid-structure-interaction problems using radial basis functions. *Aerospace Sci. Technol.*, 2:125–134, 2001.

[26] F. Bélahcène and J.A. Désidéri. Paramétrisation de Bézier adaptative pour l'optimisation de forme en Aérodynamique. Technical Report 4943, INRIA Sophia Antipolis, 2003.

[27] J.A. Bello, E. Fernandez-Cara, J. Lemoine, and J. Simon. The differentiability of the drag with respect to the variations of a Lipschitz domain in a Navier-Stokes flow. *SIAM J. Control. Optim.*, 35(2):626–640, 1997.

[28] P. Benner and Quintana-Ortí. Model reduction based on spectral projection methods. In P. Benner, V. Mehrmann, and D.C. Sorensen, editors, *Dimension Reduction of Large-Scale Systems*, volume 45 of *Lecture Notes in Computational Science and Engineering*, pages 5–48. Springer-Verlag Berlin Heidelberg, 2005.

[29] M. Benzi, G.H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.

[30] M. Berggren. Numerical solution of a flow-control problem: vorticity reduction by dynamic boundary action. *SIAM J. Sci. Comput.*, 19(3):829–860, 1998.

[31] M. Berggren, R. Glowinski, and J.L. Lions. A computational approach to controllability issues for flow-related models. (ii): control of two-dimensional, linear advection-diffusion and Stokes models. *Int. J. Comput. Fluid Dynamics*, 6(4):253–274, 1996.

[32] M. Berggren, R. Glowinski, and J.L. Lions. A computational approach to controllability issues for flow-related models. (ii): pointwise control of the viscous burger equation. *Int. J. Comput. Fluid Dynamics*, 6(4):237–252, 1996.

[33] G. Berkooz, P. Holmes, and J.L. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annu. Rev. Fluid Mech.*, 25(1):539–575, 1993.

[34] C. Bertoglio, P. Moireau, and J.-F. Gerbeau. Sequential parameter estimation for fluid-structure problems. Application to hemodynamics. *Int. J. Num. Meth. Biomed. Engrg.*, 2011. In press (DOI: 10.1002/cnm.1476).

[35] B.K. Bharadvaj, R.F. Mabon, and D.P. Giddens. Steady flow in a model of the human carotid bifurcation. part I - flow visualization. *Journal of Biomechanics*, 15:349–362, 1982.

[36] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Convergence rates for greedy algorithms in reduced basis methods. *SIAM J. Math. Anal.*, 43(3):1457–1472, 2011.

[37] S. Boisgérault and J.P. Zolésio. Boundary variations in the Navier-Stokes equations and Lagrangian functionals. In J. Cagnol, editor, *Shape optimization and optimal design. Proceedings of the IFIP conference*, volume 216 of *Lect. Notes Pure Appl. Math.*, pages 7–26. Marcel Dekker, 2001.

[38] F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11:567–585, 1989.

[39] A. Borzì and V. Schulz. *Computational optimization of systems governed by partial differential equations*. Society for Industrial and Applied Mathematics, Philadephia, PA, 2011.

[40] M. Botsch and I. Kobbelt. Freeform shape representations for efficient geometry processing. In *Proceedings of the International Conference on Shape Modeling and Applications*. IEEE Computer Society, 2003.

[41] M. Boulakia, M.A. Fernández, J.-F. Gerbeau, and N. Zemzemi. Direct and inverse problems in electrocardiography. In *AIP Conference Proceedings*, volume 1048, page 113, 2008.

[42] J. Bourot. On the numerical computation of the optimum profile in Stokes flow. *J. Fluid Mech.*, 65(3):513–515, 1974.

[43] S. Boyaval, C. Le Bris, T. Lelièvre, Y. Maday, N.C. Nguyen, and A.T. Patera. Reduced basis techniques for stochastic problems. *Arch. Comput. Methods Engrg.*, 17(4):1–20, 2010.

[44] J. Boyd. Six strategies for defeating the runge phenomenon in Gaussian radial basis functions on a finite interval. *Comput. Math. Appl.*, 60(12):3108–3122, 2010.

[45] N. W. Bressloff. Parametric geometry exploration of the human carotid artery bifurcation. *J. Biomech.*, 40:2483–2491, 2007.

[46] N.W. Bressloff, A. Forrester, J. Banks, and V.B. Kolachalama. Shape optimization of the carotid artery bifurcation. *Proceedings of ASMO-UK*, 2004.

[47] F. Brezzi. On the existence, uniqueness, and approximation of saddle point problems arising from Lagrangian multipliers. *R.A.I.R.O., Anal. Numér.*, 2:129–151, 1974.

[48] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer Verlag, 1991.

[49] F. Brezzi, J. Rappaz, and P.A. Raviart. Finite dimensional approximation of nonlinear problems. Part I: Branches of nonsingular solutions. *Numer. Math.*, 36:1–25, 1980.

[50] M.D. Buhmann. *Radial Basis Functions: Theory and Implementations*, volume 12 of *Cambridge monographs on applied and computational mathematics*. Cambridge University Press, UK, 2003.

[51] T. Bui-Thanh, M. Damodaran, and K. Willcox. Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition. *AIAA Journal*, 42(8):1505–1516, 2004.

[52] T. Bui-Thanh, K. Willcox, and O. Ghattas. Parametric reduced-order models for probabilistic analysis of unsteady aerodynamics applications. *AIAA Journal*, 46(10), 2008.

[53] J. Burkardt, M. Gunzburger, and H.C. Lee. Centroidal voronoi tessellation-based reduced-order modeling of complex systems. *SIAM J. Sci. Comput.*, 28(2):459–484, 2006.

[54] J. Burkardt, M. Gunzburger, and H.C. Lee. POD and CVT-based reduced-order modeling of Navier-Stokes flows. *Comput. Meth. Appl. Mech. Engrg.*, 196(1-3):337–355, 2006.

[55] E. Burman and M.A. Fernández. Continuous interior penalty finite element method for the time-dependent Navier–Stokes equations: space discretization and convergence. *Numer. Math.*, 107(1):39–77, 2007.

[56] G. Caloz and J. Rappaz. Numerical analysis for nonlinear and bifurcation problems. In P.G. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis, Vol. V*, Techniques of Scientific Computing (Part 2), pages 487–637. Elsevier Science B.V., 1997.

[57] F. Campolongo, J. Cariboni, and A. Saltelli. An effective screening design for sensitivity analysis of large models. *Environ. Model. & Software*, 22(10):1509–1518, 2007.

[58] C.G. Caro, T.J. Pedley, R.C. Schroter, W.A. Seed, and K.H. Parker. *The Mechanics of the Circulation*. Cambridge University Press, 2nd edition, 2011.

[59] J. Céa. Conception optimale ou identification de formes: calcul rapide de la dérivée directionelle de la fonction cout. *Mathematical Modelling and Numerical Analysis*, 20(3):371–402, 1986.

[60] J. Céa, S. Garreau, P. Guillaume, and M. Masmoudi. The shape and topological optimizations connection. *Comput. Methods Appl. Mech. Engrg.*, 188:713–726, 2000.

[61] Y. Chen, J. Hesthaven, Y. Maday, and J. Rodriguez. A monotonic evaluation of lower bounds for inf-sup stability constants in the frame of reduced basis approximations. *C. R. Acad. Sci. Paris, Ser. I*, 346:1295–1300, 2008.

[62] F. Chinesta, P. Ladeveze, and E. Cueto. A short review on model order reduction based on proper generalized decomposition. *Arch. Comput. Methods Engrg.*, 18:395–404, 2011.

[63] E.A. Christensen, M. Brøns, and J.N. Sørensen. Evaluation of proper orthogonal decomposition–based decomposition techniques applied to parameter-dependent nonturbulent flows. *SIAM J. Sci. Comput.*, 21:1419–1434, 1999.

[64] J. Cole, J. Watterson, and M. O'Reilly. Is there a haemodynamic advantage associated with cuffed arterial anastomoses? *J. Biomech*, 35:1337–1446, 2002.

[65] J. Cole, J. Watterson, and M. O'Reilly. Numerical investigation of the haemodynamics at a patched arterial bypass anastomosis. *Medical Engineering & Physics*, 24:393–401, 2002.

[66] P.G. Constantine and Q. Wang. Input subspace detection for dimension reduction in high dimensional approximation. *submitted*, 2012.

[67] P. Crosetto. *Fluid-structure interaction problems in hemodynamics: parallel solvers, preconditioners, and applications.* PhD thesis, N. 5109, École Polytechnique Fédérale de Lausanne, 2011.

[68] P. Crosetto, S. Deparis, G. Fourestey, and A. Quarteroni. Parallel algorithms for fluid-structure interaction problems in haemodynamics. *SIAM J. Sci. Comput.*, 33(4):1598–1622, 2011.

