The variational multiscale method for discontinuous Galerkin type finite element formulations

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Klaas Franciscus Stein Stoter

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Prof. Dominik Schillinger

December, 2019

Abstract

In the committee of the doctoral exam have served:

↪ Prof. Dominik Schillinger, as the main advisor

↪ Prof. Henryk K. Stolarski, as the committee chair

↪ Prof. Bernardo Cockburn

↪ Prof. Thomas J.R. Hughes

↪ Prof. Sonia G. Mogilevskaya

In this dissertation, we develop the foundation for a framework that unifies variational multiscale analysis and discontinuous Galerkin type methods. We adopt the variational multiscale principles while using finite element approximation spaces that are flexible on Dirichlet boundaries, or even allowed to be completely discontinuous from element to element. We show that many classical methods follow as particular choices of scale decomposition in the variational multiscale paradigm. The methods that we derive as such are Nitsche's method for the weak enforcement of essential boundary conditions, Bassi-Rebay's 1st method, the Interior Penalty method, the Local Discontinuous Galerkin method and the Hybridizable Discontinuous Galerkin method. We derive explicitly the projection operators corresponding to each of these formulations. This illustrates that these 'non-conformal' methods and the accompanying penalty terms are in complete agreement with the variational multiscale formalism. At the same time, knowledge of these projectors guides our fine-scale modeling efforts. We develop a residual-based model that incorporates the non-vanishing fine scales at the element boundaries. Our model includes additional boundary terms with new model parameters. For all model parameters, we propose a parameter estimation strategy that is effective for both lower- and higher-order basis functions. Our numerical experiments illustrate that the classical residual-based fine-scale model exhibits overly diffusive behavior at boundaries with weakly enforced conditions. The additional terms in the new augmented model counters this defect and improves the quality of the boundary layer approximation.

Table of contents

PART III   Discontinuous Galerkin methods

List of tables

List of figures

Glossary

Abbreviations

Acronym	Full form
DG	Discontinuous Galerkin
FE	Finite element
GLS	Galerkin/least squares
HDG	Hybridizable Discontinuous Galerkin
IP	Interior Penalty
LDG	Local Discontinuous Galerkin
LES	Large eddy simulation
NIP	Nonsymmetric Interior Penalty
NURBS	Non-uniform rational B-spline
PDE	Partial differential equation
PSPG	Pressure stabilized/Petrov-Galerkin
RKDG	Runge-Kutta Discontinuous Galerkin
SUPG	Streamline upwind/Petrov-Galerkin
VMS	Variational multiscale

Subscripts, superscripts and markings

Symbol	Description	Example
$\partial\{\cdot\}$	A boundary.	$\partial\Omega$, ∂K
$\partial_n\{\cdot\}$	Normal derivative.	$\partial_n\phi$
$\{\cdot\}^+$	Outflow or zero-flux section of the boundary or solution.	$\partial\Omega^+$, ϕ^+
$\{\cdot\}^-$	Inflow section of the boundary, or the contiguous trace.	$\partial\Omega^-$, ϕ^-

$\{\cdot\}_D$	Dirichlet boundary or data.	$\partial\Omega_D,\ \phi_D$
$\{\cdot\}_N$	Neumann boundary or data.	$\partial\Omega_N,\ g_N$
$\{\cdot\}^h$	Coarse-scale (mesh dependent) function or variable.	$\phi^h,\ \mathcal{W}^h$
$\{\cdot\}'$	Fine-scale component.	$\phi',\ \mathcal{W}'$
$\bar{\{\cdot\}}$	Set closure, volume including its boundary.	$\bar{\Omega},\ \bar{K}$
$\hat{\{\cdot\}}$	Coefficient of polynomial representation, see Eq. (6.20).	$\hat{w}$
$\|\cdot\|_{\{\cdot\}}$	Norm, see Definition 2.2.	$\|\phi\|_{H^1(\Omega)}$
$(\cdot,\cdot)_{\{\cdot\}}$	Inner product, see Definition 2.4.	$(\phi,w)_{L^2(\Omega)}$
$\langle\cdot,\cdot\rangle$	Duality pairing, see Definition 2.12.	$\langle\lambda,\phi\rangle$
$[\![\cdot]\!]$	Jump operator, see Eq. (8.7).	$[\![\phi]\!],\ [\![\boldsymbol{\sigma}]\!]$
$\{\!\{\cdot\}\!\}$	Average operator, see Eq. (8.8).	$\{\!\{\phi\}\!\}$

Greek symbols

Symbol	Description	Reference
Γ	Interface skeleton, sum of element boundaries.	Eq. (9.3)
Γ_0	Internal element boundaries.	Eq. (9.5)
Γ_D	Internal element boundaries and Dirichlet boundary.	Eq. (9.5)
Γ_N	Internal element boundaries and Neumann boundary.	Eq. (9.5)
Ω	Domain of definition of the PDE.	Eq. (5.1)
$\tilde{\Omega}$	Broken domain, sum of open subsets.	Eq. (9.3)
β	Penalty (switch) parameter in LDG and HDG.	Table 8.1
δ	Dirac delta distribution.	Eq. (3.24)
η	Penalty parameter in Nitsche, IP, LDG and HDG.	Table 8.1
γ	Element averaged surface Green's function.	Eq. (6.24)
κ	Diffusion coefficient.	Eq. (5.1)
λ	Lagrange multiplier.	Eq. (4.10)
μ	Constraint functional.	Eq. (3.28)
ϕ	Primal solution variable.	Eq. (5.1)
$\boldsymbol{\sigma}$	Auxiliary solution variable.	Eq. (10.1)
τ	Element averaged volumetric Green's function.	Eq. (6.24)
$\xi(Pe)$	Upwind function.	Eq. (6.35)
$\zeta(Pe_n)$	Boundary function.	Eq. (6.39)

Latin symbols

Symbol	Description	Reference
$\mathcal{C}^n$	Set of continuous function of order n.	-
F	Element face.	Eq. (6.23)
$\mathcal{G}$	Green's operator.	Eq. (3.12)
$H^s = W^{s,2}$	Hilbertian Sobolev space.	Def. 2.7
$\mathbb{H}$	Green's function boundary operator.	Eq. (6.16)
$\mathscr{I}$	Identity operator.	Eq. (5.7)
K	Open element domain.	Eq. (3.45)
L^2	Hilbertian Lebesgue space.	Def. 2.6
$\mathcal{L}$	Differential operator.	Eq. (2.30)
$\mathcal{L}^*$	Adjoint differential operator.	Eq. (3.23)
$\mathscr{P}$	Projection operator.	Def. 2.24
Pe	Element Péclet number.	Eq. (6.35)
Pe_n	Boundary Péclet number.	Eq. (6.39)
$\mathbb{P}^p$	Space of polynomials up to order p.	Eq. (4.19)
$\mathcal{R}$	Residual.	Eq. (3.11)
$\mathbb{R}$	Set of real numbers.	-
$\mathcal{T}$	Mesh, the set of open element domains.	Eq. (9.2)
T_1, T_2	Length scales from inverse estimates.	Eq. (6.50)
$\mathcal{V}, \mathcal{W}, \mathcal{Q}$	Arbitrary function spaces.	-
$\boldsymbol{a}$	Advective velocity vector.	Eq. (5.1)
c_1, c_2, c_s	Scaling coefficients for the parameter estimations.	Eq. (6.42)
d	Spatial dimension of the domain.	Eq. (5.1)
f	Forcing/source function.	Eq. (5.1)
$\hat{f}, \hat{\boldsymbol{g}}$	Numerical fluxes.	Eq. (8.27)
$g(x,y)$	Green's function.	Eq. (3.26)
g_N	Neumann boundary data.	Eq. (5.1)
h	Element diameter.	Eq. (6.24)
$\boldsymbol{n}$	Normal vector.	Eq. (5.1)
p	Polynomial order of primal variable basis functions.	Eq. (10.10)
q	Polynomial order of auxiliary variable basis functions.	Eq. (10.10)

Thesis introduction and outline

Finite element analysis has become an indispensable tool in the engineer's toolbox. Yet, the finite element method is not infallible and it has to be used with care. We illustrate this in Figure 1.1, which shows finite element approximations of two differential equations. One is purely second-order and a first-order term has been added to the other. The boundary conditions and source functions are such that both exact solutions are the same. These differential equations are linear, one-dimensional and steady-state. They are close to the simplest conceivable problem formulations. While the finite element approximation of Figure 1.1a is nodally exact, the solution shown in Figure 1.1b is such a poor approximation that it is effectively useless. Why does this happen? The variational multiscale paradigm is a product of the answer to this question.

(a) Poisson equation: $-0.01\frac{\mathrm{d}^2}{\mathrm{d}x^2}\phi = \frac{100}{e^{30}-1}e^{100x}$.

(b) Including first-order term: $\frac{\mathrm{d}}{\mathrm{d}x}\phi - 0.01\frac{\mathrm{d}^2}{\mathrm{d}x^2}\phi = 0$.
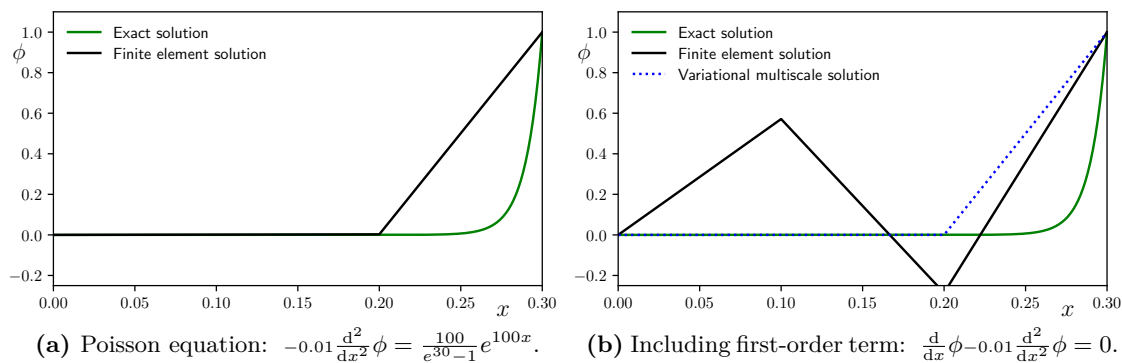
Fig. 1.1: *Exact and finite element solution of two ordinary differential equations.*

It is not that the finite element basis functions are not capable of representing the solution: the solutions of both differential equations are the same and Figure 1.1a shows a perfect result. It is also not that the weak formulation of the latter is flawed: it still defines the exact solution and if we would refine the finite element mesh we would converge to it. The answer is a combination of both these points: the necessary interpolation error of the coarse approximation is blown up due to the unstable nature[1] of the the differential equation. It is thus the lack of "fine scales" in the approximation that causes the "coarse scales" to behave erratic. This suggests a pathway for remedying the situation: the *effect* of the missing scales on the coarse-scale solution needs to be included. The variational multiscale method offers a strategy for determining exactly what that fine-scale effect has to be. For the example of Figure 1.1b this means that a slight modification of the finite element formulation allows us to retrieve the nodally exact solution shown with the dotted blue line.

1.1   Research objective

In this dissertation, we explore these ideas for non-standard finite element methods. We consider Nitsche's method of weak imposition of essential conditions, and various discontinuous Galerkin methods. Examples of computations that make use of such finite element formulations for the problem of Figure 1.1b are shown in Figure 1.2. The figure illustrates the release of the constraint on the Dirichlet boundary for Nitsche's method, and the release of all inter-element continuity for the Interior Penalty method. The constraints are weakly enforced in the finite element formulation via the addition of penalty and consistency terms on the relevant boundaries. Both these finite element approximations show an increased solution quality compared to the plain finite element approximation. Based on this increased stability, these discontinuous Galerkin type formulations have had significant impact on the application of finite element methods for fluid mechanics. Nevertheless, the solutions in Figure 1.2 are not perfect, and especially the solution of Figure 1.2a still requires some variational multiscale type stabilization. Currently, mergers between these methods and the variational multiscale method are limited to a straight-forward combination of both formulations. The goal of this dissertation is to develop a more rigorous unification.

---

[1]With the functional analysis theory from Chapter 2 we can quantify this "instability" as the ratio of "coercivity" over "boundedness" of the relevant operators.

(a) Nitsche's formulation with penalty parameter $\eta = 6$.

(b) Interior Penalty formulation with penalty parameter $\eta = 6$.
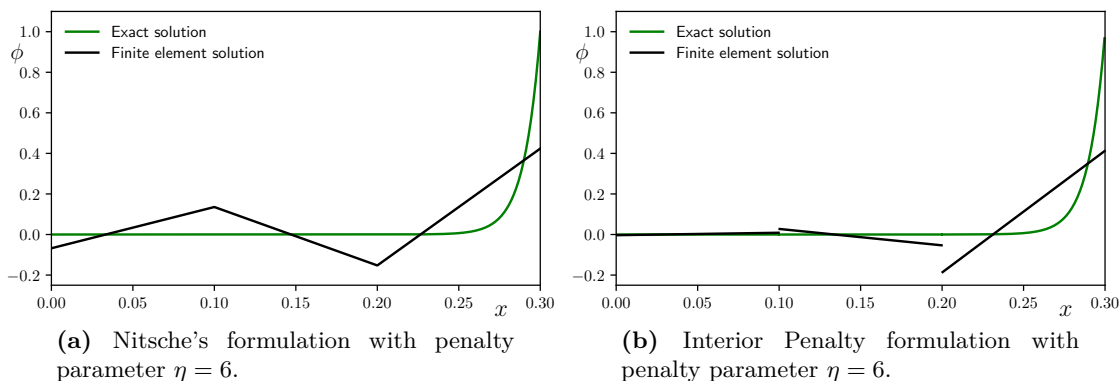
Fig. 1.2: *Illustrations of two 'non-standard' methods for the problem of Figure 1.1b.*

We have two main objectives:

↪ To make the variational multiscale paradigm applicable for discontinuous Galerkin type formulations.

↪ To devise a residual-based model for the fine scales, suitable for these finite element frameworks.

The first point focuses on theory and formalism. The second on approximation and validation by numerical experiments.

There are clear challenges involved with both objectives. The variational multiscale method is based on a decomposition of a well-posed weak formulation, but discontinuous Galerkin type methods do not conform to these formulations. Additionally, the penalty terms required to stabilize discontinuous Galerkin methods appear to be purely 'numerical', which disagrees with the essence of the variational multiscale methodology. In regard to the fine-scale modeling, the order of continuity of the approximation field has always been a delicate issue during the design of the residual-based models. A reevaluation of the model approximations is required when completely discontinuous spaces are used.

1.2   Outline

This dissertation is structured in three parts. Part I focuses on the variational multiscale method on familiar terrain: conformal finite element approximation. This part is largely literature based, although we present some novel ideas and provide our own perspective.

Parts II and III concern the core of the research. In those parts, we extend the variational multiscale ideas to non-standard finite element formulations. In Part II, we focus on the weak enforcement of essential conditions through Nitsche's method, and in Part III on the use of discontinuous Galerkin formulations. We show that these finite element formulations emerge naturally from the variational multiscale framework, and we develop suitable fine-scale models.

PART I

Conformal finite element spaces

Introduction

In this introductory part, we provide an overview of all the terminology, notation and concepts that we require in this dissertation. The core ideas of the variational multiscale method could be explained with just a few illustrative examples in Chapter 1, but to build from them a rigorous framework of analysis requires careful construction of the foundation and attention to detail. An appreciation of these details will be crucial for the extension of the variational multiscale principles to the 'non-standard' finite element methods that we focus on in Parts II and III.

Part I is structured as follows: in Chapter 2 we summarize some definitions and theorems from functional analysis that we use repeatedly. We discuss various function spaces, their duals and the pairings between them, different types of mappings, projections and space decompositions, weak formulations of partial differential equations and the conditions for their well-posedness. This is the core theory on which all the following relies. Chapter 3 is a literature review of the variational multiscale method. We discuss the decomposition of the weak statement, the inversion of the fine-scale problem and different fine-scale modeling efforts. Finally, in Chapter 4, we extend this theory to mixed formulations. There, we derive some preliminary results that are of importance when we consider discontinuous Galerkin methods.

Functional analysis background

In this chapter, we provide a brief summary of the required functional analysis background compiled from various sources [3, 27, 76, 81, 159]. It embodies the core theory on which the variational multiscale principles developed in the subsequent chapters are built: we discuss function spaces, projections, space decompositions, weak formulations of partial differential equations, and their well-posedness. Frequent reference to this chapter will be made in all the following work.

2.1 Function spaces

In the context of functional analysis, a "space" refers to a set of "elements". The notation that we will use to define a set is as follows:

$$\mathcal{S} = \{1, x \mapsto x^2, \mathbb{R}^3\}, \tag{2.1}$$

where, in this example, $\mathcal{S}$ is the set composed of the number 1, the square function $x \mapsto x^2$, and the set of all the real thee-dimensional vectors.

We are typically interested in spaces that possess a linear structure. These are called "vector spaces".

Definition 2.1 (Vector space)**.** *The set $\mathcal{V}$ is a real vector space when any linear combination of each of its entries is also a member of $\mathcal{V}$:*

$$u \in \mathcal{V}, v \in \mathcal{V} \to a\,u + b\,v \in \mathcal{V} \quad \forall\, a \in \mathbb{R}, b \in \mathbb{R}\,. \tag{2.2}$$

The $\mathcal{C}^n(\Omega)$ family of function spaces may be an intuitive example. These are the sets of all functions defined on Ω such that their n-th derivative is continuous. $\mathcal{C}^0(\Omega)$ functions have no continuous derivatives, and $\mathcal{C}^\infty(\Omega)$ functions are infinitely differentiable. Since differentiation is a linear operation, any weighted combination of a function in $\mathcal{C}^n(\Omega)$ is still in $\mathcal{C}^n(\Omega)$. Hence these are vector spaces. These are natural spaces to consider when dealing with the classical forms of partial differential equations.

In the analysis of weak formulations of partial differential equations (to be specified later), we do not only require our spaces to satisfy this linearity requirement, we also require them to be "complete". Completeness means that the limit of any converging sequence (a Cauchy sequence) of elements of $\mathcal{V}$ is also in $\mathcal{V}$. This notion requires a measure to define convergence, i.e., some definition of the magnitude of the (difference between the) members of $\mathcal{V}$. We call this measure of magnitude a "norm":

Definition 2.2 (Norm)**.** *An operator that acts on the elements of a set and produces a scalar:*

$$\begin{aligned} \|\cdot\| : \mathcal{V} &\to \mathbb{R} \\ u &\mapsto \|u\|\,, \end{aligned} \tag{2.3}$$

while satisfying the following properties:

$$\|u\| = 0 \Leftrightarrow u = 0\,, \tag{2.4a}$$

$$\|au\| = |a|\|u\| \qquad \forall\, u \in \mathcal{V}, a \in \mathbb{R}\,, \tag{2.4b}$$

$$\|u + v\| \le \|u\| + \|v\| \quad \forall\, u \in \mathcal{V}, v \in \mathcal{V}\,. \tag{2.4c}$$

With this definition of 'magnitude' at hand, we call a complete vector space a Banach space:

Definition 2.3 (Banach space)**.** *A vector space endowed with a norm, that is complete with respect to that norm.*

Often, it is not only useful to have a measure for the magnitude of the elements of $\mathcal{V}$, but also a measure for the interrelation between these elements. For instance, we wish to be able to establish concepts such as orthogonality. We do so by means of an inner product:

Definition 2.4 (Inner product). *An operator that takes two entries of a set and produces a scalar:*

$$
\begin{aligned}
(\cdot, \cdot) : \mathcal{V} \times \mathcal{V} &\to \mathbb{R} \\
u, v &\mapsto (u, v),
\end{aligned}
\tag{2.5}
$$

while satisfying the following properties:

$$
(au + v, w) = a\,(u, w) + (v, w) \quad \forall\, u, v, w \in \mathcal{V} \times \mathcal{V} \times \mathcal{V} \ \text{ and } \ \forall\, a \in \mathbb{R}, \tag{2.6a}
$$

$$
(u, v) = (v, u) \qquad\qquad\qquad \forall\, u, v \in \mathcal{V} \times \mathcal{V}, \tag{2.6b}
$$

$$
(u, u) = \|u\|^2 \qquad\qquad\qquad \forall\, u \in \mathcal{V}. \tag{2.6c}
$$

So, the inner product of $u \in \mathcal{V}$ with itself represents the square of its magnitude, i.e., the square of its norm.

This new construct allows us to define another type of vector space, namely the Hilbert space:

Definition 2.5 (Hilbert space). *A Banach space where its norm is induced by an inner product.*

Vector, Banach and Hilbert spaces are quite general. They apply to sets of various types of objects. As we are interested in solving partial differential equations, we concern ourselves with spaces of functions. The functions that we consider require different types of regularities to ensure that the different types of differential operators act on them in meaningful ways. This leads to the definition of the Lebesgue and Sobolev spaces.

Definition 2.6 (Lebesgue space). *The Lebesgue spaces $L^p(\Omega)$ (where we only consider $1 \leq p \leq \infty$) are Banach spaces of functions defined on domain Ω which are p-power integrable:*

$$
L^p(\Omega) = \{ u(\Omega) : \left(\int_\Omega |u|^p \right)^{1/p} < \infty \}. \tag{2.7}
$$

This space is endowed with a norm, under which it is complete:

$$\|u\|_{L^p(\Omega)} = \left(\int_\Omega |u|^p \right)^{1/p}. \tag{2.8}$$

For $p = 2$, this norm is induced by the following inner product, making $L^2(\Omega)$ a Hilbert space:

$$(u, v)_{L^p(\Omega)} = \int_\Omega u\, v. \tag{2.9}$$

Definition 2.7 (Sobolev space)**.** *The Sobolev spaces $W^{s,p}(\Omega) \subset L^p(\Omega)$ (where we only consider $s \geq 0$ and $1 \leq p \leq \infty$) are Banach spaces with the following regularity requirements:*

$$W^{s,p}(\Omega) = \left\{ u \in L^p(\Omega) : \frac{\partial^{\|\boldsymbol{\alpha}\|_1} u}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} \in L^p(\Omega) \ \forall\, \boldsymbol{\alpha} \in \mathbb{N}_0^d \ with \ \|\boldsymbol{\alpha}\|_1 \leq s \right\}, \tag{2.10}$$

where d is the spatial dimension of Ω (i.e., $\Omega \subset \mathbb{R}^d$), $\mathbb{N}_0^d$ is the set of zero and the natural numbers, and $\|\cdot\|_1$ is the 1-norm, defined as the sum of the absolute values of the vector.

These (partial) derivatives should be understood in a weak sense.

Definition 2.8 (Weak derivative)**.** *The weak derivatives of a function are defined in the sense of their action on a test function, and can be determined via integration by parts rules:*

$$\int_\Omega \frac{\partial^{|\boldsymbol{\alpha}|_1} u}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} v = (-1)^{|\boldsymbol{\alpha}|_1} \int_\Omega u\, \frac{\partial^{|\boldsymbol{\alpha}|_1} v}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} \qquad \forall\, v \in C_0^\infty(\Omega), \tag{2.11}$$

where $C_0^\infty(\Omega)$ is the space of infinitely (strongly) differentiable functions that are zero on the boundary of Ω.

The weak derivative is defined for a much larger class of function than the strong derivative. Indeed, only in the sense of weak derivatives do the Sobolev spaces form a complete space.

The norm associated with this completeness is:

$$\|u\|_{W^{s,p}(\Omega)} = \sum_{|\boldsymbol{\alpha}|_1 \le s} \|\frac{\partial^{|\boldsymbol{\alpha}|_1} u}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}\|_{L^p(\Omega)} . \tag{2.12}$$

Just like for the Lebesgue spaces this norm is induced by the following inner product in the special case of $p = 2$:

$$(u, v)_{W^{s,p}(\Omega)} = \sum_{|\boldsymbol{\alpha}|_1 \le s} \int_{\Omega} \frac{\partial^{|\boldsymbol{\alpha}|_1} u}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} \frac{\partial^{|\boldsymbol{\alpha}|_1} v}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} . \tag{2.13}$$

This makes the $W^{s,2}(\Omega)$ spaces Hilbert spaces, which we henceforth refer to as $H^s(\Omega)$. Other spaces that can be relevant in the context of partial differential equations are:

$$H(\text{div}; \Omega) := \{\boldsymbol{u} \in [L^2(\Omega)]^d : \nabla \cdot \boldsymbol{u} \in L^2(\Omega)\}, \tag{2.14}$$

and

$$H(\text{curl}; \Omega) := \{\boldsymbol{u} \in [L^2(\Omega)]^d : \nabla \times \boldsymbol{u} \in [L^2(\Omega)]^d\}. \tag{2.15}$$

So far, all these function spaces have the (open) domain $\Omega \subset \mathbb{R}^d$ associated to them. The functions in these spaces are required to display a certain regularity in this d-dimensional domain, defined through integrability. As a result, they are typically undefined on $(d-1)$-dimensional sets (which have a "Lebesgue measure" of zero). This is somewhat problematic as we often need to manipulate or prescribe quantities on the $(d-1)$-dimensional boundary of Ω, denoted $\partial\Omega$. If the boundary is sufficiently smooth (which we assume in all the following) then it turns out that there exists an *unambiguous* extension of some $u \in H^1(\Omega)$ to a *unique* function $u|_{\partial\Omega} \in L^2(\partial\Omega)$. The operator that takes u and produces its "trace" $u|_{\partial\Omega}$ is called the "trace operator":

Definition 2.9 (Trace operator). *The unique bounded linear operator:*

$$T : H^1(\Omega) \to L^2(\partial\Omega)$$
$$u \quad \mapsto Tu = u|_{\partial\Omega}, \tag{2.16}$$

that defines the trace of u on $\partial\Omega$ in such a way that it returns the restriction of u on $\partial\Omega$ when u it is a member of $\mathcal{C}^1(\bar{\Omega})$.

As implied by this definition, functions in $L^2(\Omega)$ do not have well-defined traces on $\partial\Omega$. Similarly, for functions in $H^1(\Omega)$ the derivatives can not be uniquely defined on the boundary. This will be a recurring limitation that has to be dealt with in this dissertation. On the other hand, both the functions and the derivatives of functions in $H^2(\Omega)$ are elements of $H^1(\Omega)$, such that we may determine their trace and the traces of their derivatives. We then define the space of functions $H^{s-\frac{1}{2}}(\partial\Omega) \subset L^2(\partial\Omega)$ as all the possible traces of functions in $H^s(\Omega)$ for $s \geq 1$. Subsequently, we may define the subspaces $H_0^s(\Omega) \subset H^s(\Omega)$ as all functions with zero traces and zero traces of their derivatives.

Now that we have defined various function spaces, we may ask ourselves which objects can act on the functions in these spaces. The just described trace operator is an example, and so are differential operators. Both of these may be interpreted as mappings between Sobolev spaces. Similarly, we can define functionals as mappings to the real number line:

Definition 2.10 (Linear functional). *A mapping from a vector space to $\mathbb{R}$:*

$$\ell : \mathcal{V} \to \mathbb{R}$$
$$u \mapsto \ell(u) \, ,$$

(2.17)

that is linear:

$$\ell(au + v) = a\ell(u) + \ell(v) \qquad \forall\, u, v \in \mathcal{V} \times \mathcal{V} \ and \ \forall\, a \in \mathbb{R} \, .$$

(2.18)

Again, we can collect functionals into sets. Of relevance to the Lebesgue and the Sobolev spaces are sets of all possible linear functionals that may act on all functions in $L^p(\Omega)$ or $H_0^s(\Omega)$ to produce a real number. These sets of functionals are the so-called dual spaces of $L^p(\Omega)$ or $H_0^s(\Omega)$.

Definition 2.11 (Dual space). *The dual space of a vector space $\mathcal{V}$ is the set of all bounded linear functionals that can act on each of the members of $\mathcal{V}$.*

We will typically denote the dual space of some arbitrary set $\mathcal{V}$ as $\mathcal{V}^*$ (in most literature it is denoted with a prime but we have reserved this symbol for fine-scale quantities), although in the particular case of $H_0^s(\Omega)$ we write them as $H^{-s}(\Omega)$. We may then connect functionals in dual spaces with functions in the respective space. This is what we call a duality pairing.

Definition 2.12 (Duality pairing)**.** *The action of some functional in $\mathcal{V}^*$ on some function in $\mathcal{V}$. We denote this as $_{\mathcal{V}^*}\langle \ell, v \rangle_{\mathcal{V}}$. The result of this operation is a scalar.*

The writing notation for duality pairings looks suspiciously much like that of an inner product. Indeed, an important theorem in functional analysis makes a connection between duality pairings and inner products. This is the Riesz-representation theorem.

Theorem 2.1 (Riesz-representation)**.** *If $\mathcal{V}$ is a Hilbert space with norm $(\cdot, \cdot)_{\mathcal{V}}$, then for every functional $\ell \in \mathcal{V}^*$ there exists a unique function $k \in \mathcal{V}$ such that:*

$$\ell(v) = {}_{\mathcal{V}^*}\langle \ell, v \rangle_{\mathcal{V}} =: (k, v)_{\mathcal{V}} \qquad \forall\, v \in \mathcal{V}\,. \tag{2.19}$$

This theorem permits us to write duality pairings that involve Hilbert spaces simply as $\langle \ell, v \rangle$ which reduces the notational clutter.

2.2 Mappings, projections and decompositions

In the previous section, the term "mapping" was already used, and indeed *norms*, *inner products*, the *trace operator* and *functionals* are all examples of mappings. In this section, we collect some nomenclature used to describe different types of mappings. Then we state some definitions and basic results for the particular mappings that are central to variational multiscale analysis: projections.

A mapping is quite general: it is some operation that takes an element from some space and produces (or "maps to") an element in some other space. The input space is called its domain, and the output space its codomain.

Definition 2.13 (Domain and codomain)**.** *For the mapping $\mathscr{M} : \mathcal{V} \to \mathcal{W}$, the space $\mathcal{V}$ is its domain, and $\mathcal{W}$ its codomain.*

There is no guarantee, however, that every element in the codomain is being mapped to, nor that every element in the domain maps to a nontrivial (non-zero) element. This calls for the definitions of the range (or image) and the kernel (or nullspace) of a mapping. Figure 2.1 illustrates the meanings of domain, codomain, range and nullspace.

Definition 2.14 (Range or image)**.** *The range of $\mathscr{M} : \mathcal{V} \to \mathcal{W}$ is the set of elements $w \in \mathcal{W}$ for which there exists $v \in \mathcal{V}$ such that $\mathscr{M} v = w$. This set is denoted $\mathrm{ran}(\mathscr{M})$.*
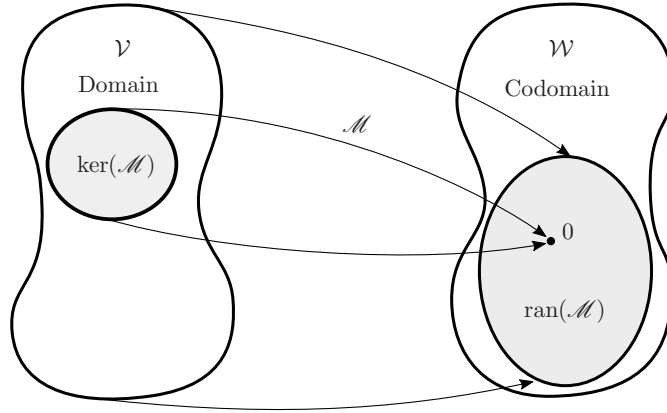
Fig. 2.1: *Illustration of basic mapping terminology.*

Definition 2.15 (Kernel or nullspace). *The kernel of $\mathscr{M} : \mathcal{V} \to \mathcal{W}$ is the set of elements $v \in \mathcal{V}$ such that $\mathscr{M}v = 0 \in \mathcal{W}$. This set is denoted* $\ker(\mathscr{M})$.

Clearly, the definition of the nullspace requires that there is some element in $\mathcal{W}$ that can be identified as the null-element. If $\mathcal{W}$ is a Banach space, then this element corresponds to the $\|w\|_{\mathcal{W}} = 0 \Leftrightarrow w = 0$ requirement.

With the foundation of mappings laid out, we can define special kinds of mappings.

Definition 2.16 (Injection). *A mapping $\mathscr{M} : \mathcal{V} \to \mathcal{W}$ is injective (or "one-to-one") when $\mathscr{M}u = \mathscr{M}v \Rightarrow u = v \, \forall u \in \mathcal{V}, v \in \mathcal{V}$ (i.e., every element in $\mathcal{V}$ maps to a unique element in $\mathscr{M}$). This mapping is said to map "into" $\mathcal{W}$.*

Definition 2.17 (Surjection). *A mapping $\mathscr{M} : \mathcal{V} \to \mathcal{W}$ is surjective when $\forall w \in \mathcal{W} \, \exists u \in \mathcal{V} : \mathscr{M}u = w$ (i.e., every element in $\mathcal{W}$ is being mapped to by at least one element in $\mathcal{V}$). This mapping is said to map "onto" $\mathcal{W}$.*

Definition 2.18 (Bijection). *A mapping $\mathscr{M} : \mathcal{V} \to \mathcal{W}$ is bijective when it is both injective and surjective (i.e., there is a "one-to-one correspondence" between each element in $\mathcal{V}$ and each element in $\mathcal{W}$).*

The concepts of injective, surjective and bijective mappings are illustrated in Figure 2.2.

(a) Injective but not surjective.

(b) Surjective but not injective.
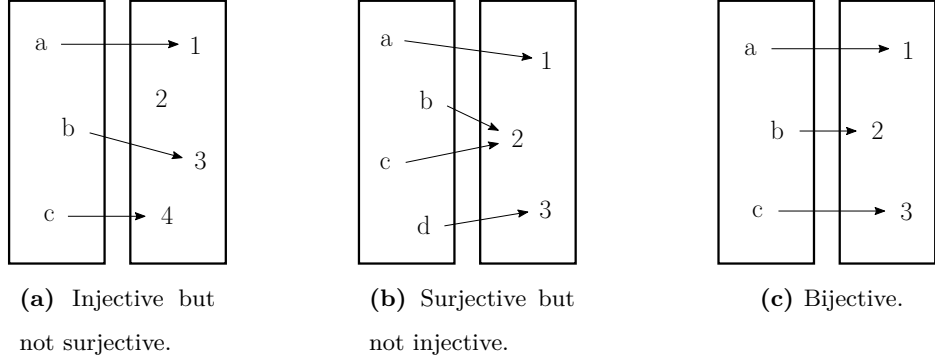
(c) Bijective.

Fig. 2.2: Three classes of mappings: injective, surjective and bijective.

So far, the spaces in relation to our mappings have not been categorized. They may be arbitrary sets. If we map between Banach spaces, where the concept of 'magnitude' is meaningful, some useful properties may be inferred. Let us define a neighborhood around v_0 with 'radius' d_v as $V_0 = \{v \in \mathcal{V} : \|v_0 - v\|_{\mathcal{V}} < d_v\}$. Similarly we define a neighborhood around $w_0 = \mathcal{M}v_0$ with radius d_w as $W_0 = \{w \in \mathcal{W} : \|w_0 - w\|_{\mathcal{W}} < d_w\}$. We can then define continuous mappings, bounded mappings and coercive mappings:

Definition 2.19 (Continuous mapping)**.** *Let $\mathcal{M} : \mathcal{V} \to \mathcal{W}$ map between normed vector spaces. $\mathcal{M}$ is continuous if for any output radius d_w around $w_0 = \mathcal{M}v_0$ we can find an input radius d_v such that $\mathcal{M}v \in W_0 \ \forall v \in V_0$. That is, the mapping of the subdomain V_0 falls completely in the region W_0.*

Definition 2.20 (Bounded mapping)**.** *Let $\mathcal{M} : \mathcal{V} \to \mathcal{W}$ be a linear mapping between normed vector spaces. $\mathcal{M}$ is bounded if there exists some $C > 0$ such that:*

$$\|\mathcal{M}v\|_{\mathcal{W}} \leq C \|v\|_{\mathcal{V}} \qquad \forall v \in \mathcal{V}. \tag{2.20}$$

A linear operator is bounded if and only if it is continuous. The smallest C for which Eq. (2.20) holds is called the boundedness coefficient. This boundedness coefficient satisfies all the norm-axioms (see Definition 2.2) and may be identified as the norm of the operator:

Definition 2.21 (Operator norm)**.** *The norm of a bounded operator is:*

$$\|\mathcal{M}\| := \sup_{v \in \mathcal{V} \setminus \{0\}} \frac{\|\mathcal{M}v\|_{\mathcal{W}}}{\|v\|_{\mathcal{V}}}. \tag{2.21}$$

15

Boundedness, and the resulting definition of the operator norm, concerns the upper limit to the 'scaling' induced by the mapping. In a similar sense, mappings can have a lower limit to this magnitude change. A mapping with this property is said to be "coercive":

Definition 2.22 (Coercive mapping). *Let $\mathscr{M} : \mathcal{V} \to \mathcal{W}$ be a linear mapping between normed vector spaces. $\mathscr{M}$ is coercive if there exists some $C > 0$ such that:*

$$\|\mathscr{M}v\|_{\mathcal{W}} \geq C \|v\|_{\mathcal{V}} \qquad \forall\, v \in \mathcal{V}. \tag{2.22}$$

The largest C for which Eq. (2.22) holds is called the coercivity constant.

We can define mappings where the domain and codomain are the same set. This means that we can apply the mapping recursively. If the result does not change under this recursive operation, then we call this mapping "idempotent".

Definition 2.23 (Idempotency). *An idempotent mapping $\mathscr{M} : \mathcal{V} \to \mathcal{V}$ satisfies:*

$$\mathscr{M}(\mathscr{M}v) = \mathscr{M}v \qquad \forall\, v \in \mathcal{V}. \tag{2.23}$$

Finally, we are in the position to define a "projector".

Definition 2.24 (Projector). *An idempotent mapping $\mathscr{P} : \mathcal{V} \to \mathcal{V}$ between vector spaces.*

Often, we will write a projector as $\mathscr{P} : \mathcal{V} \to \mathcal{W}$ instead, where $\mathcal{W} \subset \mathcal{V}$. It is then implied that $\mathcal{W}$ is its range, meaning that $\mathscr{P}$ is also a surjective mapping. In all the following we assume that $\mathcal{W}$ must contain the zero element. This is necessarily true for linear projections, which are the ones that we will use.

The inner product was introduced to establish the notion of orthogonality among the elements in Hilbert spaces. Two functions u and v are orthogonal in some inner product when $(v, u) = 0$. We can now design projectors that map $v \in \mathcal{V}$ orthogonally onto $\mathcal{W}$, in the sense that the remainder (the difference between v and $\mathscr{P}v$) is orthogonal to all $w \in \mathcal{W}$.

Definition 2.25 (Orthogonal projection). *A projection $\mathscr{P} : \mathcal{V} \to \mathcal{W}$ that satisfies:*

$$(v - \mathscr{P}v, w) = 0 \quad \forall\, w \in \mathcal{W}, \tag{2.24}$$

for some inner product $(\cdot, \cdot)$.

Straightforward manipulation of Eq. (2.24) gives:

$$\left(v - \mathscr{P}v, \mathscr{P}v\right) - \left(v - \mathscr{P}v, w\right) = \left(v - \mathscr{P}v, \mathscr{P}v - w\right) = 0\,, \tag{2.25}$$

which may be used to write:

$$
\begin{aligned}
\|v - \mathscr{P}v\|^2 &\le \|v - \mathscr{P}v\|^2 + 2\left(v - \mathscr{P}v, \mathscr{P}v - w\right) + \|\mathscr{P}v - w\|^2 \\
&= \|v - \mathscr{P}v + \mathscr{P}v - w\|^2 \\
&= \|v - w\|^2 \quad \forall\, w \in \mathcal{W}\,.
\end{aligned}
\tag{2.26}
$$

This statement can be written alternatively as:

$$\|v - \mathscr{P}v\| = \inf_{w \in \mathcal{W}} \|v - w\|\,, \tag{2.27}$$

which says that $\mathscr{P}v$ is the solution $w \in \mathcal{W}$ that is the optimal approximation in the norm induced by the inner product. For example, the L^2-projection, denoted $\mathscr{P}_{L^2}$, is the orthogonal projection where Eq. (2.24) concerns the L^2-inner product and which results in $\mathscr{P}_{L^2}v$ that is the best approximation to v in the L^2-norm. Similarly, $\mathscr{P}_{H_0^1}$ minimizes the error in the H_0^1-norm, which is defined as the H^1-seminorm (as this is an equivalent norm in H_0^1).