[69] A. De Boer, M.S. Van der Schoot, and H. Bijl. Mesh deformation based on radial basis function interpolation. *Comput. Struct.*, 85:784–795, 2007.

[70] A. De Boer, A. Van Zuijlen, and H. Bijl. Comparison of conservative and consistent approaches for the coupling of non-matching meshes. *Comput. Meth. Appl. Mech. Engrg.*, 197(49–50):4284–4297, 2008.

[71] L. Dedè. Optimal flow control for Navier-Stokes equations: Drag minimization. *Int. J. Numer. Meth. Fluids*, 55(4):347 – 366, 2007.

[72] L. Dedè. *Adaptive and reduced basis methods for optimal control problems in environmental applications.* PhD thesis, Politecnico di Milano, 2008.

[73] L. Dedè. Reduced basis method and a posteriori error estimation for parametrized linear-quadratic optimal control problems. *SIAM J. Sci. Comput.*, 32(2):997–1019, 2010.

[74] L. Dedè. Reduced basis method and error estimation for parametrized optimal control problems with control constraints. *J. Sci. Comput.*, 50(2):287–305, 2012.

[75] M.C. Delfour and J.P. Zolésio. *Shapes and geometries - analysis, differential calculus, and optimization.* Society for Industrial and Applied Mathematics, Philadephia, PA, 2nd edition, 2011.

[76] M. D'Elia, L. Mirabella, T. Passerini, M. Perego, M. Piccinelli, C. Vergara, and A. Veneziani. Applications of variational data assimilation in computational hemodynamics. *In: D. Ambrosi, A. Quarteroni, G. Rozza (Eds.) Modeling of Physiological Flows, Springer-Verlag Italia*, 2011.

[77] L. Demkowicz. Babuška = Brezzi? Technical Report 08-06, ICE, Univ. of Texas, Austin, 2006.

[78] S. Deparis. Reduced basis error bound computation of parameter-dependent Navier-Stokes equations by the natural norm approach. *SIAM J. Num. Anal.*, 46(4):2039–2067, 2008.

[79] S. Deparis and G. Rozza. Reduced basis method for multi-parameter-dependent steady Navier-Stokes equations: Applications to natural convection in a cavity. *J. Comp. Phys.*, 228(12):4359–4378, 2009.

[80] V. Deplano, C. Bertolotti, and O. Boiron. Numerical simulations of unsteady flows in a stenosed coronary bypass graft. *Med. Biol. Eng. Comput.*, 39:488–499, 2001.

[81] M. Desai and K. Ito. Optimal controls of Navier–Stokes equations. *SIAM J. Control Optim.*, 32(5):1428–1446, 1994.

[82] J.A. Désidéri, B. Abou El Majd, and A. Janka. Nested and self-adaptive Bézier parameterizations for shape optimization. *J. Comput. Phys.*, 224(1):117–131, 2007.

[83] Z. Ding, K. Wang, J. Li, and X. Cong. Flow field and oscillatory shear stress in a tuning-fork-shaped model of the average human carotid bifurcation. *Journal of Biomechanics*, 34:1555–1562, 2001.

[84] H. Do, A. A. Owida, W. Yang, and Y. S. Morsi. Numerical simulation of the haemodynamics in end-to-side anastomoses. *Int. J. Numer. Meth Fluids*, 67(5):638–650, 2011.

[85] D. Doorly and S. Sherwin. Geometry and flow. In L. Formaggia, A Quarteroni, and A. Veneziani, editors, *Cardiovascular Mathematics*, volume 1 of *Modeling, Simulation and Applications (MS&A)*. Springer-Verlag Italia, Milano, 2009.

[86] A. Dumon, C. Allery, and A. Ammar. Proper general decomposition (PGD) for the resolution of Navier-Stokes equations. *J. Comput. Phys.*, 230:1387–1407, 2011.

[87] O. Dur, S.T. Coskun, K.O. Coskun, D. Frakes, L.B. Kara, and K. Pekkan. Computer-aided patient-specific coronary artery graft design improvements using CFD coupled shape optimizer. *Cardiovasc. Engr. Tech.*, pages 1–13, 2011.

[88] R. Duvigneau, B. Chaigne, and J.-A. Désidéri. Multi-level parameterization for shape optimization in aerodynamics and electromagnetics using a particle swarm optimization algorithm. Technical Report 6003, INRIA, 2006.

[89] R. Dziri, M. Moubachir, and J.P. Zolésio. Dynamical shape gradient for the Navier-Stokes system. *C. R. Acad. Sci. Paris. Sér. I Math.*, 338:183–186, 2004.

[90] J.L. Eftang, M. Grepl, and A.T. Patera. A posteriori error bounds for the empirical interpolation method. *C.R. Math. Acad. Sci. Paris, Série I*, 348(9–10):575–579, 2010.

[91] J.L. Eftang, M. Grepl, A.T. Patera, and E.M. Rønquist. Approximation of parametric derivatives by the empirical interpolation method. *Submitted*, 2011. Technical report NTNU Numerics 4.2011: `http://www.math.ntnu.no/preprint/numerics/`.

[92] H.C. Elman, D.J. Silvester, and A.J. Wathen. *Finite Elements and Fast Iterative Solvers with Applications in Incompressible Fluid Dynamics*. Oxford Science Publications, Series in Numerical Mathematics and Scientific Computation, Clarendon Press, Oxford, 2005.

[93] C.R. Ethier, S. Prakash, D.A. Steinman, R.L. Leask, G.G. Couch, and M. Ojha. Steady flow separation patterns in a 45 degree junction. *J. Fluid Mech.*, 411(1):1–38, 2000.

[94] C.R. Ethier, D.A. Steinman, X. Zhang, S.R. Karpik, and M. Ojha. Flow waveform effects on end-to-side anastomotic flow patterns. *J. Biomech.*, 31(7):609–617, 1998.

[95] G. Farin. *Curves and surfaces for computer-aided geometric design: a practical guide*. Morgan Kaufmann, 2001.

[96] M.A. Fernández and J.F. Gerbeau. Algorithms for fluid-structure interaction problems. In L. Formaggia, A Quarteroni, and A. Veneziani, editors, *Cardiovascular Mathematics*, volume 1 of *Modeling, Simulation and Applications (MS&A)*. Springer-Verlag Italia, Milano, 2009.

[97] J.P. Fink and W.C. Rheinboldt. On the error behavior of the reduced basis technique for nonlinear finite element approximations. *Z. Angew. Math. Mech.*, 63(1):21–28, 1983.

[98] R.K. Fisher, T.V. How, T. Carpenter, J.A. Brennan, and P.L. Harris. Optimising miller cuff dimensions. the influence of geometry on anastomotic flow patterns. *Europ. J. Vasc. Endov. Surg.*, 21(3):251–260, 2001.

[99] R. Fletcher. Modified Marquardt subroutine for non-linear least squares. Technical report, Atomic Energy Research Establishment, Harwell (England), 1971.

[100] L. Formaggia, A. Quarteroni, and A. Veneziani (Eds.). *Cardiovascular Mathematics – Modeling and simulation of the circulatory system*, volume 1 of *Modeling, Simulation and Applications (MS&A)*. Springer-Verlag Italia, Milano, 2009.

[101] M. Frangos, Y. Marzouk, K. Willcox, and B. van Bloemen Waanders. Surrogate and reduced-order modeling: A comparison of approaches for large-scale statistical inverse problems. In L. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, Y. Marzouk, B. van Bloemen Waanders, and K. Willcox, editors, *Large-Scale Inverse Problems and Quantification of Uncertainty*, pages 123–149. John Wiley & Sons, Ltd, 2010.

[102] T. Frauenfelder, E. Boutsianis, T. Schertler, L. Husmann, S. Leschka, D. Poulikakos, B. Marincek, and H. Alkadhi. Flow and wall shear stress in end-to-side and side-to-side anastomosis of venous coronary artery bypass grafts. *BioMedical Engineering OnLine*, 35(6), 2007.

[103] N. Fujii. Lower semicontinuity in domain optimization problems. *J. Optim. Theory Appl.*, 59:407–422, 1988.

[104] J. Gain and D. Bechmann. A survey of spatial deformation from a user-centered perspective. *ACM Trans. Graph.*, 27(4):107:1–107:21, 2008.

[105] J.E. Gain and N.A. Dodgson. Preventing self-intersection under free-form deformation. *IEEE Trans. Vis. Comput. Graph.*, 7(4):289–298, 2001.

[106] D. Galbally, K. Fidkowski, K. Willcox, and O. Ghattas. Nonlinear model reduction for uncertainty quantification in large-scale inverse problems. *Int. J. Numer. Methods Engrg.*, 81(12):1581–1608, 2010.

[107] Z. Gao and Y. Ma. Shape gradient of the dissipated energy functional in shape optimization for the viscous incompressible flow. *Appl. Numer. Math.*, 58(11):1720–1741, 2008.

[108] Z. Gao, Y. Ma, and H. Zhuang. Shape optimization for Stokes flow. *Appl. Numer. Math.*, 58:827–844, 2008.

[109] F. Gelsomino and G. Rozza. Comparison and combination of reduced order modelling techniques in 3D parametrized heat transfer problems. *Math. Comput. Model. Dynam. Syst.*, 17(4):371–394, 2011.

[110] A.-L. Gerner and K. Veroy. Reduced basis a posteriori error bounds for the Stokes equations in parametrized domains: a penalty approach. *Math. Models Meth. Appl. Sci.*, 21(10):2103–2134, 2010.

[111] A.-L. Gerner and K. Veroy. Certified reduced basis methods for parametrized saddle point problems. *Submitted*, 2011. Technical report AICES-2011/11-01: `http://www.aices.rwth-aachen.de:8080/aices/preprint/list2011.html`.

[112] O. Ghattas and J.H. Bark. Optimal control of two- and three-dimensional incompressible Navier-Stokes flows. *J. Comput. Phys.*, 136(2):231–244, 1997.

[113] D.P. Giddens, H.S. Bassiouny, S. Glagov, and C.K. Zahns. Measurements of velocity and wall shear stress inside a PTFE vascular graft model under steady flow conditions. *J. Biomech. Engr.*, 119:187, 1997.

[114] S. Giordana, S.J. Sherwin, J. Peiró, D.J. Doorly, J.S. Crane, K.E. Lee, N.J.W. Cheshire, and C.G. Caro. Local and global geometric influence on steady flow in distal anastomoses of peripheral bypass grafts. *J. Biomech. Engr.*, 127:1087, 2005.

[115] V. Girault and P.-A. Raviart. *Finite element methods for Navier-Stokes equations: Theory and algorithms*. Springer-Verlag, Berlin and New York, 1986.

[116] R. Glowinski and O. Pironneau. Towards the computation of minimum drag profiles in viscous laminar flow. *Appl. Math. Modelling*, 1(2):58–66, 1976/77.

[117] W.J. Gordon and C.A. Hall. Construction of curvilinear co-ordinate systems and applications to mesh generation. *Int. J. Numer. Methods Engrg.*, 7(4):461–477, 1973.

[118] C. Grandmont. Existence et unicité de solutions d'un problème de couplage fluide-structure bidimensionnel stationnaire. *C. R. Math. Acad. Sci. Paris*, 326:651–656, 1998.

[119] C. Grandmont. Existence for a three-dimensional steady state fluid-structure interaction problem. *J. Math. Fluid Mech.*, 4:76–94, 2002.

[120] M. Grepl. Certified reduced basis methods for nonaffine linear time-varying and nonlinear parabolic partial differential equations. *Math. Mod. and Meth. in Appl. Sc.*, 22(3):1150015, 2012.

[121] M. Grepl and M. Kärcher. Reduced basis a posteriori error bounds for parametrized linear-quadratic elliptic optimal control problems. *C. R. Math. Acad. Sci. Paris, Sér. I Math.*, 349(15-16):873 – 877, 2011.

[122] M. Grepl, Y. Maday, N.C. Nguyen, and A.T. Patera. Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. *ESAIM Math. Modelling Numer. Anal.*, 41(3):575–605, 2007.

[123] M. Grepl, N.C. Nguyen, K. Veroy, A.T. Patera, and G.R. Liu. Certified rapid solution of partial differential equations for real-time parameter estimation and optimization. In L. Biegler, O. Ghattas, M. Heinkenschloss, D. Keyes, and B. Van Bloemen Waanders, editors, *Real-time PDE-Constrained Optimization*, pages 197–215. Society for Industrial and Applied Mathematics, Philadephia, PA, 2007.

[124] M. Grepl and K. Veroy. A level set reduced basis approach to parameter estimation. *C.R. Math. Acad. Sci. Paris, Series I*, 349(23–24):1229–1232, 2011.

[125] P.M. Gresho and R.L. Sani. *Incompressible Flow and the Finite Element Method: Advection-Diffusion and Isothermal Laminar Flow*. John Wiley & Sons, 1998.

[126] S. Gugercin and A. Antoulas. A survey of model reduction by balanced truncation and some new results. *Int. J. Control.*, 77(8):748–766, 2004.

[127] M. D. Gunzburger. *Finite Element Methods for Viscous Incompressible Flows*. Academic Press, 1989.

[128] M.D. Gunzburger. *Perspectives in Flow Control and Optimization*. Advances in Design and Control. Society for Industrial and Applied Mathematics, Philadephia, PA, 2003.

[129] M.D. Gunzburger and P.B. Bochev. Finite element methods for optimization and control problems for the Stokes equations. *Comput. Math. Appl.*, 48(7–8):1035–1057, 2004.

[130] M.D. Gunzburger, L. Hou, and T.P. Svobodny. Boundary velocity control of incompressible flow with an application to viscous drag reduction. *SIAM J. Control Optim.*, 30:167, 1992.

[131] M.D. Gunzburger, L.S. Hou, and T.P. Svobodny. Optimal control problems for a class of nonlinear equations with an application to control of fluids. In S.S. Sritharan, editor, *Optimal control of viscous flow*, pages 43–62. Society for Industrial and Applied Mathematics, 1998.

[132] M.D. Gunzburger and H. Kim. Existence of an optimal solution of a problem for the stationary Navier-Stokes equations. *SIAM J. Control Optim.*, 36(3):895–909, 1998.

[133] M.D. Gunzburger, H. Kim, and S. Manservisi. On a shape control problem for the stationary Navier-Stokes equations. *ESAIM Math. Modelling Numer. Anal.*, 34(6):1233–1258, 2000.

[134] M.D. Gunzburger, J.S. Peterson, and J.N. Shadid. Reducer-order modeling of time-dependent PDEs with multiple parameters in the boundary data. *Comput. Methods Appl. Mech. Engrg.*, 196:1030–1047, 2007.

[135] B. Haasdonk and M. Ohlberger. Reduced basis method for finite volume approximations of parametrized linear evolution equations. *ESAIM Math. Modelling Numer. Anal.*, 42:277–302, 2008.

[136] H. Haruguchi and S. Teraoka. Intimal hyperplasia and hemodynamic factors in arterial bypass and arteriovenous grafts: a review. *J. Artif. Organs*, 6(4):227–235, 2003.

[137] J. Haslinger and R.A.E. Mäkinen. *Introduction to shape optimization: theory, approximation, and computation*. Society for Industrial and Applied Mathematics, Philadephia, PA, 2003.

[138] M. Heinkenschloss. Formulation and analysis of a sequential quadratic programming method for the optimal dirichlet boundary control of Navier-Stokes flow. In *Optimal Control: Theory, Algorithms, and Applications*, pages 178–203. Kluwer Academic Publishers B.V., 1998.

[139] A. Henrot and M. Pierre. *Variation et optimisation de formes. Une analyse géométrique*. Springer, 2005. Series Mathématiques et Applications, Vol. 48.

[140] R. Herzog and F. Schmidt. Weak lower semi-continuity of the optimal value function and applications to worst-case robust optimal control problems. *Optim.*, 2011. In press (DOI: 10.1080/02331934.2011.603322).

[141] J.S. Hesthaven, B. Stamm, and S. Zhang. Efficient greedy algorithms for high-dimensional parameter spaces with applications to empirical interpolation and reduced basis methods. Technical Report 2011-15, Scientific Computing Group, Brown University, Providence, RI, USA, 2011.

[142] M. Hintermüller, K. Kunisch, Y. Spasov, and S. Volkwein. Dynamical systems-based optimal control of incompressible fluids. *Int. J. Numer. Meth. Fluids*, 46(4):345–359, 2004.

[143] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*. Springer, 2009.

[144] P. Holmes, J.L. Lumley, and G. Berkooz. *Turbulence, coherent structures, dynamical systems and symmetry*. Cambridge Univ. Press, 1998.

[145] J.D. Humphrey. Review paper: Continuum biomechanics of soft biological tissues. *Proc. R. Soc. A*, 459(2029):3–46, 2003.

[146] D.B.P. Huynh, D.J. Knezevic, Y. Chen, J.S. Hesthaven, and A.T. Patera. A natural-norm successive constraint method for inf-sup lower bounds. *Comput. Meth. Appl. Mech. Engrg.*, 199(29–32):1963–1975, 2010.