Given a projector, we can decompose some element $v \in \mathcal{V}$ into two components: one in the range of the projector $\bar{v} := \mathscr{P}v$, and a supplement $v' := v - \bar{v}$. Due to the idempotency property of the projector this supplement satisfies $\mathscr{P}v' = \mathscr{P}v - \mathscr{P}\mathscr{P}v = \mathscr{P}v - \mathscr{P}v = 0$, meaning that it lies in the kernel of $\mathscr{P}$. Any $v \in \mathcal{V}$ may thus *uniquely* be decomposed as $v = \bar{v} + v'$ with $\bar{v} \in \operatorname{ran}(\mathscr{P})$ and $v' \in \operatorname{ker}(\mathscr{P})$. The spaces $\operatorname{ran}(\mathscr{P})$ and $\operatorname{ker}(\mathscr{P})$ form a direct sum decomposition of $\mathcal{V} = \operatorname{ran}(\mathscr{P}) \oplus \operatorname{ker}(\mathscr{P})$:

Definition 2.26 (Direct sum decomposition)**.** *The decomposition of a space $\mathcal{V}$ into two subspaces $\mathcal{W}_1 \subset \mathcal{V}$ and $\mathcal{W}_2 \subset \mathcal{V}$ such that any $v \in \mathcal{V}$ may be uniquely written as $v = w_1 + w_2$ where $w_1 \in \mathcal{W}_1$ and $w_2 \in \mathcal{W}_2$. We then write:*

$$\mathcal{V} = \mathcal{W}_1 \oplus \mathcal{W}_2\,. \tag{2.28}$$

Equivalently, two spaces $\mathcal{W}_1$ and $\mathcal{W}_2$ form a direct sum decomposition of $\mathcal{V}$ if and only if:

$$\mathcal{V} = \mathcal{W}_1 \cup \mathcal{W}_2 \,, \tag{2.29a}$$

$$\mathcal{W}_1 \cap \mathcal{W}_2 = \{0\} \,. \tag{2.29b}$$

Remark 2.1. *It will prove important to realize that the direct sum decomposition* $\mathcal{V} = \mathcal{W}_1 \oplus \mathcal{W}_2$ *is not defined purely by* $\mathcal{V}$ *and* $\mathcal{W}_1$. *Given those two spaces there are many complements* $\mathcal{W}_2$ *that satisfy the decomposition. For instance,* $\mathcal{W}_2$ *may be the kernel of any projector with* $\mathrm{ran}(\mathscr{P}) = \mathcal{W}_1$. *Definition 2.25 illustrates that such a projector is by no means unique.*

2.3 Partial differential equations and weak formulations

Partial differential equations (PDEs) may be cast in the framework of the previous sections in a number of ways. Consider the following abstract form of a partial differential equation:

$$\mathcal{L}u = f \qquad \text{in } \Omega \,, \tag{2.30a}$$

$$\text{Constraints} \,, \tag{2.30b}$$

where $\mathcal{L}$ is some differential operator acting on the function u in such a way that it produces the data f, while u satisfies the imposed constraints. We can interpret the operator $\mathcal{L}$ as a mapping. The type of data that we wish to consider represents the codomain of the mapping. The domain from which $\mathcal{L}$ maps depends on the nature of the differential operator as well as the constraints. For example, if $\mathcal{L}$ represents the Laplace operator and if we only permit sufficiently smooth solutions, then we may interpret $\mathcal{L}$ as the mapping:

$$\mathcal{L} = \Delta : \{v \in \mathcal{C}^2(\Omega) : v \text{ satisfies the constraints}\} \to \mathcal{C}^0(\Omega) \,. \tag{2.31}$$

The PDE then poses the question, "which permissible solution u corresponds to the data f?" This question revolves around the bijectivity of the operator $\mathcal{L}$. As the domain of $\mathcal{L}$ is considered to be $\mathcal{C}^2(\Omega)$, Eqs. (2.30) and (2.31) concerns "classical solutions" of the partial differential equation. This means that the differential operator is required to act on the solution in a meaningful way *pointwise*.

It turns out that problem statements of the form of Eqs. (2.30) and (2.31) are unnecessarily restrictive and, moreover, tentatively difficult to analyze. Instead, we may consider the PDE and its solutions in a weak sense. Weak solutions are only required to have meaning if they are 'tested' against other (more regular) functions. So, u is a weak solution of Eq. (2.30) if (and only if):

$$\int_\Omega (\mathcal{L}u - f)\, w = 0 \qquad \forall\, w \in \mathcal{C}_0^\infty(\Omega)\,. \tag{2.32}$$

Any classical solution is also a weak solution, but more solutions are possible since we are no longer restricted to $\mathcal{C}^n(\Omega)$ functions. Given this notion, we can reformulate Eq. (2.30) in a Sobolev space setting via the method of weighted residuals. We wish to find u such that:

$$\int_\Omega \mathcal{L}u\, w = \int_\Omega f\, w \qquad \forall\, w \in L^2(\Omega)\,. \tag{2.33}$$

Since the test functions have meaning under the operation of square integration, we are now permitted to consider $f \in L^2(\Omega)$ which is a superset of the space $\mathcal{C}^0(\Omega)$. Similarly, the differential operator can now make use of the Sobolev spaces of Definition 2.7. These are the suitable function spaces as their regularity conditions revolve around weak derivatives (see Definition 2.8). We define a new $\mathcal{L}$ as:

$$\mathcal{L} = \Delta : \{v \in H^2(\Omega) : v \text{ satisfies the constraints}\} \to L^2(\Omega)\,. \tag{2.34}$$

A solution of Eqs. (2.33) and (2.34) is called a "strong solution". Any strong solution satisfies Eq. (2.32) and, by that definition, is thus also a weak solution. Additionally, when $f \in \mathcal{C}^0(\Omega)$ the problem statement of Eqs. (2.30) and (2.34) is equivalent to that of Eqs. (2.30) and (2.31) and, under the assumption of unique solvability, we would retrieve a classical solution.

Equation (2.32) implies a different interpretation of differential operators. Since $\mathcal{L}u$ only has meaning when it acts on w we may define it in a functional sense. As it acts on functions in a specific space, namely L^2, we can say that $\mathcal{L}u$ lies in the dual space of L^2, with the duality pairing:

$$_{L^{2*}}\langle \mathcal{L}u, w \rangle_{L^2} = \int_\Omega \mathcal{L}u\, w\,. \tag{2.35}$$

In this setting, the differential operator no longer maps from a function space to another

function space, but instead from a function space to the dual of a function space. For the example of the Laplace operator this reads:

$$\mathcal{L} = \Delta : \{v \in H^2(\Omega) : v \text{ satisfies the constraints}\} \to L^{2*}. \tag{2.36}$$

Continuing along this line, we may formulate the problem in an even weaker sense. Sticking to the example of the Laplace operator, we still interpret $\mathcal{L}$ to be a mapping to a dual space. However, we now *define* that dual operator as:

$$\begin{aligned}
\mathcal{L} = \Delta : H_0^1(\Omega, \partial\Omega_D) &\to H^{-1}(\Omega, \partial\Omega_D) \\
u &\mapsto -(\nabla u, \nabla \cdot)_\Omega =: B(\cdot, u),
\end{aligned} \tag{2.37}$$

where, from now on and unless specified differently, the inner products are L^2-inner products on the indicated domains. With the use of Definition 2.9 we may identify $H_0^1(\Omega, \partial\Omega_D)$ as the functions in $H^1(\Omega)$ whose trace is zero on $\partial\Omega_D \subset \partial\Omega$. The dual of that space is denoted $H^{-1}(\Omega, \partial\Omega_D)$.

In order to make use of this formulation for $\mathcal{L}$ in Eq. (2.30), we have to change the right-hand side to incorporate the constraints. For the Laplace operator, we consider "essential" and "natural" boundary conditions. The essential boundary conditions are Dirichlet conditions, which specify the value of the trace of the solution on $\partial\Omega_D$. Suppose that we find *some* function u_d in $H^1(\Omega)$ that satisfies the Dirichlet constraints (which we assume exists, else the problem formulation is flawed), then we may decompose the solution to the PDE as $u = u_d + u_0$. The new unknown function u_0 is zero on the Dirichlet boundary, and is thus a member of $H_0^1(\Omega, \partial\Omega_D)$. The new right-hand side becomes:

$$\mathcal{F} = (f, \cdot)_\Omega - (g_N, \cdot)_{\partial\Omega_N} + (\nabla u_d, \nabla \cdot)_\Omega =: L(\cdot) \quad \in H^{-1}(\Omega, \partial\Omega_D), \tag{2.38}$$

where g_N is the data corresponding to the natural boundary condition. In the current case this is a Neumann condition, which specifies the normal derivative of the solution on $\partial\Omega_N$.

With these definitions of the duality pairings of $\mathcal{L}u_0$ and $\mathcal{F}$, the PDE may be written in three equivalent ways:

$$\text{Find } u_0 \text{ in } H_0^1(\Omega, \partial\Omega_D) \text{ such that: } \mathcal{L}u_0 = \mathcal{F}, \tag{2.39}$$

which will be referred to as the "operator form". Or:

$$\text{Find } u_0 \text{ in } H_0^1(\Omega, \partial\Omega_D) \text{ such that: } \langle \mathcal{L}u_0, w \rangle = \langle \mathcal{F}, w \rangle \quad \forall\, w \in H_0^1(\Omega, \partial\Omega_D), \quad (2.40)$$

or, purely after a change of notation:

$$\text{Find } u_0 \text{ in } H_0^1(\Omega, \partial\Omega_D) \text{ such that: } B(w, u_0) = L(w) \quad \forall\, w \in H_0^1(\Omega, \partial\Omega_D), \quad (2.41)$$

which is the most common notation when it comes to finite element methods. All these formulations are analogous. A solution of Eq. (2.39), (2.40) or (2.41) is called *the* weak solution. Naturally such a solution is also a weak solution according to the definition of Eq. (2.32). All classical solutions and all strong solution also satisfy the above relations. Whether a solution *exists* for all of the possible data, and whether or not this solution is *unique*, is discussed in the following section.

2.4   Linear and bilinear forms

Let us proceed with a generalized form of Eq. (2.41):

$$\text{Find } u \in \mathcal{V} \text{ s.t. } \forall\, w \in \mathcal{V}:$$
$$B(w, u) = L(w), \quad (2.42)$$

where $L : \mathcal{V} \to \mathbb{R}$ is a "linear form", a linear functional operating on elements of the Banach space $\mathcal{V}$, and $B : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ is a "bilinear form".

Fundamentally, this problem formulation is not restricted to partial differential equations and function spaces. It merely concerns (bi)linear mappings acting on Banach spaces. The definitions concerning mappings from Section 2.2 thus also apply to these mappings. For instance, we can define bounded and/or coercive (bi)linear forms. Recall that for linear functionals boundedness implies continuity and vice versa. So, the linear form is bounded and continuous if there is some positive constant α_L such that $L(w) \le \alpha_L \|w\|_\mathcal{V} \ \forall\, w \in \mathcal{V}$, where the smallest α_L that satisfies this inequality is defined to be the norm of $L(\cdot)$. The bilinear form is bounded and continuous if there is some positive constant α_B such that $B(w, u) \le \alpha_B \|w\|_\mathcal{V} \|u\|_\mathcal{V} \ \forall\, w, u \in \mathcal{V} \times \mathcal{V}$. Additionally, $B(\cdot, \cdot)$ is coercive if there exists a constant α such that $B(w, w) \ge \alpha \|w\|_\mathcal{V}^2 \ \forall\, w \in \mathcal{V}$. This is sometimes referred to as "$\mathcal{V}$-ellipticity" or "positive definiteness" of $B(\cdot, \cdot)$.

Coercivity and boundedness are important properties as they ensure that the problem of Eq. (2.42) is well-behaved. Here, 'well-behaved' means that some small perturbation in the data (right-hand side) does not cause major changes the solution. Indeed, given coercivity of $B(\cdot,\cdot)$ and boundedness of $L(\cdot)$ we directly obtain:

$$\alpha\|u\|_{\mathcal{V}}^2 \leq B(u,u) = L(u) \leq \|L\|\,\|u\|_{\mathcal{V}} \quad \Rightarrow \quad \|u\|_{\mathcal{V}} \leq \frac{1}{\alpha}\|L\|\,, \tag{2.43}$$

which relates the magnitude of the solution u to the magnitude of the data $L(\cdot)$. A larger coercivity constant means more stable solution behavior.

Additionally, the coercivity and boundedness of $B(\cdot,\cdot)$ and $L(\cdot)$ guarantee solvability of Eq. (2.42). This is captured in the famous Lax-Milgram theorem.

Theorem 2.2 (Lax-Milgram)**.** *Consider a problem of the form:*

Find $u \in \mathcal{V}$ s.t. $\forall\, w \in \mathcal{V}$:

$$B(w,u) = L(w)\,. \tag{2.44}$$

A unique solution u exists if $B(\cdot,\cdot)$ and $L(\cdot)$ are continuous and if the bilinear form is coercive.

In many cases, the Lax-Milgram theorem is the right tool to prove existence and uniqueness of a (finite element) solution. Coercivity carries over to subspaces, such that coercivity of the continuous bilinear form also guarantees solvability of the conformal finite element formulation. However, there are also cases where the theorem does not suffice. For example, some variational problems result in $B(\cdot,\cdot) : \mathcal{V}\times\mathcal{W} \to \mathbb{R}$ (so, of "Petrov-Galerkin" type), which are not captured by Theorem 2.2. Additionally, this theorem says "if" and not "if and only if". Indeed, many problems in physics and engineering lead to non-coercive bilinear forms. Of special importance are constrained minimization principles, which lead to problems of the form:

Find $u,\lambda \in \mathcal{V} \times \mathcal{Q}$ s.t. $\forall\, w,\mu \in \mathcal{V} \times \mathcal{Q}$:
$$B\big((w,\mu),(u,\lambda)\big) = a(w,u) + b(w,\lambda) + b(u,\mu) = L(w)\,. \tag{2.45}$$

This formulation can be shown to be non-coercive with the simple counter example $B((0,\mu),(0,\mu)) = 0 \;\forall\, \mu \in \mathcal{Q}$. To address these issues, we require a more general and stronger condition. Following the naming convention proposed in [76], we refer to this as

the Banach-Nečas-Babuška condition, or BNB condition for short:

Theorem 2.3 (Banach-Nečas-Babuška). *Consider a problem of the form:*

Find $u \in \mathcal{V}$ s.t. $\forall\, w \in \mathcal{W}$:

$$B(w, u) = L(w)\,, \tag{2.46}$$

where $B(\cdot, \cdot)$ and $L(\cdot)$ are continuous operators. A unique solution u exists if and only if:

$$\exists\, \beta > 0 \ \text{s.t.} \ \inf_{u \in \mathcal{V}} \sup_{w \in \mathcal{W}} \frac{B(w, u)}{\|w\|_{\mathcal{W}} \|u\|_{\mathcal{V}}} \geq \beta\,, \tag{2.47a}$$

$$B(w, u) = 0 \ \forall\, u \in \mathcal{V} \quad \Rightarrow \quad w = 0\,. \tag{2.47b}$$

The second condition, Eq. (2.47b), may equivalently be stated as:

$$\sup_{u \in \mathcal{V}} \frac{B(w, u)}{\|u\|_{\mathcal{V}} \|w\|_{\mathcal{W}}} =: \gamma_w > 0 \quad \forall\, w \in \mathcal{W} \setminus \{0\}\,, \tag{2.48}$$

which ensures that the contradiction of Eq. (2.47b) does not hold (i.e, if $w \neq 0$, then $B(w, u) \neq 0 \ \forall\, u \in \mathcal{V}$). In Eq. (2.48), γ_w is just some w-dependent real number that the inequality requires to be positive. A slightly stronger condition that implies Eq. (2.48) and thus Eq. (2.47b) is to say that there exists a lower bound to γ_w:

$$\exists\, \gamma > 0 \ \text{s.t.} \ \inf_{w \in \mathcal{W}} \sup_{u \in \mathcal{V}} \frac{B(w, u)}{\|w\|_{\mathcal{W}} \|u\|_{\mathcal{V}}} \geq \gamma\,. \tag{2.49}$$

When $\mathcal{W} = \mathcal{V}$ we see that the Lax-Milgram theorem represents the requirement that Eqs. (2.47a) and (2.49) hold for the choices $w = u$ and $u = w$, respectively, which, again, is sufficient but not necessary. Based on Eqs. (2.47a) and (2.49), we often say that well-posedness relies on "inf-sup stability" of the bilinear form.

Theorem 2.3 poses a sufficient *and necessary* condition for the well-posedness of our variational formulations. The downside is that it is often difficult to prove. Additionally, different from coercivity, inf-sup stability does not carry over to subspaces. That means that proving Theorem 2.3 for the continuous case does not mean that a (conformal) finite element approximation is well-posed.

When applied to problems of the form Eq. (2.45), the inf-sup condition can be rephrased

in a format that is slightly easier to handle. In relation to these types of problems, it is often referred to as the Ladyžhenskaya-Babuška-Brezzi (or LBB) condition.

Theorem 2.4 (Ladyžhenskaya-Babuška-Brezzi). *Consider a problem of the form:*

Find $u, \lambda \in \mathcal{V} \times \mathcal{Q}$ s.t. $\forall\, w, q \in \mathcal{V} \times \mathcal{Q}$:

$$a(w, u) + b(w, \lambda) = L(w)\,, \tag{2.50a}$$

$$b(u, q) = 0\,, \tag{2.50b}$$

where $a(\cdot, \cdot)$, $b(\cdot, \cdot)$ and $L(\cdot)$ are continuous linear operators. Define the direct sum decomposition of $\mathcal{V} = \mathcal{V}^0 \oplus \mathcal{V}^\perp$, where $\mathcal{V}^0 = \{v \in \mathcal{V} : b(v, q) = 0 \; \forall\, q \in \mathcal{Q}\}$ and $\mathcal{V}^\perp$ is its orthogonal complement. This formulation permits a unique solution under the conditions:

$$\exists\, \alpha > 0 \;\; s.t. \;\; a(w, w) \geq \alpha \|w\|_\mathcal{V}^2 \;\; \forall\, w \in \mathcal{V}^0\,, \tag{2.51a}$$

$$\exists\, \beta > 0 \;\; s.t. \;\; \inf_{q \in \mathcal{Q}} \sup_{w \in \mathcal{V}^\perp} \frac{b(w, q)}{\|w\|_\mathcal{V} \|q\|_\mathcal{Q}} \geq \beta\,. \tag{2.51b}$$

Note that the supremum over $V^\perp$ in Eq. (2.51b) is the same as the supremum over all of $\mathcal{V}$, such that this condition can be (and in most references is) written as $\inf_{q \in \mathcal{Q}} \sup_{w \in \mathcal{V}} \frac{b(w,q)}{\|w\|_\mathcal{V} \|q\|_\mathcal{Q}} \geq \beta$. The form of Eq. (2.51b) has a more natural place in the following proof.

Proof. Different from previous definitions and theorems we will prove this theorem, as we will make use of the same line of arguments in subsequent chapters. First we prove existence and uniqueness of a solution u, then the existence of and uniqueness of λ.

If a solution u exists, then, according to Eq. (2.50b), it lies in $\mathcal{V}^0$ (which is non-empty, as it contains $u = 0$ at the very least). We then write Eq. (2.50a) as relation for u:

Find $u \in \mathcal{V}^0$ s.t. $\forall\, w \in \mathcal{V}^0$:

$$a(w, u) = L(w) - b(w, \lambda) = L(w)\,, \tag{2.52}$$

where $b(w, \lambda) = 0$ as this only concerns $w \in \mathcal{V}^0$. Under the assumption of Eq. (2.51a) and with the Lax-Milgram theorem, this permits a unique solution u.

Next, we write Eq. (2.50a) as a relation for λ for a given u:

Find $\lambda \in \mathcal{Q}$ s.t. $\forall\, w \in \mathcal{V}^\perp$:

$$b(w, \lambda) = L(w) - a(w, u) =: L^*(w) \,. \tag{2.53}$$

This weak formulation only concerns test functions in $\mathcal{V}^\perp$ as it is trivially satisfied for any $w \in \mathcal{V}^0$. For this problem statement the condition of Eq. (2.51b) is exactly the first BNB condition. The second BNB condition is naturally satisfied by construction of the space $\mathcal{V}^\perp$. Theorem 2.3 then states that Eq. (2.53) permits a unique solution λ. $\square$

Remark 2.2. *A conformal finite element approximation of Eq. (2.50) would result in the following system of equations:*

$$\begin{bmatrix} \mathrm{A} & \mathrm{b} \\ \mathrm{b}^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} \mathrm{u} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathrm{F} \\ 0 \end{bmatrix} \,, \tag{2.54}$$

where $A \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^{n \times m}$. The condition of Eq. (2.51a) concerns the positive definiteness of the submatrix A *in the kernel of* b^{T}, *and the condition of Eq. (2.51b) requires $m \leq n$ and amounts to the full-rankedness of submatrix* b.

CHAPTER 3

The variational multiscale method

In this chapter, we introduce the variational multiscale (VMS) method. Starting with a historical overview in Section 3.1, the original objectives that lead to the development of the VMS framework are reviewed. In Section 3.2, we address the formal inversion of the fine-scale problem and we show that for symmetric coercive problems this formalism leads to conformal finite element formulations. In Section 3.3 we introduce approximations to arrive at residual-based fine-scale models. We discuss its relation with stabilized methods and with turbulence models.

3.1 A paradigm for computational mechanics

While the principle ideas behind the variational multiscale method can be traced back to 1995 [109], the approach was formalized in the seminal 1998 article "The variational multiscale method - a paradigm for computational mechanics" by Hughes et al. [115]. The aim was to provide a physical rationale behind the various stabilized methods that had been introduced for fluid mechanics application in the preceding 15 years [38, 39, 116, 117, 118]. Without stabilization, Galerkin approximations of advection-based PDEs tends to yield oscillatory solutions. The oscillations become more extreme as the advection dominates. This is illustrated in Figure 3.1, which shows an example simulation of a model problem that is used repeatedly throughout this dissertation. Clearly, the finite element solution is not satisfactory. We are lead to ask the question "what *is* the optimal coarse-scale representation of an exact solution?". If we manage to formulate an answer, then we are quickly lead to the followup question "how do we obtain *that* solution rather than any other?" These are the key question around which the variational multiscale paradigm revolves.

(a) Boundary conditions and advective direction.

(b) True solution for $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.003$.

(c) Galerkin approximation on a coarse mesh.
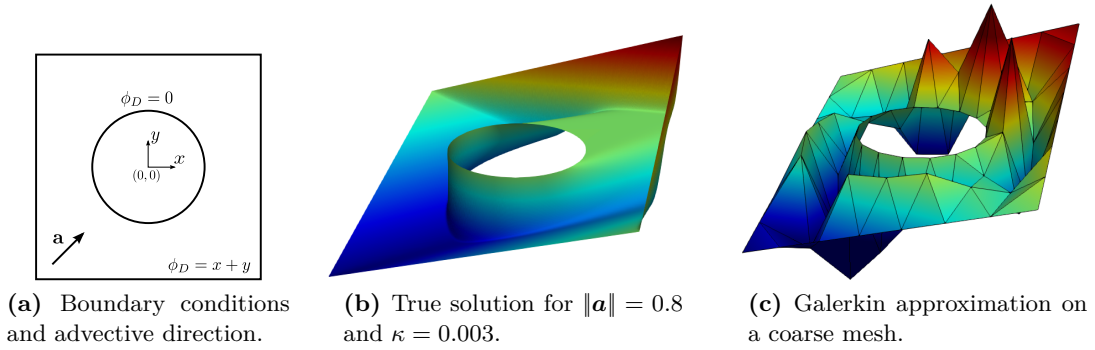
Fig. 3.1: *Example unstable behavior for a Galerkin approximation of the linear advection-diffusion equation $\boldsymbol{a} \cdot \nabla\phi - \kappa\Delta\phi = 0$ on a unit square with a hole.*

After its introduction in the 1940s [68, 105] and its naming in 1960 [45], the finite element method quickly gained momentum for solid mechanics application [140, 180, 189, 200]. In that field of study, weak formulations originate from minimization of energy functionals, leading to symmetric and coercive bilinear forms. The resulting Galerkin approximation of the continuous problem is optimal in the energy norm, which answers the earlier questions. However, these are 'a posteriori' answers in the sense that we did not impose this requirement in the construction of the finite element scheme (we will fix this in Section 3.2.3). The same Galerkin approximation applied to the just described fluid mechanics PDEs yields less favorable discrete solutions. A discrete representation of some function is not unique. We could use any projection operator (see Definition 2.24) and define the result to be the coarse-scale representation of that function. So then, how do we obtain *that* solution, rather than any other? Consider first the continuous form of some weak formulation:

Find $\phi \in \mathcal{W}$ s.t. $\forall\, w \in \mathcal{W}$:
$$B(w, \phi) = L(w) \,, \tag{3.1}$$

where $\mathcal{W}$ is an infinite-dimensional Sobolev space. Clearly, computers are not equipped for handling spaces that are infinite-dimensional, so we discretize. A Galerkin approximation follows from confining ϕ and w to a finite-dimensional subspace $\mathcal{W}^h$:

Find $\phi^h \in \mathcal{W}^h$ s.t. $\forall\, w^h \in \mathcal{W}^h$:
$$B(w^h, \phi^h) = L(w^h) \,. \tag{3.2}$$

This procedure has a somewhat ad-hoc feel to it: while Eq. (3.1) unambiguously defines

the 'true solution' as a member of $\mathcal{W}$, there is no guarantee that Eq. (3.2) with some arbitrary sampling of functions in $\mathcal{W}$ for the construction of $\mathcal{W}^h$ produces solutions that in any way represent that true solution. Indeed, for the example of Figure 3.1 it does not.

A natural conclusion, and a conclusion that drove the development of the VMS framework, is that the loss of fine-scale information induced by confining ϕ to some finite-dimensional subspace of $\mathcal{W}$ is the source of the loss of stability for advection based PDEs [109, 115, 124, 125]. An alternative approach was proposed. Rather than reducing the dimensionality of the space $\mathcal{W}$, it is *decomposed* into the finite-dimensional "coarse-scale" finite element approximation space $\mathcal{W}^h$, and an additional infinite-dimensional "fine-scale" (or sometimes "subgrid-scale") space $\mathcal{W}'$. The solution ϕ and the test function w may then be represented by a coarse-scale function in the finite element basis complemented by a fine-scale contribution:

$$\phi = \phi^h + \phi', \tag{3.3a}$$

$$w = w^h + w'. \tag{3.3b}$$

If Eq. (3.1) satisfies the conditions of Theorem 2.3, then it permits a unique solution ϕ. We now also require that Eq. (3.3a) yields a unique decomposition into ϕ^h and ϕ' for any possible solution $\phi \in \mathcal{W}$. In the language of Section 2.2, this means that $\mathcal{W}^h$ and $\mathcal{W}'$ must form a direct sum decomposition of $\mathcal{W}$:

$$\mathcal{W} = \mathcal{W}^h \oplus \mathcal{W}'. \tag{3.4}$$

By making use of these definitions and the linearity of the bilinear and linear forms, the weak formulation of Eq. (3.1) decomposes as:

Find $\phi^h, \phi' \in \mathcal{W}^h \times \mathcal{W}'$ s.t. $\forall w^h, w' \in \mathcal{W}^h \times \mathcal{W}'$:

$$B(w^h, \phi^h) + B(w^h, \phi') = L(w^h), \tag{3.5a}$$

$$B(w', \phi^h) + B(w', \phi') = L(w'). \tag{3.5b}$$

Equation (3.5a) concerns the test functions w^h that live in the finite-dimensional space $\mathcal{W}^h$. This is the "coarse-scale problem". It forms a finite set of equations and for a given ϕ' it represents a finite element formulation. Equation (3.5b) is the "fine-scale problem", which concerns all the remaining test functions in the infinite-dimensional space $\mathcal{W}'$.

If we compare Eq. (3.5a) with Eq. (3.2), then we see that there was a missing term in the

28

Galerkin approximation. This term takes into account the fine-scale effects onto the coarse-scale solution. It may represent the stabilizing effect of stabilized methods, or the subgrid-scale effects of turbulence models. It is the term that, if formulated properly, allows us to retrieve the 'optimal' solution ϕ^h according to whatever definition of 'optimality' we choose to adhere.

We must now answer the question "what is the optimal coarse-scale representation of an exact solution?" For each function $\phi \in \mathcal{W}$ we must select an optimal representation $\phi^h \in \mathcal{W}^h$ in the coarse-scale space: we are dealing with a mapping from $\mathcal{W}$ onto $\mathcal{W}^h$. This mapping is 'onto', as, naturally, we wish to map functions in $\mathcal{W}^h$ to themselves: if the true solution lies in the coarse-scale space then this is the optimal representation. Recall from Definition 2.23 that this property of a mapping is called "idempotency" and such mappings are called "projectors" by Definition 2.24. Looking ahead, idempotency of the mapping involves the variational consistency of the resulting finite element formulation. Another property that we might require is linearity: if $w^h \in \mathcal{W}^h$ is the optimal representation of $w \in \mathcal{W}^h$ and $v^h \in \mathcal{W}^h$ is the optimal representation of $v \in \mathcal{W}$, then it seems natural to define the optimal representation of $\phi = av + bw$ as $\phi^h = av^h + bw^h$ (a and b being real numbers). So, we are dealing with linear projectors.

We thus proceed under the assumption of having selected a linear projection operator of the form:

$$\mathscr{P} : \mathcal{W} \to \mathcal{W}^h \,. \tag{3.6}$$

Common choices of linear projectors are the "orthogonal projectors" of Definition 2.25. These project orthogonally with respect to some inner product such that the resulting coarse-scale solution is optimal in the corresponding norm.

In any case, the projection defines the sought-after solution pair ϕ^h, ϕ' for a given ϕ:

$$\phi^h := \mathscr{P}\phi \qquad \in \mathcal{W}^h \,, \tag{3.7a}$$
$$\phi' := \phi - \mathscr{P}\phi \quad \in \mathcal{W}' \,. \tag{3.7b}$$

From Eq. (3.7b) we may also infer the definition of the fine-scale space $\mathcal{W}'$. Due to the assumed linearity of the projector, Eq. (3.7b) requires $\mathscr{P}\phi = 0 \; \forall\, \phi \in \mathcal{W}'$. The fine-scale space is thus defined as the kernel of the projector:

$$\mathcal{W}' = \ker(\mathscr{P}) \,. \tag{3.8}$$

This definition of $\mathcal{W}'$ satisfies the direct sum decomposition of Eq. (3.4). Consequently, the multiscale weak formulation of Eq. (3.5) is well-posed, and the coarse-scale solution obtained from that weak formulation will be the one defined in Eq. (3.7a).

3.2 Formal inversion of the fine-scale problem

Of course, the goal is to solve for the coarse-scale solution ϕ^h without having to explicitly compute the fine-scale solution ϕ'. To do so, we need to substitute some form of closure in place of the fine scales that occur in the coarse-scale equation. This closure, in turn, should originate from the fine-scale equation. In the following, we express the fine scales in terms of the coarse scales, and substitute this result into the coarse-scale equation. To a very large extent, this section is based on [80, 122], although Sections 3.2.2 and 3.2.3 offer an original perspective on the theory and the proposed derivation.

3.2.1 The fine-scale Green's operator

At the end of Section 2.3, three equivalent notations were introduced to describe the same weak formulation. Adopting the duality pairing notation of Eq. (2.40) to the multiscale formulation of Eq. (3.5) gives:

Find $\phi^h, \phi' \in \mathcal{W}^h \times \mathcal{W}'$ s.t. $\forall\, w^h, w' \in \mathcal{W}^h \times \mathcal{W}'$:

$$\langle \mathcal{L}\phi^h, w^h \rangle + \langle \mathcal{L}\phi', w^h \rangle = \langle \mathcal{F}, w^h \rangle, \tag{3.9a}$$

$$\langle \mathcal{L}\phi^h, w' \rangle + \langle \mathcal{L}\phi', w' \rangle = \langle \mathcal{F}, w' \rangle, \tag{3.9b}$$

where $\mathcal{L}\phi$ and $\mathcal{F}$ are the functionals corresponding to $B(\phi, \cdot)$ and $L(\cdot)$. We assume that the bilinear forms $\langle \mathcal{L}\phi^h, w^h \rangle$ and $\langle \mathcal{L}\phi', w' \rangle$ are inf-sup stable, i.e., they satisfy the conditions of Theorem 2.3.

By moving all the terms that do not involve fine-scale trial functions in Eq. (3.9b) to the right-hand side, we obtain:

Find $\phi' \in \mathcal{W}'$ s.t. $\forall\, w' \in \mathcal{W}'$:

$$\langle \mathcal{L}\phi', w' \rangle = \langle \mathcal{F}, w' \rangle - \langle \mathcal{L}\phi^h, w' \rangle = \langle \mathcal{F} - \mathcal{L}\phi^h, w' \rangle, \tag{3.10}$$

where the coarse-scale solution ϕ^h is treated as data. This problem permits a unique fine-scale solution for each coarse-scale function by virtue of the assumed inf-sup stability.

We associate with the left-hand side the fine-scale differential operator $\mathcal{L}' : \mathcal{W}' \to \mathcal{W}'^*$ and with the right-hand side the coarse-scale residual $\mathcal{R}_{\phi^h} \in \mathcal{W}'^*$. By using these operators we rephrase Eq. (3.10) once more:

Find $\phi' \in \mathcal{W}'$ s.t. :
$$\mathcal{L}'\phi' = \mathcal{R}_{\phi^h} \,. \tag{3.11}$$

This 'operator form' of the weak statement was introduced in Eq. (2.39), and will alleviate notional difficulty in the following.

As this problem is invertible, meaning that the operator $\mathcal{L}'$ is bijective, we can associate a new operator to this inversion. We call this operator the fine-scale Green's operator, denoted by $\mathcal{G}' : \mathcal{W}'^* \to \mathcal{W}'$. The fine-scale solution follows from the coarse-scale residual as:

$$\phi' = \mathcal{G}'\mathcal{R}_{\phi^h} \,. \tag{3.12}$$

The operator $\mathcal{G}'$ is somewhat elusive: it inverts the differential operator but only maps to the fine scales. In contrast, the more familiar Green's operator $\mathcal{G}$ inverts $\mathcal{L}$ for all of $\mathcal{W}$. To get a better understanding of $\mathcal{G}'$, and to find its relation with $\mathcal{G}$, we can reformulate the fine-scale problem.

From Section 3.1 we know that the fine-scale space consists of all functions in $\mathcal{W}$ for which the projection yields the zero function. $\mathcal{W}'$ is thus a constrained space. An alternative formulation of the fine-scale problem follows from releasing this constraint from $\mathcal{W}'$ and incorporating it weakly with the aid of a Lagrange multiplier:

Find ϕ' in $\mathcal{W}$ and $\lambda^h \in \mathcal{W}^{h,*}$ s.t. :
$$\mathcal{L}\phi' + \mathscr{P}^T\lambda^h = \mathcal{R}_{\phi^h} \,, \tag{3.13a}$$
$$\mathscr{P}\phi' = 0 \,. \tag{3.13b}$$

This equation involves $\mathcal{L}$, which maps from $\mathcal{W}$ to $\mathcal{W}^*$. The inverse calls for the classical Green's operator $\mathcal{G}$. Using $\mathcal{G}$ to relate ϕ' to λ^h gives:

$$\phi' = \mathcal{G}(\mathcal{R}_{\phi^h} - \mathscr{P}^T\lambda^h) \,. \tag{3.14}$$

Substituting the result in Eq. (3.13b) yields the following requirement:

$$\mathscr{P}\mathcal{G}\mathcal{R}_{\phi^h} - \mathscr{P}\mathcal{G}\mathscr{P}^T\lambda^h = 0\,. \tag{3.15}$$

This may be rewritten as an expression for λ^h:

$$\lambda^h = (\mathscr{P}\mathcal{G}\mathscr{P}^T)^{-1}\mathscr{P}\mathcal{G}\mathcal{R}_{\phi^h}\,, \tag{3.16}$$

which, in turn, may be substituted in Eq. (3.14) to arrive at an expression for ϕ' that only involves the classical Green's function and the coarse-scale residual:

$$\phi' = \mathcal{G}\mathcal{R}_{\phi^h} - \mathcal{G}\mathscr{P}^T(\mathscr{P}\mathcal{G}\mathscr{P}^T)^{-1}\mathscr{P}\mathcal{G}\mathcal{R}_{\phi^h}\,. \tag{3.17}$$

By factoring out the residual it is clear that the fine-scale Green's operator relates to the classical Green's operator via:

$$\mathcal{G}' = \mathcal{G} - \mathcal{G}\mathscr{P}^T(\mathscr{P}\mathcal{G}\mathscr{P}^T)^{-1}\mathscr{P}\mathcal{G}\,. \tag{3.18}$$

Substitution of Eq. (3.12) into the coarse-scale problem of Eq. (3.9a) while using Eq. (3.18) to define the fine-scale Green's operator results in a well-posed discrete formulation that will yield the optimal coarse-scale solution:

$$\text{Find } \phi^h \in \mathcal{W}^h \text{ s.t. } \forall\, w^h \in \mathcal{W}^h :$$
$$\langle \mathcal{L}\phi^h, w^h\rangle - \langle \mathcal{L}\mathcal{G}'\mathcal{L}\phi^h, w^h\rangle = \langle \mathcal{F}, w^h\rangle - \langle \mathcal{G}'\mathcal{F}, w^h\rangle\,. \tag{3.19}$$

3.2.2 The fine-scale Green's function

In some cases, the fine-scale Green's operator may be represented by a convolution integral with a "fine-scale Green's function". This function will take a central role during the fine-scale modeling later on. In the following, we show that (under certain conditions) such a function exists and we show how it relates to the classical Green's function. Atop the assumptions that were made in the previous sections, we also assume:

↪ The bilinear part of the weak formulation is coercive (which is more restrictive than the inf-sup stability assumed earlier).

↪ $\mathcal{W}'$ is sufficiently regular to permit point evaluations (this requires $\mathcal{W}' \subset H^s(\mathbb{R}^d)$

with $s > d/2$, e.g., a subspace of $H^1(\Omega)$ when Ω is one-dimensional [1]).

As an alternative to the second requirement we might replace point evaluations by a highly localized averaging operation, although that suggestion is more of practical relevance than of interest to this theoretical foundation.

Consider again the (constrained) fine-scale problem:

Find $\phi' \in \mathcal{W}'$ s.t. $\forall w' \in \mathcal{W}'$:

$$\langle \mathcal{L}\phi', w' \rangle = \langle \mathcal{R}_{\phi^h}, w' \rangle . \tag{3.20}$$

Recall that we interpret the operator $\mathcal{L}$ as the mapping $\mathcal{L} : \mathcal{W}' \to \mathcal{W}'^*$ defined as:

$$\mathcal{L}\phi' = B(\cdot, \phi') , \tag{3.21}$$

with $B(\cdot, \cdot)$ the bilinear form of the weak formulation. We may now define an adjoint operator that satisfies:

$$\langle \mathcal{L}\phi', w' \rangle = \langle \mathcal{L}^* w', \phi' \rangle \qquad \forall \phi' \in \mathcal{W}' \text{ and } \forall w' \in \mathcal{W}' . \tag{3.22}$$

Based on Eq. (3.21), the adjoint operator $\mathcal{L}^*$ is simply defined as:

$$\mathcal{L}^* w' = B(w', \cdot) . \tag{3.23}$$

As a consequence of the second assumption, the following linear functional is bounded and lies in $\mathcal{W}'^*$:

$$\langle \delta_x, \phi' \rangle = \phi'(x) , \tag{3.24}$$

where δ_x is the Dirac delta distribution at location x.

We now wish to find a test function $w' \in \mathcal{W}'$ such that Eq. (3.23) induces this functional:

Find $w' \in \mathcal{W}'$ s.t. $\forall \phi' \in \mathcal{W}'$:

$$\langle \mathcal{L}^* w', \phi' \rangle = \langle \delta_x, \phi' \rangle . \tag{3.25}$$

Since we assumed that all the involved operators are bounded and the bilinear form $B(\cdot, \cdot)$ is coercive, we may conclude from Theorem 2.2, the Lax-Milgram theorem, that this

weak statement permits a unique $w'(y) \in \mathcal{W}'$. We denote this function $g'(x, y)$, the fine-scale Green's function. Note that x, the first argument, corresponds to the sampling in Eq. (3.24), and that y, the second argument, is the variable for a function in $\mathcal{W}'$. All differentiation and integration implied by Eq. (3.25) thus involves the y coordinate.