[147] D.B.P. Huynh, D.J. Knezevic, and A.T. Patera. Certified reduced basis model characterization: a frequentistic uncertainty framework. *Comput. Methods Appl. Mech. Engrg.*, 201–204:13–24, 2012.

[148] D.B.P. Huynh, D.J. Knezevic, J.W. Peterson, and A.T. Patera. High-fidelity real-time simulation on deployed platforms. *Comp. Fluids*, 43(1):74–81, 2011.

[149] D.B.P. Huynh, N.C. Nguyen, A.T. Patera, and G. Rozza. Rapid reliable solution of the parametrized partial differential equations of continuum mechanics and transport, 2008-2012. Website: `http://augustine.mit.edu`.

[150] D.B.P Huynh, G. Rozza, S. Sen, and A.T. Patera. A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability costants. *C. R. Acad. Sci. Paris. Sér. I Math.*, 345:473–478, 2007.

[151] L. Iapichino, A. Quarteroni, and G. Rozza. A reduced basis hybrid method for the coupling of parametrized domains represented by fluidic networks. *Comput. Methods Appl. Mech. Engrg.*, 221–222:63–82, 2012.

[152] K. Ito and S. Ravindran. A reduced basis method for control problems governed by PDEs. *In W. Desch, F. Kappel and K. Kunisch eds. Control and Estimation of Distributed Parameter System*, pages 153–168, 1998.

[153] K. Ito and S.S. Ravindran. A reduced order method for simulation and control of fluid flows. *J. Comput. Phys.*, 143(2):403–425, 1998.

[154] K. Ito and S.S. Ravindran. Reduced basis method for optimal control of unsteady viscous flows. *Int. J. Comput. Fluid Dynamics*, 15(2):97–113, 2001.

[155] S. Jakobsson and O. Amoignon. Mesh deformation using radial basis functions for gradient-based aerodynamic shape optimization. *Comput. Fluids*, 36:1119–1136, 2007.

[156] S. Jakobsson, M. Patriksson, J. Rudholm, and A. Wojciechowski. A method for simulation based optimization using radial basis functions. *Optim. Eng.*, 11:501–532, 2010.

[157] S. James Press. *Subjective and objective Bayesian statistics: principles, models and applications*, volume 590 of *Wiley Series in Probability and Statistics*. John Wiley & Sons, 2009.

[158] A. Jameson. Aerodynamic design via control theory. *J. Sci. Comput.*, 3:233–260, 1988.

[159] A. Jameson. Optimum aerodynamic design using cfd and control theory. In *Proceedings of the 12th AIAA Computational Fluid Dynamics Conference 1995*, pages 926–949, 1995. AIAA Paper 95-1729.

[160] A. Jameson and J.C. Vassberg. Computational fluid dynamics for aerodynamic design: its current and future impact. In *Proceedings of the 39th AIAA Aerospace Sciences Meeting and Exibit, Reno NV*, 2001. AIAA paper 2001-0538.

[161] P.S. Johansson, H.I. Andersson, and E.M. Rønquist. Reduced-basis modeling of turbulent plane channel flow. *Compu. Fluids*, 35(2):189–207, 2006.

[162] J. Kaipio and E. Somersalo. *Statistical and computational inverse problems*, volume 160 of *Applied Mathematical Sciences*. Springer Science+Business Media, Inc., 2005.

[163] H. Kasumba and K. Kunisch. Shape design optimization for viscous flows in a channel with a bump and an obstacle. In *15th International Conference on Methods and Models in Automation and Robotics (MMAR), 2010, Miedzyzdroje, Poland*, pages 284–289. IEEE, 2010.

[164] R.S. Keynton, M.M. Evancho, R.L. Sims, N.V. Rodway, A. Gobin, and S.E. Rittgers. Intimal hyperplasia and wall shear in arterial bypass graft distal anastomoses: an in vivo model study. *J. Biomech. Engr.*, 123:464, 2001.

[165] B.N. Khoromskij and C. Schwab. Tensor-structured galerkin approximation of parametric and stochastic elliptic pdes. *SIAM J. Sci. Comput.*, 33(1):364–385, 2011.

[166] H. Kim. Penalized approach and analysis of an optimal shape control problem for the stationary Navier-Stokes equations. *J. Korean Math. Soc.*, 38:1–23, 2001.

[167] H. Kim and O.K. Kwon. On a vorticity minimization problem for the stationary 2D Stokes equations. *J. Korean Math. Soc.*, 43(1):45–63, 2006.

[168] D.J. Knezevic, N.C. Nguyen, and A.T. Patera. Reduced basis approximation and a posteriori error estimation for the parametrized unsteady Boussinesq equations. *Math. Mod. and Meth. in Appl. Sc.*, 21(7):1415–1442, 2011.

[169] V. Kolachalama, N. Bressloff, and P. Nair. Mining data from hemodynamic simulations via bayesian emulation. *BioMedical Engineering OnLine*, 6(1):47, 2007.

[170] D. Kressner and C. Tobler. Low-rank tensor krylov subspace methods for parametrized linear systems. *SIAM J. Matrix Anal Appl.*, 32(4):1288–1316, 2011.

[171] D.N. Ku and D.P. Giddens. Pulsatile flow in a model carotid bifurcation. *Arterioscler. Thromb. Vasc. Biol.*, 3(1):31–39, 1983.

[172] D.N. Ku, D.P. Giddens, C.K. Zarins, and S. Glagov. Pulsatile flow and atherosclerosis in the human carotid bifurcation. positive correlation between plaque location and low oscillating shear stress. *Arterioscler. Thromb. Vasc. Biol.*, 5(3):293–302, 1985.

[173] J.P. Ku, C.J. Elkins, and C.A. Taylor. Comparison of cfd and mri flow and velocities in an in vitro large artery bypass graft model. *Ann. Biomed. Engr.*, 33(3):257–269, 2005.

[174] K. Kunisch and B. Vexler. Optimal vortex reduction for instationary flows based on translation invariant cost functionals. *SIAM J. Control Optim*, 46(4):1368–1397, 2007.

[175] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM J. Numer. Anal.*, 40(2):492–515, 2003.

[176] K. Kunisch and S. Volkwein. Proper orthogonal decomposition for optimality systems. *ESAIM Math. Modelling Numer. Anal.*, 42(1):1–23, 2008.

[177] K. Kunisch and S. Volkwein. Optimal snapshot location for computing pod basis functions. *ESAIM Math. Modelling Numer. Anal.*, 44(3):509–529, 2010.

[178] O. Ladyzhenskaya. *The mathematical theory of viscous incompressible flow*. Gordon and Breach Science Publishers, New York, 2nd edition, 1969.

[179] H.J. Lamousin and W.N. Waggenspack. NURBS-based free-form deformations. *IEEE Comput. Graph. Appl.*, 14(6):59–65, 1994.

[180] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza. A reduced computational and geometrical framework for inverse problems in haemodynamics. *Submitted*, 2011. Technical report MATHICSE 12.2011: `http://mathicse.epfl.ch/`.

[181] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza. Boundary control and shape optimization for the robust design of bypass anastomoses under uncertainty. *Submitted*, 2012. Technical report MATHICSE 03.2012: `http://mathicse.epfl.ch/`.

[182] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza. Shape parameterization techniques in model reduction, optimal shape design, and inverse problems. *In preparation*, 2012.

[183] T. Lassila, A. Manzoni, and G. Rozza. On the approximation of stability factors for general parametrized partial differential equations with a two-level affine decomposition. *Submitted*, 2011. Technical report MATHICSE 08.2011: `http://mathicse.epfl.ch/`.

[184] T. Lassila, A. Quarteroni, and G. Rozza. A reduced basis model with parametric coupling for fluid-structure interaction problem. *SIAM J. Sci. Comput.*, 2012. Accepted.

[185] T. Lassila and G. Rozza. Parametric free-form shape design with PDE models and reduced basis method. *Comput. Methods Appl. Mech. Engrg.*, 199(23–24):1583–1592, 2010.

[186] P. Le Tallec and J. Mouro. Fluid structure interaction with large structural displacements. *Comput. Methods. Appl. Mech. Engrg.*, 190:3039–3067, 2001.

[187] S.W. Lee, L. Antiga, J. D. Spence, and D. A. Steinman. Geometry of the carotid bifurcation predicts its exposure to disturbed flow. *Stroke*, 39(8):2341–2347, 2008.

[188] S. Lefteriu and A.C. Antoulas. Topics in model order reduction with applications to circuit simulation. In P. Benner, M. Hinze, and E.J.W. ter Maten, editors, *Model Reduction for Circuit Simulation*, volume 74 of *Lecture Notes in Electrical Engineering*, pages 85–107. Springer Netherlands, 2011.