With some final manipulation we see that the fine-scale solution may be obtained from a convolution integral with the fine-scale Green's function:

$$
\begin{aligned}
\phi'(x) &= \langle \mathcal{L}^* g'(x, y), \phi' \rangle = \langle \mathcal{L} \phi', g'(x, y) \rangle = \langle \mathcal{R}_{\phi^h}, g'(x, y) \rangle \\
&= L(g'(x, y)) - B(g'(x, y), \phi^h) \,.
\end{aligned} \tag{3.26}
$$

The fine-scale Green's function is thus the function in $\mathcal{W}'$ that solves Eq. (3.25) for any $\phi' \in \mathcal{W}'$. To find a relation with the classical Green's functions, which solves the same problem but in an unconstrained space, we can rephrase and enforce the constraints in $\mathcal{W}'$ with Lagrange multipliers.

In this dissertation, all projectors will involve optimality conditions of error potentials of the form:

$$
\begin{aligned}
\mathscr{P} : \mathcal{W} &\to \mathcal{W}^h \\
\phi &\mapsto \underset{\phi^h \in \mathcal{W}^h}{\arg\inf} \tfrac{1}{2} E(\phi - \phi^h, \phi - \phi^h) \,,
\end{aligned} \tag{3.27}
$$

where $\frac{1}{2} E(\cdot, \cdot)$ denotes the symmetric and coercive bilinear form that represents the error potential. Additionally, all coarse-scale spaces will be spanned by a finite number of bases (say, n). The constraints on the fine-scale space imposed by Eq. (3.27) may then be inferred from the Gâteaux derivative of the potential. They may be written in any of the following equivalent forms:

$$
E(w^h, \phi - \phi^h) = 0 \qquad \forall\, w^h \in \mathcal{W}^h := \operatorname{span}\{N_i\}_{i=1}^n \,, \tag{3.28a}
$$

$$
E(\sum_{i=1}^n \hat{w}_i N_i, \phi - \phi^h) = 0 \quad \forall\, [\hat{w}_1, \cdots, \hat{w}_n]^{\mathrm{T}} \in \mathbb{R}^n \,, \tag{3.28b}
$$

$$
E(N_i, \phi - \phi^h) = 0 \qquad \text{for } i = 1, \cdots, n \,, \tag{3.28c}
$$

$$
\langle \mu_i, \phi - \phi^h \rangle = 0 \qquad \text{for } i = 1, \cdots, n \,, \tag{3.28d}
$$

where $\mu_i : \mathcal{W} \to \mathbb{R}$ are the functionals induced by $E(N_i, \cdot)$. They thus define the set of functional constraints $\langle \mu_i, \phi' \rangle = 0$ satisfied by any $\phi' \in \mathcal{W}'$.

In unconstrained form, the adjoint fine-scale problem reads:

Find $w', \boldsymbol{\lambda} \in \mathcal{W} \times \mathbb{R}^n$ s.t. $\forall \phi', \boldsymbol{q} \in \mathcal{W} \times \mathbb{R}^n$:

$$\langle \mathcal{L}^* w', \phi' \rangle + \langle \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{\mu}, \phi' \rangle = \langle \delta_x, \phi' \rangle \,, \tag{3.29a}$$

$$\langle \boldsymbol{q}^{\mathrm{T}} \boldsymbol{\mu}, w' \rangle = 0 \,, \tag{3.29b}$$

where we introduced the vector of functionals $\boldsymbol{\mu} = [\mu_1, \cdots, \mu_n]^{\mathrm{T}}$ for ease of notation. Note the analogy with Eq. (3.13).

In more familiar 'Green's function notation', we are looking for the fine-scale Green's function that satisfies:

$$\begin{cases} w' \in \mathcal{W}' \,, \\ \mathcal{L}^* w' = \delta_x - \sum\limits_{i=1}^{n} \lambda_i \mu_i \ \in \mathcal{W}^* \,. \end{cases} \tag{3.30}$$

The first requirement in Eq. (3.30) follows from Eq. (3.29b), while the second requirement still involves the unconstrained dual space due to Eq. (3.29a).

To obtain an expression for w' independent of the Lagrange multipliers $\boldsymbol{\lambda}$, we use a reduction procedure analogous to the one for the fine-scale Green's operator. From Eqs. (3.29b) and (3.30) we obtain the following expression for the Lagrange multipliers:

$$w' = \mathcal{L}^{*-1} \delta_x - \boldsymbol{\lambda}^{\mathrm{T}} \mathcal{L}^{*-1} \boldsymbol{\mu} \,, \tag{3.31a}$$

$$\langle \boldsymbol{q}^{\mathrm{T}} \boldsymbol{\mu}, \mathcal{L}^{*-1} \delta_x - \boldsymbol{\lambda}^{\mathrm{T}} \mathcal{L}^{*-1} \boldsymbol{\mu} \rangle = 0 \quad \Rightarrow$$
$$\langle \boldsymbol{q}^{\mathrm{T}} \boldsymbol{\mu}, \boldsymbol{\lambda}^{\mathrm{T}} \mathcal{L}^{*-1} \boldsymbol{\mu} \rangle = \langle \boldsymbol{q}^{\mathrm{T}} \boldsymbol{\mu}, \mathcal{L}^{*-1} \delta_x \rangle \qquad \forall \boldsymbol{q} \in \mathbb{R}^n \,, \tag{3.31b}$$

$$\lambda_j \langle \mu_i, \mathcal{L}^{*-1} \mu_j \rangle = \langle \mu_i, \mathcal{L}^{*-1} \delta_x \rangle \qquad \text{for } i = 1, \cdots, n \,. \tag{3.31c}$$

These steps are analogous to Eqs. (3.14) to (3.16).

The right-hand side of Eq. (3.31c) forms a set of n scalars, i.e., a (row) vector. Similarly, the left-hand side represents a multiplication between a vector of scalar Lagrange multiplier and a matrix of scalars. Post multiplying with the inverse of the matrix of scalars gives an expression for the vector of Lagrange multipliers:

$$\boldsymbol{\lambda}^{\mathrm{T}} = \langle \boldsymbol{\mu}, \mathcal{L}^{*-1} \delta_x \rangle \left[\langle \boldsymbol{\mu}^{\mathrm{T}}, \mathcal{L}^{*-1} \boldsymbol{\mu} \rangle \right]^{-1} . \tag{3.32}$$

We substitute this back into Eq. (3.31a) to solve for $w' = g'(x, y)$:

$$g'(x, y) = \mathcal{L}^{*-1}\delta_x - \langle \boldsymbol{\mu}^{\mathrm{T}}, \mathcal{L}^{*-1}\delta_x \rangle \left[\langle \boldsymbol{\mu}^{\mathrm{T}}, \mathcal{L}^{*-1}\boldsymbol{\mu} \rangle \right]^{-1} \mathcal{L}^{*-1}\boldsymbol{\mu}. \tag{3.33}$$

The first term in this expression is the classical Green's function. The remainder is a vector-matrix-vector multiplication that ensures that $g'(x, y) \in \mathcal{W}'$.

3.2.3 The Galerkin approximation for symmetric coercive problems

The formalism of the previous two sections allows us to obtain the optimal coarse-scale representation of the true solution of a PDE, directly from the given data. Again, 'optimal' refers to the chosen projector. For PDEs that result in a symmetric and coercive bilinear form, such as compressible linear elasticity and heat diffusion, Galerkin's method yields solutions that are perfectly satisfactory coarse-scale representations. For such problems, there is no need for including fine-scale effects. In this section, we interpret this important result from a variational multiscale point of view.

In the case of a symmetric and coercive bilinear form, the weak formulation stems from a potential minimization:

$$\phi = \arg\min_{\phi \in \mathcal{W}} \left\{ \frac{1}{2} B(\phi, \phi) - L(\phi) \right\}. \tag{3.34}$$

Often, $\frac{1}{2}B(\phi, \phi)$ represents the system's internal energy and $L(\phi)$ the external work. The continuous weak formulation for this problem may be stated in operator form as:

Find $\phi \in \mathcal{W}$ s.t. :
$$\mathcal{B}\phi = \mathcal{F}. \tag{3.35}$$

After decomposition of $\mathcal{W}$, the multiscale weak formulation becomes:

Find $\phi^h, \phi' \in \mathcal{W}^h \times \mathcal{W}'$ s.t. :
$$\mathcal{B}^h\phi^h + \tilde{\mathcal{B}}\phi' = \mathcal{F}^h, \tag{3.36a}$$
$$\tilde{\mathcal{B}}^{\mathrm{T}}\phi^h + \mathcal{B}'\phi' = \mathcal{F}', \tag{3.36b}$$

where the domains and codomains of the various instantiations of the $\mathcal{B}$ operator and the $\mathcal{F}$ functional should be self evident.

Suppose also that we choose the following projection operator for the definition of the coarse scales:

$$\mathscr{P} : \mathcal{W} \to \mathcal{W}^h$$

$$\phi \; \mapsto \; \underset{\phi^h \in \mathcal{W}^h}{\arg\inf} \int_{\tilde{\Omega}} \frac{1}{2} B(\phi - \phi^h, \phi - \phi^h) = \underset{\phi^h \in \mathcal{W}^h}{\arg\inf} \int_{\tilde{\Omega}} \frac{1}{2} B(\phi', \phi') \,. \tag{3.37}$$

This projection returns the optimal approximation in the energy norm. The variational formulation of this minimization represents the constraints on ϕ'. In operator form it says:

$$\tilde{\mathcal{B}}\phi' = 0 \,. \tag{3.38}$$

The unconstrained fine-scale problem then reads:

Find ϕ' in $\mathcal{W}$ and $\lambda^h \in \mathcal{W}^{h,*}$ s.t. :

$$\mathcal{B}\phi' + \tilde{\mathcal{B}}^{\mathrm{T}}\lambda^h = \mathcal{R}_{\phi^h} \,, \tag{3.39a}$$

$$\tilde{\mathcal{B}}\phi' = 0 \,. \tag{3.39b}$$

Taking $\mathcal{G}$ to be the inverse of $\mathcal{B}$, we obtain from Eq. (3.18) the fine-scale Green's function $\mathcal{G}'$ as:

$$\mathcal{G}' = \mathcal{G} - \mathcal{G}\tilde{\mathcal{B}}^T (\tilde{\mathcal{B}}\mathcal{G}\tilde{\mathcal{B}}^T)^{-1} \tilde{\mathcal{B}}\mathcal{G} \,, \tag{3.40}$$

and the closed coarse-scale formulation becomes:

Find $\phi^h \in \mathcal{W}^h$ s.t. :

$$\mathcal{B}^h\phi^h + \tilde{\mathcal{B}}\left(\mathcal{G} - \mathcal{G}\tilde{\mathcal{B}}^T (\tilde{\mathcal{B}}\mathcal{G}\tilde{\mathcal{B}}^T)^{-1} \tilde{\mathcal{B}}\mathcal{G}\right)\mathcal{R}_{\phi^h} = \mathcal{F}^h \,. \tag{3.41}$$

Focusing on the impact of the fine scales, we may simplify the expression:

$$\tilde{\mathcal{B}}\mathcal{G} - \tilde{\mathcal{B}}\mathcal{G}\tilde{\mathcal{B}}^T (\tilde{\mathcal{B}}\mathcal{G}\tilde{\mathcal{B}}^T)^{-1} \tilde{\mathcal{B}}\mathcal{G} = \tilde{\mathcal{B}}\mathcal{G} - \tilde{\mathcal{B}}\mathcal{G} = 0 \,, \tag{3.42}$$

such that the coarse-scale formulation simplifies to:

Find $\phi^h \in \mathcal{W}^h$ s.t. :

$$\mathcal{B}^h\phi^h = \mathcal{F}^h \,. \tag{3.43}$$

We are left with the finite element formulation corresponding to the Galerkin approximation of the continuous problem.

This analysis reproduces a well-known fact: plain finite element solutions of symmetric coercive problems are optimal in the energy norm [110, 180, 200]. We have merely put this knowledge in the context of the variational multiscale method. Arguably, this is the better perspective: the question "what is the optimal coarse-scale representation of an exact solution?" is nontrivial and fundamental. "The best approximation in the energy norm" should be an a priori answer from which the finite element formulation follows, rather than a coincidental a posteriori result.

3.3 Fine-scale models and residual-based stabilized methods

From the first section of this chapter we know in what sense the fine scales impact the coarse-scale equation, and from the second section we know how to compute that impact. Turning this theory into practice, is, alas, not entirely trivial. For PDEs with some level of complexity, analytical expressions for the Green's function are not known, and solving the Green's function problem numerically largely defeats the point of the variational multiscale methodology (although efforts in this direction have been made in various contexts [62, 63, 128, 145, 146, 147]). Even if the fine-scale Green's function is known, incorporating the expression for the fine scales in the coarse-scale equations would result in an extremely costly double integration. All of this becomes even less feasible if we are dealing with nonlinear PDEs, where this process would have to be repeated after each linearization. Clearly, we are looking for ways to *approximate* the effect of the fine scales. In this section, we discuss a number of modeling efforts that have been made in this regard. Often, these focus on the numerical solution of turbulent flow, i.e., the Navier-Stokes equations at moderate to high Reynolds numbers.

Shortly after its introduction, the principle ideas behind the variational multiscale method were recast in the framework of large eddy simulation (LES) of turbulent flow [119]. Large eddy simulation relies on a 'filtering operation' to define the filtered solution fields and the filtered Navier-Stokes equations [164]. The aim of the numerical method, which is most often a finite volume method, is to approximate these filtered fields. The filtered equations include subgrid-scale effects, which need to be modeled in order for the numerical method to capture the turbulent nature of the flow. This procedure strongly reminds of the variational multiscale method, but the subtle difference of using a filtering operation in LES introduces some conceptual difficulties [155, 164]. Firstly, the filtering procedure, which is often left completely implicit, becomes ambiguous towards the domain

boundary. Secondly, there are now two levels of approximation: the filtering step and numerical approximation step. Both of these problems vanish in the variational multiscale framework, where the coarse scales are *defined* as some projection onto the finite-dimensional approximation space and we are only left with the numerical approximation step.

Nevertheless, the filtered equations in LES and the coarse-scale equations in VMS bear close resemblance. Hence, the subgrid-scale models that are used in LES have been explored as fine-scale models in the VMS framework. These subgrid-scale models are inspired by the physical nature of turbulence. Turbulence theory tells us that the scale interaction that occurs is mostly of a dissipative nature [188]. Based on this observation, typical practice in LES modeling is to add a so-called "eddy viscosity". Early versions of the model are 'static' in the sense that the diffusion coefficient is kept constant [137, 172], and later versions are 'dynamic' with a diffusion coefficient that depends on flow characteristics [88, 138]. Similarly, the fine-scale occurrences in the VMS framework were first modeled as appropriately scaled diffusion terms [112, 119, 120, 144].

We also know that the energy dissipation in turbulent flow has a cascading character: kinetic energy from larger scales trickles down to neighboring smaller scales. At the bottom of the cascade are the Kolmogorov scales, where diffusion dominates and the energy dissipates as heat [99, 134, 135]. Replacing the subgrid-scale terms in the filtered LES equations by an eddy viscosity does not model this effect. However, in a finite element framework and with the variational multiscale method at hand we can mimic this behavior. If we use a hierarchical basis for our finite element discretization, then we can incorporate the eddy viscosity only for the bases with the highest polynomial order or the most local support [64, 104, 119, 120, 126, 127]. This is sometimes called a "three-scale model": there are coarse scales that exchange energy with intermediate scales which dissipate energy to the fine scales. The fine scales are those that are modeled with the diffusive term. Such models have also been used in combination with discontinuous Galerkin finite element formulations [65, 66, 67, 157], where the hierarchical integrated Legendre basis is used regularly [47].

A problem with eddy viscosity models, and thus also the equivalent diffusion models in the VMS framework, is that they purely act as energy sinks. They do not account for backscatter, which is the feeding of energy from the fine-scale solution back into the coarse-scale solution. Additionally, they are not variationally consistent: the exact solution no longer satisfies the discrete formulation. More complete, and variationally consistent, fine-scale models can be derived by making use of the formal inversion of the fine-scales discussed in Section 3.2. We saw that the fine-scale problem is analogous to the coarse-

scale problem, but with a new source term. This source term involves the residual of the coarse-scale solution. This realization lead to the development of the so-called "residual-based fine-scale models" [2, 18, 19, 41, 131, 142].

If we assume sufficient smoothness of the coarse-scale space $\mathcal{W}^h$, then we may perform integration by parts such that the duality pairing with $\mathcal{R}_{\phi^h}$ in Eq. (3.26) may be written as:

$$
\begin{aligned}
\phi'(x) = \langle \mathcal{R}_{\phi^h}, g'(x,y) \rangle &= L(g'(x,y)) - B(g'(x,y), \phi^h) \\
&= \int_\Omega (f - \mathcal{L}\phi^h) \, g'(x,y) \, \mathrm{d}y =: \int_\Omega \mathcal{R}_{\phi^h} \, g'(x,y) \, \mathrm{d}y \,,
\end{aligned}
\tag{3.44}
$$

where $\mathcal{L}$ now represents the strong form differential operator which maps between Sobolev spaces.

The weighted occurrence of the fine-scale solution in the coarse-scale equation suggests the following approximation:

$$
\begin{aligned}
\int_\Omega \mathcal{L}^* w^h(x) \, \phi'(x) \, \mathrm{d}x &= \sum_K \int_K \mathcal{L}^* w^h(x) \int_\Omega g'(x,y) \mathcal{R}_{\phi^h}(y) \, \mathrm{d}y \, \mathrm{d}x \\
&\approx \sum_K \int_K \mathcal{L}^* w^h(x) \int_K g'(x,y) \mathcal{R}_{\phi^h}(y) \, \mathrm{d}y \, \mathrm{d}x \\
&\approx \sum_K \int_K \mathcal{L}^* w^h(x) \, \frac{1}{|K|} \int_K \int_K g'(\hat{x},y) \, \mathrm{d}y \, \mathrm{d}\hat{x} \, \mathcal{R}_{\phi^h}(x) \, \mathrm{d}x \\
&= \sum_K \int_K \mathcal{L}^* w^h(x) \, \tau \, \mathcal{R}_{\phi^h}(x) \, \mathrm{d}x \,,
\end{aligned}
\tag{3.45}
$$

where τ approximates the operator $\mathcal{G}'$ in a 'homogenized' sense over an element K. It is defined as the element average of the fine-scale Green's function:

$$
\tau := \frac{1}{|K|} \int_K \int_K g'_{H_0^1}(\hat{x}, y) \, \mathrm{d}y \, \mathrm{d}\hat{x} \,.
\tag{3.46}
$$

The introductory work on the residual-based turbulence model makes use of higher-order NURBS basis functions [2, 19]. These satisfy the required order of continuity that is assumed in Eq. (3.44). For conventional $\mathcal{C}^0$-continuous finite element methods, the expression of Eq. (3.44) introduces inter-element Dirac-layers [97, 115, 124]. These can be taken into account in the model of Eq. (3.45), as illustrated for instance in [129]. In practice, however, the residual-based model of Eq. (3.45) is also directly implemented in nodal finite element frameworks [87, 91, 181, 198]. The implicit assumption is that the

fine scales vanish on element boundaries [109, 115, 125]. This assumption is valid in the one-dimensional case when the H_0^1-projector is employed [123]. In almost all other cases this is an approximation.

If the differential operator $\mathcal{L}$ involves an advective operator, i.e., an operator of the form $\boldsymbol{a} \cdot \nabla$, then Eq. (3.45) introduces a terms $\left(\tau \boldsymbol{a} \cdot \nabla w^h, \boldsymbol{a} \cdot \nabla \phi^h\right)_\Omega$. This is a diffusive term that acts in the streamline direction. The empirically motivated eddy viscosity thus emerges naturally from the VMS framework [156, 183, 184]. The associated diffusion tensor is $\tau \boldsymbol{a} \otimes \boldsymbol{a}$, and τ has a mathematical foundation as the average fine-scale Green's function. More so, since the Navier-Stokes equations are nonlinear, mixed and partly vector valued, Eq. (3.45) implies the addition of many more modeling terms. Different from the extra diffusive term, which is coercive and stabilizes the formulation, some of these other terms tend to have a destabilizing effect. This reminds of the back-scatter of energy from the fine scales to the coarse scales that was missing in the eddy viscosity models. Finally, based on the choice of projector that defines the variational multiscale split, the fine-scale Green's function satisfies an orthogonality relation with the lower scales [122]. In the example of higher-order discretization of a one-dimensional linear advection-diffusion equation and an H_0^1-projector, the parameter τ only has a nonzero value when it involves the bases of the highest order polynomials. This is in complete agreement with the three-scale models discussed earlier, which were motivated from a physical perspective.

At this stage, a reference back to the stabilized methods mentioned at the beginning of this chapter is in order. Many stabilization techniques, such as the streamline upwind/Petrov-Galerkin (SUPG) stabilization of advective terms, pressure stabilized/Petrov-Galerkin (PSPG) stabilization of the incompressibility constraint, Galerkin/least squares (GLS) stabilization, and now also variational multiscale stabilization, may be written in the following form [72, 109, 124]:

Find $\phi^h \in \mathcal{W}^h$ s.t. $\forall\, w^h \in \mathcal{W}^h$:
$$B(w^h, \phi^h) - \sum_{K \in \mathcal{T}} \left(\tau \, \mathcal{L}_{\mathrm{s}} w^h, \mathcal{L} \phi^h\right)_K = L(w^h) - \sum_{K \in \mathcal{T}} \left(\tau \, \mathcal{L}_{\mathrm{s}} w^h,\, f\right)_K, \tag{3.47}$$

where the parameter τ is then called the stabilization parameter. Different stabilization methods are obtained for different choices of the differential operator $\mathcal{L}_{\mathrm{s}}$. The variational multiscale method suggests that $\mathcal{L}_{\mathrm{s}}$ should be chosen as the adjoint differential operator. Depending on the problem at hand, we can make a choice of projector to arrive at SUPG or PSPG. In retrospect, we see that these formulations actually incorporate the multiscale nature of the solution [19, 62, 121, 122, 125, 183, 184]. The strength of this perspective is that it implies a rationale behind the construction of τ via Eq. (3.46) [35, 98, 187].

The variational multiscale method for mixed formulations

Most of the analysis proposed in the previous chapter implicitly dealt with partial differential equations in their "primal form". They concern a single solution field which may or may not be tensor valued. The sought after solution is called the "primal" variable. Examples of primal variables are the temperature in the heat equation and the displacement vector in the equations of linear elasticity. Often, we are interested in physical systems where multiple variables interrelate: the velocity vector and the pressure for the Stokes equations, or the heat and heat-flux in a temperature problem. Loosely speaking, a weak formulation that involves multiple variables is called a "mixed formulation". Their relevance to this thesis is twofold: variational multiscale analysis mostly focuses on fluid mechanics where mixed formulations are ubiquitous, and, secondly, the discontinuous Galerkin formulations of Part III are often derived by introducing additional variables.

4.1 Mixed formulations in science and engineering

Mixed formulations arise in roughly one of two ways: either the strong form of the partial differential equation involves multiple (physical) states, or the additional variables are introduced via weak enforcement of constraints.

4.1.1 Systems of partial differential equations

We need as many equations as we have unknowns. If the strong form of the PDE concerns multiple variables, we thus need multiple PDEs to describe them: we are dealing with a *systems* of partial differential equations. This would occur whenever a multi-physics problem is considered: the chemical constituents and temperature in combustion, the electric potential and displacement fields in piezoelectric solids, or the displacement and velocity fields in fluid structure interaction. Sometimes, the multi-physics aspect is so fundamental that the sets of equations have been given their own names. Examples are Maxwell's equations of electromagnetism, Euler's equations of gas dynamics, and, of course, the Navier-Stokes and Stokes equations of incompressible flow. The latter is often used as the prototypical example when it comes to mixed formulations:

$$-\nabla \cdot (2\mu\nabla^s \boldsymbol{u} - p\boldsymbol{I}) = \boldsymbol{f} \qquad \text{in} \quad \Omega\,, \tag{4.1a}$$

$$\nabla \cdot \boldsymbol{u} = 0 \qquad \text{in} \quad \Omega\,, \tag{4.1b}$$

$$\boldsymbol{u} = \boldsymbol{u}_D \qquad \text{on} \quad \partial\Omega_D\,, \tag{4.1c}$$

$$2\mu\nabla^s \boldsymbol{u} \cdot \boldsymbol{n} - p\boldsymbol{n} = \boldsymbol{t} \qquad \text{on} \quad \partial\Omega_N\,, \tag{4.1d}$$

where $\boldsymbol{u}$ is the velocity vector, p the pressure and $\boldsymbol{f}$ the body force, each defined in the d-dimensional spatial domain $\Omega \subset \mathbb{R}^d$. The boundary data consists of $\boldsymbol{u}_D$, the enforced displacement on the Dirichlet boundary $\partial\Omega_D$, and $\boldsymbol{t}$, the prescribed traction on the Neumann boundary $\partial\Omega_N$.

A mixed formulation would follow from such a system of partial differential equations by multiplying each equation by a suitable test function, integration over the domain Ω and integration by parts wherever suitable. For example, if we consider homogeneous Dirichlet boundary data:

Find $\boldsymbol{u}, p \in \left[H_0^1(\Omega, \partial\Omega_D)\right]^d \times L^2(\Omega)$ s.t. $\forall \boldsymbol{v}, q \in \left[H_0^1(\Omega, \partial\Omega_D)\right]^d \times L^2(\Omega)$:

$$\int_\Omega 2\mu\nabla^s \boldsymbol{u} : \nabla^s \boldsymbol{v} - \int_\Omega p\, \nabla \cdot \boldsymbol{v} = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} + \int_{\partial\Omega_D} \boldsymbol{t} \cdot \boldsymbol{v}\,, \tag{4.2a}$$

$$-\int_\Omega \nabla \cdot \boldsymbol{u}\, q = 0\,. \tag{4.2b}$$

If $\partial\Omega_N$ is an empty set, then p would only be defined up to an arbitrary additive constant which would need to be removed from $L^2(\Omega)$ to make the formulation well-posed. This weak formulation is of the form of Eq. (2.45), and we can use Theorem 2.4 to prove well-posedness.

Even if the physics involves only a single state, we may still choose to formulate the PDE as a system of PDEs by introducing "auxiliary variables". Take, for instance, creeping flow through a porous medium. If we assume that the permeability tensor is isotropic such that it may be written as $\kappa \boldsymbol{I}$, then the pressure field $\phi(\boldsymbol{x})$ (or more accurately the "piezometric head") in the medium will reach a steady state described by a Laplace equation:

$$-\nabla \cdot (\kappa \nabla \phi) = 0 \qquad \text{in } \Omega\,, \tag{4.3a}$$

$$\phi = \phi_D \qquad \text{on } \partial\Omega_D\,, \tag{4.3b}$$

$$\kappa \nabla \phi \cdot \boldsymbol{n} = g_N \qquad \text{on } \partial\Omega_N\,. \tag{4.3c}$$

According to Darcy's law [70], the velocity profile may be derived from the pressure field as:

$$\boldsymbol{u} = -\kappa \nabla \phi\,. \tag{4.4}$$

When we introduce this definition of the velocity as an auxiliary variable, then Eq. (4.3) can be restated as:

$$\boldsymbol{u} = -\kappa \nabla \phi \qquad \text{in } \Omega\,, \tag{4.5a}$$

$$\nabla \cdot \boldsymbol{u} = 0 \qquad \text{in } \Omega\,, \tag{4.5b}$$

$$\phi = \phi_D \qquad \text{on } \partial\Omega_D\,, \tag{4.5c}$$

$$-\boldsymbol{u} \cdot \boldsymbol{n} = g_N \qquad \text{on } \partial\Omega_N\,. \tag{4.5d}$$

Equation (4.5) is called a "reduced order form" of Eq. (4.3). This term refers to the fact that we stated a second order differential equation as a set of first order differential equations. For homogeneous Neumann data, a mixed formulation for this set of PDEs is:

Find $\boldsymbol{u}, \phi \in H_0(\text{div}; \Omega, \partial\Omega_N) \times L^2(\Omega)$ s.t. $\forall \boldsymbol{v}, q \in H_0(\text{div}; \Omega, \partial\Omega_N) \times L^2(\Omega)$:

$$\int_\Omega \kappa^{-1} \boldsymbol{u} \cdot \boldsymbol{v} - \int_\Omega \phi \nabla \cdot \boldsymbol{v} = -\int_{\partial\Omega_D} \phi_D \, \boldsymbol{v} \cdot \boldsymbol{n}\,, \tag{4.6a}$$

$$-\int_\Omega \nabla \cdot \boldsymbol{u} \, q = 0\,, \tag{4.6b}$$

where $H_0(\text{div}; \Omega, \partial\Omega_N)$ is the space of vectors with square integrable divergence and vanishing normal components on $\partial\Omega_N$.

From an engineering perspective, the underlying physics is more directly represented in

44

Eq. (4.5) than it was in Eq. (4.3). The first line represents the modeling (constitutive) equation. In the case of Darcy's law it is a reformulation of Newton's second law for porous media flow. The second line is the incompressibility constraint. The Neumann condition can be identified as the imposed inflow/outflow. Based on these observations, there are a number of reasons to favor Eqs. (4.5) and (4.6) over Eq. (4.3) as the basis of analysis [27].

Firstly, in many applications the auxiliary variable is the quantity of interest. Also in this example of porous media flow the velocity profile is likely to be of more interest than the piezometric head. Other examples include the stress/strain in solid mechanics applications to determine material failure and the heat flux in the (convection-)diffusion equation to understand the necessity of installing insulatory devices. In all these examples, the variable of interest does not occur in the primal formulation. To obtain it, the finite element solution would need to be post-processed with numerical differentiation (i.e., via Eq. (4.4)), which reduces the order of accuracy of the scheme.

Another important reason lies in the explicit occurrence of the constraint in the finite element formulation. For the Stokes equation, the alternative would require the use of 'divergence free' finite element approximation spaces. Construction of such spaces on arbitrary meshes is highly complex [101]. For the Poisson equation of Eq. (4.3), the primal formulation would result in a finite element scheme that is not inherently mass conservative [141], for which it would require specialized post-processing [139]. The weak incorporation of the incompressibility constraint serves as a handle for crafting finite element formulations that naturally satisfy such constraints. Suitably chosen spaces would produce solutions that are pointwise divergence free [77, 78, 79, 132]. This is a vivid branch of research that goes under various names: structure preserving methods [82], compatible discretization [6, 26], mimetic methods [36], discrete exterior calculus [71, 102] and finite element exterior calculus [4, 10].

A third reason concerns the imposition of boundary conditions. For a Poisson problem, we typically think of the Dirichlet conditions as the essential conditions. Indeed, these are the ones that we may impose on $H^1(\Omega)$ through the trace operator of Definition 2.9 to produce $H_0^1(\Omega, \partial\Omega_D)$. In the mixed form of Eq. (4.6), the primal variable is a member of $L^2(\Omega)$, for which functions do not admit a unique trace. We are thus not permitted to incorporate Dirichlet data in the space definition. Rather, it shows up naturally as right-hand side data in the weak statement. The space for the auxiliary variable *does* permit a trace operation: one that defines the normal component on the boundary. This is precisely the operation required to impose the Neumann condition, which thus becomes that essential condition that we can include strongly in the definition of the function space. The swap of the essential and natural conditions may result in higher quality finite element

approximation. For instance, when the Dirichlet conditions result in sharp layers, strong imposition tends to produce oscillatory solutions. Weakly imposed conditions are more 'forgiving' due to the relaxation of the constraint.

Finally, a reduced order form enables the use of finite element approximation spaces that would otherwise fail. For example, PDEs that involve high orders of derivatives, such as the fourth order biharmonic operator used to describe the deflection of thin shells or the fourth order Cahn-Hilliard equation to describe phase-field evolution, cannot directly be solved by using off-the-shelf $\mathcal{C}^0$-continuous finite element function spaces. By writing the partial differential equation in a reduced order form, the existing technology is once again suitable [25, 110, 166, 174, 175, 199].

4.1.2 Weak enforcement of constraints

Classically, variational formulations correspond to minimization problems:

$$\phi = \underset{\phi \in \mathcal{W}}{\arg\inf}\, J(\phi) := \underset{\phi \in \mathcal{W}}{\arg\inf} \left\{ \frac{1}{2} a(\phi, \phi) - L(\phi) \right\}, \tag{4.7}$$

where $a(\cdot, \cdot)$ is some symmetric coercive bilinear form and $L(\cdot)$ some linear form. Suppose that we are looking for the function that minimizes the potential functional $J(\phi)$, under the condition that ϕ satisfies a constraint. Examples could be essential boundary conditions [11], coupling conditions [12, 93] and incompressibility. We could incorporate these constraints in the minimization through the admissible functions:

$$\phi = \underset{\substack{\phi \in \mathcal{W} \\ B\phi = 0}}{\arg\inf}\, J(\phi), \tag{4.8}$$

where B is the operator corresponding to the pairing $\langle B\phi, q \rangle = b(\phi, q) = 0 \; \forall\, q \in \mathcal{Q}$ that represents the constraint. This is perfectly valid, and often leads to well-posed formulations. In some cases, this is precisely the adopted procedure: essential conditions are often imposed nodally in finite element implementations. For other constraints, Eq. (4.8) leads to very complex definitions of the function spaces for which it is challenging to find finite element subspaces.

Instead, we may incorporate the constraint weakly. We rephrase the constrained

minimization [34]:

$$\phi = \arg\inf_{\phi \in \mathcal{W}} \left\{ \sup_{q \in \mathcal{Q}} \left(J(\phi) + b(\phi, q) \right) \right\}. \tag{4.9}$$

For any function that does not satisfy the constraint, the supremum over $\mathcal{Q}$ will produce an infinite value. The infimum over $\mathcal{W}$ must thus be some function that *does* satisfy the constraint. Note that this only defines the function ϕ; the functions $q \in \mathcal{Q}$ just represent more independent test functions that ensure the constraint is enforced. To arrive at a 'square' system, we must introduce another solution variable in $\mathcal{Q}$. We can compose a "dual problem" as:

$$\lambda = \arg\sup_{\lambda \in \mathcal{Q}} \left\{ \inf_{w \in \mathcal{W}} \left(J(w) + b(w, \lambda) \right) \right\}. \tag{4.10}$$

The solution to this dual problem is the Lagrange multipliers associated with the constraint.

We define a new potential $J^*(w, q) = J(w) + b(w, q)$. The convexity of the potential $J(w)$ and the linearity of the constraint $b(w, q)$ allows us to make the following statement about the ϕ and λ that solve Eqs. (4.9) and (4.10) [27]:

$$J^*(\phi, q) \leq J^*(\phi, \lambda) \leq J^*(w, \lambda) \qquad \forall\, w \in \mathcal{W} \text{ and } \forall\, q \in \mathcal{Q}. \tag{4.11}$$

The solution pair ϕ, λ thus represents a "saddle point" or "equilibrium point" of $J^*(w, q)$. At an equilibrium point the Gâteaux derivative of $J^*(w, q)$ vanishes for both the w and the q fields. We are thereby lead to search for the ϕ and λ that correspond to the variational statement:

Find $\phi, \lambda \in \mathcal{W} \times \mathcal{Q}$ s.t. $\forall\, w, q \in \mathcal{W} \times \mathcal{Q}$:

$$a(w, \phi) + b(w, \lambda) = L(w), \tag{4.12a}$$

$$b(\phi, q) = 0. \tag{4.12b}$$

We are again left with a weak formulation of the form of Eq. (2.45), for which well-posedness depends on Theorem 2.4.

The example of the Stokes equations and Darcy equations in the previous section were not by accident. Both mixed formulations may also be derived from a constrained

minimization perspective. In both cases, the incompressibility constraint gives:

$$b(\boldsymbol{v}, q) = -\int_\Omega \nabla \cdot \boldsymbol{v} \; q \,. \tag{4.13}$$

For the Stokes equations:

$$J(\boldsymbol{v}) = \int_\Omega \mu \nabla^s \boldsymbol{v} : \nabla^s \boldsymbol{v} - \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} - \int_{\partial\Omega_N} \boldsymbol{t} \cdot \boldsymbol{v} \,, \tag{4.14}$$

and for the Darcy equations:

$$J(\boldsymbol{v}) = \int_\Omega \frac{1}{2} \kappa^{-1} \boldsymbol{v} \cdot \boldsymbol{v} - \int_{\partial\Omega_D} \phi_D \, \boldsymbol{v} \cdot \boldsymbol{n} \,, \tag{4.15}$$

where the admissible functions $\boldsymbol{u}$ must satisfy the essential Neumann condition, Eq. (4.5d).

This alternative derivation of the same formulations illustrates the tight relation between variational principles and PDEs, also in the context of mixed formulations.

Remark 4.1. *We will repeatedly need to define the saddle points ϕ, λ of some functional $J^*(w, q)$. To avoid the cumbersome notation of Eqs. (4.9) and (4.10), we make use of the following simplified notation:*

$$\phi, \lambda = \arg\inf_{\phi \in \mathcal{W}} \, \arg\sup_{\lambda \in \mathcal{Q}} J^*(\phi, \lambda) \,, \tag{4.16}$$

which should thus be understood in the sense of:

$$\phi = \arg\inf_{\phi \in \mathcal{W}} \left\{ \sup_{q \in \mathcal{Q}} J^*(\phi, q) \right\}, \tag{4.17a}$$

$$\lambda = \arg\sup_{\lambda \in \mathcal{Q}} \left\{ \inf_{w \in \mathcal{W}} J^*(w, \lambda) \right\}. \tag{4.17b}$$

4.1.3 Finite element spaces

Finite element methods based on reduced order forms of PDEs were first proposed in the 1960s [84, 85, 86]. It soon became clear that stable results could not be guaranteed without a careful selection of the discrete subspaces in the Galerkin approximation. For the examples of Eqs. (4.2) and (4.6), this means that we are restricted in our choices of

finite element spaces for the pressure and velocity, or else spurious oscillations spoil the numerical results [27, 29, 83]. We now know that the pair of spaces must satisfy the inf-sup condition from Theorem 2.4.

Early research thus focused on finding suitable pairs of approximation spaces. In the case of the Poisson problem of Eq. (4.6), this required constructing finite-dimensional subspaces of $H(\text{div}, \Omega)$, which lead to a number of new finite elements. The famous examples are the Raviart-Thomas finite element [158]:

$$
\begin{aligned}
RT_p(g_N) = \{\boldsymbol{v}^h \in H(\text{div}, \Omega) : \boldsymbol{v}^h\big|_K \in [\mathbb{P}^p(K)]^d + \boldsymbol{x}\, \tilde{\mathbb{P}}^p(K) \;\; \forall K \in \mathcal{T}, \\
\boldsymbol{v} \cdot \boldsymbol{n} = \mathcal{I}_h g_N \text{ on } \partial\Omega_N\},
\end{aligned}
\tag{4.18}
$$

and the Brezzi-Douglas-Marini family of finite elements [31, 32]:

$$
\begin{aligned}
BDM_p(g_N) = \{\boldsymbol{v}^h \in H(\text{div}, \Omega) : \boldsymbol{v}^h\big|_K \in [\mathbb{P}^p(K)]^d \;\; \forall K \in \mathcal{T}, \\
\boldsymbol{v} \cdot \boldsymbol{n} = \mathcal{I}_h g_N \text{ on } \partial\Omega_N\}.
\end{aligned}
\tag{4.19}
$$

In these space definitions, $\mathbb{P}^p(K)$ is the space of polynomials and $\tilde{\mathbb{P}}$ the space of homogeneous polynomials, both of order p and defined on the volume K. Each K represents an element of the mesh $\mathcal{T}$. The operator $\mathcal{I}_h$ is defined as the interpolation associated with the definition of the degrees of freedom in the finite element triplet [44, 76]. Early error and stability analysis of finite element formulations with these approximation spaces was performed in [7].

The mixed Stokes formulation also involves gradient operators, such that the velocity field must lie in $\left[H_0^1(\Omega, \partial\Omega_D)\right]^d$. Various pressure/velocity approximation pairs have been experimented with [27, 110]. Among those with proven stability characteristics are the Taylor-Hood elements [28, 33, 182] and the Crouzeix-Raviart elements [43, 69]. Pairs that show spurious oscillations in the pressure field may still be used if the solution is post-processed [111], or by adding pressure stabilization [116].

4.2 Multiscale mixed formulations and suitable projectors

The mixed formulation of the Stokes and Poisson problems of Eqs. (4.2) and (4.6) are well-posed, but the use of an improper finite element discretization gives unstable results. It is thus again the loss of fine-scale information that results in poor solution behavior. We turn to the variational multiscale method to remedy the situation.

In Section 3.2.3, we showed that a Galerkin approximation of an energy minimization yields the same finite element formulation as a variational multiscale formulation with as a choice of projector the minimizer of the energy norm. Similarly, we can interpret the Galerkin approximations of mixed formulations as variational multiscale formulations with specific projection operators. Stable formulations for 'bad' finite elements may still be obtained by altering the projection operator. We illustrate this procedure for the Stokes problem. The first step is a variational multiscale decomposition of the weak form. Some care must be taken in regards to the definition of the fine scales, as emphasized by the following lemma.

Lemma 4.1. *Let $\mathscr{P}$ be an arbitrary surjective projector that maps as follows:*

$$\mathscr{P} : \mathcal{W} \times \Sigma \to (\mathcal{W} \times \Sigma)^h := \mathcal{W}^h \times \Sigma^h . \tag{4.20}$$

Despite the decoupling of the range of $\mathscr{P}$ into $\mathcal{W}^h$ and Σ^h, it can not be guaranteed that the kernel decouples as such. That is, in general:

$$\nexists \ \mathcal{W}', \Sigma' \ s.t. \ \ker(\mathscr{P}) = \mathcal{W}' \times \Sigma' . \tag{4.21}$$

Proof. The projection of some pair $(w, 0)$ maps to a w-dependent point in the range of $\mathscr{P}$:

$$\mathscr{P}(w, 0) =: (w_w^h, \sigma_w^h) \in \mathcal{W}^h \times \Sigma^h . \tag{4.22}$$

The fine-scale supplement lies in the kernel of $\mathscr{P}$:

$$(\mathscr{I} - \mathscr{P})(w, 0) = (w - w_w^h, -\sigma_w^h) \in \ker(\mathscr{P}) . \tag{4.23}$$

From Eq. (4.22) we know that $\sigma_w^h \in \Sigma^h$ and thus that $(0, \sigma_w^h) \in \mathcal{W}^h \times \Sigma^h$. Due to the idempotency and surjectivity of $\mathscr{P}$, this solution pair maps to itself:

$$\mathscr{P}(0, \sigma_w^h) = (0, \sigma_w^h) . \tag{4.24}$$

If $\ker(\mathscr{P}) = \mathcal{W}' \times \Sigma'$, then from Eq. (4.23) $-\sigma_w^h \in \Sigma'$, which would mean:

$$\mathscr{P}(0, \sigma_w^h) = (0, 0) \,. \tag{4.25}$$

Equation (4.25) contradicts with Eq. (4.24) except for the special case that $\sigma_w^h = 0$. This would only hold for all w if $\mathscr{P}$ decouples as:

$$\mathscr{P}(w, \sigma) = (\mathscr{P}_w \, w, \mathscr{P}_\sigma \sigma) \,, \tag{4.26}$$

which is only a special case of projector $\mathscr{P}$. $\qquad\square$

With the conclusion of this lemma in mind, we obtain the following multiscale decomposition of the Stokes formulation:

Find $\boldsymbol{u}^h, p^h \in \boldsymbol{\mathcal{V}}^h \times \mathcal{Q}^h$ and $\boldsymbol{u}', p' \in (\boldsymbol{\mathcal{V}} \times \mathcal{Q})'$
$\qquad$ s.t. $\forall \, \boldsymbol{v}^h, q^h \in \boldsymbol{\mathcal{V}}^h \times \mathcal{Q}^h$ and $\boldsymbol{v}', q' \in (\boldsymbol{\mathcal{V}} \times \mathcal{Q})'$:

$$\int_\Omega 2\mu \nabla^s (\boldsymbol{u}^h + \boldsymbol{u}') : \nabla^s \boldsymbol{v}^h - \int_\Omega (p^h + p') \, \nabla \cdot \boldsymbol{v}^h = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}^h + \int_{\partial\Omega_D} \boldsymbol{t} \cdot \boldsymbol{v}^h \,,$$

$$\int_\Omega 2\mu \nabla^s (\boldsymbol{u}^h + \boldsymbol{u}') : \nabla^s \boldsymbol{v}' - \int_\Omega (p^h + p') \, \nabla \cdot \boldsymbol{v}' = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}' + \int_{\partial\Omega_D} \boldsymbol{t} \cdot \boldsymbol{v}' \,,$$

$$- \int_\Omega \nabla \cdot (\boldsymbol{u}^h + \boldsymbol{u}') \, q^h = 0 \,,$$

$$- \int_\Omega \nabla \cdot (\boldsymbol{u}^h + \boldsymbol{u}') \, q' = 0 \,,$$

where:

$$\boldsymbol{\mathcal{V}} \times \mathcal{Q} := \left[H_0^1(\Omega, \partial\Omega_D) \right]^d \times L^2(\Omega) = (\boldsymbol{\mathcal{V}}^h \times \mathcal{Q}^h) \oplus (\boldsymbol{\mathcal{V}} \times \mathcal{Q})' \,. \tag{4.28}$$

Suppose that we choose the following optimality condition as our projection operator:

$$\mathscr{P} : \boldsymbol{\mathcal{V}} \times \mathcal{Q} \to \boldsymbol{\mathcal{V}}^h \times \mathcal{Q}^h$$

$$\boldsymbol{u}, p \ \mapsto \ \underset{\boldsymbol{u}^h \in \boldsymbol{\mathcal{V}}^h}{\arg\inf} \underset{p^h \in \mathcal{Q}^h}{\arg\sup} \int_\Omega \mu \nabla^s (\boldsymbol{u} - \boldsymbol{u}^h) : \nabla^s (\boldsymbol{u} - \boldsymbol{u}^h) \qquad (4.29)$$

$$- \int_\Omega \nabla \cdot (\boldsymbol{u} - \boldsymbol{u}^h) \, (p - p^h) \,.$$

Following the procedure established in Section 3.2.3, we rewrite this projector as a set of constraint functionals. These correspond to the variational statement that defines the equilibrium point of the saddle-point problem:

$$- \left(2\mu\nabla^s \boldsymbol{\nu}^h, \nabla^s(\boldsymbol{u} - \boldsymbol{u}^h)\right)_\Omega + \left(\nabla \cdot \boldsymbol{\nu}^h, p - p^h\right)_\Omega = 0 \qquad \forall \boldsymbol{\nu}^h \in \boldsymbol{\mathcal{V}}^h \,, \tag{4.30a}$$

$$\left(\nabla \cdot (\boldsymbol{u} - \boldsymbol{u}^h), \rho^h\right)_\Omega = 0 \qquad\qquad\qquad \forall \rho^h \in \mathcal{Q}^h \,. \tag{4.30b}$$

These functionals define the fine-scale space $(\boldsymbol{\mathcal{V}} \times \mathcal{Q})' = \ker \mathscr{P}$ as all pairs $\boldsymbol{u}', p'$ that satisfy:

$$- \left(2\mu\nabla^s \boldsymbol{\nu}^h, \nabla^s \boldsymbol{u}'\right)_\Omega + \left(\nabla \cdot \boldsymbol{\nu}^h, p'\right)_\Omega = 0 \qquad\qquad \forall \boldsymbol{\nu}^h \in \boldsymbol{\mathcal{V}}^h \,, \tag{4.31a}$$

$$\left(\nabla \cdot \boldsymbol{u}', \rho^h\right)_\Omega = 0 \qquad\qquad\qquad\qquad \forall \rho^h \in \mathcal{Q}^h \,. \tag{4.31b}$$

Since *all* fine-scale pairs $\boldsymbol{u}', p'$ satisfy these relations for *all* test functions $\boldsymbol{\nu}^h, \rho^h$, we may choose $\boldsymbol{\nu}^h = \boldsymbol{v}^h$ and $\rho^h = q^h$ and add Eq. (4.31a) to Eqs. (4.27a) and (4.27c). This elimination of the fine scales, directly implied by the chosen projector, leaves us with a conformal Galerkin approximation of the Stokes problem:

Find $\boldsymbol{u}^h, p^h \in \boldsymbol{\mathcal{V}}^h \times \mathcal{Q}^h$ s.t. $\forall \boldsymbol{v}^h, q^h \in \boldsymbol{\mathcal{V}}^h \times \mathcal{Q}^h$:

$$\int_\Omega 2\mu\nabla^s \boldsymbol{u}^h : \nabla^s \boldsymbol{v}^h - \int_\Omega p^h \, \nabla \cdot \boldsymbol{v}^h = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}^h + \int_{\partial\Omega_D} \boldsymbol{t} \cdot \boldsymbol{v}^h \,, \tag{4.32a}$$

$$- \int_\Omega \nabla \cdot \boldsymbol{u}^h \, q^h = 0 \,. \tag{4.32b}$$

Stated inversely, solutions to the conformal Galerkin approximation of the Stokes problem are projections of the exact solution with Eq. (4.29) as the projection operator. Since the continuous weak form is well-posed, well-posedness of the coarse-scale problem depends on well-posedness of the projection operation. Apparently, this is not the case for arbitrary sets of finite element approximation spaces. We have two options: either we follow the approach of Section 4.1.3 and search for stable discretization pairs, or we change the projection operator.

For many unstable solution pairs the problem lies in the pressure field. If it permits a 'zero-energy' pressure solution, then the stiffness matrix becomes singular. If, with the use of iterative solvers, we still manage to find an inverse, then these zero-energy modes show up as (checkerboard-like) spurious oscillations. Suppose that we wish to use an equal-order $\mathcal{C}^0$-continuous interpolation pair for the velocities and the pressures. Such a discretization

is notorious for suffering from spurious oscillations. We can remove this zero-energy mode by adding an H_0^1-weighting in the projection operator:

$$\mathscr{P} : \mathcal{V} \times \mathcal{Q} \to \mathcal{V}^h \times \mathcal{Q}^h$$

$$\boldsymbol{u}, p \quad \mapsto \underset{\boldsymbol{u}^h \in \mathcal{V}^h}{\arg\inf} \underset{p^h \in \mathcal{Q}^h}{\arg\sup} \int_\Omega \mu \nabla^s (\boldsymbol{u} - \boldsymbol{u}^h) : \nabla^s (\boldsymbol{u} - \boldsymbol{u}^h)$$

$$- \int_\Omega \nabla \cdot (\boldsymbol{u} - \boldsymbol{u}^h) \, (p - p^h) - \int_\Omega \tfrac{1}{2} \gamma \nabla (p - p^h) \cdot \nabla (p - p^h) \, . \tag{4.33}$$

In order for the gradient to be a valid operation on every $p \in \mathcal{Q}$ we must change the space $\mathcal{Q}$ to $H^1(\Omega)$. Since this is more constrained than the original $L^2(\Omega)$, the second LBB condition is still valid and the original weak statement is still well-posed.

Going through the same steps would result in the following coarse-scale equations:

Find $\boldsymbol{u}^h, p^h \in \mathcal{V}^h \times \mathcal{Q}^h$ s.t. $\forall \boldsymbol{v}^h, q^h \in \mathcal{V}^h \times \mathcal{Q}^h$:

$$\int_\Omega 2\mu \nabla^s \boldsymbol{u}^h : \nabla^s \boldsymbol{v}^h - \int_\Omega p^h \, \nabla \cdot \boldsymbol{v}^h = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}^h + \int_{\partial\Omega_D} \boldsymbol{t} \cdot \boldsymbol{v}^h \, , \tag{4.34a}$$

$$- \int_\Omega \nabla \cdot \boldsymbol{u}^h \, q^h + \int_\Omega \gamma \, \nabla p' \cdot \nabla q^h \ = 0 \, , \tag{4.34b}$$

where a fine-scale pressure solution remains. We can model this with a residual-based model originating from the momentum equations. This results in the 'stabilized' weak formulation of the Stokes equation [116]:

Find $\boldsymbol{u}^h, p^h \in \mathcal{V}^h \times \mathcal{Q}^h$ s.t. $\forall \boldsymbol{v}^h, q^h \in \mathcal{V}^h \times \mathcal{Q}^h$:

$$\int_\Omega 2\mu \nabla^s \boldsymbol{u}^h : \nabla^s \boldsymbol{v}^h - \int_\Omega p^h \, \nabla \cdot \boldsymbol{v}^h = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}^h + \int_{\partial\Omega_D} \boldsymbol{t} \cdot \boldsymbol{v}^h \, , \tag{4.35a}$$

$$- \int_\Omega \nabla \cdot \boldsymbol{u}^h \, q^h + \int_\Omega \tau_{\mathrm{PSPG}} \boldsymbol{\mathcal{R}}_{\mathrm{M}} \cdot \nabla q^h \ = 0 \, , \tag{4.35b}$$

where $\boldsymbol{\mathcal{R}}_{\mathrm{M}}$ is the residual of the momentum equations and τ_{PSPG} is the pressure-stabilized/Petrov-Galerkin stabilization parameter. It relates to the homogenized fine-scale Green's operator (see Section 3.3) and to the weighting factor γ in Eq. (4.33).

These steps exemplify a procedure that we use repeatedly in Parts II and III of this dissertation. We decompose a weak formulation into coarse-scale and fine-scale components, and find the projector that leads to the Galerkin approximation in the elliptic case.

This projector may not be suitable for the approximation space under consideration, so we add a weighting of an error to the error potential. Then, we use the optimality conditions of the error potential, equivalently interpretable as the constraint equations that define the fine-scale space, to eliminate many of the fine scales in the coarse-scale equations.

As we focus on advection-diffusion model problems, the relevant elliptic case is the Poisson equation. For the mixed formulation, the multiscale decomposition reads:

Find $\boldsymbol{u}^h, \phi^h \in \boldsymbol{\mathcal{V}}^h \times \mathcal{Q}^h$ and $\boldsymbol{u}', \phi' \in (\boldsymbol{\mathcal{V}} \times \mathcal{Q})'$

$\qquad$ s.t. $\forall \boldsymbol{v}^h, q^h \in \boldsymbol{\mathcal{V}}^h \times \mathcal{Q}^h$ and $\boldsymbol{v}', q' \in (\boldsymbol{\mathcal{V}} \times \mathcal{Q})'$:

$$\int_\Omega \kappa^{-1}(\boldsymbol{u}^h + \boldsymbol{u}') \cdot \boldsymbol{v}^h - \int_\Omega (\phi^h + \phi') \, \nabla \cdot \boldsymbol{v}^h = - \int_{\partial\Omega_D} \phi_D \, \boldsymbol{v}^h \cdot \boldsymbol{n} \,, \tag{4.36a}$$

$$\int_\Omega \kappa^{-1}(\boldsymbol{u}^h + \boldsymbol{u}') \cdot \boldsymbol{v}' - \int_\Omega (\phi^h + \phi') \, \nabla \cdot \boldsymbol{v}' = - \int_{\partial\Omega_D} \phi_D \, \boldsymbol{v}' \cdot \boldsymbol{n} \,, \tag{4.36b}$$

$$- \int_\Omega \nabla \cdot (\boldsymbol{u}^h + \boldsymbol{u}') \, q^h = 0 \,, \tag{4.36c}$$

$$- \int_\Omega \nabla \cdot (\boldsymbol{u}^h + \boldsymbol{u}') \, q' = 0 \,, \tag{4.36d}$$

where now:

$$\boldsymbol{\mathcal{V}} \times \mathcal{Q} := H_0(\mathrm{div}; \Omega, \partial\Omega_N) \times L^2(\Omega) = (\boldsymbol{\mathcal{V}}^h \times \mathcal{Q}^h) \oplus (\boldsymbol{\mathcal{V}} \times \mathcal{Q})' \,. \tag{4.37}$$

We obtain a conformal finite element formulation from the following optimality projector:

$$\mathscr{P} : \boldsymbol{\mathcal{V}} \times \mathcal{Q} \to \boldsymbol{\mathcal{V}}^h \times \mathcal{Q}^h$$

$$\boldsymbol{u}, p \;\; \mapsto \;\; \underset{\boldsymbol{u}^h \in \boldsymbol{\mathcal{V}}^h}{\arg\inf} \, \underset{p^h \in \mathcal{Q}^h}{\arg\sup} \int_\Omega \tfrac{1}{2}\kappa^{-1}(\boldsymbol{u} - \boldsymbol{u}^h) \cdot (\boldsymbol{u} - \boldsymbol{u}^h)$$

$$- \int_\Omega \nabla \cdot (\boldsymbol{u} - \boldsymbol{u}^h)(\phi - \phi^h) \,. \tag{4.38}$$

This projection operation is well-posed for a suitable chosen pair of approximation spaces. These may involve the Raviart-Thomas or Brezzi-Douglas-Marini type elements discussed in Section 4.1.3.

By multiplying the error potential by negative one, we are permitted to switch the spaces

associated with the infimum and the supremum:

$$
\begin{aligned}
\mathscr{P} : \mathcal{Q} \times \mathcal{V} &\to \mathcal{Q}^h \times \mathcal{V}^h \\
\phi, \boldsymbol{u} &\mapsto \underset{\phi^h \in \mathcal{Q}^h}{\arg\inf} \underset{\boldsymbol{u}^h \in \mathcal{V}^h}{\arg\sup} \, J^*(\phi', \boldsymbol{u}') \, .
\end{aligned}
\tag{4.39}
$$

In subsequent chapters, we assume κ to be constant such that we may divide the error potential in Eq. (4.38) by it. Additionally, we will add an advection operator such that we require the space $\mathcal{Q}$ to be $H^1(\Omega)$. Through integration by parts, the error potential may then be written as:

$$
J^*(\phi', \boldsymbol{u}') = \int_\Omega \tfrac{1}{2} \|\kappa \nabla \phi'\|^2 - \int_\Omega \tfrac{1}{2} \|\boldsymbol{u}' + \kappa \nabla \phi'\|^2 + \int_{\partial\Omega_D} \boldsymbol{u}' \cdot \boldsymbol{n} \, \kappa \phi' \, ,
\tag{4.40}
$$

where the norms are Euclidean vector norms.

This representation of the projector and the error potential provides an interpretation from an optimization perspective. The first term and the infimum over ϕ concern an H_0^1-norm optimality, and the supremum over $\boldsymbol{u}^h$ concerns an L^2-norm optimality via the second term. For both optimality conditions the boundary term acts like a source of error.

Conclusion

In this part of the dissertation, we have laid out the foundation on which we build in Parts II and III. There, we adopt the variational multiscale principles for 'non-standard' formulations. In Part II we focus on the weak imposition of essential conditions by means of Nitsche's method, and in Part III we consider a multitude of discontinuous Galerkin formulations. These finite element methods are often referred to as "non-conformal", as the approximation spaces are not subspaces of the function spaces relevant to the continuous weak formulation. Additionally, the added flexibility of the approximation spaces at element or domain boundaries raises question regarding the inversion of the fine-scale problem and the ensuing residual-based models. With the detailed understanding of the original variational multiscale method developed in Part I we can pinpoint exactly where the challenges lie and proceed to tackle them.

PART II

Weak imposition of essential boundary conditions

Introduction

From the previous introductory chapters it should be clear that the variational multiscale paradigm plays a central role for approximating the solution of flow related PDEs with finite element methods. The resulting residual-based models encapsulate stabilized methods and turbulence models. These principles have stood at the forefront of making the finite element method applicable for fluid-mechanics applications [72].

Another methodology that has been shown to yield favorable results for CFD applications is the use of weakly imposed Dirichlet boundary conditions [22, 23, 24, 65, 157]. Typically, Nitsche's method is the method of choice for weakly enforcing essential boundary conditions. While Nitsche's method was initially proposed in relation to energy minimization functionals [152], both its symmetric and nonsymmetric variants have since been studied extensively in fluid-mechanics applications [22, 23, 24, 40]. One of the main drivers for the significant recent interest in weakly enforced boundary conditions is their importance in immersed finite element methods [163, 165, 167, 168]. Notable references in the context of fluid mechanics include [21, 103, 106, 107, 131, 197, 198]. For immersed finite element methods, the approximation space is no longer tailored to fit the domain boundary and the boundary conditions can not easily be enforced strongly in the approximation space. Hence, there is a need for the weak enforcement of the essential boundary conditions by manipulation of the variational formulation.

At first glance, the variational multiscale method and Nitsche's method appear to be at conflict: the basis for a variational multiscale decomposition is a well-posed continuous weak formulation, but Nitsche's method requires flexible spaces at Dirichlet boundaries and involves penalty terms that become unbounded in the continuous limit. Additionally, the fine-scale solution, by design, does not vanish on the Dirichlet boundary, which, as discussed in Section 3.3, violates one of the key assumptions on which traditional residual-based fine-scale models are built. The next few chapters embody our recent efforts for mitigating these issues [177, 178, 179].

This part of the dissertation is structured as follows. In Chapter 5, we derive a variational multiscale finite element formulation of the advection-diffusion equation and we show that Nitsche's method arises from a particular choice of fine-scale closure. In Chapter 6, we develop the fine-scale model that takes into account the non-vanishing fine scales at the Dirichlet boundary and provide estimates for the involved model parameters. The complete formulation is summarized in Section 6.6, where we also show that the resulting bilinear form is coercive and hence that the finite element formulation permits a unique solution. In Section 7.1, we verify the theory for a one-dimensional model problem, and in Section 7.2 we computationally investigate the performance for a two-dimensional model problem that involves multiple boundary layers.

The continuous problem and its variational multiscale decomposition

5.1 Multiscale weak formulation with flexible spaces at Dirichlet boundaries

The classical model problem for variational multiscale analysis is the steady advection-diffusion equation. Let $\Omega \subset \mathbb{R}^d$ denote the spatial domain with boundary $\partial\Omega$. The governing equations in strong form read:

$$\boldsymbol{a} \cdot \nabla\phi - \nabla \cdot \kappa\nabla\phi = f \quad \text{in} \quad \Omega, \tag{5.1a}$$

$$\phi = \phi_D \quad \text{on} \quad \partial\Omega_D, \tag{5.1b}$$

$$\kappa\,\partial_n\phi = g_N \quad \text{on} \quad \partial\Omega_N^+, \tag{5.1c}$$

$$\kappa\,\partial_n\phi - \boldsymbol{a} \cdot \boldsymbol{n}\phi = g_N \quad \text{on} \quad \partial\Omega_N^-, \tag{5.1d}$$

where the dependent variable $\phi = \phi(\boldsymbol{x})$ maps Ω into $\mathbb{R}$. The source function $f : \Omega \to \mathbb{R}$, the Dirichlet data $\phi_D : \partial\Omega_D \to \mathbb{R}$ and the Neumann (or Robin) data $g_N : \partial\Omega_N \to \mathbb{R}$ are exogenous functions that are assumed to be members of $H^{1/2}(\cdot)$ on their respective domains. The advective velocity $\boldsymbol{a} = \boldsymbol{a}(\boldsymbol{x})$ is a given solenoidal vector field ($\nabla \cdot \boldsymbol{a} = 0$) and the diffusivity κ is strictly positive. The Dirichlet and Neumann (or Robin) boundaries $\partial\Omega_D$ and $\partial\Omega_N$ are complementary subsets of the boundary $\partial\Omega$, i.e., $\partial\Omega = \overline{\partial\Omega_D} \cup \overline{\partial\Omega_N}$,

and the superscripts $+$ and $-$ indicate outflow $(\boldsymbol{a} \cdot \boldsymbol{n} \geq 0)$ and inflow parts $(\boldsymbol{a} \cdot \boldsymbol{n} < 0)$ respectively. On the boundary, the normal gradient is denoted $\partial_n \phi = \boldsymbol{n} \cdot \nabla \phi$. By convention, $\boldsymbol{n}$ denotes the outward facing unit normal vector.

To obtain the weak formulation we multiply by a test function and integrate by parts wherever suitable. Different from classical functional (and variational multiscale) analysis we keep the traces of our function spaces on the domain boundary variable. This requires the use of Lagrange multipliers for the enforcement of the Dirichlet boundary conditions. To ensure inf-sup stability of the resulting bilinear form, we substitute the known data on the inflow part of the Dirichlet boundary in the advective term. We then obtain the following weak formulation:

Find $\phi \in \mathcal{W}$ and $\lambda \in \mathcal{Q}$ s.t. $\forall \, w \in \mathcal{W}$ and $q \in \mathcal{Q}:$

$$
\begin{aligned}
- \left(\boldsymbol{a} \cdot \nabla w, \phi\right)_{\Omega} + \left\langle \boldsymbol{a} \cdot \boldsymbol{n} \, w, \phi \right\rangle_{\partial\Omega^+} &+ \left(\nabla w, \kappa \nabla \phi\right)_{\Omega} + \left\langle \lambda, w \right\rangle_{\partial\Omega_D} \\
&= (w, f)_{\Omega} + \left\langle w, g_N \right\rangle_{\partial\Omega_N} - \left\langle \boldsymbol{a} \cdot \boldsymbol{n} \, w, \phi_D \right\rangle_{\partial\Omega_D^-},
\end{aligned}
\tag{5.2a}
$$

$$
\left\langle q, \phi \right\rangle_{\partial\Omega_D} = \left\langle q, \phi_D \right\rangle_{\partial\Omega_D},
\tag{5.2b}
$$

where $(\,\cdot\,,\,\cdot\,)_{\Omega}$ denotes the L^2-inner product on domain Ω, and $\left\langle\,\cdot\,,\,\cdot\,\right\rangle_{\partial\Omega}$ denotes a duality pairing (see Definition 2.12), which is to be interpreted in the sense of an L^2-inner product on the surface. In its current form, the suitable functional spaces may be identified as $\mathcal{W} = H^1(\Omega)$ and $\mathcal{Q} = H^{-1/2}(\partial\Omega_D)$. Equivalence of the strong and weak forms, Eqs. (5.1) and (5.2), dictates that $\lambda = -\kappa\partial_n\phi$. This is stated in the following lemma, which also concerns the well-posedness of the weak formulation.

Lemma 5.1. *Equation* (5.2) *is well-posed, given sufficient regularity of the data and the domain it is equivalent to the strong form of Eq.* (5.1)*, and from this equivalence it follows that* $\lambda = -\kappa\partial_n\phi$.

Proof. This is a classical example of a mixed formulation, and the theory outlined in Section 2.4 largely applies. Well-posedness (or "inf-sup stability" in the current context) follows from Theorem 2.4, the LBB theorem. To make this theorem completely applicable we restate Eq. (5.2) as a mixed formulation with homogeneous essential conditions. We replace ϕ by $\phi_d + \phi_0$, where ϕ_d is *some* function in $H^1(\Omega)$ whose trace is equal to ϕ_D on $\partial\Omega_D$. Such a function exists, as ϕ_D was assumed to be a member of $H^{1/2}(\partial\Omega_D)$, a trace

space of $H^1(\Omega)$. The new weak formulation may be written as:

Find $\phi_0, \lambda \in \mathcal{W} \times \mathcal{Q}$ s.t. $\forall\, w, q \in \mathcal{W} \times \mathcal{Q}$:

$$a(w, \phi_0) + b(w, \lambda) = L(w) - a(w, \phi_D)\,,$$

$$b(\phi_0, q) = 0\,.$$

The LBB theorem requires that $a(\cdot, \cdot)$ is coercive on the kernel of $b(\cdot, \cdot)$ and that $b(\cdot, \cdot)$ is inf-sup stable. The inf-sup stability of $b(\cdot, \cdot)$ follows from the choice $\phi_0 = q$ on $\partial\Omega_D$. The kernel of $b(\cdot, \cdot)$ is denoted $H_0^1(\Omega, \partial\Omega_D)$ and consists of all functions in $H^1(\Omega)$ whose trace is zero on $\partial\Omega_D$. The bilinear form $a(\cdot, \cdot)$ is coercive on this space: the diffusive term itself is coercive as the H^1-seminorm is equivalent to the H^1-norm for the $H_0^1(\Omega, \partial\Omega_D)$ space of functions, and the advective terms only positively contribute to the coercive as is proven later in Lemma 6.2. Equation (5.2) is thus a well-posed weak formulation.

Equivalence with the strong form follows from integration by parts while making use of $\nabla \cdot \boldsymbol{a} = 0$. After clustering of the integrals we obtain:

$$\left(w, \boldsymbol{a}\nabla\phi - \nabla \cdot \kappa\nabla\phi - f\right)_\Omega - \left\langle w, g_N + \boldsymbol{a}\cdot\boldsymbol{n}\,\phi - \kappa\partial_n\phi\right\rangle_{\partial\Omega_N^-} - \left\langle w, g_N - \kappa\partial_n\phi\right\rangle_{\partial\Omega_N^+}$$

$$+ \left\langle w, \boldsymbol{a}\cdot\boldsymbol{n}\left(\phi_D - \phi\right) + \lambda + \kappa\partial_n\phi\right\rangle_{\partial\Omega_D^-} + \left\langle w, \lambda + \kappa\partial_n\phi\right\rangle_{\partial\Omega_D^+} = 0\,. \qquad (5.4)$$

As this holds for all $w \in H^1(\Omega)$, each individual term must vanish. This leads to the volumetric and Neumann parts of the strong form: Eqs. (5.1a), (5.1c) and (5.1d). Additionally, from the choice of $q = \phi - \phi_D$ in Eq. (5.2b) it follows that $\phi = \phi_D$ on $\partial\Omega_D$. The terms in Eq. (5.4) on the Dirichlet boundary then produce $\lambda = -\kappa\partial_n\phi$. $\qquad\square$

The variational multiscale approach splits the trial solution and test function spaces into coarse and fine scales. The coarse scales live on the finite element grid, whereas the fine scales are determined via a model equation. This decomposition may be written as:

$$\mathcal{W} = \mathcal{W}^h \oplus \mathcal{W}',\qquad\qquad (5.5)$$

where $\mathcal{W}^h$ is the space spanned by the finite-dimensional discretization and the fine-scale space $\mathcal{W}'$ is an infinite-dimensional complement in $\mathcal{W}$. The components of the solutions

and test functions decouple as

$$\phi = \phi^h + \phi', \tag{5.6a}$$

$$w = w^h + w', \tag{5.6b}$$

with coarse scales $\phi^h, w^h \in \mathcal{W}^h$ and fine scales $\phi', w' \in \mathcal{W}'$. The direct sum decomposition in Eq. (5.5) is associated with a projection operator:

$$w^h = \mathscr{P}w \in \mathcal{W}^h, \tag{5.7a}$$

$$w' = (\mathscr{I} - \mathscr{P})\, w \in \mathcal{W}', \tag{5.7b}$$

where $\mathscr{P} : \mathcal{W} \to \mathcal{W}^h$ is the projector and $\mathscr{I} : \mathcal{W} \to \mathcal{W}$ is the identity operator. Formally, this projector is incorporated in the weak formulation through the definition of the fine-scale space $\mathcal{W}'$. It follows from Eq. (5.7b) that $\mathcal{W}' = \mathrm{Im}(\mathscr{I} - \mathscr{P}) = \ker\mathscr{P}$. Then, the direct sum decomposition of Eq. (5.5) ensures the unique decomposition that satisfies Eqs. (5.6) and (5.7). Using this multiscale split we arrive at the following alternative – equivalent – weak statement:

Find $\phi^h \in \mathcal{W}^h, \phi' \in \mathcal{W}', \lambda \in \mathcal{Q}$ s.t. $\forall w^h \in \mathcal{W}^h, w' \in \mathcal{W}', q \in \mathcal{Q}$:

$$\begin{aligned}
-\left(\boldsymbol{a} \cdot \nabla w^h, \phi^h + \phi'\right)_\Omega &+ \left(\nabla w^h, \kappa\nabla\phi^h + \kappa\nabla\phi'\right)_\Omega + \left\langle \boldsymbol{a}\cdot\boldsymbol{n}\, w^h, \phi^h + \phi'\right\rangle_{\partial\Omega^+} \\
&+ \left\langle\lambda, w^h\right\rangle_{\partial\Omega_D} = \left(w^h, f\right)_\Omega + \left\langle w^h, g_N\right\rangle_{\partial\Omega_N} - \left\langle\boldsymbol{a}\cdot\boldsymbol{n}\,w^h, \phi_D\right\rangle_{\partial\Omega_D^-},
\end{aligned} \tag{5.8a}$$

$$\begin{aligned}
-\left(\boldsymbol{a} \cdot \nabla w', \phi^h + \phi'\right)_\Omega &+ \left(\nabla w', \kappa\nabla\phi^h + \kappa\nabla\phi'\right)_\Omega + \left\langle \boldsymbol{a}\cdot\boldsymbol{n}\, w', \phi^h + \phi'\right\rangle_{\partial\Omega^+} \\
&+ \left\langle\lambda, w'\right\rangle_{\partial\Omega_D} = \left(w', f\right)_\Omega + \left\langle w', g_N\right\rangle_{\partial\Omega_N} - \left\langle\boldsymbol{a}\cdot\boldsymbol{n}\,w', \phi_D\right\rangle_{\partial\Omega_D^-},
\end{aligned} \tag{5.8b}$$

$$\left\langle q, \phi^h + \phi'\right\rangle_{\partial\Omega_D} = \left\langle q, \phi_D\right\rangle_{\partial\Omega_D}. \tag{5.8c}$$

Equation (5.8a) is the "coarse-scale problem", and can be interpreted as a relation for ϕ^h for a given λ and ϕ'. The Lagrange multiplier is taken care of by substituting the known expression in the continuous case, viz. $\lambda = -\kappa\partial_n\phi^h - \kappa\partial_n\phi'$. Similarly, the "fine-scale problem" of equation (5.8b) can be conceived as a relation for the fine-scale component $\phi' \in \mathcal{W}'$. This space, however, is infinite-dimensional and is thus not amenable to discrete implementation. Hence, in equation (5.8a) a closure model will be substituted in place of the fine-scale solution.

5.2 Nitsche's method as a partial fine-scale closure

Next, we show that a particular fine-scale closure condition leads to Nitsche's classical formulation. Our goal is to illustrate that Nitsche's method and the accompanying penalty terms are not in conflict with the VMS theory, but rather can be interpreted naturally in the VMS framework as a particular choice of fine-scale closure. Consider the following projection operator, which we will refer to as the Nitsche projector:

$$\mathscr{P}_N : \mathcal{W} \to \mathcal{W}^h$$

$$\phi \;\mapsto\; \underset{\phi^h \in \mathcal{W}^h}{\arg\inf} \int_\Omega \frac{1}{2}\kappa(\nabla\phi - \nabla\phi^h)\cdot(\nabla\phi - \nabla\phi^h) - \int_{\partial\Omega_D} \kappa(\partial_n\phi - \partial_n\phi^h)(\phi - \phi^h)$$
$$+ \int_{\partial\Omega} \frac{1}{2}\kappa\eta^*(\phi - \phi^h)^2\,. \tag{5.9}$$

where η^* is a factor that weighs the added boundary term. We 'choose' to define it as:

$$\eta^* = \begin{cases} 0 & \text{on } \partial\Omega_N^- \,, \\ \frac{\boldsymbol{a}\cdot\boldsymbol{n}}{\kappa} & \text{on } \partial\Omega_N^+ \,, \\ \eta & \text{on } \partial\Omega_D^- \,, \\ \eta + \frac{\boldsymbol{a}\cdot\boldsymbol{n}}{\kappa} & \text{on } \partial\Omega_D^+ \,, \end{cases} \tag{5.10}$$

for which the reason will become clear momentarily.

The functional that is minimized involves an integral of the normal derivative of functions in $\mathcal{W}$ on the boundary of Ω. Recall from Definition 2.9 and the surrounding discussion that this operation is not permitted for functions in $H^1(\Omega)$. In order for these integrals to be well-defined for all $w \in \mathcal{W}$, we require a redefinition of $\mathcal{W}$. We set $\mathcal{W} = \mathcal{V} \cup \mathcal{C}^0_{\text{FE}}(\Omega)$. The space $\mathcal{C}^0_{\text{FE}}(\Omega)$ is whatever $\mathcal{C}^0$-continuous finite element approximation space we identified as $\mathcal{W}^h$, and $\mathcal{V}$ is defined as $\mathcal{V} = \{\phi \in H^1(\Omega) : \Delta\phi \in L^2(\Omega)\}$. The following operation is allowed for functions in $\mathcal{V}$:

$$\int_{\partial\Omega} \partial_n\phi\, v = \int_\Omega \nabla\cdot(\nabla\phi\, v) = \int_\Omega \Delta\phi\, v + \int_\Omega \nabla\phi\cdot\nabla v \quad \forall\, v \in H^1(\Omega)\,, \tag{5.11}$$

which defines $\partial_n\phi$ on the boundary. As we are considering data of the form $f \in L^2(\Omega)$, the true solution ϕ will automatically be an element of this new space.

Given a function $\phi \in \mathcal{W}$, the function $\phi^h \in \mathcal{W}^h$ that minimizes the functional in Eq. (5.9)

may be determined by taking the Gâteaux derivative. This results in the following discrete weak statement

$$- \left(\nabla v^h, \kappa \nabla \phi - \nabla \phi^h \right)_\Omega + \left\langle v^h, \kappa \partial_n \phi - \kappa \partial_n \phi^h \right\rangle_{\partial \Omega_D} + \left\langle \kappa \partial_n v^h, \phi - \phi^h \right\rangle_{\partial \Omega_D}$$
$$- \left\langle \kappa \eta v^h, \phi - \phi^h \right\rangle_{\partial \Omega_D} - \left\langle \boldsymbol{a} \cdot \boldsymbol{n} v^h, \phi - \phi^h \right\rangle_{\partial \Omega^+} = 0 \qquad \forall v^h \in \mathcal{W}^h \,. \tag{5.12}$$

Before continuing, let us first assess that (or when) Eq. (5.9) indeed defines a projection.

Lemma 5.2. *Under the assumption that η is 'large enough', the mapping $\mathscr{P}_N$ is surjective and idempotent, and hence a projection onto $\mathcal{W}^h$ (see Definition 2.24).*

Proof. The above conditions may be written as:

1. $\operatorname{ran}(\mathscr{P}_N) = \mathcal{W}^h$.

2. $\mathscr{P}_N \mathscr{P}_N \phi = \mathscr{P}_N \phi \ \forall \phi \in \mathcal{W}$.

They may be combined into a single statement: $\mathscr{P}_N \phi^h = \phi^h \ \forall \phi^h \in \mathcal{W}^h$.

The mapping $\mathcal{W} \to \mathcal{W}^h$ is equivalently defined by Eq. (5.12), which may be written as:

$$B_N(v, \phi^h) = B_N(v, \phi) \qquad \forall v \in \mathcal{W}^h \,. \tag{5.13}$$

Let us first assume that $B_N(\cdot, \cdot)$ is coercive on $\mathcal{W}^h$. By the assumed regularities, the operators are also bounded. From the Lax-Milgram theorem it then follows that Eq. (5.13) permits a unique solution. The projection of some $\varphi^h \in \mathcal{W}^h$ may be obtained by solving:

$$B_N(v, \phi^h) = B_N(v, \varphi^h) \Rightarrow B_n(v, \phi^h - \varphi^h) = 0 \qquad \forall v \in \mathcal{W}^h \,. \tag{5.14}$$

By choosing $v = \phi^h - \varphi^h$, and from the coercivity property $B(v, v) \geq 0$:

$$B_n(\phi^h - \varphi^h, \phi^h - \varphi^h) = 0 \Rightarrow \phi^h - \varphi^h = 0 \Rightarrow \phi^h = \varphi^h \,. \tag{5.15}$$

The $\mathscr{P}_N$-operator from Eq. (5.9) is thus a projection if $B_N(\cdot, \cdot)$, defined by Eqs. (5.12) and (5.13), is coercive on $\mathcal{W}^h$. Coercivity of $B_N(\cdot, \cdot)$ is a requirement on η, as will be

65

discussed in a different context in the next chapter. Refer to Lemma 6.3 for the proof, as well as for an expression for a minimal η that is 'large enough'. □

If we replace $\phi - \phi^h$ by ϕ' in Eq. (5.12), then we obtain the following associated optimality condition for the fine scales:

$$- \left(\nabla v^h, \kappa \nabla \phi'\right)_\Omega + \left\langle v^h, \kappa \partial_n \phi'\right\rangle_{\partial\Omega_D} + \left\langle \kappa \partial_n v^h, \phi'\right\rangle_{\partial\Omega_D}$$
$$- \left\langle \kappa \eta v^h, \phi'\right\rangle_{\partial\Omega_D} - \left\langle \boldsymbol{a} \cdot \boldsymbol{n} v^h, \phi'\right\rangle_{\partial\Omega^+} = 0 \qquad \forall v^h \in \mathcal{W}^h. \tag{5.16}$$

Typically, in the variational multiscale framework, one would attempt to invert the fine-scale problem of Eq. (5.8b) while satisfying the requirement posed by Eq. (5.16). However, the particular structures of the coarse-scale problem and the Nitsche projector allow for a more direct inversion of (part of) the fine scales in the coarse-scale equation [178]. First, we recognize that the fine-scale terms in Eq. (5.16) on the Dirichlet boundary may be written in terms of the coarse-scale solution through the definition $\phi' = \phi - \phi^h = \phi_D - \phi^h$. Then, since Eq. (5.16) holds for all $v^h \in \mathcal{W}^h$, we may choose $v^h = w^h$ and add the obtained equality to the coarse-scale problem of Eq. (5.8a). By following this procedure we are left with Nitsche's formulation of the advection-diffusion problem:

Find $\phi^h \in \mathcal{W}^h$ s.t. $\forall\, w^h \in \mathcal{W}^h$:

$$- \left(\boldsymbol{a} \cdot \nabla w^h, \phi^h + \phi'\right)_\Omega + \left\langle \boldsymbol{a} \cdot \boldsymbol{n}\, w^h, \phi^h\right\rangle_{\partial\Omega^+} + \left(\nabla w^h, \kappa \nabla \phi^h\right)_\Omega - \left\langle w^h, \kappa \partial_n \phi^h\right\rangle_{\partial\Omega_D}$$
$$- \left\langle \kappa \partial_n w^h, \phi^h\right\rangle_{\partial\Omega_D} + \left\langle w^h, \kappa \eta \phi^h\right\rangle_{\partial\Omega_D} = \left(w^h, f\right)_\Omega - \left\langle \boldsymbol{a} \cdot \boldsymbol{n}\, w^h, \phi_D\right\rangle_{\partial\Omega_D^-}$$
$$+ \left\langle w^h, g_N\right\rangle_{\partial\Omega_N} - \left\langle \kappa \partial_n w^h, \phi_D\right\rangle_{\partial\Omega_D} + \left\langle w^h, \kappa \eta \phi_D\right\rangle_{\partial\Omega_D}, \tag{5.17}$$

where only one fine-scale term remains. We emphasize that both the symmetric term of Nitsche's method as well as its penalty term originate from the fine-scale inversion per Eq. (5.16).