[189] T. Lehnhäuser and M. Schäfer. A numerical approach for shape optimization of fluid flow domains. *Comput. Methods Appl. Mech. Engrg.*, 194:5221–5241, 2005.

[190] M. Lei, J. Archie, and C. Kleinstreuer. Computational design of a bypass graft that minimizes wall shear stress gradients in the region of the distal anastomosis. *J. Vasc. Surg.*, 25(4):637–646, 1997.

[191] C. Lieberman, K. Willcox, and O. Ghattas. Parameter and state model reduction for large-scale statistical inverse problems. *SIAM J. Sci. Comput.*, 32(5):2523–2542, 2010.

[192] J.L. Lions. *Optimal Control of Systems governed by Partial Differential Equations.* Springer-Verlag, 1971.

[193] F. Loth, P.F. Fischer, and H.S. Bassiouny. Blood flow in end-to-side anastomoses. *Annu. Rev. Fluid. Mech.*, 40:367–393, 2008.

[194] A. E. Løvgren, Y. Maday, and E. M. Rønquist. The reduced basis element method for fluid flows. In *Analysis and simulation of fluid dynamics. Advances in mathematical fluid dynamics. Birkhäuser, Boston*, pages 129–154, 2007.

[195] A. E. Løvgren, Y. Maday, and E. M. Rønquist. The reduced basis element method: offline-online decomposition in the nonconforming, nonaffine case. In *Spectral and High Order Methods for Partial Differential Equations. Selected papers from the ICOSAHOM '09 conference, June 22-26, Trondheim, Norway*, pages 247–254. Springer, Series: Lecture Notes in Computational Science and Engineering, vol. 76, J.S. Hesthaven, E.M. Rønquist (Eds.), 2011.

[196] A.E. Løvgren, Y. Maday, and E.M. Rønquist. A reduced basis element method for the steady Stokes problem. *ESAIM Math. Modelling Numer. Anal.*, 40(3):529–552, 2006.

[197] X. Ma and G.E.M. Karniadakis. A low-dimensional model for simulating three-dimensional cylinder flow. *J. Fluid. Mech*, 458:181–190, 2002.

[198] L. Machiels, Y. Maday, A. T. Patera, and D. Rovas. A blackbox reduced basis output bound method for shape optimization. In T. Chan, T. Kako, H. Kawarada, and O. Pironneau, editors, *Proceedings of the 12th International Conference on Domain Decomposition Methods*, pages 429–436, 2001.

[199] Y. Maday. Analysis of coupled models for fluid-structure interaction of internal flows. In L. Formaggia, A Quarteroni, and A. Veneziani, editors, *Cardiovascular Mathematics*, volume 1 of *Modeling, Simulation and Applications (MS&A)*. Springer-Verlag Italia, Milano, 2009.

[200] Y. Maday, N.C. Nguyen, A.T. Patera, and G.S.H. Pau. A general multipurpose interpolation procedure: the magic points. *Commun. Pure Appl. Anal.*, 8(1), 2009.

[201] Y. Maday and E.M. Rønquist. A reduced-basis element method. *J. Sci. Comput.*, 17:447–459, 2002.

[202] Y. Maday and E.M. Rønquist. The reduced basis element method: Application to a thermal fin problem. *SIAM J. Sci. Comput.*, 26(1):240–258, 2004.

[203] A.C.I. Malossi, P. Blanco, S. Deparis, and A. Quarteroni. Algorithms for the partitioned solution of weakly coupled fluid models for cardiovascular flows. *Int. J. Num. Methods Biomed. Engrg.*, 27:2035–2057, 2011.

[204] A. Manbachi, Y. Hoi, B.A. Wasserman, E.G. Lakatta, and D.A. Steinman. On the shape of the common carotid artery with implications for blood velocity profiles. *Phys. Meas.*, 32(12):1885–1897, 2011.

[205] A. Manzoni, A. Quarteroni, and G. Rozza. Model reduction techniques for fast blood flow simulation in parametrized geometries. *Int. J. Numer. Methods Biomed. Engrg.*, 2011. In press (`DOI: 10.1002/cnm.1465`).

[206] A. Manzoni, A. Quarteroni, and G. Rozza. Shape optimization of cardiovascular geometries by reduced basis methods and free-form deformation techniques. *Int. J. Numer. Methods Fluids*, 2011. In press (`DOI: 10.1002/fld.2712`).

[207] A. Manzoni, A. Quarteroni, and G. Rozza. Computational reduction for parametrized pdes: strategies and applications. *Milan J. Math.*, In press, 2012. Technical report MATHICSE 15.2012: `http://mathicse.epfl.ch/`.

[208] A. Manzoni and G. Rozza. Model order reduction by geometrical parametrization for shape optimization in computational fluid dynamics. In J.C.F. Pereira and A. Sequeira, editors, *Proceedings of the V European Conf. Computat. Fluid Dynamics, ECCOMAS CFD 2010, June 14-17, Lisbon, Portugal*, 2010.

[209] A. Manzoni and G. Rozza. A posteriori error estimation for quadratic outputs in parametrized Stokes and Navier-Stokes problems. *In preparation*, 2012.

[210] A.L. Marsden, J.A. Feinstein, and C.A. Taylor. A computational framework for derivative-free optimization of cardiovascular geometries. *Comput. Methods Appl. Mech. Engrg.*, 197(21–24):1890–1905, 2008.

[211] D.B. McDonald, W.J. Grantham, W.L. Tabor, and M.J. Murphy. Global and local optimization using radial basis function response surface models. *Applied Mathematical Modelling*, 31(10):2095–2110, 2007.

[212] K. McLeod, A. Caiazzo, M.A. Fernandez, T. Mansi, I.E. Vignon-Clementel, M. Sermesant, X. Pennec, Y. Boudjemline, and J.F. Gerbeau. Atlas-based reduced models of blood flows for fast patient-specific simulations. In *Statistical Atlases and Computational Models of the Heart. Proceedings from the First International Workshop, STACOM 2010, and Cardiac Electrophysical Simulation Challenge, CESC 2010, September 20, 2010, Beijing, China*, pages 95–104. Springer, Series: Lecture Notes in Computer Science and Engineering, vol. 6364, O. Camara, M. Pop, K. Rhode, M. Sermesant, N. Smith, A. Young (Eds.), 2010.

[213] J.S. Milner, J.A. Moore, B.K. Rutt, and D.A. Steinman. Hemodynamics of human carotid artery bifurcations: Computational studies with models constructed with magnetic resonance imaging of normal subjects. *Journal of Vascular Surgery*, 28:143–156, 1998.

[214] B. Mohammadi and O. Pironneau. *Applied shape optimization for fluids*. Oxford University Press, 2001.

[215] B. Mohammadi and O. Pironneau. Optimal shape design for fluids. *Annual Rev. Fluids Mech.*, 36:255–279, 2004.

[216] J. Montagnat, H. Delingette, and N. Ayache. A review of deformable surfaces: topology, geometry and deformation. *Image and Vision Computing*, 19(14):1023 – 1040, 2001.

[217] A.M. Morris, C.B. Allen, and T.C.S. Rendall. CFD-based optimization of aerofoils using radial basis functions for domain element parameterization and mesh deformation. *Int. J. Numer. Methods Fluids*, 58:827–860, 2008.

[218] M.D. Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161–174, 1991.

[219] K. Mosegaard and A. Tarantola. Probabilistic approach to inverse problems. In P. Jennings W. Lee, H. Kanamori and C. Kisslinger, editors, *International Handbook of Earthquake and Engineering Seismology*, volume 81, Part A, pages 237–265. Academic Press, 2002.

[220] M. Moubachir and J.P. Zolésio. *Moving shape analysis and control: applications to fluid structure interactions*. Chapman & Hall/CRC, 2006.

[221] F. Murat and J. Simon. *Sur le contrôle par un domaine géométrique*, volume 74015. Université de Paris 6, 1976. Pré-publication du Laboratoire d'Analyse Numérique.

[222] F. Mut, R. Löhner, A. Chien, S. Tateshima, F. Viñuela, C. Putman, and J.R. Cebral. Computational hemodynamics framework for the analysis of cerebral aneurysms. *International Journal for Numerical Methods in Biomedical Engineering*, 2010.

[223] F. Negri, G. Rozza, A. Manzoni, and A. Quarteroni. Certified reduced basis method for parametrized elliptic optimal control problems in saddle-point formulation. *In preparation*, 2012.

[224] N. C. Nguyen, K. Veroy, and A. T. Patera. Certified real-time solution of parametrized partial differential equations. In S. Yip, editor, *Handbook of Materials Modeling*, pages 1523–1558. Springer, The Netherlands, 2005.