Remark 5.1. *If we write the remaining fine-scale term as $\left(\mathcal{L}_{stab} w^h, \phi'\right)_\Omega$, then we observe that the appropriate fine-scale model (read: stabilization technique) in conjunction with Nitsche's method involves the SUPG operator $\mathcal{L}_{stab} = -\boldsymbol{a} \cdot \nabla$. Typically, in variational multiscale stabilized methods, the adjoint differential operator $\mathcal{L}_{stab} = -\boldsymbol{a} \cdot \nabla - \nabla \cdot \kappa \nabla$ is proposed. However, since the diffusive part of the fine-scale terms has already been incorporated via the introduction of the Nitsche terms, the diffusive operator no longer occurs in the fine-scale term.*

5.3 A short study on the Nitsche projector

The projector defines the scale decomposition of Eq. (5.5) via Eq. (5.7). This means that it impacts the appropriate modeling choices in the development of the fine-scale model. Before continuing the fine-scale modeling, we thus first dedicate a short study on the Nitsche projector. The ensuing analysis will heavily rely on the work by Hughes and Sangalli in [122], which is rehearsed in Section 3.2, who performed the same analysis using the H_0^1- and L^2-projectors.

A projector may be defined by a finite number of functional constraints; as many as the dimension of $\mathcal{W}^h$:

$$\mathscr{P}_N \phi = \phi^h \text{ such that:}$$
$$\langle \mu_i, \phi - \phi^h \rangle = 0 \qquad i = 1, ..., \dim(\mathcal{W}^h) \,. \tag{5.18}$$

As was proposed in Section 3.2, the μ_i functionals for the Nitsche projector may be inferred from the weak statement of Eq. (5.16). By performing integration by parts on its first term while interpreting the resulting integral in the sense of distributions, the functional constraints follow from substituting the various candidates for v^h. For a one-dimensional domain with a set of nodes $\Gamma = \{x_1, x_2, .., x_n\}$ and element domains $\mathcal{T} = \{[x_1, x_2], .., [x_{n-1}, x_n]\}$ and a typical nodal finite element construction of $\mathcal{W}^h$ with polynomial order p, we find:

$$\mathscr{P}_N \phi = \phi^h \text{ such that:}$$
$$\langle \mu, \phi - \phi^h \rangle = \phi(x_i) - \phi^h(x_i) = 0 \qquad \text{for } x_i \in \Gamma \setminus \partial\Omega_D \,, \tag{5.19a}$$
$$\langle \mu, \phi - \phi^h \rangle = \int_K (\phi - \phi^h) x^q = 0 \qquad \text{for } K \in \mathcal{T} \text{ and } 0 \le q \le p-2 \,, \tag{5.19b}$$
$$\langle \mu, \phi - \phi^h \rangle = \kappa \partial_n (\phi - \phi^h)\big|_{x_i} - \kappa\eta(\phi - \phi^h)\big|_{x_i} = 0 \qquad \text{for } x_i \in \partial\Omega_D^- \,, \tag{5.19c}$$
$$\langle \mu, \phi - \phi^h \rangle = \kappa \partial_n (\phi - \phi^h)\big|_{x_i} - (\kappa\eta + \boldsymbol{a} \cdot \boldsymbol{n})(\phi - \phi^h)\big|_{x_i} = 0 \quad \text{for } x_i \in \partial\Omega_D^+ \,. \tag{5.19d}$$

The first requirement dictates nodal exactness of the finite element formulation, and together with the second requirement these define the H_0^1-projector [122]. The last two originate from the extra degrees of freedom on the Dirichlet boundary.

Remark 5.2. *The central role of the projector can be useful for the interpretation of the obtained finite element solution. For instance, for the current example of Nitsche's method, a better approximation of the true diffusive flux on the Dirichlet boundary could*

be obtained by rewriting Eqs. (5.19c) and (5.19d):

$$-\kappa\partial_n\phi = -\kappa\partial_n\phi^h + \kappa\eta(\phi^h - \phi_D) \qquad on \quad \partial\Omega_D^-, \tag{5.20a}$$

$$-\kappa\partial_n\phi = -\kappa\partial_n\phi^h + (\kappa\eta + \boldsymbol{a}\cdot\boldsymbol{n})(\phi^h - \phi_D) \quad on \quad \partial\Omega_D^+. \tag{5.20b}$$

The same expressions were also proposed in [22] by Bazilevs et al., although motivated based on discrete conservation laws.

Remark 5.3. *In [94], Harari and Albocher perform a spectral analysis of Nitsche's formulation. Their work shows that its spectrum consists of* i) *traditional modes that are independent of the Nitsche parameter and which vanish on Dirichlet boundaries, and* ii) *modes that depend on the Nitsche parameter and are locally supported in a layer along the Dirichlet boundaries. In Eq. (5.19), we observe a similar split in the (functional) constraints imposed by the Nitsche projector.*

A residual-based multiscale model with non-vanishing fine-scale boundary values

6.1 Inversion of the fine-scale problem

For our finite element scheme to yield solutions close to $\phi^h = \mathscr{P}_N \phi$, the weighted integral of ϕ' in Eq. (5.17) needs to be accurately modeled. The model for the remaining fine-scale quantity originates from the inversion of the fine-scale problem of Eq. (5.8b).

Consider the following form of Eq. (5.8b), the fine-scale problem:

Find $\phi' \in \mathcal{W}'$ s.t. $\forall w' \in \mathcal{W}'$:
$$a(w', \phi') = (w', f)_\Omega + \langle w', g_N \rangle_{\partial\Omega_N} - \langle \boldsymbol{a} \cdot \boldsymbol{n}\, w', \phi_D \rangle_{\partial\Omega_D^-} - a(w', \phi^h) - b(w', \lambda). \quad (6.1)$$

As was proposed in Section 3.2.2, we change to duality pairing notation and introduce the adjoint operator $\mathcal{L}^*$:

Find $\phi' \in \mathcal{W}'$ s.t. $\forall w' \in \mathcal{W}'$:
$$\langle \mathcal{L}^* w', \phi' \rangle = \langle \mathcal{R}_{\phi^h}, w' \rangle - b(w', \lambda). \quad (6.2)$$

where, from Eq. (5.8b), these pairings are defined as:

$$\langle \mathcal{L}^* w', \cdot \rangle = -(\boldsymbol{a} \cdot \nabla w', \cdot)_\Omega + \langle \boldsymbol{a} \cdot \boldsymbol{n}\, w', \cdot \rangle_{\partial\Omega^+} + (\nabla w', \kappa \nabla \cdot)_\Omega \,, \tag{6.3a}$$

$$\langle \mathcal{R}_{\phi^h}, w' \rangle = (w', f)_\Omega + \langle w', g_N \rangle_{\partial\Omega_N} - \langle \boldsymbol{a} \cdot \boldsymbol{n}\, w', \phi_D \rangle_{\partial\Omega_D^-}$$

$$+ (\boldsymbol{a} \cdot \nabla w', \phi^h)_\Omega - \langle \boldsymbol{a} \cdot \boldsymbol{n}\, w', \phi^h \rangle_{\partial\Omega^+} - (\nabla w', \kappa \nabla \phi^h)_\Omega \tag{6.3b}$$

$$=: (w', f)_\Omega + \langle w', g_N \rangle_{\partial\Omega_N} - \langle \boldsymbol{a} \cdot \boldsymbol{n}\, w', \phi_D \rangle_{\partial\Omega_D^-} - \langle \mathcal{L}\phi^h, w' \rangle \,,$$

$$b(w', \lambda) = \langle \lambda, w' \rangle_{\partial\Omega_D} \,. \tag{6.3c}$$

Next, we look for the test function w' such that the pairing with $\mathcal{L}^* w'$ induces a point sampling at x, that is:

Find $w' \in \mathcal{W}'$ s.t. $\forall \phi' \in \mathcal{W}'$:

$$\langle \mathcal{L}^* w', \phi' \rangle = \phi'(x) = \langle \delta_x, \phi' \rangle \,, \tag{6.4}$$

where we thus assume that δ_x lies in the dual space of $\mathcal{W}'$. This is the case when $\mathcal{W}' \subset H^s(\mathbb{R}^d)$ with $s > d/2$, with in particular $\mathcal{W}' \subset H^1(\mathbb{R})$ [1]. Clearly, this assumption is rather restrictive, and in practice it may be circumvented by replacing δ_x with a localized averaging operator.

The weak statement of Eq. (6.4) with the bilinear form of Eq. (6.3a) represents an advection-diffusion equation for w' with Robin boundary conditions on $\partial\Omega^+$. The negative sign in front of the velocity vector means that the advective field is effectively reversed. The Robin condition is thus imposed on the inflow part of $\partial\Omega$, which makes this is a well-posed problem.

Recall from Section 3.2.2, and specifically Eq. (3.30), that Eq. (6.4) defines the "fine-scale Green's function" as:

$$\begin{cases} g'(x,y) \in \mathcal{W}' & \text{for fixed } x \,, \\ \mathcal{L}^* g'(x,y) = \delta_x + \displaystyle\sum_{i=1}^{\dim(\mathcal{W}^h)} c_i(x)\, \mu_i \,. \end{cases} \tag{6.5}$$

For ease of notation later on, $\mathcal{L}^*$ acts on the y-dependence of $g'(x,y)$. The μ_i functions refer to the functional constraints imposed by the projection operation. For the Nitsche projector in the one-dimensional case they are those of Eq. (5.18). Effectively, the sum provides a relaxation such that $g'(x,y)$ may be found in the constrained space $\mathcal{W}'$.

When we substitute this fine-scale Green's function as $w'(y)$ in (6.2), we obtain:

$$\langle \mathcal{L}^* g'(x,y), \phi' \rangle = \langle \delta_x, \phi' \rangle + \sum_{i=1}^{\dim(\mathcal{W}^h)} \langle c_i \mu_i, \phi' \rangle \tag{6.6}$$

$$= \phi'(x) = \langle \mathcal{R}_{\phi^h}, g'(x,y) \rangle - \langle \lambda, g'(x,y) \rangle_{\partial \Omega_D},$$

where the summation vanishes due to Eq. (5.18). This relation determines ϕ' from a given coarse scale $\phi^h = \mathscr{P}_N \phi$ and a given Lagrange multiplier λ.

6.2   Adoption of the H_0^1 fine-scale Green's function

Problematically, the solution for the fine scales still depends on the Lagrange multiplier, which we wish to remove from the final formulation. Additionally, the fine-scale Green's function in Eqs. (6.5) and (6.6) corresponds to the Nitsche projector but most literature on the variational multiscale method focuses on a scale decomposition by means of the H_0^1-projector:

$$\mathscr{P}_{H_0^1} : H_0^1 \rightarrow \mathcal{W}^h \cap H_0^1$$

$$\phi \mapsto \underset{\phi^h \in \mathcal{W}^h \cap H_0^1}{\arg\inf} \int_\Omega \frac{1}{2} \kappa (\nabla \phi - \nabla \phi^h) \cdot (\nabla \phi - \nabla \phi^h). \tag{6.7}$$

To solve the issue with the Lagrange multiplier, and to maintain the connection with existing fine-scale models, we reintroduce the H_0^1 fine-scale Green's function as follows:

$$g'_N(x,y) = g'_{H_0^1}(x,y) + \tilde{g}'(x,y). \tag{6.8}$$

The newly added subscripts indicate the projector with which the fine-scale Green's function is associated. The $g'_{H_0^1}$ Green's function corresponds to the solution of the following weak statement:

Find $g'_{H_0^1} \in \ker(\mathscr{P}_{H_0^1})$ s.t. $\forall \phi' \in \ker(\mathscr{P}_{H_0^1})$:
$$\langle \mathcal{L}^* w', \phi' \rangle = \langle \delta_x, \phi' \rangle. \tag{6.9}$$

Following the procedure outlined in Section 3.2.2, we rephrase this as a formulation where the constraints are released from the function spaces and are enforced weakly with Lagrange multipliers. As emphasized in Section 5.3, the set of functional constraints μ_i

corresponding to the projector $\mathscr{P}_{H_0^1}$ is a subset of those of $\mathscr{P}_N$. If we order the set μ_i such that coinciding occurrences come first, and if we collect the coinciding occurrences in the vector $\boldsymbol{\mu}$, then we may write in place of Eq. (6.9):

$$\text{Find } g'_{H_0^1}, \boldsymbol{d}, \lambda_0 \in \mathcal{W} \times \mathbb{R}^{\dim(\mathcal{W}^h \cap H_0^1)} \times H^{-1/2}(\partial\Omega_D)$$
$$\text{s.t. } \forall \phi', \hat{\boldsymbol{d}}, q_0 \in \mathcal{W} \times \mathbb{R}^{\dim(\mathcal{W}^h \cap H_0^1)} \times H^{-1/2}(\partial\Omega_D):$$

$$\langle \mathcal{L}^* g'_{H_0^1}, \phi' \rangle + \langle \boldsymbol{d}^{\mathrm{T}} \boldsymbol{\mu}, \phi' \rangle + \langle \lambda_0, \phi' \rangle_{\partial\Omega_D} = \langle \delta_x, \phi' \rangle\,, \tag{6.10a}$$

$$\langle \hat{\boldsymbol{d}}^{\mathrm{T}} \boldsymbol{\mu}, g'_{H_0^1} \rangle = 0\,, \tag{6.10b}$$

$$\langle q_0, g'_{H_0^1} \rangle_{\partial\Omega_D} = 0\,. \tag{6.10c}$$

Note that we also added the Lagrange multiplier λ_0 to enforce the homogeneous Dirichlet constraint. Thereby we are allowed to use the space $\mathcal{W}$ in this alternative, yet equivalent, formulation. From Eq. (6.10a) we thus observe:

$$\mathcal{L}^* g'_{H_0^1} = \delta_x - \lambda_0 - \sum_{i=1}^{\dim(\mathcal{W}^h \cap H_0^1)} d_i(x)\, \mu_i \ \in \mathcal{W}^*\,. \tag{6.11}$$

Using this relation, as well as the analogous relation for g'_N from Eq. (6.5), we derive a condition for the 'remainder' $\tilde{g}' = g'_N - g'_{H_0^1}$:

$$\mathcal{L}^* \tilde{g}'(x,y) = \lambda_0 + \sum_{i=1}^{\dim(\mathcal{W}^h \cap H_0^1)} (c_i(x) + d_i(x))\, \mu_i + \sum_{i=\dim(\mathcal{W}^h \cap H_0^1)+1}^{\dim(\mathcal{W}^h)} c_i(x)\, \mu_i\,. \tag{6.12}$$

We can now derive the relation for the fine scales while making use of the H_0^1 fine-scale Green's function. We substitute the decomposition of Eq. (6.8) into Eq. (6.6), and rewrite as follows:

$$\phi'(x) = \langle \mathcal{R}_{\phi^h}, g'_{H_0^1} + \tilde{g}' \rangle - \langle \lambda, g'_{H_0^1} + \tilde{g}' \rangle_{\partial\Omega_D}\,, \tag{6.13a}$$

$$\phi'(x) = \langle \mathcal{R}_{\phi^h}, g'_{H_0^1} \rangle - \langle \lambda, g'_{H_0^1} \rangle_{\partial\Omega_D} + \langle \mathcal{R}_{\phi^h}, \tilde{g}' \rangle - \langle \lambda, \tilde{g}' \rangle_{\partial\Omega_D}\,, \tag{6.13b}$$

$$\phi'(x) = \langle \mathcal{R}_{\phi^h}, g'_{H_0^1} \rangle - \langle \lambda, g'_{H_0^1} \rangle_{\partial\Omega_D} + \langle \mathcal{L}^* \tilde{g}', \phi' \rangle\,, \tag{6.13c}$$

$$\phi'(x) = \langle \mathcal{R}_{\phi^h}, g'_{H_0^1} \rangle + \langle \lambda_0, \phi' \rangle_{\partial\Omega_D}\,. \tag{6.13d}$$

From Eq. (6.13b) to Eq. (6.13c) we make use of Eq. (6.2). From Eq. (6.13c) to Eq. (6.13d) we substitute the known value of $g'_{H_0^1} = 0$ on Dirichlet boundaries in the first Lagrange

multiplier term, and we substitute the relation of Eq. (6.12) in place of $\mathcal{L}^* \tilde{g}'$. Recall from Eq. (5.18) that all the terms associated with the summations in Eq. (6.12) vanish when $\phi^h = \mathscr{P}_N \phi$. If we also substitute the known fine-scale solution $\phi' = \phi_D - \phi^h$ on the Dirichlet boundary, then the expression for $\phi'(x)$ becomes:

$$\phi'(x) = \langle \mathcal{R}_{\phi^h}, g'_{H_0^1} \rangle + \langle \lambda_0, \phi_D - \phi^h \rangle_{\partial \Omega_D} . \tag{6.14}$$

Remark 6.1. *The inversion posed by Eq. (6.14) is no longer unique; it is satisfied for solutions $\phi^h = \mathscr{P}_N \phi$, but also for solutions $\phi^h = \mathscr{P}_{H_0^1} \phi$. However, the partial fine-scale closure discussed in Section 5.2 is not satisfied by $\phi^h = \mathscr{P}_{H_0^1} \phi$. The formulation obtained after substitution of Eq. (6.14) into Nitsche's coarse-scale formulation of Eq. (5.17) will hence be uniquely satisfied by $\phi^h = \mathscr{P}_N \phi$.*

We are now able to use the familiar H_0^1 fine-scale Green's function. Additionally, the Lagrange multiplier λ no longer occurs in the expression for the fine scales. Instead, we are left with the pairing with a new Lagrange multiplier, λ_0. This term takes into account the non-zero fine-scale boundary value at the boundary with weakly enforced essential conditions. The original residual-based model assumes that the fine-scale solution vanishes on element boundaries [19]. Since we aim not to make this assumption, this will be the key term that we retain to obtain a more suitable fine-scale model.

Recall that λ_0 is the Lagrange multiplier of the adjoint problem of Eq. (6.9) that enforces the zero Dirichlet condition on the H_0^1 fine-scale Green's function on $\partial \Omega_D$. The pairing with λ_0 may thus be written as:

$$\langle \lambda_0, \phi_D - \phi^h \rangle_{\partial \Omega_D} = -\int_{\partial \Omega_D} \mathbb{H} g'_{H_0^1}(x, y) \left(\phi(y) - \phi^h(y) \right) \mathrm{d}y, \tag{6.15}$$

with $\mathbb{H}$ being some differential operator. The operator $\mathbb{H}$ follows from the equivalence between the weak form of Eq. (6.9) and its strong form, similar to what was derived in Lemma 5.1. We obtain:

$$\mathbb{H} = -\kappa \, \boldsymbol{n} \cdot \nabla_y . \tag{6.16}$$

After integration by parts, the duality pairing with the 'weak' residual may be written as

a volumetric integral. Then, Eq. (6.14) becomes:

$$\phi'(x) = \int_\Omega g'_{H_0^1}(x,y)\,\mathcal{R}_{\phi^h}(y)\,\mathrm{d}y - \int_{\partial\Omega_D} \mathbb{H}g'_{H_0^1}(x,y)\,(\phi_D(y) - \phi^h(y))\,\mathrm{d}y\,. \tag{6.17}$$

6.3 The classical one-dimensional case

Even though we just derived an exact expression for the fine-scale solution, we are merely interested in its effect on the coarse scales. The only term in the coarse-scale formulation of Eq. (5.17) in which ϕ' appears is the advective term. Substitution of Eq. (6.17) gives the fine-scale contribution to the coarse-scale equation as:

$$
\begin{aligned}
-\int_\Omega \boldsymbol{a}\cdot\nabla w^h\,\phi'\,\mathrm{d}x = &-\int_\Omega\int_\Omega \boldsymbol{a}\cdot\nabla w^h(x)\,g'_{H_0^1}(x,y)\,\mathcal{R}_{\phi^h}(y)\,\mathrm{d}y\,\mathrm{d}x \\
&-\int_\Omega\int_{\partial\Omega_D} \boldsymbol{a}\cdot\nabla w^h(x)\mathbb{H}g'_{H_0^1}(x,y)\,(\phi_D(y)-\phi^h(y))\,\mathrm{d}y\,\mathrm{d}x\,.
\end{aligned}
\tag{6.18}
$$

The double integration and the limited availability of Green's functions makes the closure relation of Eq. (6.18), albeit exact, unsuitable for computational use. Simplifications via approximations need to be introduced. The strategy that we will employ repeatedly in this dissertation is to reformulate such that exactness is maintained in the case of constant physical parameters on a one-dimensional domain, while ease of implementation is established in the general case.

In the one-dimensional case, the nodal exactness induced by the $\mathscr{P}_{H_0^1}$ projector results in an element-local fine-scale Green's function [122]. The double integrals in Eq. (6.18) can thus be split in contributions of individual elements. The newly added term only affects elements that lie adjacent to the Dirichlet boundary, where the precise value of the fine-scale solution is known as $\phi' = \phi_D - \phi^h$. Let us consider the contribution of one element that shares a node with the Dirichlet boundary:

$$
\begin{aligned}
-\int_K \boldsymbol{a}\cdot\nabla w^h\,\phi'\,\mathrm{d}x = &-\int_K\int_K \boldsymbol{a}\cdot\nabla w^h(x)\,g'_{H_0^1}(x,y)\,\mathcal{R}_{\phi^h}(y)\,\mathrm{d}y\,\mathrm{d}x \\
&-\int_K\int_{\partial K\cap\partial\Omega_D} \boldsymbol{a}\cdot\nabla w^h(x)\mathbb{H}g'_{H_0^1}(x,y)\,(\phi_D(y)-\phi^h(y))\,\mathrm{d}y\,\mathrm{d}x\,.
\end{aligned}
\tag{6.19}
$$

To simplify, we use the polynomial representation of the test function $w^h\big|_K$ and the residual $\mathcal{R}_{\phi^h}\big|_K$. When $\mathcal{W}^h$ is constructed with p-order nodal elements and the source

function f is at most polynomial order $p - 1$, then:

$$- \boldsymbol{a} \cdot \nabla w^h(x) = -a \sum_{i=1}^{p} \hat{w}_i \, i \, x^{i-1} \qquad \text{for } x \in K \,, \tag{6.20a}$$

$$\mathcal{R}_{\phi^h}(y) = \sum_{j=1}^{p} \hat{R}_j \, y^{j-1} \qquad \text{for } y \in K \,. \tag{6.20b}$$

We also have the following properties of the fine-scale Green's function $g'_{H_0^1}(x, y)$:

$$\int_K \int_K x^q \, g'_{H_0^1}(x, y) \, y^r \mathrm{d}y \, \mathrm{d}x = 0 \qquad \text{when } q < p - 1 \text{ or } r < p - 1 \,, \tag{6.21a}$$

$$\int_K \int_{\partial K} x^q \, \mathbb{H} g'_{H_0^1}(x, y) \mathrm{d}y \, \mathrm{d}x = 0 \qquad \text{when } q < p. \tag{6.21b}$$

Property (6.21a) is shown in [122], and we prove (6.21b) in Theorem 6.1 in he following section. Substituting Eq. (6.20) into Eq. (6.19) while using Eq. (6.21) yields:

$$- \int_K \boldsymbol{a} \cdot \nabla w^h \, \phi' \, \mathrm{d}x = - \int_K \int_K a \left(\hat{w}_p \, p \, x^{p-1} \right) g'_{H_0^1}(x, y) \left(\hat{R}_p \, y^{p-1} \right) \mathrm{d}y \, \mathrm{d}x$$
$$- \int_K \int_{\partial K \cap \partial \Omega_D} a \left(\hat{w}_p \, p \, x^{p-1} \right) \mathbb{H} g'_{H_0^1}(x, y) \left(\phi_D(y) - \phi^h(y) \right) \mathrm{d}y \, \mathrm{d}x \,. \tag{6.22}$$

We assume a to be constant in K, and we extract a and all other constants from the double integrations. We then integrate the two right-hand side terms over K and $F := \partial K \cap \partial \Omega_D$, respectively, and we divide them by $|K|$ and $|F|$:

$$- \int_K \boldsymbol{a} \cdot \nabla w^h \, \phi' \, \mathrm{d}x = - \int_K a \, \hat{w}_p p \, h^{p-1} \left[\frac{1}{|K|} \int_K \int_K \frac{x^{p-1}}{h^{p-1}} g'_{H_0^1}(x, y) \frac{y^{p-1}}{h^{p-1}} \mathrm{d}y \, \mathrm{d}x \right] \hat{R}_p h^{p-1} \, \mathrm{d}\hat{x}$$
$$- \int_F a \, \hat{w}_p p \, h^{p-1} \left[\frac{1}{|F|} \int_K \int_F \frac{x^{p-1}}{h^{p-1}} \mathbb{H} g'_{H_0^1}(x, y) \, \mathrm{d}y \, \mathrm{d}x \right] \left(\phi_D(\hat{x}) - \phi^h(\hat{x}) \right) \mathrm{d}\hat{x} \,, \tag{6.23}$$

where we can identify the following model parameters:

$$\tau = \frac{1}{|K|} \int_K \int_K \frac{x^{p-1}}{h^{p-1}} g'_{H_0^1}(x, y) \frac{y^{p-1}}{h^{p-1}} \mathrm{d}y \, \mathrm{d}x \,, \tag{6.24a}$$

$$\gamma = \frac{1}{|F|} \int_K \int_F \frac{x^{p-1}}{h^{p-1}} \mathbb{H} g'_{H_0^1}(x, y) \, \mathrm{d}y \, \mathrm{d}x \,. \tag{6.24b}$$

The multiplication and division by h^{p-1} in Eq. (6.24) (h being the element size) ensures that the parameters in Eq. (6.24) remain dimensionally consistent with varying polynomial order.

6.4 Fine-scale Green's functions and γ's vanishing moments

The fine-scale Green's function associated to the H_0^1-projector has been studied extensively in [122]. The authors prove the element local nature of $g'_{H_0^1}(x,y)$ in the one-dimensional case. They also show that in a single element, the fine-scale Green's function can then be obtained from the element local classical Green's function:

$$
g'^p_{H_0^1}(x,y) = g(x,y) - \left[\int\limits_0^h g(x,y)\,\mathrm{d}y \quad \cdots \quad \int\limits_0^h y^{p-2}g(x,y)\,\mathrm{d}y \right]
$$

$$
\left[\begin{array}{ccc} \int\limits_0^h\int\limits_0^h g(x,y)\,\mathrm{d}x\,\mathrm{d}y & \cdots & \int\limits_0^h\int\limits_0^h y^{p-2}g(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ \vdots & \ddots & \vdots \\ \int\limits_0^h\int\limits_0^h x^{p-2}g(x,y)\,\mathrm{d}x\,\mathrm{d}y & \cdots & \int\limits_0^h\int\limits_0^h x^{p-2}y^{p-2}g(x,y)\,\mathrm{d}x\,\mathrm{d}y \end{array} \right]^{-1} \left[\begin{array}{c} \int\limits_0^h g(x,y)\,\mathrm{d}x \\ \vdots \\ \int\limits_0^h x^{p-2}g(x,y)\,\mathrm{d}x \end{array} \right], \quad (6.25)
$$

which was also derived in Section 3.2.2, and specifically Eq. (3.33). Refer to [98, 124] or [179] for the expression for $g(x,y)$. The resulting functions for $p = 1, 2$ and 3 are plotted in Figure 6.1. For this particular case, the fine-scale Green's function for $p = 1$ is exactly the element local classical Green's function $g(x,y)$.

The derivation in Section 6.3 requires vanishing 'moments' of the fine-scale Green's function, as stated in Eq. (6.21b). We prove that this holds for $g'^p_{H_0^1}(x,y)$ in the following theorem.

(a) $g'^1_{H_0^1}(x,y) = g(x,y)$.   **(b)** $g'^2_{H_0^1}(x,y)$.   **(c)** $g'^3_{H_0^1}(x,y)$.
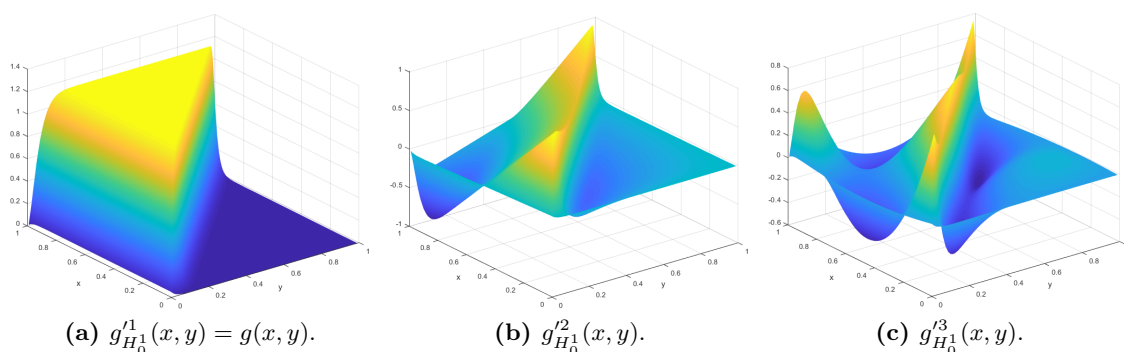
Fig. 6.1: *Fine-scale Green's functions on one element for different polynomial coarse-scale basis functions. Using $\kappa = 0.02$ and $a = 0.8$ on an element of size $h = 1$.*

Theorem 6.1. *Define a γ-like parameter that depends on coarse-scale polynomial order p and a Qth moment:*

$$\gamma^{Q,p} := \frac{1}{|F|} \int\limits_0^h \int\limits_F \frac{x^{Q-1}}{h^{Q-1}} \mathbb{H} g'^p_{H_0^1}(x,y) \, \mathrm{d}y \, \mathrm{d}x \,, \tag{6.26}$$

then from the definition of $g'^p_{H_0^1}(x,y)$ in Eq. (6.25) it follows that:

$$\gamma^{Q,p} = \begin{cases} 0 & \text{if } Q < p\,, \\ \\ \gamma & \text{if } Q = p\,. \end{cases} \tag{6.27}$$

Proof. The equality $\gamma^{Q,p} = \gamma$ for $Q = p$ follows directly from the definition of γ. For $Q < p$ we substitute the definition of the fine-scale Green's function from Eq. (6.25). After carrying out the integrations in Eq. (6.25), the first vector becomes independent of the y-variable, and the last vector independent of the x-variable. The center matrix is filled with constants. This means that the differential operator and integration from Eq. (6.26) act on different vectors and they can thus be separated. After re-ordering of derivatives and integrals we obtain:

$$\gamma^{Q,p} = \gamma^{Q,1} - \left[\int\limits_0^h\int\limits_0^h x^{Q-1} g(x,y) \, \mathrm{d}x \, \mathrm{d}y \quad \cdots \quad \int\limits_0^h\int\limits_0^h x^{Q-1} y^{p-2} g(x,y) \, \mathrm{d}x \, \mathrm{d}y \right]$$

$$\begin{bmatrix} \int\limits_0^h\int\limits_0^h g(x,y) \, \mathrm{d}x \, \mathrm{d}y & \cdots & \int\limits_0^h\int\limits_0^h y^{p-2} g(x,y) \, \mathrm{d}x \, \mathrm{d}y \\ \vdots & \ddots & \vdots \\ \int\limits_0^h\int\limits_0^h x^{p-2} g(x,y) \, \mathrm{d}x \, \mathrm{d}y & \cdots & \int\limits_0^h\int\limits_0^h x^{p-2} y^{p-2} g(x,y) \, \mathrm{d}x \, \mathrm{d}y \end{bmatrix}^{-1} \begin{bmatrix} \gamma^{1,1} \\ \vdots \\ \gamma^{p-2,1} \end{bmatrix}. \tag{6.28}$$

For ease of notation we denote the involved vectors and matrix $\boldsymbol{\zeta}^T$, $\mathbf{C}^{-1}$ and $\boldsymbol{\xi}$. By recognizing that $\boldsymbol{\zeta}^T$ is the Qth row of $\mathbf{C}$ we can write $\boldsymbol{\zeta}^T = \mathbf{e}_Q^T \mathbf{C}$, where $\mathbf{e}_Q$ is a vector of zeros with a 1 at the Qth row. Substitution into the matrix-vector multiplication yields:

$$\gamma_F^{Q,p} = \gamma_F^{Q,1} - \boldsymbol{\zeta}^T \mathbf{C}^{-1} \boldsymbol{\xi} = \gamma_F^{Q,1} - \mathbf{e}_Q^T \mathbf{C}\, \mathbf{C}^{-1} \boldsymbol{\xi} = \gamma_F^{Q,1} - \mathbf{e}_Q^T \boldsymbol{\xi} = \gamma_F^{Q,1} - \gamma_F^{Q,1} = 0\,. \tag{6.29}$$

Note that this only holds for $Q < p$, since $\mathbf{C}$ has $p - 1$ rows. $\qquad\square$

6.5 Fine-scale closure generalization

Up until now, all derivations have been exact. To make use of the integral expressions in Eq. (6.23) on multidimensional domains, we approximate them by the following inner products:

$$-\int_K a\,\hat{w}_p p\,h^{p-1}\tau\hat{R}_p h^{p-1}\,\mathrm{d}\hat{x} \quad \approx -\left(\boldsymbol{a}\cdot\nabla w^h, \tau_{\text{eff}}\mathcal{R}_{\phi^h}\right)_K, \tag{6.30a}$$

$$-\int_F a\,\hat{w}_p p\,h^{p-1}\gamma\left(\phi_D - \phi^h\right)\mathrm{d}\hat{x} \quad \approx -\left\langle\boldsymbol{a}\cdot\nabla w^h, \gamma_{\text{eff}}\left(\phi_D - \phi^h\right)\right\rangle_{F+}. \tag{6.30b}$$

As Eq. (6.30b) indicates, we only make use of the newly proposed term at the outflow Dirichlet boundary $F^+ := \partial K \cap \partial\Omega_D^+$. This is where the boundary layers occur, and where the weak enforcement of the Dirichlet conditions results in impactful fine-scale boundary values.

All the approximations involved in the final finite element formulation may be traced back to these two equations. Essentially, they shift the modeling effort onto the effective stabilization parameters τ_{eff} and γ_{eff}. We propose to design τ_{eff} and γ_{eff} such that these are approximations of τ and γ that take into account the change of (bi)linear forms, while being suitable for multidimensional computations for arbitrary order polynomial basis functions.

6.5.1 Estimation of operator impact

In the one-dimensional case, the bilinear forms of the left-hand and right-hand sides of Eq. (6.30) may be written as:

$$B_{\text{vol}}(w^h, \phi^h) = \left(a\left(\frac{h^{p-1}}{(p-1)!}\right)\frac{\partial^p}{\partial x^p}w^h, \tau\left(\frac{h^{p-1}}{(p-1)!}\right)\frac{\partial^{p-1}}{\partial x^{p-1}}\mathcal{L}\phi^h\right)_K, \tag{6.31a}$$

$$\tilde{B}_{\text{vol}}(w^h, \phi^h) = \left(a\,\frac{\partial}{\partial x}w^h, \tau_{\text{eff}}\mathcal{L}\phi^h\right)_K, \tag{6.31b}$$

$$B_{\text{bdy}}(w^h, \phi^h) = \left\langle a\left(\frac{h^{p-1}}{(p-1)!}\right)\frac{\partial^p}{\partial x^p}w^h, \gamma\,\phi^h\right\rangle_{\partial K \cap \partial\Omega_D^+}, \tag{6.31c}$$

$$\tilde{B}_{\text{bdy}}(w^h, \phi^h) = \left\langle a\,\frac{\partial}{\partial x}w^h, \gamma_{\text{eff}}\,\phi^h\right\rangle_{\partial K \cap \partial\Omega_D^+}. \tag{6.31d}$$

In Definition 2.21 we defined the norm of an operator as the maximal 'magnitude change' between input and output. We use this measure to quantify the 'impact' of the above

bilinear forms:

$$\|B\| := \sup_{\frac{\partial}{\partial x} w^h \neq 0, \frac{\partial}{\partial x} \phi^h \neq 0} \frac{|B(w^h, \phi^h)|}{\|\frac{\partial}{\partial x} w^h\|_{L^2(K)} \|\frac{\partial}{\partial x} \phi^h\|_{L^2(K)}} \,, \tag{6.32}$$

where we have chosen to define the norm of the bilinear forms with respect to the H^1-seminorm of its arguments, as this seminorm is one of the terms in the optimality condition induced by the Nitsche projector according to Eq. (5.9).

We choose τ_{eff} and γ_{eff} such that the impact of these bilinear forms equal: $\|B_{\text{vol}}\| = \|\tilde{B}_{\text{vol}}\|$ and $\|B_{\text{bdy}}\| = \|\tilde{B}_{\text{bdy}}\|$. If we assume constant parameters in K, and the advective dominant case such that $\mathcal{L}\phi^h$ may be approximated by $a\frac{\partial}{\partial x}\phi^h$, then we obtain:

$$\tau_{\text{eff}} \approx \tau \left(\frac{h^{p-1}}{(p-1)!}\right)^2 \frac{\|(\frac{\partial^p}{\partial x^p} \cdot, \frac{\partial^p}{\partial x^p} \cdot)_K\|}{\|(\frac{\partial}{\partial x} \cdot, \frac{\partial}{\partial x} \cdot)_K\|} = \tau \left(\frac{h^{p-1}}{(p-1)!}\right)^2 \sup_{\frac{\partial}{\partial x} w^h \neq 0} \frac{\|\frac{\partial^p}{\partial x^p} w^h\|_{L^2(K)}^2}{\|\frac{\partial}{\partial x} w^h\|_{L^2(K)}^2} \,, \tag{6.33a}$$

$$\gamma_{\text{eff}} \approx \gamma \left(\frac{h^{p-1}}{(p-1)!}\right) \frac{\|\langle\frac{\partial^p}{\partial x^p} \cdot, \cdot\rangle_F\|}{\|\langle\frac{\partial}{\partial x} \cdot, \cdot\rangle_F\|} = \gamma \left(\frac{h^{p-1}}{(p-1)!}\right) \frac{\sup_{\frac{\partial}{\partial x} w^h \neq 0} \left(\left|\frac{\partial^p}{\partial x^p} w^h\right|_F \Big/ \|\frac{\partial}{\partial x} w^h\|_{L^2(K)}\right)}{\sup_{\frac{\partial}{\partial x} w^h \neq 0} \left(\left|\frac{\partial}{\partial x} w^h\right|_F \Big/ \|\frac{\partial}{\partial x} w^h\|_{L^2(K)}\right)} \,. \tag{6.33b}$$

The inverse inequalities in these expressions are computable by hand [95]. For linear, quadratic and cubic coarse-scale basis function, the relations between the parameters τ_{eff} and τ, and γ_{eff} and γ become:

$$\tau_{\text{eff}}, \gamma_{\text{eff}} \approx \begin{cases} \tau, & \gamma & \text{for } p = 1\,, \\ 12\,\tau, & \sqrt{3}\,\gamma & \text{for } p = 2\,, \\ 180\,\tau, & 2\sqrt{5}\,\gamma & \text{for } p = 3\,. \end{cases} \tag{6.34}$$

6.5.2 τ-parameter approximation for $p \in \{1, 2, 3\}$

In literature, we find that the τ-parameter that is used with higher-order basis functions is often the same as that for $p = 1$ (i.e., obtained from the element local Green's function), sometimes with a p-dependent mesh size scaling. In this section, we propose an approximation of τ for linear, quadratic and cubic elements based on the actual fine-scale Green's functions. These are devised such that they limit to the exact expressions in the advection (τ_a) or diffusion (τ_d) dominated cases. Using the definition of τ from Eq. (6.24a), together with the fine-scale Green's functions from 6.4, the following exact expressions for τ may

be computed for $p = 1, 2$ and 3, respectively [39, 109, 122]:

$$\tau_1 = \frac{h}{2\|\boldsymbol{a}\|} \left(\frac{2 + Pe - (2 - Pe)\exp(Pe)}{-Pe + Pe\exp(Pe)} \right) =: \frac{h}{2\|\boldsymbol{a}\|} \xi_1(Pe), \tag{6.35a}$$

$$\tau_2 = \frac{h}{72\|\boldsymbol{a}\|} \left(\frac{12 + 6Pe + Pe^2 - (12 - 6Pe + Pe^2)\exp(Pe)}{-2Pe - Pe^2 + (2Pe - Pe^2)\exp(Pe)} \right) =: \frac{h}{2\|\boldsymbol{a}\|} \xi_2(Pe), \tag{6.35b}$$

$$\tau_3 = \frac{h}{1800\|\boldsymbol{a}\|} \left(\frac{120 + 60Pe + 12Pe^2 + Pe^3 - (120 - 60Pe + 12Pe^2 - Pe^3)\exp(Pe)}{-12Pe - 6Pe^2 - Pe^3 + (12Pe - 6Pe^2 + Pe^3)\exp(Pe)} \right)$$

$$=: \frac{h}{2\|\boldsymbol{a}\|} \xi_3(Pe), \tag{6.35c}$$

where $Pe = \frac{\|\boldsymbol{a}\| h}{\kappa}$ is the element Péclet number and ξ is the *upwind function*. From these equations we obtain the following advective and diffusive limits:

$$\tau_{1,a} := \lim_{Pe \to \infty} \tau_1 = \frac{h}{2\|\boldsymbol{a}\|}, \qquad \tau_{1,d} := \lim_{Pe \to 0^+} \tau_1 = \frac{h^2}{12\kappa}, \tag{6.36a}$$

$$\tau_{2,a} = \frac{h}{72\|\boldsymbol{a}\|}, \qquad \tau_{2,d} = \frac{h^2}{720\kappa}, \tag{6.36b}$$

$$\tau_{3,a} = \frac{h}{1800\|\boldsymbol{a}\|}, \qquad \tau_{3,d} = \frac{h^2}{25200\kappa}. \tag{6.36c}$$

The following approximation strategy for τ is used frequently in stabilized methods [19, 116, 170, 186, 187]:

$$\tau \approx \frac{1}{\sqrt{\tau_a^{-2} + \tau_d^{-2}}}. \tag{6.37}$$

To determine the effectiveness of the scaling for the various polynomial orders, we substitute Eq. (6.36) into (6.37). In all cases, we can rewrite the expression to obtain the effective approximate upwind function. For example, for linear elements:

$$\tau_1 \approx \frac{1}{\sqrt{\frac{4\|\boldsymbol{a}\|^2}{h^2} + \frac{144\kappa^2}{h^4}}} = \frac{h}{2\|\boldsymbol{a}\|} \frac{1}{\sqrt{1 + 36Pe^{-2}}} =: \frac{h}{2\|\boldsymbol{a}\|} \tilde{\xi}_1(Pe). \tag{6.38}$$

Figure 6.2a illustrates how the approximate upwind functions $\tilde{\xi}(Pe)$ relates to the exact upwind functions of Eq. (6.35). The figure shows that the approximation of τ according to Eq. (6.37) has the correct asymptotic limits, and converges to these limits at the correct rates. We observe that this holds true for each polynomial order.

6.5.3 γ-parameter approximation for $p \in \{1, 2, 3\}$

In a similar sense, we wish to construct an approximate γ, based on generic (spatial dimension independent) parameters, that share the asymptotic scaling behavior of the exact one-dimensional expression. The exact expressions of γ may be computed from Eq. (6.24b) as:

$$\gamma_1 = \frac{h}{2}\left(\frac{2 + 2Pe_n - 2\exp(Pe_n)}{Pe_n - Pe_n\exp(Pe_n)}\right) =: \frac{h}{2}\,\zeta_1(Pe_n)\,, \tag{6.39a}$$

$$\gamma_2 = \frac{h}{12}\left(\frac{12 + 8Pe_n + 2Pe_n^2 - (12 - 4Pe_n)\exp(Pe_n)}{2Pe_n + Pe_n^2 - (2Pe_n - Pe_n^2)\exp(Pe_n)}\right) =: \frac{h}{2}\,\zeta_2(Pe_n)\,, \tag{6.39b}$$

$$\gamma_3 = \frac{h}{60}\left(\frac{120 + 72Pe_n + 18Pe_n^2 + 2Pe_n^3 - (120 - 48Pe_n + 6Pe_n^2)\exp(Pe_n)}{12Pe_n + 6Pe_n^2 + Pe_n^3 - (12Pe_n - 6Pe_n^2 + Pe_n^3)\exp(Pe_n)}\right)$$
$$=: \frac{h}{2}\,\zeta_3(Pe_n), \tag{6.39c}$$

for $p = 1, 2$ and 3, respectively. $Pe_n = \frac{\boldsymbol{a}\cdot\boldsymbol{n}\,h}{\kappa}$ is a boundary type element Péclet number. Since we only make use of γ on the outflow boundary $(\boldsymbol{a} \cdot \boldsymbol{n} \geq 0)$, we exclusively consider $Pe_n \geq 0$ in the following. The advective and diffusive limits of these expressions are:

$$\gamma_{1,a} := \lim_{Pe_n\to\infty}\gamma_1 = 0\,, \qquad \gamma_{1,d} := \lim_{Pe_n\to 0^+}\gamma_1 = \frac{h}{2}\,, \tag{6.40a}$$

$$\gamma_{2,a} = 0\,, \qquad\qquad \gamma_{2,d} = \frac{h}{12}\,, \tag{6.40b}$$

$$\gamma_{3,a} = 0\,, \qquad\qquad \gamma_{3,d} = \frac{h}{60}\,. \tag{6.40c}$$
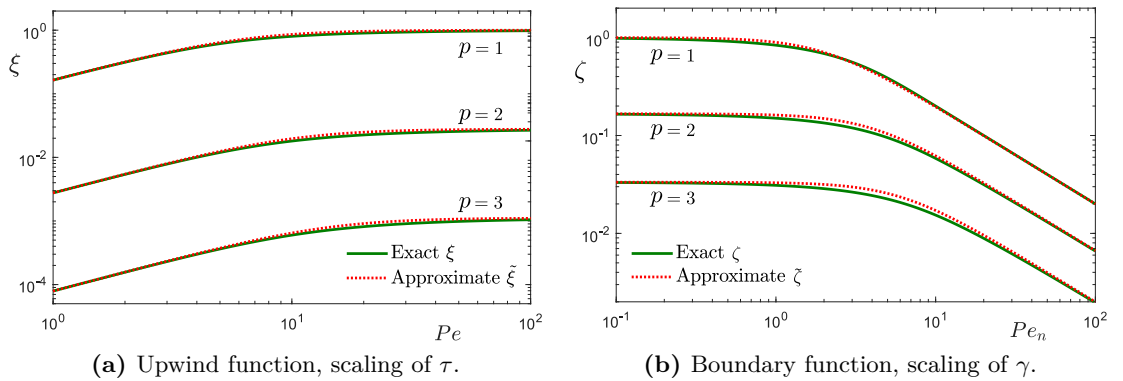
(a) Upwind function, scaling of τ.   **(b)** Boundary function, scaling of γ.

Fig. 6.2: *Exact ξ and ζ functions and their approximations for the one-dimensional case. Showing the correct scaling behavior.*

Since these limits do not depend on $\boldsymbol{a}$ nor on κ, the approximation strategy of Eq. (6.37) is not viable. Instead, we deduce the following scaling of γ based on its definition from Eq. (6.24b):

$$\gamma \propto |\mathbb{H}| \, \hat{\tau}, \tag{6.41}$$

where $|\mathbb{H}|$ represents the scaling induced by $\mathbb{H}$ as defined in Eq. (6.16), and $\hat{\tau}$ is a modified τ-like quantity that takes into account the difference in domain of integration between γ and τ in Eq. (6.24). $\mathbb{H}$ scales linearly with κ and inversely with length, such that we may choose $|\mathbb{H}| \propto \sqrt{\kappa \tau_d^{-1}}$. By adopting a similar approximation for $\hat{\tau}$ as Eq. (6.37), we obtain:

$$\gamma \approx c_s \sqrt{\kappa \tau_d^{-1}} \sqrt{\frac{1}{c_1 \tau_a^{-2} + c_2 \tau_d^{-2}}} = c_s \sqrt{\frac{\kappa}{c_1 \tau_d \tau_a^{-2} + c_2 \tau_d^{-1}}} \, . \tag{6.42}$$

c_s takes into account the shape effect of the element as the measure of the relevant boundary versus the measure of the element interior:

$$c_s := \frac{h|F|}{|K|} \, , \tag{6.43}$$

where the multiplication with h ensures a mesh size independent scaling, and h should be the representative element size that is used in the approximation of τ through τ_a and τ_d. Throughout this dissertation we use the longest element edge. In the one-dimensional case $|F| = 1$ and $|K| = h$, such that $c_s = 1$.

The coefficients c_1 and c_2 are introduced to capture the difference in scaling between τ and $\hat{\tau}$. They can be determined from Eq. (6.42) by ensuring the correct limiting behavior of γ in the one-dimensional case. For example, for linear elements, substituting the limits of Eq. (6.36a) results in:

$$\gamma_1 \approx \frac{h}{2} \sqrt{\frac{1}{\frac{1}{12} c_1 Pe_n^2 + 3 \, c_2}} =: \frac{h}{2} \tilde{\zeta}_1(Pe_n) \, . \tag{6.44}$$

By ensuring that Eq. (6.44) has the same asymptotic limits as Eq. (6.39) and also has the same convergence rate towards zero, we obtain $c_1 = 3$ and $c_2 = 1/3$. The same strategy results in $c_1 = 1.25$ and $c_2 = 0.2$ for $p = 2$ and $c_1 = 7/9$ and $c_2 = 1/7$ for $p = 3$. Figure 6.2b shows the approximate and exact boundary functions and confirms that the approximation of γ displays the correct asymptotic scaling behavior for all polynomial orders.

6.6 Summary of the finite element formulation and analysis of existence and uniqueness of a solution

As touched upon in Remark 6.1 in Section 6.2, the uniqueness of the fine-scale inversion is not a trivial issue. Additionally, the required simplification and modeling steps discussed in Section 6.5 may raise further questions regarding the existence of the approximate coarse-scale solution. In a finite-dimensional functional setting, existence and uniqueness follow directly from the coercivity of the bilinear form, which we analyze in this section. We first summarize the complete finite element formulation, including all the modeling terms and parameters.

With all the modeling terms included, the finite element formulation becomes:

$$
\begin{aligned}
\text{Find } \phi^h &\in \mathcal{W}^h \text{ s.t. } \forall\, w^h \in \mathcal{W}^h : \\
B(w^h,\phi^h) &= B_{\mathrm{A}}(w^h,\phi^h) + B_{\mathrm{D}}(w^h,\phi^h) + B_{\mathrm{VMS},\tilde{\Omega}}(w^h,\phi^h) + B_{\mathrm{VMS},\partial\Omega_D^+}(w^h,\phi^h) \\
&= (w^h, f)_\Omega + \langle w^h, g_N \rangle_{\partial\Omega_N} - \langle \boldsymbol{a}\cdot\boldsymbol{n}\, w^h, \phi_D \rangle_{\partial\Omega_D^-} \\
&\quad - \langle \kappa\partial_n w^h, \phi_D \rangle_{\partial\Omega_D} + \langle w^h, \kappa\eta\phi_D \rangle_{\partial\Omega_D} \\
&\quad + (\boldsymbol{a}\cdot\nabla w^h\, \tau_{\mathrm{eff}}, f)_\Omega + \langle \boldsymbol{a}\cdot\nabla w^h\, \gamma_{\mathrm{eff}}, \phi_D \rangle_{\partial\Omega_D^+},
\end{aligned}
\tag{6.45}
$$

where the advection and diffusion parts of the bilinear form are:

$$
B_{\mathrm{A}}(w^h,\phi^h) = -(\boldsymbol{a}\cdot\nabla w^h, \phi^h)_\Omega + \langle \boldsymbol{a}\cdot\boldsymbol{n}\, w^h, \phi^h \rangle_{\partial\Omega^+},
\tag{6.46a}
$$

$$
\begin{aligned}
B_{\mathrm{D}}(w^h,\phi^h) = (\kappa\nabla w^h, \nabla\phi^h)_\Omega &- \langle \kappa w^h, \partial_n\phi^h \rangle_{\partial\Omega_D} \\
&- \langle \kappa\partial_n w^h, \phi^h \rangle_{\partial\Omega_D} + \langle \kappa\eta w^h, \phi^h \rangle_{\partial\Omega_D},
\end{aligned}
\tag{6.46b}
$$

and where the two variational multiscale components are:

$$
B_{\mathrm{VMS},\tilde{\Omega}}(w^h,\phi^h) = (\boldsymbol{a}\cdot\nabla w^h\, \tau_{\mathrm{eff}}, \boldsymbol{a}\cdot\nabla\phi^h - \nabla\cdot\kappa\nabla\phi^h)_{\tilde{\Omega}},
\tag{6.47a}
$$

$$
B_{\mathrm{VMS},\partial\Omega_D^+}(w^h,\phi^h) = \langle \boldsymbol{a}\cdot\nabla w^h\, \gamma_{\mathrm{eff}}, \phi^h \rangle_{\partial\Omega_D^+},
\tag{6.47b}
$$

with $\tilde{\Omega}$ the sum of open element domains.

Expressions for the parameters τ_{eff} and γ_{eff} are collected in Table 6.1. These expressions take into account all the considerations discussed in Sections 6.5.1 to 6.5.3. As the table shows, we have formulated all model parameters such that they depend exclusively on $\tau_{1,a}$
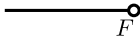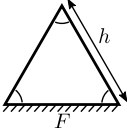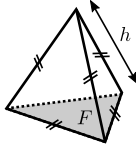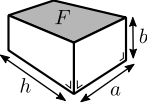
and $\tau_{1,d}$, i.e., those relating to linear elements. The exact approach for computing these limiting values remains flexible. For example, one could incorporate the Jacobian of the element mapping [19, 23], use element local length-scales and Péclet numbers [187], or use the analytical expressions of Eq. (6.36a) and the element diameter (which is the adopted approach throughout this dissertation).

In the coercivity analysis of this finite element formulation, we assume that $\boldsymbol{a}$ and κ are constant in Ω and that the grid is (quasi) uniform, such that also η, τ_a and τ_d can be chosen as global constants. The derivations can trivially be modified for non-uniform grids or non-constant $\boldsymbol{a}$ and κ and elementwise parameters η, τ_a and τ_d. We further assume that the parameter τ_{eff} satisfies:

$$\tau_{\text{eff}} \leq C_d \tau_{1,d} \leq \inf_{w^h \in \mathcal{W}^h} \frac{1}{2\kappa} \frac{\|\nabla w^h\|_\Omega^2}{\|\Delta w^h\|_{\hat{\Omega}}^2} \propto \frac{h^2}{\kappa}, \tag{6.48}$$

The first inequality is satisfied due to the harmonic mean structure of τ_{eff}, where the C_d's follow directly from Table 6.1 as 1, $\frac{1}{5}$ and $\frac{3}{35}$ for linear, quadratic and cubic basis functions, respectively. On a one-dimensional mesh, this means that $C_d \tau_{1,d} = \frac{h^2}{12\kappa}$, $\frac{h^2}{60\kappa}$ and $\frac{h^2}{140\kappa}$. In this one-dimensional case, the inverse estimate in Eq. (6.48) may be explicitly computed, resulting in maxima of ∞, $\frac{h^2}{24\kappa}$ and $\frac{h^2}{120\kappa}$ for polynomial orders of 1, 2 and 3. The condition of Eq. (6.48) is thus satisfied. We expect similar results in multiple spatial dimensions for meshes with reasonable quality.

Table 6.1: *Overview of τ_{eff} and γ_{eff} expressions for different elements and polynomial degrees.*

	$p = 1$	$p = 2$	$p = 3$
τ_{eff} :	$\sqrt{\dfrac{1}{\tau_{1,a}^{-2} + \tau_{1,d}^{-2}}}$	$\sqrt{\dfrac{1}{9\,\tau_{1,a}^{-2} + 25\,\tau_{1,d}^{-2}}}$	$\sqrt{\dfrac{1}{25\,\tau_{1,a}^{-2} + \frac{1225}{9}\,\tau_{1,d}^{-2}}}$
γ_{eff} :	$c_s \sqrt{\dfrac{\kappa}{3\tau_{1,d}\tau_{1,a}^{-2} + \frac{1}{3}\tau_{1,d}^{-1}}}$	$c_s \sqrt{\dfrac{\kappa}{9\tau_{1,d}\tau_{1,a}^{-2} + 4\tau_{1,d}^{-1}}}$	$c_s \sqrt{\dfrac{\kappa}{15\tau_{1,d}\tau_{1,a}^{-2} + 15\tau_{1,d}^{-1}}}$
Element:			
c_s (Eq. (6.43)):	1	$\frac{4}{\sqrt{3}}$	$2\sqrt{\frac{2}{3}}$	h/b

Finally, we require that η satisfies:

$$\eta \geq 4\left(T_1 + c_s^2 T_2\right) \propto \frac{1}{h}, \tag{6.49}$$

with:

$$T_1 = \sup_{w^h \in \mathcal{W}^h} \frac{\|\partial_n w^h\|_{\partial\Omega_D}^2}{\|\nabla w^h\|_\Omega^2} \propto \frac{1}{h}, \tag{6.50a}$$

$$T_2 = \sup_{w^h \in \mathcal{W}^h} \frac{\|\boldsymbol{a} \cdot \nabla w^h\|_{\partial\Omega_D^+}^2}{\|\boldsymbol{a} \cdot \nabla w^h\|_\Omega^2} \propto \frac{1}{h}. \tag{6.50b}$$

We carry out the coercivity proof for the different components of the bilinear forms separately, and then look at the formulation as a whole. We start with determining a relation between the model parameters γ_{eff} and τ_{eff}.

Lemma 6.1. *For $p = 1$, 2 or 3, the expressions from Table 6.1 are such that γ_{eff} is bounded by τ_{eff} according to:*

$$\gamma_{\text{eff}}^2 \leq 3c_s^2 \, \kappa \tau_{\text{eff}}. \tag{6.51}$$

Proof. We write the γ_{eff} and τ_{eff} expressions from Table 6.1 in the following general form:

$$\tau_{\text{eff}} = \sqrt{\frac{1}{(C_a \tau_{1,a})^{-2} + (C_d \tau_{1,d})^{-2}}}, \tag{6.52}$$

$$\gamma_{\text{eff}} = c_s \sqrt{\frac{\kappa}{\tau_{1,d}}} \sqrt{\frac{1}{(C_1 \tau_{1,a})^{-2} + (C_2 \tau_{1,d})^{-2}}}. \tag{6.53}$$

Dividing γ_{eff} by τ_{eff} and squaring gives:

$$\left(\frac{\gamma_{\text{eff}}}{\tau_{\text{eff}}}\right)^2 = c_s^2 \frac{\kappa}{\tau_{1,d}} \frac{(C_a \tau_{1,a})^{-2} + (C_d \tau_{1,d})^{-2}}{(C_1 \tau_{1,a})^{-2} + (C_2 \tau_{1,d})^{-2}}. \tag{6.54}$$

After multiplying both sides by $C_d \tau_{1,d}$ and using $C_d \tau_{1,d} \geq \tau$, we obtain:

$$C_d \tau_{1,d} \left(\frac{\gamma_{\text{eff}}}{\tau_{\text{eff}}}\right)^2 = c_s^2 \kappa \frac{C_d C_a^{-2} \tau_{1,a}^{-2} + C_d^{-1} \tau_{1,d}^{-2}}{(C_1 \tau_{1,a})^{-2} + (C_2 \tau_{1,d})^{-2}} \geq \tau_{\text{eff}} \left(\frac{\gamma_{\text{eff}}}{\tau_{\text{eff}}}\right)^2 = \frac{\gamma_{\text{eff}}^2}{\tau_{\text{eff}}}. \tag{6.55}$$

The fraction may be bound from above as:

$$\frac{\gamma_{\text{eff}}^2}{\tau_{\text{eff}}} \leq c_s^2 \kappa \frac{\max(\frac{C_d}{C_a^2}, \frac{1}{C_d})(\tau_{1,a}^{-2} + \tau_{1,d}^{-2})}{\min(\frac{1}{C_1^2}, \frac{1}{C_2^2})(\tau_{1,a}^{-2} + \tau_{1,d}^{-2})} = c_s^2 \kappa \frac{\max(\frac{C_d}{C_a^2}, \frac{1}{C_d})}{\min(\frac{1}{C_1^2}, \frac{1}{C_2^2})} = \begin{cases} 3\,c_s^2\kappa & \text{for } p = 1\,, \\ \frac{5}{4}\,c_s^2\kappa & \text{for } p = 2\,, \\ \frac{7}{9}\,c_s^2\kappa & \text{for } p = 3\,. \end{cases} \quad (6.56)$$

Equation (6.51) follows from the maximum of the three cases. $\square$

Lemma 6.2. *The bilinear form in Eq. (6.46a) satisfies the following coercivity result:*

$$B_A(w^h, w^h) \geq \tfrac{1}{2}\|\sqrt{|\boldsymbol{a}\cdot\boldsymbol{n}|}w^h\|_{\partial\Omega}^2 \quad \forall\, w^h \in \mathcal{W}^h\,, \quad (6.57)$$

where the norms are L^2-norms on the indicated domains.

Proof. Direct substitution of $\phi^h = w^h$ into Eq. (6.46a) results in:

$$B_A(w^h, w^h) = -\big(\boldsymbol{a} \cdot \nabla w^h, w^h\big)_\Omega + \big\langle \boldsymbol{a} \cdot \boldsymbol{n}\, w^h, w^h \big\rangle_{\partial\Omega^+}\,. \quad (6.58)$$

Making use of the property $\nabla \cdot \boldsymbol{a} = 0$, the first term may be rewritten as follows:

$$-\int_\Omega \nabla \cdot (\tfrac{1}{2}\boldsymbol{a}\,(w^h)^2) = -\tfrac{1}{2}\int_{\partial\Omega} \boldsymbol{a}\cdot\boldsymbol{n}\,(w^h)^2 = \tfrac{1}{2}\int_{\partial\Omega^-} |\boldsymbol{a}\cdot\boldsymbol{n}|\,(w^h)^2 - \tfrac{1}{2}\int_{\partial\Omega^+} |\boldsymbol{a}\cdot\boldsymbol{n}|\,(w^h)^2\,. \quad (6.59)$$

Substitution into Eq. (6.58) completes the proof. $\square$

Lemma 6.3. *Under the condition posed by Eq. (6.48), the bilinear form in Eq. (6.46b) satisfies the following coercivity result:*

$$B_D(w^h, w^h) \geq \tfrac{1}{2}\kappa\,\|\nabla w^h\|_\Omega^2 + (\eta - 2T_1)\kappa\|w^h\|_{\partial\Omega_D}^2 \quad \forall\, w^h \in \mathcal{W}^h\,, \quad (6.60)$$

where the norms are L^2-norms on the indicated domains, and T_1 is given in Eq. (6.50a).

Proof. Direct substitution of $\phi^h = w^h$ into Eq. (6.46b) results in:

$$B_{\mathrm{D}}(w^h, w^h) = \kappa \|\nabla w^h\|_{\Omega}^2 - 2 \left\langle \kappa\, \partial_n w^h, w^h \right\rangle_{\partial\Omega_D} + \eta\kappa \|w^h\|_{\partial\Omega_D}^2 \,. \tag{6.61}$$

By using Young's inequality to bound the nonsymmetric term, we obtain:

$$-2 \left\langle \kappa\, \partial_n w^h, w^h \right\rangle_{\partial\Omega_D} \geq -\epsilon\kappa \|\partial_n w^h\|_{\partial\Omega_D}^2 - \frac{\kappa}{\epsilon} \|w^h\|_{\partial\Omega_D}^2 \,,$$

$$\geq -\epsilon T_1 \kappa \|\nabla w^h\|_{\Omega}^2 - \frac{\kappa}{\epsilon} \|w^h\|_{\partial\Omega_D}^2 \,, \tag{6.62}$$

where T_1 is defined in Eq. (6.50a). Choosing the parameter from Young's inequality as $\epsilon = 1/(2T_1)$ completes the proof. $\qquad\square$

Lemma 6.4. *Under the condition posed by Eq. (6.48), the volumetric variational multiscale term in Eq. (6.47a) satisfies the following coercivity result:*

$$B_{VMS,\tilde{\Omega}}(w^h, w^h) \geq \tfrac{1}{2}\tau_{eff} \|\boldsymbol{a} \cdot \nabla w^h\|_{\Omega}^2 - \tfrac{1}{4}\kappa \|\nabla w^h\|_{\Omega}^2 \quad \forall\, w^h \in \mathcal{W}^h \,, \tag{6.63}$$

where the norms are L^2-norms on the indicated domains.

Proof. Direct substitution of $\phi^h = w^h$ results in:

$$B_{\mathrm{VMS},\tilde{\Omega}}(w^h, w^h) = \tau_{\mathrm{eff}} \|\boldsymbol{a} \cdot \nabla w^h\|^2 + \left(\sqrt{\tau_{\mathrm{eff}}}\, \boldsymbol{a} \cdot \nabla w^h, \sqrt{\tau_{\mathrm{eff}}}\, \kappa \Delta w^h\right)_{\tilde{\Omega}} \,. \tag{6.64}$$

With Young's inequality we bound the second term from below:

$$B_{\mathrm{VMS},\tilde{\Omega}}(w^h, w^h) \geq \tau_{\mathrm{eff}} \|\boldsymbol{a} \cdot \nabla w^h\|^2 - \tfrac{1}{2}\tau_{\mathrm{eff}} \|\boldsymbol{a} \cdot \nabla w^h\|^2 - \tfrac{1}{2}\tau_{\mathrm{eff}}\kappa^2 \|\Delta w^h\|^2 \,. \tag{6.65}$$

Using the assumed bound of τ_{eff} from Eq. (6.48) completes the proof. $\qquad\square$

Lemma 6.5. *The boundary variational multiscale term in Eq. (6.47b) satisfies the following coercivity result:*

$$B_{VMS,\partial\Omega_D^+}(w^h, \phi^h) \geq -\tfrac{1}{4}\tau_{eff} \|\boldsymbol{a} \cdot \nabla w^h\|_{\Omega}^2 - 3\,c_s^2 T_2 \kappa \|w^h\|_{\partial\Omega_D^+}^2 \quad \forall\, w^h \in \mathcal{W}^h \,, \tag{6.66}$$

where the norms are L^2-norms on the indicated domains.

Proof. After substitution of $\phi^h = w^h$, we obtain:

$$B_{\text{VMS},\partial\Omega_D^+}(w^h, w^h) = \langle \boldsymbol{a} \cdot \nabla w^h, \gamma_{\text{eff}} w^h \rangle_{\partial\Omega_D^+} \geq -\tfrac{1}{2}\varepsilon\|\boldsymbol{a} \cdot \nabla w^h\|_{\partial\Omega_D^+}^2 - \tfrac{1}{2}\frac{\gamma_{\text{eff}}^2}{\varepsilon}\|w^h\|_{\partial\Omega_D^+}^2$$

$$\geq -\tfrac{1}{2}\varepsilon T_2\|\boldsymbol{a} \cdot \nabla w^h\|_\Omega^2 - \tfrac{1}{2}\frac{\gamma_{\text{eff}}^2}{\varepsilon}\|w^h\|_{\partial\Omega_D^+}^2 . \tag{6.67}$$

The first inequality follows from Young's inequality with parameter ε, and the second inequality as well as the parameter T_2 originate from the inverse estimate of Eq. (6.50b). Choosing the parameter $\varepsilon = \tau_{\text{eff}}/(2T_2)$ and using the result of *Lemma* 6.1 to relate τ_{eff} and γ_{eff} completes the poof. $\square$

Theorem 6.2. *The combined bilinear form of Eq.* (6.45) *satisfies the following coercivity result:*

$$B(w^h, w^h) \geq \tfrac{1}{4}\tau_{\text{eff}}\|\boldsymbol{a} \cdot \nabla w^h\|_\Omega^2 + \tfrac{1}{2}\|\sqrt{|\boldsymbol{a}\cdot\boldsymbol{n}|}w^h\|_{\partial\Omega}^2 + \tfrac{1}{4}\kappa\|\nabla w^h\|_\Omega^2 + \tfrac{1}{4}\eta\kappa\|w^h\|_{\partial\Omega_D}^2 , \tag{6.68}$$

which holds $\forall\, w^h \in \mathcal{W}^h$. The norms are L^2-norms on the indicated domains.

Proof. Direct substitution of $\phi^h = w^h$ in the bilinear form, and using the results of *Lemmas* 6.2 to 6.5, results in:

$$B(w^h, w^h) = B_{\text{A}}(w^h, w^h) + B_{\text{D}}(w^h, w^h) + B_{\text{VMS},\tilde{\Omega}}(w^h, w^h) + B_{\text{VMS},\partial\Omega_D^+}(w^h, w^h)$$

$$\geq \tfrac{1}{2}\|\sqrt{|\boldsymbol{a}\cdot\boldsymbol{n}|}w^h\|_{\partial\Omega}^2 + \tfrac{1}{2}\kappa\|\nabla w^h\|_\Omega^2 + (\eta - 2T_1)\kappa\|w^h\|_{\partial\Omega_D}^2 + \tfrac{1}{2}\tau_{\text{eff}}\|\boldsymbol{a}\cdot\nabla w^h\|_\Omega^2$$

$$- \tfrac{1}{4}\kappa\|\nabla w^h\|_\Omega^2 - \tfrac{1}{4}\tau_{\text{eff}}\|\boldsymbol{a}\cdot\nabla w^h\|_\Omega^2 - 3\,c_s^2 T_2\kappa\|w^h\|_{\partial\Omega_D^+}^2$$

$$= \tfrac{1}{4}\tau_{\text{eff}}\|\boldsymbol{a}\cdot\nabla w^h\|_\Omega^2 + \tfrac{1}{2}\|\sqrt{|\boldsymbol{a}\cdot\boldsymbol{n}|}w^h\|_{\partial\Omega}^2 + \tfrac{1}{4}\kappa\|\nabla w^h\|_\Omega^2$$

$$+ (\eta - 2T_1 - 3\,c_s^2 T_2)\kappa\|w^h\|_{\partial\Omega_D}^2 . \tag{6.69}$$

Using the assumption on η from Eq. (6.49) completes the proof. $\square$

Remark 6.2. *Note that both Nitsche's method (Lemma 6.3) and the VMS method (Lemma 6.4) rely on the first term in Eq. (6.61) for their stability. As a result, the combined use of VMS and weakly enforced boundary conditions requires a larger η/T_1 ratio compared to the typical choice of $\eta = 2T_1$ for the standard Nitsche's method [74].*

CHAPTER 7

Numerical experiments

7.1 Numerical verification for a one-dimensional model problem

We present a number of numerical experiments to verify the derivation from Sections 6.1 to 6.3, and to investigate the accuracy improvement that may be achieved by using the new residual-based fine-scale model of sections 6.5 and 6.6.

7.1.1 Linear basis functions

Figure 7.1 shows the result for a simulation with $\boldsymbol{a} = 0.8$, $\kappa = 0.02$ and $\eta = 2/h$ on the domain $\Omega = [0, 0.3]$, discretized with three linear elements. The solid green line shows the exact solution ϕ. With the current discretization, the boundary layer falls completely within a single element. We obtain the exact coarse-scale solution, indicated with the black line, by projecting the exact solution onto the finite element mesh with the Nitsche projector from Section 5.2. The blue dotted line is obtained by only using the classical VMS term, equivalent to $\gamma = 0$, whereas the red dashed solution incorporates the exact augmented VMS model from Eq. (6.23) with the exact parameter definitions

from Eq. (6.24). Finally, the line with the circular markers shows the result when the generalized model of sections 6.5 and 6.6.

It is a celebrated fact that for this model problem the VMS method with strongly enforced boundary conditions results in nodally exact solutions [115]. The results of Figure 7.1a show that this property is lost when the boundary conditions are enforced weakly, and that the under-resolved boundary layer affects the approximation on a large part of the domain. Figure 7.1b shows that the magnitude of the fine-scale solution on the outflow boundary of the domain is considerable. Hence, the assumption of vanishing fine scales, which is critical in the derivation of the classical VMS model (emphasized in Section 3.3), is severely violated. In contrast, the augmented VMS model is exactly the Nitsche projection of the exact solution. As a result, the nodal exactness of the computational solution is retrieved and the adverse effect of the boundary layer is constrained to a single element. Due to the relative simplicity of this problem, solution obtained with the generalized model is only affected by the estimation of the model parameters. As it is nearly identical to the exact coarse-scale solution we conclude that, at least for this simple case, the estimation strategy is effective.

Remark 7.1. *It is well known that for the current case the classical VMS term simplifies to a (consistent) diffusion term. Interestingly, in a similar sense the augmented term in the VMS formulation simplifies to a reduced diffusion in the symmetric part of Nitsche's formulation. In this context, the solution obtained by using the classical VMS model may be interpreted as excessively diffusive in the boundary layer, which is (consistently) counteracted by the augmented VMS term.*

(a) Coarse-scale solutions.　　　　**(b)** Fine-scale solutions (errors).
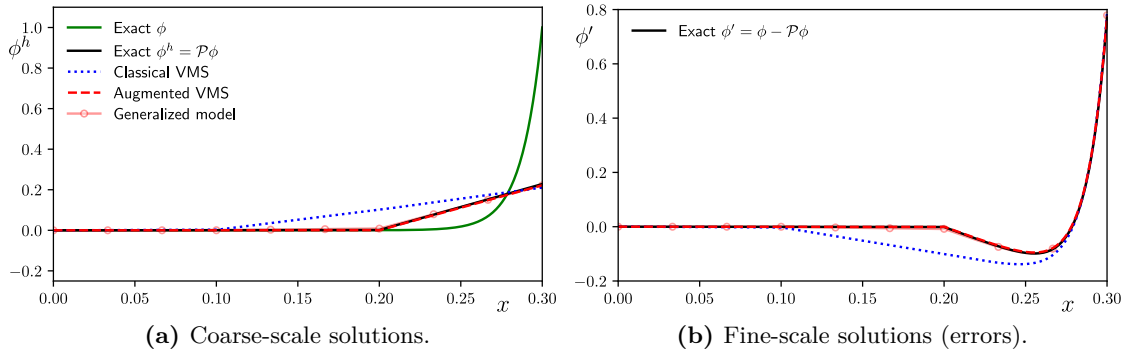
Fig. 7.1: *One-dimensional results using three linear elements.*

7.1.2 Higher-order basis functions

We use the same problem formulation but discretize with three higher-order elements. With quadratic basis functions and $\eta = 3/h$ we obtain the solutions from Figures 7.2a and 7.2b, and with cubic basis function and $\eta = 6/h$ we obtain the solutions from Figures 7.2c and 7.2d. We can largely draw the same conclusions as for the linear basis functions: the fine-scale solution deviates significantly from zero at the outflow boundary. As a result, the solution quality of the classical VMS model is spoiled. By using the augmented VMS model we obtain the Nitsche projection of the exact solution. This gives a nodally exact solution, where the adverse effect of the boundary layer is contained within the boundary element. These points are all compliant with the theory of Section 5.2.

Additionally, we observe that the solution for the approximate augmented VMS model is very close to the exact coarse-scale solution. For the linear basis functions we concluded that the estimations of the model parameters are effective. We can now also conclude that the approximation of the differential operator described Section 6.5.1 is effective, at least for this one-dimensional case.

(a) Coarse-scale solutions, $p = 2$.

(b) Fine-scale solutions (errors), $p = 2$.

(c) Coarse-scale solutions, $p = 3$.

(d) Fine-scale solutions (errors), $p = 3$.
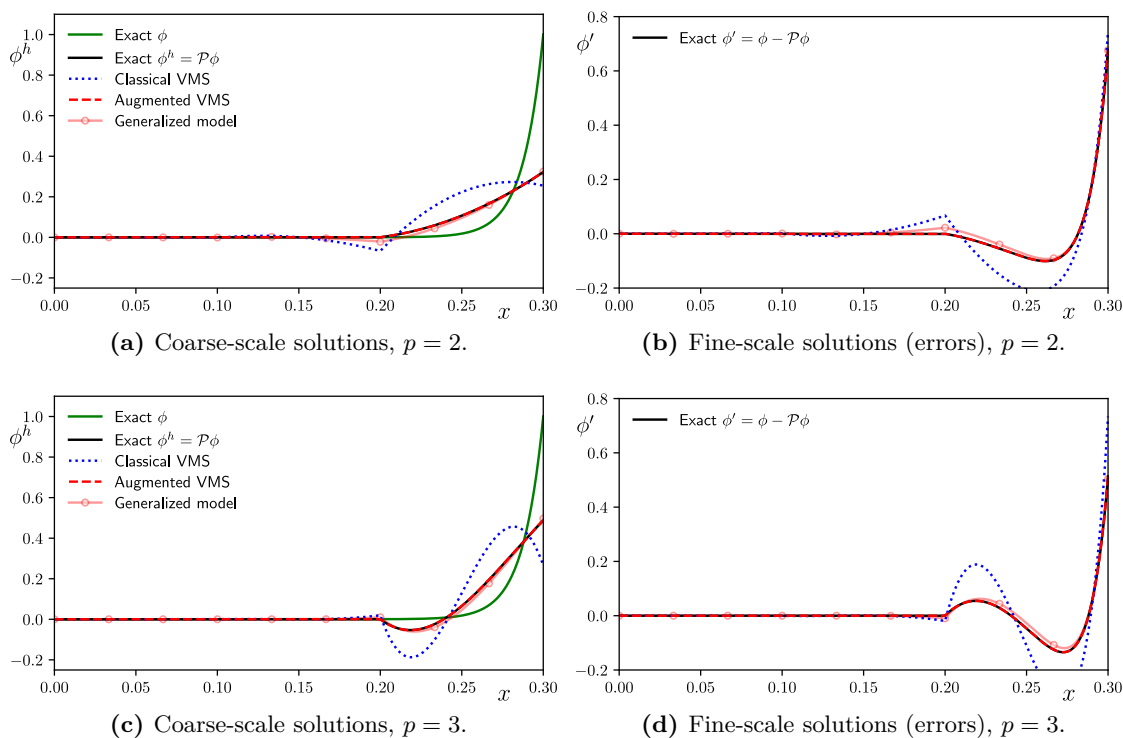
Fig. 7.2: *One-dimensional results using three quadratic or cubic elements.*

7.2 Numerical experiments for a two-dimensional model problem

Next, we present numerical experiments for a two-dimensional domain. All the model approximations become important, and their effectiveness can be assessed.

7.2.1 Linear basis functions, high and low advective dominance

We consider a model problem of a unit square with a circular hole of radius 0.24 in the center. Dirichlet conditions are enforced on all boundaries; $\phi_D = 0$ around the circular cut-out, and $\phi_D = x + y$ around the square. The advective field acts across the diagonal with a magnitude of 0.8, the diffusivity is $\kappa = 0.01$ or $\kappa = 0.003$, and we use $\eta = 10/h$. Figure 7.3a schematically illustrates the model problem, and Figure 7.3b shows the solution for $\kappa = 0.01$ obtained with a highly refined mesh. The solution features multiple boundary layers at various orientations.