[225] N.C. Nguyen. *Reduced-basis approximation and a posteriori error bounds for nonaffine and nonlinear partial differential equations: application to inverse analysis*. PhD thesis, Singapore-MIT Alliance, National University of Singapore, 2005.

[226] N.C. Nguyen. A posteriori error estimation and basis adaptivity for reduced-basis approximation of nonaffine-parametrized linear elliptic partial differential equations. *J. Comp. Phys.*, 227:983–1006, 2007.

[227] N.C. Nguyen, G. Rozza, D.B.P. Huynh, and A.T. Patera. Reduced basis approximation and a posteriori error estimation for parametrized parabolic pdes; application to real-time bayesian parameter estimation. In L. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, Y. Marzouk, B. van Bloemen Waanders, and K. Willcox, editors, *Large-Scale Inverse Problems and Quantification of Uncertainty*, pages 151–178. John Wiley & Sons, Ltd, 2010.

[228] N.C. Nguyen, G. Rozza, and A.T. Patera. Reduced basis approximation and a posteriori error estimation for the time-dependent viscous Burgers' equation. *Calcolo*, 46(3):157–185, 2009.

[229] J. Nocedal and S.J. Wright. *Numerical optimization*. Springer Verlag, 1999.

[230] A.K. Noor. Recent advances in reduction methods for nonlinear problems. *Comput. Struct.*, 13:31–44, 1981.

[231] A.K. Noor. On making large nonlinear problems small. *Comput. Meth. Appl. Mech. Engrg.*, 34:955–985, 1982.

[232] A. Nouy. Proper generalized decompositions and separated representations for the numerical solution of high dimensional stochastic problems. *Arch. Comput. Methods Engrg.*, 17:403–434, 2010.

[233] J.T. Ottesen, M.S. Olufsen, and J.K. Larsen. *Applied Mathematical Models in Human Physiology*. Society for Industrial and Applied Mathematics, Philadephia, PA, 2004.

[234] R.L. Panton. *Incompressible Flow*. Wiley & Sons, 3rd edition, 2005.

[235] P. Parang and R. Arora. Coronary vein graft disease: Pathogenesis and prevention. *Can. J. Cardiol.*, 25(2):e57–e62, 2009.

[236] A.T. Patera. Private communication.

[237] A.T. Patera and G. Rozza. *Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations*. to appear in MIT Pappalardo Graduate Monographs in Mechanical Engineering, 2009. ©Massachusetts Institute of Technology, Version 1.0.

[238] J. Peiró and A. Veneziani. Reduced models of the cardiovascular system. In L. Formaggia, A Quarteroni, and A. Veneziani, editors, *Cardiovascular Mathematics*, volume 1 of *Modeling, Simulation and Applications (MS&A)*. Springer-Verlag Italia, Milano, 2009.

[239] M. Perego, A. Veneziani, and C. Vergara. Variational approach for estimating the compliance of the cardiovascular tissue: an inverse fluid-structure interaction problem. *SIAM J. Sci. Comput.*, 33(3):1181–1211, 2011.

[240] K. Perktold and D. Hilbert. Numerical simulation of pulsatile flow in a carotid bifurcation model. *J. Biomed. Eng.*, 8:193–199, 1986.

[241] K. Perktold and M. Resch. Numerical flow studies in human carotid artery bifurcations: Basic discussion of the geometric factor in atherogenesis. *J. Biomed. Eng.*, 12:111–123, 1990.

[242] J.S. Peterson. The reduced basis method for incompressible viscous flow calculations. *SIAM J. Sci. Stat. Comput.*, 10(4):777–786, 1989.

[243] N.A. Pierce and M. B. Giles. Adjoint recovery of superconvergent functionals from PDE approximations. *SIAM Review*, 42(2):247–264, 2000.

[244] N.A. Pierce and M. B. Giles. Adjoint and defect error bounding and correction for functional estimates. *J. Comp. Phys.*, 200:769–794, 2004.

[245] R. Pinnau. Model reduction via proper orthogonal decomposition. In W.H.A. Schilder and H. van der Vorst, editors, *Model Order Reduction: Theory, Research Aspects and Applications*, pages 96–109. Springer, 2008.

[246] O. Pironneau. On optimum profiles in Stokes flow. *J. Fluid Mech.*, 59:117–128, 1973.

[247] O. Pironneau. On optimum design in fluid mechanics. *J. Fluid Mech.*, 64:97–110, 1974.

[248] O. Pironneau. *Optimal shape design for elliptic systems*. Springer-Verlag, 1984.

[249] O. Pironneau. Optimal shape design by local boundary variations. In B. Kawohl, O. Pironneau, L. Tartar, Jean-Paul Zolésio, and Olivier Pironneau, editors, *Optimal Shape Design*, volume 1740 of *Lecture Notes in Mathematics*, pages 343–384. Springer Berlin / Heidelberg, 2000. Lectures given at the joint C.I.M./C.I.M.E. Summer School held in Tróia, Portugal, June 1-6, 1998.

[250] T. A. Porsching. Estimation of the error in the reduced basis method solution of nonlinear equations. *Mathematics of Computation*, 45(172):487–496, 1985.

[251] T.A. Porsching and M.Y. Lin Lee. The reduced-basis method for initial value problems. *SIAM J. Numer. Anal.*, 24:1277–1287, 1987.

[252] M. Probst, M. Lülfesmann, M. Nicolai, H.M. Bücker, M. Behr, and C.H. Bischof. Sensitivity of optimal shapes of artificial grafts with respect to flow parameters. *Comput. Meth. Appl. Mech. Engrg.*, 199(17-20):997–1005, 2010.

[253] C. Prud'homme, D. Rovas, K. Veroy, Y. Maday, A.T. Patera, and G. Turinici. Reliable real-time solution of parametrized partial differential equations: reduced-basis output bounds methods. *Journal of Fluids Engineering*, 124(1):70–80, 2002.

[254] A. Quarteroni and G. Rozza. Optimal control and shape optimization of aorto-coronaric bypass anastomoses. *Math. Models Meth. Appl. Sci.*, 13(12):1801–1823, 2003.

[255] A. Quarteroni and G. Rozza. Numerical solution of parametrized Navier-Stokes equations by reduced basis methods. *Numer. Methods Partial Differential Equations*, 23(4):923–948, 2007.

[256] A. Quarteroni, G. Rozza, and A. Manzoni. Certified reduced basis approximation for parametrized partial differential equations in industrial applications. *J. Math. Ind.*, 1(3), 2011.

[257] A. Quarteroni, G. Rozza, and A. Quaini. Reduced basis methods for optimal control of advection-diffusion problem. In *Advances in Numerical Mathematics, W. Fitzgibbon, R. Hoppe, J. Periaux, O. Pironneau, and Y. Vassilevski, Editors*, pages 193–216, 2007.

[258] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical Mathematics*, volume 37 of *Texts in Applied Mathematics*. Springer, New York, 2000.

[259] A. Quarteroni, M. Tuveri, and A. Veneziani. Computational vascular fluid dynamics: problems, models and methods. *Comp. Vis. Sci.*, 2:163–197, 2000.

[260] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations (1st Ed.)*. Springer-Verlag, Berlin-Heidelberg, 1994.

[261] A. Quarteroni and A. Veneziani. Analysis of a geometrical multiscale model based on the coupling of PDEs and ODEs for blood flow simulations. *SIAM J. on Multiscale Model. Simul.*, 1(2):173–195, 2003.

[262] C.M. Quick, D.S. Berger, R.H. Stewart, G.A. Laine, C.J. Hartley, and A. Noordergraaf. Resolving the hemodynamic inverse problem. *IEEE Trans. Biomed. Eng*, 53(3):361–368, 2006.

[263] J.O. Ramsay and B.W. Silverman. *Applied functional data analysis: methods and case studies*. Springer Verlag, 2002.

[264] S. S. Ravindran. Reduced-order adaptive controllers for fluid flows using pod. *J. Sci. Comput.*, 15(4):457–478, 2000.

[265] S. S. Ravindran. Adaptive reduced-order controllers for a thermal flow system using proper orthogonal decomposition. *SIAM J. Sci. Comput.*, 23(6):1924–1942, 2002.

[266] S.S. Ravindran. Numerical approximation of optimal flow control problems by sqp method. In S.S. Sritharan, editor, *Optimal control of viscous flow*, pages 181–198. Society for Industrial and Applied Mathematics, 1998.

[267] S.S. Ravindran. A reduced-order approach for optimal control of fluids using proper orthogonal decomposition. *Int. J. Numer. Meth. Fluids*, 34:425–448, 2000.

[268] T. Rees, H.S. Dollar, and A. J. Wathen. Optimal solvers for PDE-constrained optimization. *SIAM J. Sci. Comput.*, 32(1):271–298, 2008.