The performance of the models can most clearly be assessed by investigating the resulting fine-scale solutions. These are shown in Figures 7.4 and 7.5 for $\kappa = 0.01$ and $\kappa = 0.003$, respectively. In Figures 7.4a and 7.5a classical VMS stabilization is used (i.e., $\gamma = 0$), and Figures 7.4b and 7.5b show the results for the augmented model. Additionally, we project the overrefined solution onto the coarse-scale function space using the Nitsche projector, and show the resulting fine-scale solution in Figures 7.4c and 7.5c. This represents the 'exact' fine-scale solution.

(a) Problem specification. (b) Solution for $\|a\| = 0.8$ and $\kappa = 0.01$.
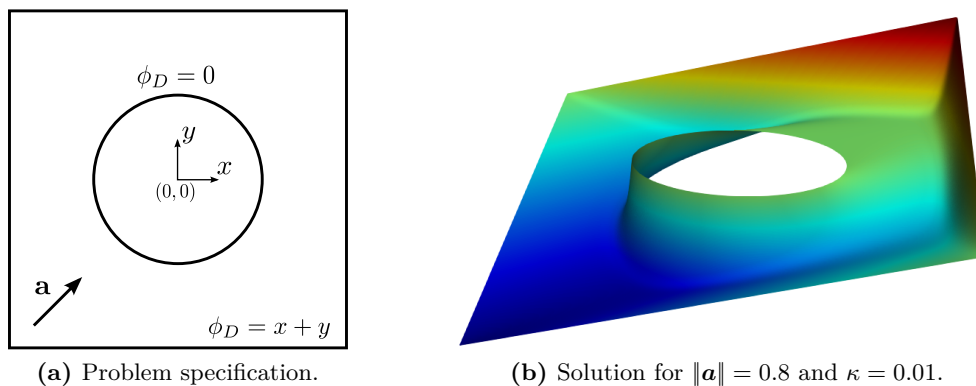
Fig. 7.3: *Two-dimensional model problem for linear basis functions.*
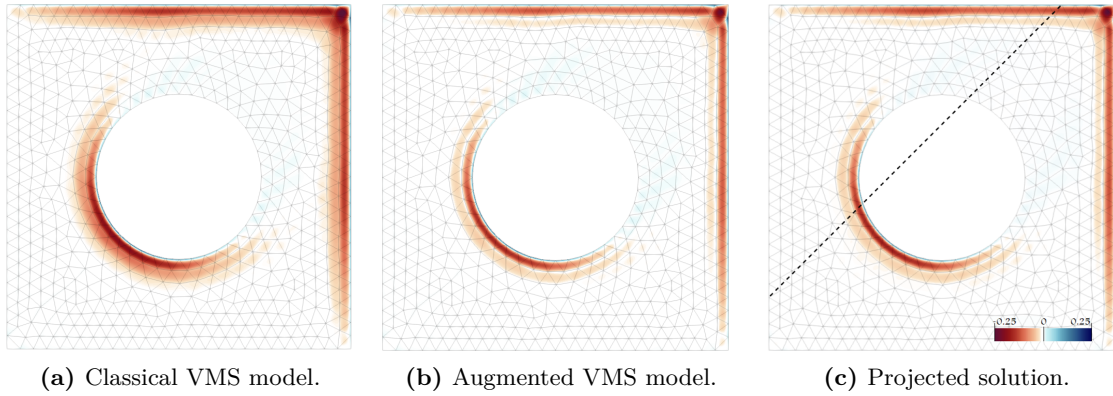
(a) Classical VMS model.　　(b) Augmented VMS model.　　(c) Projected solution.

Fig. 7.4: *Fine-scale solutions $\phi - \phi^h$ (errors) for linear basis functions, $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.01$.*

(a) Classical VMS model.　　(b) Augmented VMS model.　　(c) Projected solution.
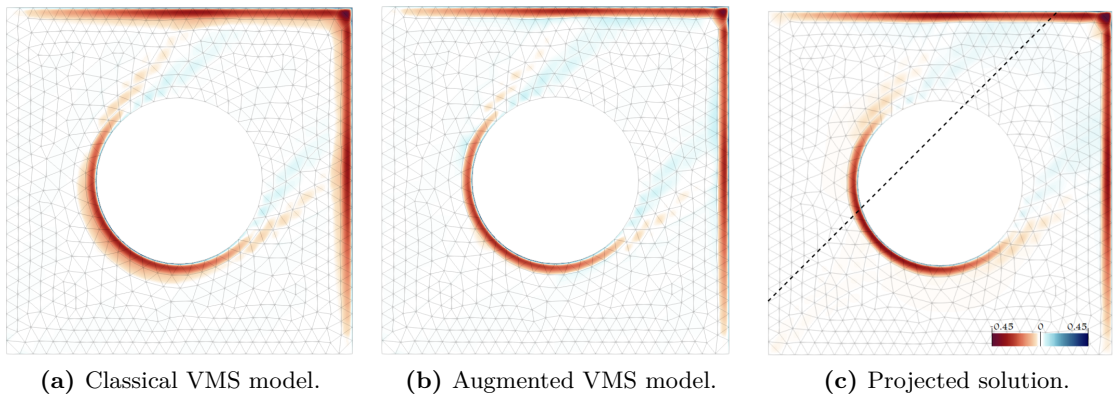
Fig. 7.5: *Fine-scale solutions $\phi - \phi^h$ (errors) for linear basis functions, $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.003$.*
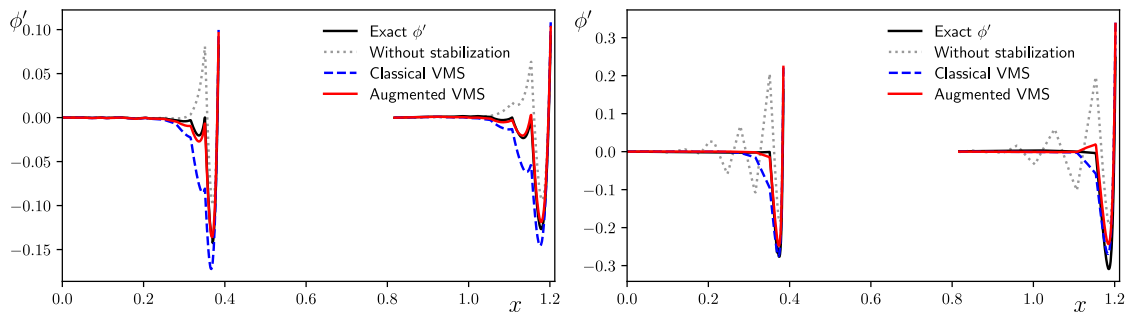
(a) For $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.01$.　　(b) For $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.003$.

Fig. 7.6: *Fine-scale solutions on the cut-planes from Figures 7.4c and 7.5c.*

The observations made for the one-dimensional case almost directly transfer to this two-dimensional problem. The Nitsche projector aims to constrain the impact of the high gradients to the boundary layer elements, without spoiling the results further into the domain. This is illustrated by large fine-scale solutions in only a single row of elements adjacent to the outflow boundary. When the classical VMS model is used, we observe a significant thickening of the range of nonzero fine scales; interpretable as excessive diffusion in the coarse-scale solution. When we add the additional modeling term this thickening is decreased, which leads to nearly the same solution quality as that obtained with the Nitsche projection. We observe these effects irrespective of the Péclet number.

To further illustrate the significance of the change, we show all three solutions on a cut-plane in Figure 7.6. Note, in particular, the similarity of Figure 7.6b and the corresponding figure for the one-dimensional case (Figure 7.1b). The fine-scale solution corresponding to a completely non-stabilized computation is also plotted to put the overall improvement of the solution quality into context.

Convergence in the L^2- or H^1-(semi)norms are not indicative of solution quality for the current case; L^2-projections of shocks lead to highly oscillatory solutions such that the non-stabilized solution often achieves the lowest L^2-error, and neither the classical nor the augmented VMS model aims to achieve optimality in the H^1-seminorm as the boundary conditions are not enforced strongly. Rather, the use of weakly enforced boundary conditions implies the optimality condition of Eq. (5.16), satisfied by solutions that minimize Eq. (5.9). The error of interest is thus the one with respect to the 'optimal' solution; the exact coarse-scale solution obtained with the Nitsche projector. This error indicates the performance of the fine-scale model. This error is plotted for different mesh densities in Figures 7.7a and 7.7b. Figures 7.4 to 7.6 correspond to the third data point in these conver-

(a) For $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.01$.   **(b)** For $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.003$.
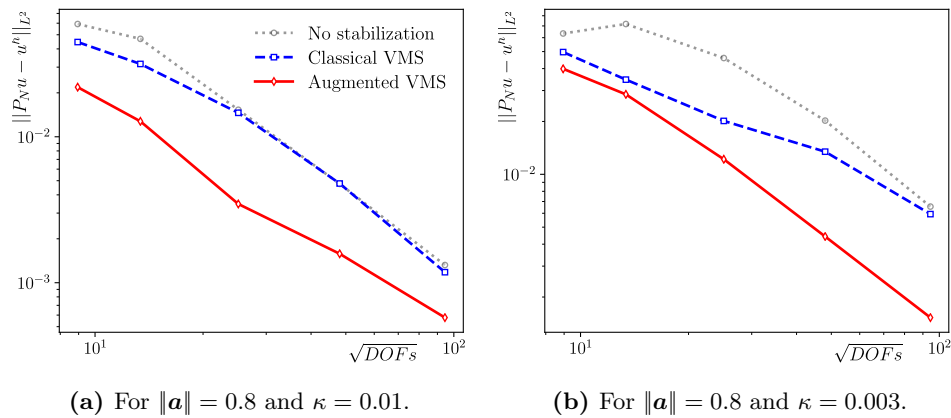
Fig. 7.7: *Error with respect to the exact coarse-scale solution, using linear basis functions.*

gence graphs. Both graphs show a considerable reduction in error when the augmented VMS model is used, which persists throughout mesh refinement. We observe that the $(L^2\text{-})$error reduction from classical to augmented VMS is often of the same magnitude, if not larger, than from non-stabilized to classical VMS.

Remark 7.2. *Analogous to the one-dimensional case, the augmented VMS term decreases the diffusivity in the symmetry part of the Nitsche formulation. Different from the one-dimensional case, this becomes vector-valued and the formulation becomes a streamline directed modified diffusion on the boundary. One could interpret this as a boundary equivalent of the streamline diffusion that the classical VMS terms revert to for the same case.*

7.2.2 Higher-order basis functions

Next, we change the geometry to a square with a polygonal exclusion, as depicted in Figure 7.8. An exact geometry representation can be achieved, which, for these higher-order basis functions, is important for accurately computing the boundary integrals for the Nitsche projection $\mathscr{P}_N\phi$. We focus on the advection dominated case of $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.003$, and we use $\eta = 4p^2/h$.

Figures 7.9a to 7.9c show the fine-scale solutions for quadratic basis functions obtained with the classical VMS model, the augmented VMS model and the Nitsche projector, respectively. We observe that the classical VMS model with the parameter estimation

(a) Problem specification.

(b) Solution for $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.003$.
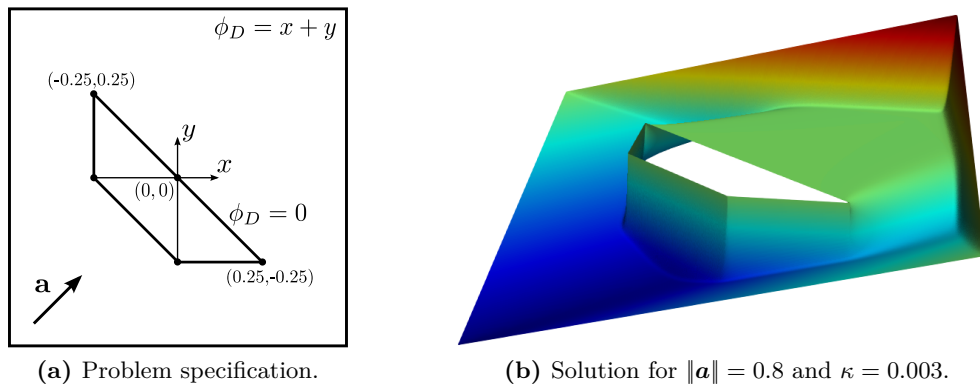
Fig. 7.8: *Two-dimensional model problem for higher-order basis functions.*

developed in Sections 6.5.1 and 6.5.2 already performs remarkably well. The boundary layers are almost exclusively contained in a single row of elements.

When we add the augmented term in the VMS model, the obtained error field qualitatively more closely resembles the true fine-scale solution shown in Figure 7.9c. We do, however, also observe some small oscillations. This is consistent with the decreased diffusion interpretation proposed in Remarks 7.1 and 7.2. A more detailed analysis of the resulting error confirms that the solution obtained with the augmented VMS model more closely resembles the true coarse-scale solution defined by the Nitsche projector. This is shown in Figure 7.10a, where the resulting fine-scale solution is plotted along a cut-plane, as well as in Figure 7.10b, which plots the L^2-error with respect to the true coarse-scale solution for various mesh densities.

(a) Classical VMS model.   **(b)** Augmented VMS model.   **(c)** Projected solution.
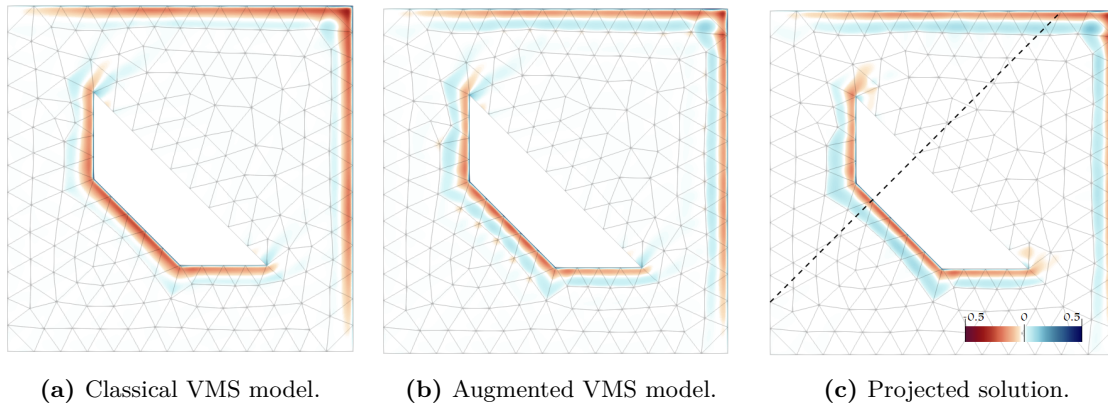
Fig. 7.9: *Fine-scale solutions $\phi - \phi^h$ (errors) for quadratic basis functions, $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.003$.*

(a) Solutions on the cut-plane illustrated in Figure 7.9c.   **(b)** Convergence with respect to exact coarse-scale solution.
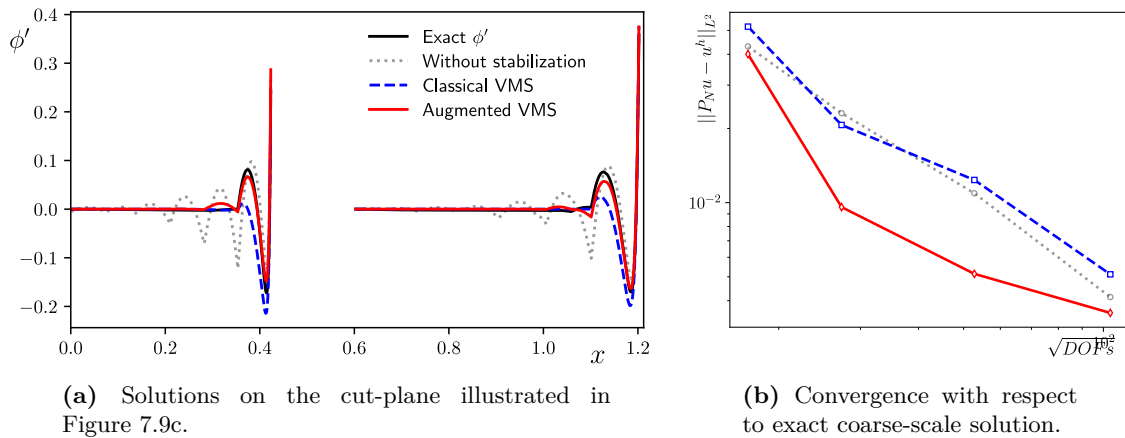
Fig. 7.10: *Detailed error behavior for quadratic basis functions.*
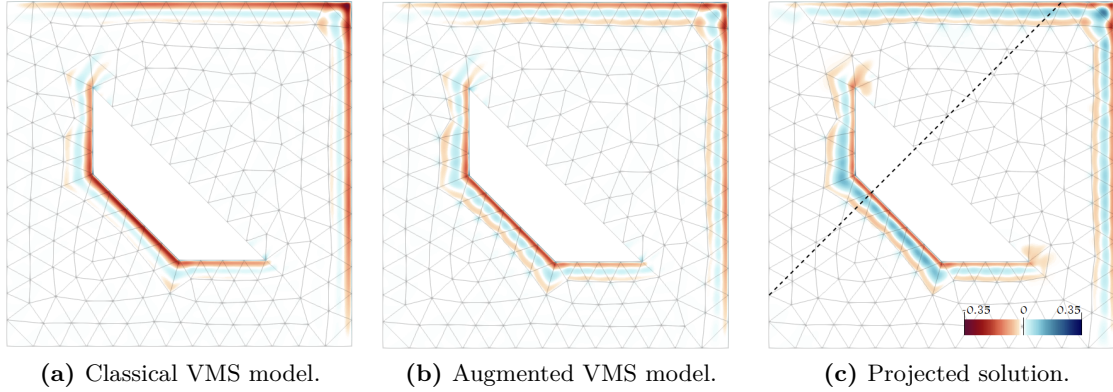
96

(a) Classical VMS model.   **(b)** Augmented VMS model.   **(c)** Projected solution.

Fig. 7.11: *Fine-scale solutions $\phi - \phi^h$ (errors) for cubic basis functions, $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.003$.*

(a) Solutions on the cut-plane illustrated in Figure 7.11c.

(b) Convergence with respect to exact coarse-scale solution.
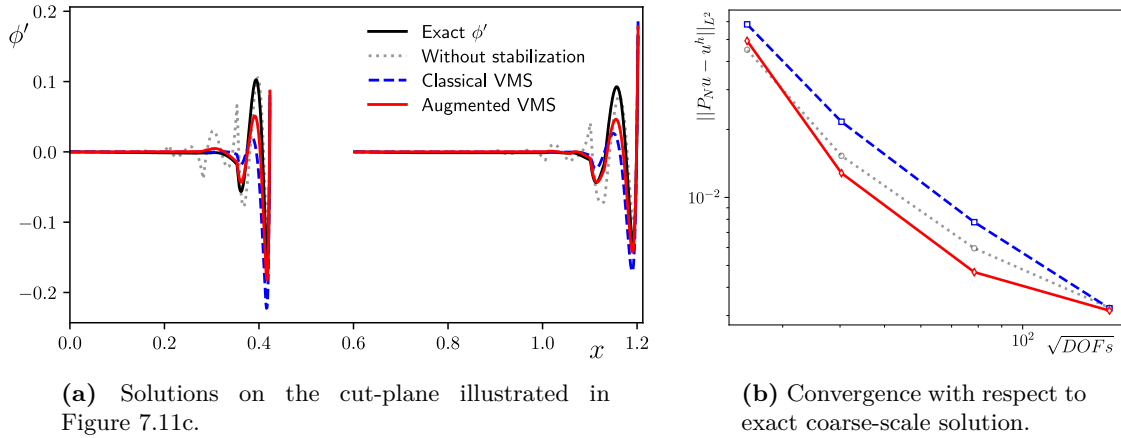
Fig. 7.12: *Detailed error behavior for cubic basis functions.*

Finally, if we use cubic basis functions, we obtain the results from Figures 7.11 and 7.12. Similar conclusions may be drawn as for the case of quadratic basis functions: the classical model with the parameters from Table 6.1 leads to a coarse-scale solution where the error is contained in the first row of elements. Adding the augmented term results in a solution that exhibits small oscillations, but nonetheless bears closer resemblance to the true coarse-scale solution, as measured qualitatively in Figure 7.12a and quantitatively in Figure 7.12b.

It should also be noted that the Dirichlet boundary conditions are more closely satisfied with these cubic basis functions, as shown in Figure 7.12a. This is, at least in part, due to the larger penalty parameter $\eta \propto p^2$. The near-strong enforcement of the Dirichlet condition leaves a small fine-scale boundary value. The new term in the augmented

model becomes almost inoperative, and the classical VMS model suffices. Indeed, the difference between the augmented and classical models is not as pronounced as it was in earlier simulations. These results convey that the augmented model provides fine-scale corrections in the pre-asymptotic regime, and vanishes (asymptotically) when such fine-scale corrections cease to be relevant.

Conclusion

In this part of the dissertation, we have unified the theories of variational multiscale analysis and weakly enforced essential boundary conditions into one consistent framework. Individually, these instrumental numerical methods have shown great value in the context of fluid mechanics. With their merger, we are in a position to develop a fine-scale model that is appropriate for use in combination with Nitsche's method.

When the Dirichlet boundary conditions are enforced weakly, the standard H_0^1-projector is no longer applicable for the scale decomposition around which the variational multiscale method revolves. Instead, we propose a new projector, which we call the Nitsche projector. We show that adoption of this projector in the multiscale formulation naturally leads to Nitsche's formulation. That is, both the penalty term and the symmetry term in Nitsche's formulation automatically fall into place as part of the scale decomposition.

The model for the remaining fine-scale terms is based on the inversion of the fine-scale problem, which, in turn, is formally posed in the kernel space of the projector. We show that the functional constraints that define the H_0^1-projector are a subset of those corresponding to the Nitsche projector, such that we can largely base the inversion of the fine-scale problem on existing theory. An important difference in the context of weakly enforced boundary conditions is that the assumption of vanishing fine scales on element boundaries is no longer applicable for elements adjacent to the Dirichlet boundary. As a result, the fine-scale model that we obtain is the classical VMS model plus an additional boundary term. This 'augmented' term takes into account the non-vanishing fine scales on the Dirichlet boundary. It may be interpreted as a consistent streamline diffusion in the symmetry term of Nitsche's formulation.

Additionally, we develop approximations for the modeling parameters τ and γ based on the fine-scale Green's functions. These expressions and approximation strategies are also suitable for discretization with higher-order basis functions.

With this new model, and these new parameter definitions, we retrieve nodally exact solutions on one-dimensional meshes for all polynomial orders. This is an important

property of the classical VMS model, which was lost when the boundary conditions are enforced weakly. On two-dimensional domains, we observe that the augmented model more closely resembles the actual coarse-scale solution defined by the Nitsche projector, as measured in an L^2 sense. This holds for all polynomial orders. For quadratic and cubic basis functions, the model without the augmented term already performs very well with the newly developed τ approximations. The error due to the boundary layer is contained in a single row of elements. For linear basis functions, however, the classical VMS model leads to a boundary layer that is too thick. This is almost completely mitigated when the augmented model is added to the formulation.

PART III

Discontinuous Galerkin methods

Introduction

Reed and Hill were studying neutron transport when they introduced the first discontinuous Galerkin finite element method in 1973 [160]. The hyperbolic PDE that they were aiming to solve would yield unstable results had they used a conventional finite element approach, and, at the time, residual-based stabilized methods had not been developed. Their alternative approach of using discontinuous basis functions coupled by numerical fluxes appeared not to need any additional stabilization. Of course, we now know that the numerical fluxes *are* a form of stabilization [30]. Based on this property, it has become one of the prevalent finite element methods when it comes to solving hyperbolic problems.

The equations that govern the realm of fluid mechanics are also often of a hyperbolic type. While the pressure field in the incompressible Navier-Stokes equations behaves elliptically, the behavior of the velocity field limits to either that of a parabolic or a hyperbolic PDE, depending on the Reynolds number. The compressible Navier-Stokes equations are completely parabolic/hyperbolic, and the Euler equations of gas dynamics are purely hyperbolic. The natural degree of stability that well designed discontinuous Galerkin formulations demonstrate is one of the reasons for their popularity in fluid mechanics applications, especially when higher-order approximation is required [195]. The second reason is that fluid mechanics revolves around conservation laws: mass, momentum, energy, etc. As the support of the test functions is limited to individual elements, many discontinuous Galerkin methods satisfy element local conservation of the discrete approximations of these quantities [52, 54] (an important side note here is that continuous Galerkin methods can also be proven to satisfy element-local conservation of derived quantities [113, 114]).

As both the variational multiscale paradigm and the discontinuous Galerkin framework have had a substantial impact on the application of finite element methods for fluid mechanics, it seems natural to explore synergies between them. Various efforts in this direction have been made. In [65, 66, 67], Collis et al. devise a 'three-scale' VMS turbulence model in a discontinuous Galerkin framework, which they call the "local variational multiscale" method. In [61], Codina et al. add modeling terms to discontinuous Galerkin formulation inspired by the remaining fine scales on element boundaries. In [63], Coley and

Evans propose to use discontinuous 'bubble functions' to discretize the fine-scale problem. Finally, in [123], Hughes et al. use a VMS decomposition of a solution defined with a discontinuous Galerkin method to separate continuous and discontinuous scales, thereby reducing the computational cost of the original discontinuous Galerkin method. In all of these approaches, the VMS procedure is added 'on top' of a discontinuous Galerkin framework (or vice-versa). A consistent unification of the paradigms is still left to be devised. In this part of the dissertation we thus aim to answer the key questions "what is the optimal coarse-scale representation of an exact solution?" and "how do we obtain that solution" with discontinuous coarse scales in mind. As we saw in Section 3.3, the answers to these questions are intimately related with the design of turbulence models, which is just as relevant for discontinuous Galerkin methods as it is for continuous Galerkin methods, see, e.g., [14, 53, 65, 66, 136, 143, 148, 154, 190, 195]. Secondly, much like how residual-based stabilized methods were developed in the '80s and understood in the context of VMS decades later, just so might it be possible to put the stabilizing effects of discontinuous Galerkin methods under the same umbrella.

Even more than Nitsche's method, which was the focus of Part II, the very nature of discontinuous Galerkin methods appears to conflict with the fundamentals of the variational multiscale method. Discontinuous Galerkin methods are often referred to as "non-conformal" methods, as the approximation spaces are not subspaces of those relevant to the original weak formulation. Variational multiscale analysis relies on a decomposition of the spaces of that original weak formulation. Additionally, the penalty terms required to ensure that discontinuous Galerkin formulation are inf-sup stable become unbounded in the limit of infinite-dimensional function spaces. Even the modeling of the fine scales in the variational multiscale framework hinges on sufficient continuity of the finite element basis: the introductory work on the residual-based turbulence model made use of higher-order NURBS basis functions [2, 19]. To make the theory applicable for conventional $\mathcal{C}^0$-continuous finite element basis functions, the fine scales must incorporate inter-element Dirac-layers [97, 115, 124] or mimic their effect [129]. In practice, the standard residual-based model is also implemented in nodal finite element frameworks [87, 91, 181, 198], which effectively means that the fine scales are assumed to vanish on element boundaries [109, 115, 125]. In the original 1995 paper this was already called a "rather strong assumption" [109]. Certainly in a discontinuous Galerkin framework, where the coarse scales jump, these assumptions are not valid.

In the following, we build on the ideas developed in Part II and extend them to tackle all the above mentioned challenges. This will allow us to develop fine-scale models suitable for discontinuous Galerkin methods. At the same time, we will see that we may interpret

these discontinuous Galerkin formulations as particular choices of scale decompositions. Just like for Nitsche's method, this illustrates that the penalty terms involved in these finite element formulations are not purely numerical: they follow from the earlier two key questions.

The remainder of this part is structured as follows: Chapter 8 is a review of the standard approach for deriving discontinuous Galerkin methods. This will help put the alternative variational multiscale perspective into context, which we develop in Chapter 9 and in Chapter 10. These chapters focus on primal and mixed formulations respectively. In both cases, we show that the VMS perspective can serve as a unifying framework from which a wide variety of discontinuous Galerkin finite element formulations can be derived as special cases of projection operators. Knowledge of these projectors guides the development of the fine-scale model in Chapter 11. We use this model in numerical experiments in Chapter 12, where we explore the impact of the various modeling terms.

Crafting discontinuous Galerkin methods

The most suitable flavor of finite element method for approximating the solution of a partial differential equation naturally depends on the nature of the equation. Partial differential equations may roughly be categorized as "elliptic", "parabolic" or "hyperbolic". The prototypical examples being the Laplace equation, the heat equation and the wave equation. While the precise definition of this classification is quite delicate, we can roughly distinguish between these classes by considering the flow of information. Figure 8.1 shows example solutions of the three prototypical PDEs in two (or 'one-plus-one') dimensions. In the case of the Laplace equation, information flows in all directions. For the heat equation, the information flows in both directions in space but only moves forward in time. For the purely hyperbolic wave equation, the information flows forward in time and along a direction in space that depends on initial conditions.

(a) Elliptic, e.g., $\phi_{xx} + \phi_{yy} = 0$.

(b) Parabolic, e.g., $\phi_t - \phi_{xx} = 0$.

(c) Hyperbolic, e.g., $\phi_{tt} - \phi_{xx} = 0$.
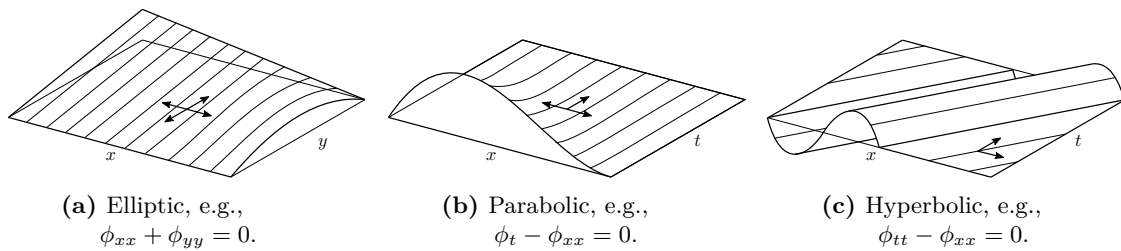
Fig. 8.1: *Different types of partial differential equations.*

(a) Continuous Galerkin. **(b)** Space-continuous, time-discontinuous. **(c)** Discontinuous Galerkin.
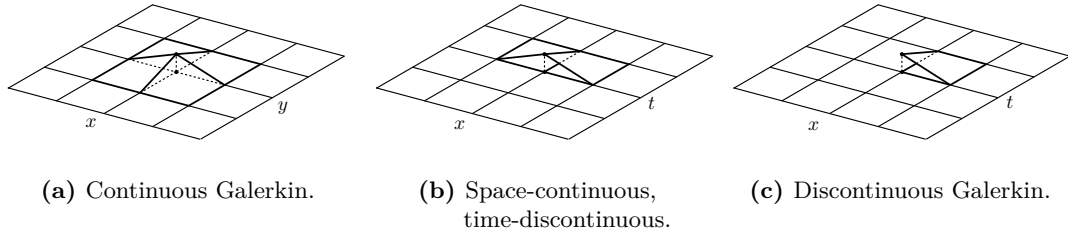
Fig. 8.2: *Different discretization techniques.*

Based on these observations, it makes intuitive sense to discretize the domain of elliptic equations with basis functions that span multiple elements in all directions. This is the case for the classical $\mathcal{C}^0$-continuous finite element method, as illustrated in Figure 8.2a. For parabolic equations this does not appear to be a natural choice: basis functions that root in the future have no business affecting the solution in the past. Instead, one could make use of a space-continuous time-discontinuous finite element method. A basis function associated with such a method is illustrated in Figure 8.2b. Extrapolating, the use of completely discontinuous basis functions appears to be a natural choice for hyperbolic problems, see Figure 8.2c. This motivates the use of discontinuous Galerkin methods for treating hyperbolic PDEs.

8.1 Hyperbolic conservation laws

The general form of a system of hyperbolic conservation laws reads:

$$\frac{\mathrm{d}}{\mathrm{d}t}u_i + \nabla \cdot \boldsymbol{f}_i(\boldsymbol{u}) = 0 \qquad \text{for } i = 1, \cdots, s\,, \tag{8.1}$$

where $\boldsymbol{u}$ is a vector of the s state variables and $\boldsymbol{f}_i(\boldsymbol{u})$ are the vector valued "flux functions" for each of the states. For linear equations, the flux functions take the general form $\boldsymbol{f}_i(\boldsymbol{u}) = \boldsymbol{D}_i\boldsymbol{u}$, where $\boldsymbol{D}_i$ is a (possibly spatially dependent) correlation matrix in $\mathbb{R}^d \times \mathbb{R}^s$. Examples are the scalar advection equation where $\boldsymbol{D}$ becomes the advective field, and the wave equation of Figure 8.1c for which $\boldsymbol{u} = [u_t, u_x]^{\mathrm{T}}$, $\boldsymbol{D}_1 = [0, -1]$ and $\boldsymbol{D}_2 = [-1, 0]$.

To show that this is indeed a conservation law, we integrate over some arbitrary volume Ω:

$$\int_\Omega \frac{\mathrm{d}}{\mathrm{d}t}u_i + \int_\Omega \nabla \cdot \boldsymbol{f}_i(\boldsymbol{u}) = \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega u_i + \int_{\partial\Omega} \boldsymbol{f}_i(\boldsymbol{u}) \cdot \boldsymbol{n} = 0\,, \tag{8.2}$$

106

which states that the total u_i in Ω evolves purely based on the flux through the boundary surface. If $\boldsymbol{f}_i(\boldsymbol{u})$ is continuous in space (and hence single valued on $\partial\Omega$), then the same relation on some neighboring volume would dictate that whatever flows out of Ω flows into its neighbor, implying conservation of u_i.

Developing a discontinuous Galerkin method that approximates Eq. (8.1) begins with multiplying with a test function and integrating over an open element domain K:

$$\int\limits_K \frac{\mathrm{d}}{\mathrm{d}t} u_i\, v_i + \int\limits_K \nabla \cdot \boldsymbol{f}_i(\boldsymbol{u})\, v_i = 0\,. \tag{8.3}$$

We integrate the flux term by parts and sum over all n elements in the mesh $\mathcal{T} = \{K_j\}_{j=1}^n$, which we assume to be a tessellation of the complete domain:

$$\sum_{j=1}^n \int\limits_{K_j} \frac{\mathrm{d}}{\mathrm{d}t} u_i\, v_i - \int\limits_{K_j} \boldsymbol{f}_i(\boldsymbol{u}) \cdot \nabla v_i + \int\limits_{\partial K_j} \boldsymbol{f}_i(\boldsymbol{u}) \cdot \boldsymbol{n}\, v_i = 0\,. \tag{8.4}$$

Next, we define the 'broken' domain $\tilde{\Omega}$ as the union of all open element domains, and the mesh skeleton Γ as the union of all element interfaces:

$$\tilde{\Omega} := \bigcup_{i=1}^n K_j\,, \tag{8.5a}$$

$$\Gamma := \bigcup_{i=1}^n \partial K_j\,, \tag{8.5b}$$

$$\Gamma_0 := \Gamma \setminus \partial\Omega\,. \tag{8.5c}$$

We also make use of the following interface identity:

$$\sum_{j=1}^n \int\limits_{\partial K_j} \boldsymbol{w} \cdot \boldsymbol{n}\, v = \int\limits_\Gamma [\![\boldsymbol{w}]\!]\{\!\{v\}\!\} + \int\limits_{\Gamma_0} \{\!\{\boldsymbol{w}\}\!\} \cdot [\![v]\!]\,, \tag{8.6}$$

where we define the jump operators as:

$$[\![v]\!] := \begin{cases} v^+ \boldsymbol{n}^+ + v^- \boldsymbol{n}^- & \text{on } \Gamma_0\,, \\ v\boldsymbol{n} & \text{on } \partial\Omega\,, \end{cases} \tag{8.7a}$$

$$[\![\boldsymbol{w}]\!] := \begin{cases} \boldsymbol{w}^+ \cdot \boldsymbol{n}^+ + \boldsymbol{w}^- \cdot \boldsymbol{n}^- & \text{on } \Gamma_0\,, \\ \boldsymbol{w} \cdot \boldsymbol{n} & \text{on } \partial\Omega\,, \end{cases} \tag{8.7b}$$

and the average operator as:

$$\{v\} := \begin{cases} \frac{1}{2}(v^+ + v^-) & \text{on } \Gamma_0 \,, \\ v & \text{on } \partial\Omega \,. \end{cases} \tag{8.8}$$

The pluses and minuses in these definitions arbitrarily refer to the solution fields on either side of the interface. Note that the jump of a scalar becomes a vector, and that the jump of a t-th order tensor becomes a tensor of order $t - 1$ (e.g., a vector becomes a scalar).