[269] T.C.S. Rendall and C.B. Allen. Efficient mesh motion using radial basis functions with data reduction algorithms. *J. Comput. Phys.*, 228(17):6231–6249, 2009.

[270] W. C. Rheinboldt. On the theory and error estimation of the reduced basis method for multi-parameter problems. *Nonlinear Analysis, Theory, Methods and Applications*, 21(11):849–858, 1993.

[271] A.M. Robertson, A. Sequeira, and M.V. Kameneva. Hemorheology. *Hemodynamical Flows*, pages 63–120, 2008.

[272] D.V. Rovas. *Reduced-Basis Output Bound Methods for Parametrized Partial Differential Equations*. PhD thesis, Massachusetts Institute of Technology, 2003.

[273] G. Rozza. On optimization, control and shape design of an arterial bypass. *Int. J. Numer. Meth. Fluids*, 47(10–11):1411–1419, 2005.

[274] G. Rozza. Real-time reduced basis techniques for arterial bypass geometries. In K.J. Bathe, editor, *Computational Fluid and Solid Mechanics*, pages 1283–1287. Elsevier, 2005. Proceedings of the Third M.I.T. Conference on Computational Fluid and Solid Mechanics, June 14-17, 2005.

[275] G. Rozza. Reduced-basis methods for elliptic equations in sub-domains with *a posteriori* error bounds and adaptivity. *Appl. Numer. Math.*, 55(4):403–424, 2005.

[276] G. Rozza. *Shape design by optimal flow control and reduced basis techniques: applications to bypass configurations in haemodynamics*. PhD thesis, N. 3400, École Polytechnique Fédérale de Lausanne, 2005.

[277] G. Rozza. Real-time reduced basis solutions for Navier-Stokes equations: optimization of parametrized bypass configurations. In *Proceedings of the ECCOMAS CFD 2006, IV European Conference on Computational Fluid Dynamics, Egmond aan Zee, The Netherlands, P. Wesseling, E. Onate, J. Periaux (Eds.)*, 2006.

[278] G. Rozza. Reduced basis methods for Stokes equations in domains with non-affine parameter dependence. *Comput. Vis. Sci.*, 12(1):23–35, 2009.

[279] G. Rozza, D.B.P. Huynh, and A. Manzoni. Reduced basis approximation and error bounds for Stokes flows in parametrized geometries: roles of the inf-sup stability constants. *Submitted*, 2011. Technical report MATHICSE 22.2010: `http://mathicse.epfl.ch/`.

[280] G. Rozza, D.B.P. Huynh, and A.T. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Arch. Comput. Methods Engrg.*, 15:229–275, 2008.

[281] G. Rozza, T. Lassila, and A. Manzoni. Reduced basis approximation for shape optimization in thermal flows with a parametrized polynomial geometric map. In E.M. Rønquist J.S. Hesthaven, editor, *Spectral and High Order Methods for Partial Differential Equations. Selected papers from the ICOSAHOM '09 conference, June 22-26, Trondheim, Norway*, volume 76 of *Lecture Notes in Computational Science and Engineering*, pages 307–315. Springer, 2011.

[282] G. Rozza and K. Veroy. On the stability of reduced basis methods for Stokes equations in parametrized domains. *Comput. Meth. Appl. Mech. Engrg.*, 196(7):1244–1260, 2007.

[283] W. Rudin. *Principles of Mathematical Analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., 3rd edition, 1976.

[284] J. Sacks, W.J. Welch, T.J. Mitchell, and H.P. Wynn. Design and analysis of computer experiments. *Stat. Science*, 4(4):409–435, 1989.

[285] F. Saleri, P. Gervasio, G. Rozza, M. Discacciati, and A. Manzoni. `MLIfe`, A Matlab LIbrary for Finite Elements, tutorial (in progress), 2000–2012. MOX, Politecnico di Milano and CMCS, Ecole Polytechnique Fédérale de Lausanne. ©Politecnico di Milano.

[286] J.A. Samareh. A survey of shape parametrization techniques. Technical Report NASA/CP-1999-209136, NASA, 1999.

[287] J.A. Samareh. Aerodynamic shape optimization based on free-form deformation. In *Proc. 10th AIAA/ISSMO Multidisciplinary Anal. Optim. Conf.*, volume 6, pages 3672–3683, 2004.

[288] S. Sankaran and A.L. Marsden. The impact of uncertainty on shape optimization of idealized bypass graft models in unsteady flow. *Phys. Fluids*, 22:121902, 2010.

[289] T.J. Santner, B.J. Williams, and W.Notz. *The design and analysis of computer experiments*. Springer-Verlag, New York, 2003.

[290] S.S. Sarakinos, E. Amoiralis, and I.K. Nikolos. Exploring freeform deformation capabilities in aerodynamic shape parameterization. In *Proceedings of the International Conference on Computer as a tool, November 21-24, 2005*, volume 1, pages 535–538, 2005.

[291] R. Schaback. Error estimates and condition numbers for radial basis function interpolation. *Adv. Comput. Math.*, 3:251–264, 1994.

[292] R. Schaback and H. Wendland. Adaptive greedy techniques for approximate solution of large rbf systems. *Numer. Algor.*, 24:239–254, 2000.

[293] R. Schaback and H. Wendland. Numerical techniques based on radial basis functions. In *Curve and Surface Fitting: Saint-Malo 1999*, pages 359–374. Vanderbilt University Press, 2000.

[294] W. Schilder. Introduction to model order reduction. In W. Schilder and H. van der Vorst, editors, *Model Order Reduction: Theory, Research Aspects and Applications,*, pages 3–32. Springer, 2008.

[295] C. Schillings, S. Schmidt, and V. Schulz. Efficient shape optimization for certain and uncertain aerodynamic design. *Comp. Fluids*, 46(1):78–87, 2011.

[296] V. Schulz and C. Schillings. Problem formulations and treatment of uncertainties in aerodynamic design. *AIAA Journal*, 47(3):646–654, 2009.

[297] T.W. Sederberg and S.R. Parry. Free-form deformation of solid geometric models. *Comput. Graph.*, 20(4):151–160, 1986.

[298] S. Sen, K. Veroy, D.B.P. Huynh, S. Deparis, N.C. Nguyen, and A.T. Patera. "Natural norm" a posteriori error estimators for reduced basis approximations. *J. Comp. Phys.*, 217(1):37–62, 2006.

[299] A.M. Shaaban and A.J. Duerinckx. Wall shear stress and early atherosclerosis: a review. *Am. J. Roentgenology*, 174(6):1657, 2000.

[300] S. Shan and G.G. Wang. Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions. *Struct. Multidiscip. Optim.*, 41(2):219–241, 2010.

[301] J. Simon. Domain variation for drag Stokes flows. *In: Bermudez, A. (Ed.) Lecture Notes in Control and Inform. Sci.*, 114:277–283, 1987.

[302] L. Sirovich. Turbulence and the dynamics of coherent structures, part i: Coherent structures. *Quart. Appl. Math.*, 45(3):561–571, 1987.

[303] J. Sokolowski and J.-P. Zolésio. *Introduction to shape optimization: shape sensitivity analysis*. Springer, 1992.

[304] P.B. Stark and L. Tenorio. A primer of frequentist and bayesian inference in inverse problems. In L. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, Y. Marzouk, B. van Bloemen Waanders, and K. Willcox, editors, *Large-Scale Inverse Problems and Quantification of Uncertainty*, pages 151–178. John Wiley & Sons, Ltd, 2010.

[305] A.M. Stuart. Inverse problems: a Bayesian perspective. *Acta Numerica*, 19(1):451–559, 2010.

[306] C.A. Taylor and C.A. Figueroa. Patient-specific modeling of cardiovascular mechanics. *Annu. Rev. Biomed. Eng.*, 11(1):109–134, 2009.

[307] R. Temam. *Navier-Stokes Equations*. AMS Chelsea, Providence, Rhode Island, 2001.

[308] M. Thiriet and K.H. Parker. Physiology and pathology of the cardiovascular system: a physical perspective. In L. Formaggia, A Quarteroni, and A. Veneziani, editors, *Cardiovascular Mathematics*, volume 1 of *Modeling, Simulation and Applications (MS&A)*. Springer-Verlag Italia, Milano, 2009.

[309] T. Tonn, K. Urban, and S. Volkwein. Optimal control of parameter-dependent convection-diffusion problems around rigid bodies. *SIAM J. Sci. Comput.*, 32(3):1237–1260, 2010.

[310] T. Tonn, K. Urban, and S. Volkwein. Comparison of the reduced basis and pod a posteriori error estimators for an elliptic linear-quadratic optimal control problem. *Math. Comput. Model. Dynam. Syst.*, 17(4):355–369, 2011.

[311] F. Troltzsch. *Optimal control of partial differential equations: theory, methods and applications*. Graduate Studies in Mathematics. American Mathematical Society, 2010.

[312] K. Veroy and A.T. Patera. Certified real-time solution of the parametrized steady incompressible Navier-Stokes equations: rigorous reduced-basis a posteriori error bounds. *Int. J. Numer. Meth. Fluids*, 47:773–788, 2005.

[313] K. Veroy, C. Prud'homme, D. V. Rovas, and A. T. Patera. *A posteriori* error bounds for reduced basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations. In *Proceedings of the 16th AIAA Computational Fluid Dynamics Conference*, 2003. Paper 2003-3847.

[314] F.A.C. Viana, C. Gogu, and R.T. Haftka. Making the most out of surrogate models: tricks of the trade. In *Proceedings of the ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, pages 587–598, 2010.

[315] S. Volkwein. Model reduction using proper orthogonal decomposition, 2011. Lecture Notes, University of Konstanz, `www.math.uni-konstanz.de/numerik/personen/volkwein/teaching/POD-Vorlesung.pdf`.

[316] G.G. Wand and S. Shan. Review of metamodeling techniques in support of engineering design optimization. *ASME J. Mech. Design*, 129(4):370–380, 2007.

[317] S. Wang, K. Lim, B.C. Khoo, and M.Y. Wang. An extended level set method for shape and topology optimization. *J. Comp. Phys.*, 221(1):395–421, 2007.

[318] S. Wang and M.Y. Wang. Radial basis functions and level set method for structural topology optimization. *Int. J. Numer. Methods Engrg.*, 65(12):2060–2090, 2006.

[319] H. Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Adv. Comput. Math.*, 4(1):389–396, 1995.

[320] H. Wendland. *Scattered Data Approximation*, volume 17 of *Cambridge monographs on applied and computational mathematics*. Cambridge, 2005.

[321] H. Wendland. Computational aspects of radial basis function approximation. In K. Jetter, M.D. Buhmann, W. Haussmann, R. Schaback, and J. Stäckler, editors, *Topics in Multivariate Approximation and Interpolation*, volume 12 of *Studies in Computational Mathematics*, pages 231–256. Elsevier, 2006.

[322] Z. Wu. Compactly supported positive definite radial functions. *Adv. Comput. Math.*, 4(1):283–292, 1995.

[323] F. Xiong and C. Chong. A parametric numerical investigation on haemodynamics in distal coronary anastomoses. *Medical Engineering & Physics*, 30(3):311–320, 2008.

[324] J. Xu and L. Zikatanov. Some observation on Babuška and Brezzi theories. *Numer. Math.*, 94(1):195–202, 2003.

[325] K. Zaw, G.R. Liu, B. Deng, and K.B.C. Tan. Rapid identification of elastic modulus of the interface tissue on dental implants surfaces using reduced-basis method and a neural network. *J. Biomech.*, 42(5):634–641, 2009.

[326] D. Zeng, Z. Ding, M.H. Friedman, and C.R. Ethier. Effects of cardiac motion on right coronary artery hemodynamics. *Ann. Biomed. Eng.*, 31(4):420–429, 2003.

# Curriculum Vitæ

Andrea Manzoni was born in Bollate (Milano), Italy, on December 31$^{st}$, 1984. He got his High School Diploma (grade: 100/100) at Liceo Scientifico "Luigi Cremona" in Milan in 2003, specializing in science education. Then, he pursued his education at Politecnico di Milano, where he got his Bachelor Degree in Mathematical Engineering (grade: 110/110 cum laude) in 2006 and his Master Degree in Mathematical Engineering, with a specialization in Scientific Computing (grade: 110/110 cum laude) in 2008, carrying on his Master Thesis at MOX - Modeling and Scientific Computing Laboratory. He was awarded with the Best Graduate Student Prize of the year (2005-2006) for the Bachelor Degree in Mathematical Engineering and of the Best Graduate Student Prize of the year (2007-2008) for the Master Degree in Mathematical Engineering.
Since January 2009 he has been teaching assistant at CMCS - Chair de Modélisation et Calcul Scientifique at EPFL, where he started his Ph.D. thesis under the direction of Prof. A. Quarteroni and Dr. Gianluigi Rozza, within a Research Project funded by the Swiss National Science Foundation (SNSF). In these years, he has also taken part to the activities of the ERC Mathcard Project, granted by European Union.
In January, 2011 he has attended the SIAM CSE 11 Conference on Computational Science and Engineering, where he was awarded of the Bavarian Graduate School in Computational Engineering (BGCE) Student Paper Prize for his work on model order reduction by reduced basis for optimal control and shape optimization problems. He has visited the two institutions establishing the prize, Technische Universität München and Universität Erlangen-Nürnberg, Germany. He attended several international conferences and he has been invited to the Journées Lions-Magenes, a workshop organized in Fall 2011 by the LJLL Laboratory of Université Pierre et Marie Curie, Paris VI, celebrating the memory of Prof. Enrico Magenes and Prof. Jacques-Louis Lions. He is author of the following publications:

**Referred publications in International Journals**

1. T. Lassila, A. Manzoni, A. Quarteroni, G. Rozza. A reduced computational and geometrical framework for inverse problems in haemodynamics. Submitted, 2011. Preprint: Technical Report MATHICSE 12.2011.
2. T. Lassila, A. Manzoni, G. Rozza. On the approximation of stability factors for general parametrized partial differential equations with a two-level affine decomposition. Submitted, 2011. Preprint: Technical Report MATHICSE 08.2011.
3. T. Lassila, A. Manzoni, A. Quarteroni, G. Rozza. Boundary control and shape optimization for the robust design of bypass anastomoses under uncertainty. Submitted, 2012. Preprint: Technical Report MATHICSE 03.2012.
4. A. Manzoni. Model order reduction by reduced basis for optimal control and shape optimization. Paper awarded of the $3^{rd}$ BGCE Student Paper Prize at the 2011 SIAM Computational Science & Engineering Conference, Reno, NV, USA.
5. A. Manzoni, A. Quarteroni, G. Rozza. Model reduction techniques for fast blood flow simulation in parametrized geometries. Int. J. Numer. Meth. Biomed. Engng. 2011. Published online in Wiley InterScience, DOI: 10.1002/cnm.1465 (`www.interscience.wiley.com`).
6. A. Manzoni, A. Quarteroni, G. Rozza. Shape optimization for viscous flows by reduced basis methods and free-form deformation. Int. J. Numer. Meth. Fluids 2011. Published online in Wiley InterScience, DOI: 10.1002/fld.2712 (`www.interscience.wiley.com`).
7. A. Manzoni, A. Quarteroni, G. Rozza. Computational reduction for parametrized PDEs: strategies and applications. Accepted for publication in Milan J. Math. 2012. Preprint: Mathicse Report 15.2012.

8. A. Quarteroni, G. Rozza, A. Manzoni. Certified Reduced Basis Approximation for Parametrized Partial Differential Equations and Applications. Journal of Mathematics in Industry 2011, 1:3.
9. G. Rozza, D.B.P Huynh, A. Manzoni. Reduced basis approximation and error bounds for Stokes flows in parametrized geometries: roles of the inf-sup stability constants. Submitted, 2010. Preprint: Technical Report MATHICSE 22.2010.

**Conference Proceedings/Paper (with review)**

10. A. Quarteroni, A. Manzoni, G. Rozza. Model order reduction by reduced basis methods and free-form deformation for shape optimization. In Schnelle Löser für Partielle Differentialgleichungen, R. Bank and W. Hackbusch (Eds.), Mathematisches Forschungsinstitut Oberwolfach, Germany, May 22-28, 2011.
11. 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*, selected papers from the ICOSAHOM '09 conference, June 22-26, Trondheim, Norway. Springer, Series: Lecture Notes in Computational Science and Engineering, vol. 76, J.S. Hesthaven, E.M. Rønquist (Eds.), pp. 307–315, 2011.
12. G. Rozza, A. Manzoni. Model order reduction by geometrical parametrization for shape optimization in computational fluid dynamics. Proceedings of ECCOMAS CFD 2010, V European Conference on Computational Fluid Dynamics, Lisbon, Portugal, J.C.F. Pereira and A. Sequeira (Eds), 2010.
13. G. Rozza, A. Manzoni, T. Lassila. Reduction strategies for shape dependent inverse problems in haemodynamics. Proceedings of 25th IFIP TC7 Conference on System Modeling and Optimization, Berlin (Germany), September 12-16, 2011. Submitted. Preprint: Technical Report MATHICSE 09.2012.