By using the definitions from Eq. (8.5) and the interface identity from Eq. (8.6) we can write Eq. (8.4) as:

$$\int_{\tilde{\Omega}} \frac{\mathrm{d}}{\mathrm{d}t} u_i \, v_i - \int_{\tilde{\Omega}} \boldsymbol{f}_i(\boldsymbol{u}) \cdot \nabla v_i + \int_{\Gamma} [\![\boldsymbol{f}_i(\boldsymbol{u})]\!] \, \{v_i\} + \int_{\Gamma_0} \{\boldsymbol{f}_i(\boldsymbol{u})\} \cdot [\![v_i]\!] = 0 \,. \tag{8.9}$$

Next, we approximate with a finite element method. If we allow solution fields to be discontinuous from element to element, then the evaluation of the flux function on the element boundary becomes ambiguous. Various formulations may be obtained by placing more or less weight on the solution from either side of the interface. Different choices affect the properties of the concluding approximation method. To make this explicit, we replace the flux function with a "numerical flux", denoted $\hat{\boldsymbol{f}}_i^{\pm}$. We are left with the following discrete representation:

$$\int_{\tilde{\Omega}} \frac{\mathrm{d}}{\mathrm{d}t} u_i^h \, v_i - \int_{\tilde{\Omega}} \boldsymbol{f}_i(\boldsymbol{u}^h) \cdot \nabla v_i + \int_{\Gamma} [\![\hat{\boldsymbol{f}}_i]\!] \, \{v_i\} + \int_{\Gamma_0} \{\hat{\boldsymbol{f}}_i\} \cdot [\![v_i]\!] = 0 \,. \tag{8.10}$$

The numerical flux is a function of $\boldsymbol{u}^h$ from either side of the interface:

$$\hat{\boldsymbol{f}}_i^{\pm} = \hat{\boldsymbol{f}}_i(\boldsymbol{u}^{h,\pm}, \boldsymbol{u}^{h,\mp}) \,. \tag{8.11}$$

Typically, a suitable numerical flux satisfies two properties that mimic the real flux function: conservativity and consistency. A conservative flux ensures that Eq. (8.10) conserves u_i across interfaces. This requires that the jump of the numerical flux vanishes, and thus that the normal component of the numerical flux is single valued on element boundaries. Based on the definition of Eq. (8.11), this means that the normal component of $\hat{\boldsymbol{f}}_i^{\pm}$ is symmetric with respect to its arguments:

$$[\![\hat{\boldsymbol{f}}_i]\!] = 0 \;\; \Rightarrow \;\; \hat{\boldsymbol{f}}_i^+ \cdot \boldsymbol{n}^+ = -\hat{\boldsymbol{f}}_i^- \cdot \boldsymbol{n}^- \;\; \Rightarrow \;\; \hat{\boldsymbol{f}}_i(\boldsymbol{u}^+, \boldsymbol{u}^-) \cdot \boldsymbol{n}^+ = -\hat{\boldsymbol{f}}_i(\boldsymbol{u}^-, \boldsymbol{u}^+) \cdot \boldsymbol{n}^- . \tag{8.12}$$

Consistency requires that the numerical flux equals the true flux if the true (entropy [81]) solution is substituted as its arguments:

$$\hat{\boldsymbol{f}}_i(\boldsymbol{u}, \boldsymbol{u}) = \boldsymbol{f}_i(\boldsymbol{u}) \,. \tag{8.13}$$

Thus far, the setup of the method is reminiscent a finite volume approach. Accordingly, many of the numerical fluxes that were originally designed for finite volume methods have also been investigated for discontinuous Galerkin formulations. The following are the most cited ones [46, 47, 100]:

The Lax-Friedrichs flux:

$$\hat{\boldsymbol{f}}_i(\boldsymbol{u}^\pm, \boldsymbol{u}^\mp) \cdot \boldsymbol{n}^\pm = \left[\tfrac{1}{2} \boldsymbol{f}_i(\boldsymbol{u}^\pm) + \tfrac{1}{2} \boldsymbol{f}_i(\boldsymbol{u}^\mp) + C[\![u_i]\!] \right] \cdot \boldsymbol{n}^\pm \,,$$
$$\text{with: } C = \max_{\boldsymbol{w} \in \boldsymbol{S}(\boldsymbol{u}^+, \boldsymbol{u}^-)} \left| \frac{\partial \boldsymbol{f}_i}{\partial u_i}(\boldsymbol{w}) \cdot \boldsymbol{n}^\pm \right| \,, \tag{8.14}$$

where $\boldsymbol{S}(\boldsymbol{u}^+, \boldsymbol{u}^-)$ is the set of all states in between $\boldsymbol{u}^+$ and $\boldsymbol{u}^-$:

$$\boldsymbol{S}(\boldsymbol{u}^+, \boldsymbol{u}^-) = \{ \boldsymbol{w} \in \mathbb{R}^s : \min(u_i^+, u_i^-) \le w_i \le \max(u_i^+, u_i^-) \text{ for } i = 1, \cdots, s \} \,. \tag{8.15}$$

The Godunov flux [89]:

$$\hat{\boldsymbol{f}}_i(\boldsymbol{u}^\pm, \boldsymbol{u}^\mp) \cdot \boldsymbol{n}^\pm = \begin{cases} \displaystyle\min_{\boldsymbol{w} \in \boldsymbol{S}(\boldsymbol{u}^+, \boldsymbol{u}^-)} \boldsymbol{f}_i(\boldsymbol{w}) \cdot \boldsymbol{n}^\pm & \text{if } u_i^+ \le u_i^- \,, \\ \displaystyle\max_{\boldsymbol{w} \in \boldsymbol{S}(\boldsymbol{u}^+, \boldsymbol{u}^-)} \boldsymbol{f}_i(\boldsymbol{w}) \cdot \boldsymbol{n}^\pm & \text{if } u_i^+ > u_i^- \,. \end{cases} \tag{8.16}$$

The Engquist-Osher flux for scalar conservation laws [75]:

$$\hat{\boldsymbol{f}}_i(u^\pm, u^\mp) \cdot \boldsymbol{n}^\pm = \int_0^{u^\pm} \max(0, \frac{\partial \boldsymbol{f}}{\partial u}(w) \cdot \boldsymbol{n}^\pm) \, \mathrm{d}w$$
$$+ \int_0^{u^\mp} \min(0, \frac{\partial \boldsymbol{f}}{\partial u}(w) \cdot \boldsymbol{n}^\pm) \, \mathrm{d}w \,, \tag{8.17}$$

and the Roe flux [162]:

$$\hat{\boldsymbol{f}}_i(\boldsymbol{u}^\pm, \boldsymbol{u}^\mp) \cdot \boldsymbol{n}^\pm = \left[\tfrac{1}{2} \boldsymbol{f}_i(\boldsymbol{u}^\pm) + \tfrac{1}{2} \boldsymbol{f}_i(\boldsymbol{u}^\mp) \right] \cdot \boldsymbol{n}^\pm - \tfrac{1}{2} \sum_{j=1}^s \alpha^j(\boldsymbol{u}^\pm, \boldsymbol{u}^\mp) \, |\lambda^j| \, r_i^j \,, \tag{8.18}$$

where $\boldsymbol{r}^j$ and λ^j are the j-th right-eigenvector and right-eigenvalue of the linearized flux

function. The $\alpha^j(\boldsymbol{u}^\pm, \boldsymbol{u}^\mp)$ coefficients correspond to the representation of $\boldsymbol{u}^\pm \pm \boldsymbol{u}^\mp$ in the $\boldsymbol{r}$-basis:

$$\alpha^j(\boldsymbol{u}^\pm, \boldsymbol{u}^\mp) \text{ s.t.: } \boldsymbol{u}^\pm - \boldsymbol{u}^\mp = \sum_{j=1}^s \alpha^j \boldsymbol{r}^j. \tag{8.19}$$

The Roe flux is what is called an "approximate Riemann solver" [162]. Such fluxes are designed to approximate the solution behavior of a "Riemann problem", which is a hyperbolic conservation law with discontinuous initial data [81]. Their relevance is obvious: every time step in a discontinuous Galerkin method concerns the propagation of discontinuous data. The "van Leer", "Harten, Lax and van Leer" (HLL) and "Harten, Lax and van Leer - Contact" (HLLC) [96, 191, 192, 193, 194] fluxes are other examples of approximate Riemann solvers that are used for discontinuous Galerkin methods.

Remark 8.1. *All the flux formulations that have just been cited lead to the same expression for the linear advection equation:*

$$\hat{\boldsymbol{f}}^\pm \cdot \boldsymbol{n}^\pm = \boldsymbol{a} \cdot \boldsymbol{n}^\pm \{\!\{u\}\!\} + \tfrac{1}{2}|\boldsymbol{a} \cdot \boldsymbol{n}| \,[\![u]\!], \tag{8.20}$$

which is conservative and consistent. After rewriting, we can see that this flux exclusively concerns the upwind value of u. Hence, this is called the upwind flux:

$$\begin{aligned}
\hat{\boldsymbol{f}}_i^\pm \cdot \boldsymbol{n}^\pm &= \boldsymbol{a} \cdot \boldsymbol{n}^\pm \{\!\{u\}\!\} + \tfrac{1}{2}|\boldsymbol{a} \cdot \boldsymbol{n}| \,[\![u]\!] \cdot \boldsymbol{n}^\pm \\
&= \tfrac{1}{2}\boldsymbol{a} \cdot \boldsymbol{n}^\pm(u^{up} + u^{down}) + \tfrac{1}{2}\,\boldsymbol{a} \cdot \boldsymbol{n}^{up}\,(u^{up}\boldsymbol{n}^{up} + u^{down}\boldsymbol{n}^{down}) \cdot \boldsymbol{n}^\pm \\
&= \tfrac{1}{2}(u^{up}\boldsymbol{a} + u^{down}\boldsymbol{a}) \cdot \boldsymbol{n}^\pm + \tfrac{1}{2}(u^{up}\boldsymbol{a} - u^{down}\boldsymbol{a}) \cdot \boldsymbol{n}^\pm \\
&= u^{up}\,\boldsymbol{a} \cdot \boldsymbol{n}^\pm,
\end{aligned} \tag{8.21}$$

where

$$u^{up} = \lim_{\epsilon \to 0^+} u(\boldsymbol{x} - \epsilon \boldsymbol{a}). \tag{8.22}$$

If $\boldsymbol{a}$ runs tangent to the boundary this definition becomes ambiguous. However, Eq. (8.21) involves a multiplication with $\boldsymbol{a} \cdot \boldsymbol{n}^\pm$ which would be zero.

All the aforementioned numerical fluxes couple the discontinuous basis functions in space. The treatment of the time derivative is a whole other story. Hyperbolic conservation laws tend to develop sharp layers and shocks. In the nonlinear case, even the question of uniqueness becomes complicated (requiring the concept of entropy solutions). Construction of numerical schemes focuses on them being "total variation diminishing" [90, 171]. Often, this requires the introduction of slope limiting procedures or shock-capturing terms. An honorable mention of the Runge-Kutta Discontinuous Galerkin (RKDG) method is in order. This method, introduced in the series of papers [50, 55, 56, 57, 59, 60], was one of the earliest generalized frameworks with proven stability characteristics. It combines a Runge-Kutta time-stepping algorithm with shock capturing mechanisms and numerical fluxes to arrive at a scheme that is total variation bounded.

8.2   The second-order elliptic operator

The natural ability to treat hyperbolic terms makes the discontinuous Galerkin method a welcome addition to the family of finite element methods. Still, the relevance of purely hyperbolic systems is limited. For increased versatility, it is imperative that we can treat elliptic (e.g., diffusive) terms [8]. The first methods were developed a few years after the work of Reed and Hill: between 1975 and 1980 a number of Interior Penalty type methods were proposed [5, 13, 73, 196]. After about 20 years of minor advances, a resurgence of methods occurred in the '90s. All these methods were put in a unified framework in the seminal 2002 work by Arnold, Brezzi, Cockburn and Marini [9]. Below follows a short summary of the theory that is relevant to this dissertation.

As a model problem for the second order elliptic operator, we will consider the steady state diffusion equation with mixed Dirichlet/Neumann boundary data:

$$-\nabla \cdot (\kappa \nabla \phi) = 0 \qquad \text{in } \Omega \,, \tag{8.23a}$$

$$\phi = \phi_D \qquad \text{on } \partial \Omega_D \,, \tag{8.23b}$$

$$\kappa \partial_n \phi = g_N \qquad \text{on } \partial \Omega_N \,. \tag{8.23c}$$

As highlighted in the previous section, different choices of numerical fluxes lead to different discontinuous Galerkin methods. By writing the PDE in a reduced order form (see Chapter 4), we will be able to introduce numerical fluxes for both the primary and the auxiliary variable, resulting in a wide range of methods. As an auxiliary variable, we

introduce $\boldsymbol{\sigma}$ as the diffusive flux:

$$\boldsymbol{\sigma} = -\kappa\nabla\phi \qquad \text{in } \Omega\,, \tag{8.24a}$$

$$\nabla\cdot\boldsymbol{\sigma} = 0 \qquad \text{in } \Omega\,, \tag{8.24b}$$

$$\phi = \phi_D \qquad \text{on } \partial\Omega_D\,, \tag{8.24c}$$

$$-\boldsymbol{\sigma}\cdot\boldsymbol{n} = g_N \qquad \text{on } \partial\Omega_N\,. \tag{8.24d}$$

Following the procedure from the previous section we obtain a weak formulation by multiplication with a test function, integration over each element separately, per-element integration by parts and summation:

$$\left(\kappa^{-1}\boldsymbol{\tau}, \boldsymbol{\sigma}\right)_{\tilde{\Omega}} = \left(\nabla\cdot\boldsymbol{\tau}, \phi\right)_{\tilde{\Omega}} - \left\langle[\![\boldsymbol{\tau}]\!], \{\!\{f(\phi)\}\!\}\right\rangle_{\Gamma} - \left\langle\{\!\{\boldsymbol{\tau}\}\!\}, [\![f(\phi)]\!]\right\rangle_{\Gamma_0}\,, \tag{8.25a}$$

$$-\left(\nabla w, \boldsymbol{\sigma}\right)_{\tilde{\Omega}} + \left\langle\{\!\{w\}\!\}, [\![\boldsymbol{g}(\boldsymbol{\sigma})]\!]\right\rangle_{\Gamma} + \left\langle[\![w]\!], \{\!\{\boldsymbol{g}(\boldsymbol{\sigma})\}\!\}\right\rangle_{\Gamma_0} = 0\,, \tag{8.25b}$$

where the exact flux functions are:

$$f(\phi) = \phi\,, \tag{8.26a}$$

$$\boldsymbol{g}(\boldsymbol{\sigma}) = \boldsymbol{\sigma}\,. \tag{8.26b}$$

In order to discretize with piecewise discontinuous basis functions, we replace these by numerical fluxes:

$$\hat{f}^{\pm} = \hat{f}(\phi^{\pm}, \boldsymbol{\sigma}^{\pm}, \phi^{\mp}, \boldsymbol{\sigma}^{\mp})\,, \tag{8.27a}$$

$$\hat{\boldsymbol{g}}^{\pm} = \hat{\boldsymbol{g}}(\phi^{\pm}, \boldsymbol{\sigma}^{\pm}, \phi^{\mp}, \boldsymbol{\sigma}^{\mp})\,. \tag{8.27b}$$

We then introduce finite-dimensional spaces for ϕ, $\boldsymbol{\sigma}$ and the associated test functions. Afterwards, we perform reversed integration by parts on $(\nabla w, \boldsymbol{\sigma})_{\Omega}$ to arrive at a formulation with symmetric operators. This general form reads:

Find $\phi^h, \boldsymbol{\sigma}^h \in \mathcal{W}^h \times \Sigma^h$ s.t. $\forall\, w, \boldsymbol{\tau} \in \mathcal{W}^h \times \Sigma^h$:

$$-\left(\kappa^{-1}\boldsymbol{\tau}, \boldsymbol{\sigma}^h\right)_{\tilde{\Omega}} + \left(\nabla\cdot\boldsymbol{\tau}, \phi^h\right)_{\tilde{\Omega}} - \left\langle[\![\boldsymbol{\tau}]\!], \{\!\{\hat{f}\}\!\}\right\rangle_{\Gamma} - \left\langle\{\!\{\boldsymbol{\tau}\}\!\}, [\![\hat{f}]\!]\right\rangle_{\Gamma_0} = 0\,, \tag{8.28a}$$

$$\left(w, \nabla\cdot\boldsymbol{\sigma}^h\right)_{\tilde{\Omega}} - \left\langle\{\!\{w\}\!\}, [\![\boldsymbol{\sigma}^h - \hat{\boldsymbol{g}}]\!]\right\rangle_{\Gamma} - \left\langle[\![w]\!], \{\!\{\boldsymbol{\sigma}^h - \hat{\boldsymbol{g}}\}\!\}\right\rangle_{\Gamma_0} = 0\,. \tag{8.28b}$$

The spaces $\mathcal{W}^h$ and Σ^h are typically spanned by polynomials restricted to the elements.

Different methods follow from different choices of $\hat{f}^{\pm}$ and $\hat{\boldsymbol{g}}^{\pm}$. Table 8.1 provides an overview of the most used ones, and those that will be considered in subsequent chapters.

It includes the Interior Penalty (IP) method and its nonsymmetric variant (NIP), the Baumann-Oden method, Bassi-Rebay's 1st method (their second method requires an additional volumetric term in Eq. (8.28a) for stabilization), the Local Discontinuous Galerkin (LDG) method and the Hybridizable Discontinuous Galerkin (HDG) method.

On the domain boundary, all methods treat the primal flux $\hat{f}$ the same:

$$\hat{f} = \begin{cases} \phi_D & \text{on } \partial\Omega_D \,, \\ \phi^h & \text{on } \partial\Omega_N \,. \end{cases} \tag{8.29a}$$

The IP, NIP and Baumann-Oden methods treat the auxiliary flux on boundary facets as:

$$\hat{\boldsymbol{g}} \cdot \boldsymbol{n} = \begin{cases} -\kappa\nabla\phi^h \cdot \boldsymbol{n} + \kappa\eta(\phi^h - \phi_D) & \text{on } \partial\Omega_D \,, \\ -g_N & \text{on } \partial\Omega_N \,, \end{cases} \tag{8.30a}$$

where η is a positive penalty parameter for the IP and NIP methods, and $\eta = 0$ for the Baumann-Oden method. For Bassi-Rebay's 1st method, the LDG method and the HDG method it is treated as:

$$\hat{\boldsymbol{g}} \cdot \boldsymbol{n} = \begin{cases} \boldsymbol{\sigma}^h \cdot \boldsymbol{n} + \kappa\eta(\phi^h - \phi_D) & \text{on } \partial\Omega_D \,, \\ -g_N & \text{on } \partial\Omega_N \,, \end{cases} \tag{8.31a}$$

where Bassi-Rebay's 1st method follows from $\eta = 0$.

Table 8.1: *Numerical fluxes on interior facets for various discontinuous Galerkin methods for the elliptic operator. Adopted from [9].*

Method name	Flux $\hat{f}^{\pm}$	Flux $\hat{\boldsymbol{g}}^{\pm}$
IP [5]	$\{\!\{\phi^h\}\!\}$	$-\{\!\{\kappa\nabla\phi^h\}\!\} + \kappa\eta[\![\phi^h]\!]$
NIP [161]	$\frac{3}{2}\phi^{h,\pm} - \frac{1}{2}\phi^{h,\mp}$	$-\{\!\{\kappa\nabla\phi^h\}\!\} + \kappa\eta[\![\phi^h]\!]$
Baumann-Oden [17]	$\frac{3}{2}\phi^{h,\pm} - \frac{1}{2}\phi^{h,\mp}$	$-\{\!\{\kappa\nabla\phi^h\}\!\}$
Bassi-Rebay's 1st [15]	$\{\!\{\phi^h\}\!\}$	$\{\!\{\boldsymbol{\sigma}^h\}\!\}$
LDG [58]	$\{\!\{\phi^h\}\!\} - \boldsymbol{\beta} \cdot [\![\phi^h]\!]$	$\{\!\{\boldsymbol{\sigma}^h\}\!\} + \boldsymbol{\beta}[\![\boldsymbol{\sigma}^h]\!] + \kappa\eta[\![\phi^h]\!]$
HDG [48]	$\{\!\{\phi^h\}\!\} - \boldsymbol{\beta} \cdot [\![\phi^h]\!] + C[\![\boldsymbol{\sigma}^h]\!]$	$\{\!\{\boldsymbol{\sigma}^h\}\!\} + \boldsymbol{\beta}[\![\boldsymbol{\sigma}^h]\!] + \kappa\eta[\![\phi^h]\!]$

Same as for the hyperbolic model problem, the fluxes of all these methods may be characterized as conservative and/or consistent (or neither). We have chosen to focus exclusively on consistent methods: when the true solutions are substituted into the expressions of Table 8.1, then we retrieve the true fluxes of Eq. (8.26). Not all methods are conservative, however. The dependency on $+/-$ in the primal fluxes of the NIP and Baumann-Oden methods make them dependent on the side of the facet on which they are evaluated. The jumps of these fluxes are nonzero and hence the methods are not conservative.

Most of the methods in Table 8.1 are stable and can be proven to converge optimally. Under refinement this requires $\eta \propto h^{-1}$. The exceptions are Bassi-Rebay's 1$^{\text{st}}$ method which is generally unstable (meaning that the inf-sup requirement of Theorem 2.3 does in general not hold), the Baumann-Oden method which is unstable for $p = 1$ polynomial bases and can only be proven to converge sub-optimally in the L^2-norm for higher order bases, and the Nonsymmetric Interior Penalty method which can also only be proven to converge suboptimally in the L^2-norm [9, 37, 40, 161].

The variational multiscale decomposition for broken spaces

Now that we are familiar with the common approach for deriving discontinuous Galerkin formulations, we proceed with offering an alternative. In this chapter, we use discontinuous approximation spaces in the variational multiscale method. We show that the Interior Penalty formulation emerges from a special choice of scale decomposition.

9.1 Multiscale weak formulation with flexible spaces at element boundaries

Recall the governing strong form of the steady advection-diffusion equation:

$$\boldsymbol{a} \cdot \nabla\phi - \nabla \cdot \kappa\nabla\phi = f \quad \text{in} \quad \Omega \,, \tag{9.1a}$$

$$\phi = \phi_D \quad \text{on} \quad \partial\Omega_D \,, \tag{9.1b}$$

$$\kappa \, \partial_n\phi = g_N \quad \text{on} \quad \partial\Omega_N^+ \,, \tag{9.1c}$$

$$\kappa \, \partial_n\phi - \boldsymbol{a} \cdot \boldsymbol{n}\phi = g_N \quad \text{on} \quad \partial\Omega_N^- \,. \tag{9.1d}$$

We again make the assumptions that the source function is square integrable and that the advective field is solenoidal: $f \in L^2(\Omega)$ and $\nabla \cdot \boldsymbol{a} = 0$. Recall that the superscripts $+$ and $-$ indicate outflow $(\boldsymbol{a} \cdot \boldsymbol{n} \geq 0)$ and inflow $(\boldsymbol{a} \cdot \boldsymbol{n} < 0)$ parts of the boundary, respectively.

Next, we decompose Ω into a collection of open subdomains $\mathcal{T} = \{K_i\}_{i=1}^n$. To ensure that $\mathcal{T}$ is a tessellation of Ω, the subdomains need to satisfy:

$$\bigcup_{i=1}^{n} \bar{K}_i = \bar{\Omega}\,, \tag{9.2a}$$

$$K_i \cap K_j = \emptyset \ \ \forall\, i \neq j\,, \tag{9.2b}$$

where $\bar{K}_i$ and $\bar{\Omega}$ are the closures of K_i and Ω. Based on this tessellation, we can define the set of open subdomains $\tilde{\Omega}$, and their interface skeleton Γ:

$$\tilde{\Omega} := \bigcup_{i=1}^{n} K_i\,, \tag{9.3a}$$

$$\Gamma := \bigcup_{i=1}^{n} \partial K_i\,. \tag{9.3b}$$

We will interpret any integral over $\tilde{\Omega}$ as the sum of the integrals over the subdomains:

$$(\phi, w)_{\tilde{\Omega}} = \sum_{i=1}^{n} (\phi, w)_{K_i}\,. \tag{9.4}$$

We denote the interior part of the interface skeleton Γ_0, the interior interfaces plus the Dirichlet boundary Γ_D, and the interior interfaces plus the Neumann boundary Γ_N:

$$\Gamma_0 := \Gamma \setminus \partial\Omega\,, \tag{9.5a}$$

$$\Gamma_D := \Gamma \setminus \partial\Omega_N = \Gamma_0 \cup \partial\Omega_D\,, \tag{9.5b}$$

$$\Gamma_N := \Gamma \setminus \partial\Omega_D = \Gamma_0 \cup \partial\Omega_N\,. \tag{9.5c}$$

We now wish to pose Eq. (9.1a) on each open domain K_i, while adding transmission conditions on Γ_0 that ensure equivalence with Eq. (9.1). The appropriate transmission conditions follow from the solution regularity that is implied and required by the PDE. Since the advective field is solenoidal, we may write Eq. (9.1a) as:

$$\nabla \cdot (\boldsymbol{a}\phi - \kappa\nabla\phi) = f \qquad \text{in } \Omega\,. \tag{9.6}$$

This description actually more closely represents the origin of the PDE as the conservation law that governs the transport of ϕ. Since $f \in L^2(\Omega)$, the strong form dictates $\nabla \cdot (\boldsymbol{a}\phi - \kappa\nabla\phi) \in L^2(\Omega)$ and thus $(\boldsymbol{a}\phi - \kappa\nabla\phi) \in H(\text{div}; \Omega)$. The regularity of this space

implies the following continuity condition:

$$[\![\boldsymbol{a}\phi - \kappa\nabla\phi]\!] = 0 \qquad \text{on } \Gamma_0 \,, \tag{9.7}$$

which is a natural condition. The essential condition follows from the regularity of ϕ itself. Ultimately, the level of continuity of the solution depends on the data, but at the very least ϕ will be a member of $H^1(\Omega)$. Functions in this space satisfy:

$$[\![\phi]\!] = \boldsymbol{0} \qquad \text{on } \Gamma_0 \,. \tag{9.8}$$

By incorporating these transmission conditions, the domain decomposed equivalent statement of Eq. (9.1) reads:

$$\boldsymbol{a} \cdot \nabla\phi - \nabla \cdot \kappa\nabla\phi = f \qquad \text{in} \quad \tilde{\Omega} \,, \tag{9.9a}$$

$$[\![\boldsymbol{a}\phi - \kappa\nabla\phi]\!] = 0 \qquad \text{on} \quad \Gamma_0 \,, \tag{9.9b}$$

$$[\![\phi]\!] = \boldsymbol{0} \qquad \text{on} \quad \Gamma_0 \,, \tag{9.9c}$$

$$\phi = \phi_D \qquad \text{on} \quad \partial\Omega_D \,, \tag{9.9d}$$

$$\kappa\, \partial_n\phi = g_N \qquad \text{on} \quad \partial\Omega_N^+ \,, \tag{9.9e}$$

$$\kappa\, \partial_n\phi - \boldsymbol{a} \cdot \boldsymbol{n}\phi = g_N \qquad \text{on} \quad \partial\Omega_N^- \,, \tag{9.9f}$$

where the PDE is thus imposed on the individual subdomains. Looking ahead, this will allow us to use discontinuous approximation spaces. In order to arrive at that point from a variational multiscale perspective, we must first rephrase Eq. (9.9) as a well-posed weak formulation. We do so by incorporating Eqs. (9.9b), (9.9e) and (9.9f) naturally, while enforcing the essential conditions of Eqs. (9.9c) and (9.9d) with the aid of Lagrange multipliers [11, 93]:

Find $\phi \in \mathcal{W}$ and $\lambda \in \mathcal{Q}$ s.t. $\forall\, w \in \mathcal{W}$ and $q \in \mathcal{Q}$:

$$-\left(\boldsymbol{a} \cdot \nabla w, \phi\right)_{\tilde{\Omega}} + \left\langle \boldsymbol{a} \cdot \boldsymbol{n}\, w, \phi\right\rangle_{\partial\Omega^+} + \left(\nabla w, \kappa\nabla\phi\right)_{\tilde{\Omega}} + \left\langle \lambda, [\![w]\!] \cdot \tilde{\boldsymbol{n}}\right\rangle_{\Gamma_D}$$
$$= (w, f)_{\tilde{\Omega}} + \left\langle w, g_N\right\rangle_{\partial\Omega_N} - \left\langle \boldsymbol{a} \cdot \boldsymbol{n}\, w, \phi_D\right\rangle_{\partial\Omega_D^-} \,, \tag{9.10a}$$

$$\left\langle q, [\![\phi]\!] \cdot \tilde{\boldsymbol{n}}\right\rangle_{\Gamma_D} = \left\langle q, \phi_D\right\rangle_{\partial\Omega_D} \,, \tag{9.10b}$$

where $\tilde{\boldsymbol{n}}$ is the outward facing normal on $\partial\Omega$, and arbitrarily refers to $\boldsymbol{n}^+$ or $\boldsymbol{n}^-$ on interior facets. Its choice will only impact the sign of the resulting solution for the Lagrange

multiplier. In its current form, the suitable function(al) spaces may be identified as:

$$\mathcal{W} = H^1(\tilde{\Omega}) := \left\{ w \in L^2(\Omega) : w\big|_K \in H^1(K) \ \forall \, K \in \mathcal{T} \right\}, \tag{9.11a}$$

$$\mathcal{Q} = H^{-1/2}(\Gamma_D). \tag{9.11b}$$

Well-posedness of this formulation and its equivalence to the strong form is proven in the following lemma.

Lemma 9.1. *The weak formulation of Equation* (9.10) *is well-posed, it is equivalent to the strong form of Eq.* (9.1), *and this equivalence dictates that* $\lambda = -\kappa \nabla \phi \cdot \tilde{\boldsymbol{n}}$ *on* $\partial \Omega_D$ *and* $\lambda = \{\!\{ \boldsymbol{a} \, \phi - \kappa \nabla \phi \}\!\} \cdot \tilde{\boldsymbol{n}}$ *on* Γ_0.

Proof. Lemma 5.1 contains a proof for the case of weakly enforced (non-homogeneous) essential conditions and non-broken spaces. In this proof we therefore focus on the broken spaces and the weak enforcement of the interface condition while assuming homogeneous essential boundary data. For ease of notation, we write Eq. (9.10) as:

Find $\phi, \lambda \in \mathcal{W} \times \mathcal{Q}$ s.t. $\forall \, w, q \in \mathcal{W} \times \mathcal{Q}$:

$$a(w, \phi) + b(w, \lambda) = L(w) - a(w, \phi_D), \tag{9.13}$$

$$b(\phi, q) = 0.$$

Theorem 2.4, the LBB theorem, states that Eq. (9.13) is a well-posed weak formulation if $a(\cdot, \cdot)$ is coercive on the kernel of $b(\cdot, \cdot)$ and if $b(\cdot, \cdot)$ is inf-sup stable. The inf-sup stability of $b(\cdot, \cdot)$ follows from choosing ϕ such that $[\![\phi]\!] \cdot \tilde{\boldsymbol{n}} = q$ on Γ_0 and $\phi = q$ on $\partial \Omega_D$. The kernel of $b(\cdot, \cdot)$ consists of all the functions ϕ for which $[\![\phi]\!] \cdot \tilde{\boldsymbol{n}} = 0$ and $\phi = 0$ on Γ_0 and $\partial \Omega_D$ respectively. The jump condition removes the added flexibility in $\mathcal{W}$ over $H^1(\Omega)$. Combined with the boundary condition this means that we require $a(\cdot, \cdot)$ to be coercive on $H_0^1(\Omega, \partial \Omega_D)$. This is indeed the case (see the proof of Lemma 5.1).

If we assume sufficient regularity of the solution, then we may perform integration by

parts on each subdomain to obtain:

$$\left(w, \boldsymbol{a}\nabla\phi - \nabla \cdot \kappa\nabla\phi - f\right)_{\tilde{\Omega}} - \left\langle w, g_N + \boldsymbol{a} \cdot \boldsymbol{n}\,\phi - \kappa\partial_n\phi\right\rangle_{\partial\Omega_N^-} - \left\langle w, g_N - \kappa\partial_n\phi\right\rangle_{\partial\Omega_N^+}$$
$$+ \left\langle [\![w]\!], \lambda\tilde{\boldsymbol{n}} + \{\!\{\kappa\nabla\phi\}\!\} - \{\!\{\boldsymbol{a}\phi\}\!\}\right\rangle_{\Gamma_0} + \left\langle\{\!\{w\}\!\}, [\![\kappa\nabla\phi]\!] - [\![\boldsymbol{a}\phi]\!]\right\rangle_{\Gamma_0} \tag{9.14}$$
$$+ \left\langle w, \lambda + \boldsymbol{a} \cdot \boldsymbol{n}\,(\phi_D - \phi) + \kappa\partial_n\phi\right\rangle_{\partial\Omega_D^-} + \left\langle w, \lambda + \kappa\partial_n\phi\right\rangle_{\partial\Omega_D^+} = 0\,.$$

As this holds for all $w \in \mathcal{W}$, equivalence with the strong form follows from equating each term to zero individually. From Eq. (9.10b) it also follows that $\phi = \phi_D$ on $\partial\Omega_D$, such that the last two terms in Eq. (9.14) yields the expression for λ on $\partial\Omega_D$. $\qquad\square$

The space $\mathcal{W}$ defined in Eq. (9.11a) requires no continuity across the interface of subdomains. A coarse-scale subspace of $\mathcal{W}$ is thus also permitted to be discontinuous from subdomain to subdomain (read: element to element). We are then finally in the position to propose a variational multiscale decomposition relating to discontinuous Galerkin methods. Consider the following decomposition:

$$\mathcal{W} = \mathcal{W}^h \oplus \mathcal{W}', \tag{9.15}$$

where the coarse-scale space $\mathcal{W}^h$ is some finite-dimensional subspace of $\mathcal{W}$, and the fine-scale space $\mathcal{W}'$ is an infinite-dimensional complement. Naturally, we identify with $\mathcal{W}^h$ some discontinuous Galerkin approximation space. The subdomains K_i may be interpreted as the elements of a finite element mesh $\mathcal{T}$. We then construct $\mathcal{W}^h$ as the span of piecewise polynomial functions of order p:

$$\mathcal{W}^h = \left\{w \in L^2(\Omega) : w|_K \in \mathbb{P}^p(K)\ \forall K \in \mathcal{T}\right\}. \tag{9.16}$$

Based on the direct sum decomposition of Eq. (9.15), any trial or test function ϕ or w may be uniquely written as some discontinuous Galerkin solution and some fine-scale supplement:

$$\phi = \phi^h + \phi' \qquad \phi^h \in \mathcal{W}^h,\ \phi' \in \mathcal{W}', \tag{9.17a}$$
$$w = w^h + w' \qquad w^h \in \mathcal{W}^h, w' \in \mathcal{W}'. \tag{9.17b}$$

We consider the case where the direct sum decomposition in Eq. (5.5) is induced by a

projection:

$$\mathscr{P} : \mathcal{W} \to \mathcal{W}^h \tag{9.18a}$$

$$w \mapsto \mathscr{P}w =: w^h \, , \tag{9.18b}$$

such that the coarse-scale and fine-scale spaces from Eq. (9.15) may be defined as the range and the kernel of the projector:

$$\mathcal{W}^h = \mathrm{ran}(\mathscr{P}) \, , \tag{9.19a}$$

$$\mathcal{W}' = \mathrm{ker}(\mathscr{P}) \, . \tag{9.19b}$$

By using this decomposition of $\mathcal{W}$ in the variational formulation of Eq. (9.10), the weak statement may equivalently be posed as:

Find $\phi^h, \phi', \lambda \in \mathcal{W}^h \times \mathcal{W}' \times \mathcal{Q}$ s.t. $\forall \, w^h, w', q \in \mathcal{W}^h \times \mathcal{W}' \times \mathcal{Q}$:

$$
\begin{aligned}
- \left(\boldsymbol{a} \cdot \nabla w^h, \phi^h + \phi' \right)_{\tilde{\Omega}} + \left(\nabla w^h, \kappa \nabla \phi^h + \kappa \nabla \phi' \right)_{\tilde{\Omega}} + \left\langle \boldsymbol{a} \cdot \boldsymbol{n} \, w^h, \phi^h + \phi' \right\rangle_{\partial \Omega^+} & \\
+ \left\langle \lambda, [\![w^h]\!] \cdot \tilde{\boldsymbol{n}} \right\rangle_{\Gamma_D} = \left(w^h, f \right)_{\tilde{\Omega}} + \left\langle w^h, g_N \right\rangle_{\partial \Omega_N} - \left\langle \boldsymbol{a} \cdot \boldsymbol{n} \, w^h, \phi_D \right\rangle_{\partial \Omega_D^-} \, , &
\end{aligned} \tag{9.20a}
$$

$$
\begin{aligned}
- \left(\boldsymbol{a} \cdot \nabla w', \phi^h + \phi' \right)_{\tilde{\Omega}} + \left(\nabla w', \kappa \nabla \phi^h + \kappa \nabla \phi' \right)_{\tilde{\Omega}} + \left\langle \boldsymbol{a} \cdot \boldsymbol{n} \, w', \phi^h + \phi' \right\rangle_{\partial \Omega^+} & \\
+ \left\langle \lambda, [\![w']\!] \cdot \tilde{\boldsymbol{n}} \right\rangle_{\Gamma_D} = \left(w', f \right)_{\tilde{\Omega}} + \left\langle w', g_N \right\rangle_{\partial \Omega_N} - \left\langle \boldsymbol{a} \cdot \boldsymbol{n} \, w', \phi_D \right\rangle_{\partial \Omega_D^-} \, , &
\end{aligned} \tag{9.20b}
$$

$$\left\langle q, [\![\phi^h]\!] \cdot \tilde{\boldsymbol{n}} + [\![\phi']\!] \cdot \tilde{\boldsymbol{n}} \right\rangle_{\Gamma_D} = \left\langle q, \phi_D \right\rangle_{\partial \Omega_D} \, . \tag{9.20c}$$

Equation (9.20a) is the "coarse-scale problem". It involves a finite number of test functions and a finite number of trial functions, both in the discontinuous Galerkin approximation space of Eq. (9.16). This means that for a given λ and ϕ', the coarse-scale problem forms a discontinuous Galerkin finite element method. The dependence of the Lagrange multiplier on ϕ^h and ϕ' is known (see Lemma 9.1). What remains is to substitute a suitable closure is place of the fine-scale solution ϕ'.

9.2 The Interior Penalty method as a partial fine-scale closure

The possible fine-scale solutions ϕ' are quite constrained. They must lie in $\mathcal{W}' = \mathrm{ker}(\mathscr{P})$ and are defined according to the projector from Eq. (9.18) as $\phi' = (\mathscr{I} - \mathscr{P})\phi$. Clearly, the fine scales are intimately related to the choice of projector. Knowledge about the

projector should guide the development of a closure model. For $\mathcal{C}^0$-continuous finite element approximation spaces, the H_0^1-projector has been shown to yield favorable properties, both in terms of solution quality and in terms of fine-scale closure model localization (see Chapter 3). When the coarse-scale basis functions are discontinuous from element to element, the H_0^1-projector is no longer well-posed. Recall its definition:

$$\mathscr{P}_{H_0^1} : \mathcal{W} \to \mathcal{W}^h$$
$$\phi \; \mapsto \; \underset{\phi^h \in \mathcal{W}^h}{\arg\inf} \int_{\tilde{\Omega}} \frac{1}{2}\kappa(\nabla\phi - \nabla\phi^h) \cdot (\nabla\phi - \nabla\phi^h) \,. \tag{9.21}$$

As the potential function only involves gradients, the solution would only be defined up to a per-element constant.

In Part II of this dissertation, we saw that Nitsche's method emerges from the VMS framework when a boundary weighting is added to this potential at boundaries where the essential conditions are enforced weakly. In a similar sense, we can retrieve the Interior Penalty method from Section 8.2 by choosing a projector that adds a specific weighting to the element interfaces. Just like for Nitsche's method, this illustrates that the penalty terms involved in these finite element formulations are not purely numerical: they follow from the key question "what is the optimal coarse-scale representation of an exact solution?"

The following projection operator is a natural extension of Eq. (9.21) for functions that are not $\mathcal{C}^0$-continuous on the domain. We call this projector the Interior Penalty projector.

$$\mathscr{P}_{IP} : \mathcal{W} \to \mathcal{W}^h$$
$$\phi \; \mapsto \; \underset{\phi^h \in \mathcal{W}^h}{\arg\inf} \int_{\tilde{\Omega}} \frac{1}{2}\kappa(\nabla\phi - \nabla\phi^h) \cdot (\nabla\phi - \nabla\phi^h) - \int_{\Gamma_D} \{\!\!\{ \kappa\nabla\phi - \kappa\nabla\phi^h \}\!\!\} \cdot [\![\phi - \phi^h]\!]$$
$$+ \int_{\Gamma} \frac{1}{2}\kappa\eta^*[\![\phi - \phi^h]\!] \cdot [\![\phi - \phi^h]\!] \,, \tag{9.22}$$

with:

$$\eta^* = \begin{cases} 0 & \text{on } \partial\Omega_N^- \,, \\ \frac{\boldsymbol{a}\cdot\boldsymbol{n}}{\kappa} & \text{on } \partial\Omega_N^+ \,, \\ \eta & \text{on } \partial\Omega_D^- \,, \\ \eta + \frac{\boldsymbol{a}\cdot\boldsymbol{n}}{\kappa} & \text{on } \partial\Omega_D^+ \,, \\ \eta + \frac{1}{2}\frac{|\boldsymbol{a}\cdot\boldsymbol{n}|}{\kappa} & \text{on } \Gamma_0 \,. \end{cases} \tag{9.23}$$

As addressed in Definition 2.9, we are not permitted to access $\partial_n \phi$ on Γ if $\phi \in H^1(\tilde{\Omega})$. For that reason, we redefine $\mathcal{W}$ as:

$$\mathcal{W} = \{w \in H^1(\tilde{\Omega}) : \Delta w \in L^2(\tilde{\Omega})\}, \tag{9.24}$$

of which the solution to Eq. (9.1) is a member.

Lemma 5.2 also applies to the Interior Penalty projector: Eq. (9.22) is a projection operator when the associated variational formulation is coercive. This, in turn, is a requirement on η. The minimal η scales inversely proportional to h and is defined by a per-element inverse estimate, in a similar sense as the parameters from Lemma 6.3 [169].

The fine-scale space $\mathcal{W}'$ consists of all the functions in $\mathcal{W}$ who's projection yields the zero function. Under the assumption that the variational formulation of the projection operator is coercive, this means that the integral terms that form the right-hand side data must be zero. By taking the Gâteaux derivative we obtain the following condition for the fine scales:

$$\phi' \in \mathcal{W}' \text{ if and only if } \forall\, v^h \in \mathcal{W}^h:$$
$$- \left(\nabla v^h, \kappa \nabla \phi' \right)_{\tilde{\Omega}} + \left\langle [\![v^h]\!], \{\!\{ \kappa \nabla \phi' \}\!\} \right\rangle_{\Gamma_D} + \left\langle \{\!\{ \kappa \nabla v^h \}\!\}, [\![\phi']\!] \right\rangle_{\Gamma_D} - \left\langle [\![v^h]\!], \eta^* [\![\phi']\!] \right\rangle_{\Gamma} = 0 \,. \tag{9.25}$$

This knowledge about the definition of the fine scales may directly be used in the coarse-scale weak formulation. When we choose $v^h = w^h$ in Eq. (9.25) and add the result to the coarse-scale problem of Eq. (9.20a), then we obtain:

$$\text{Find } \phi^h \in \mathcal{W}^h \text{ s.t. } \forall\, w^h \in \mathcal{W}^h:$$
$$- \left(\boldsymbol{a} \cdot \nabla w^h, \phi^h + \phi' \right)_{\tilde{\Omega}} + \left\langle \lambda, [\![w^h]\!] \cdot \tilde{\boldsymbol{n}} \right\rangle_{\Gamma_D} - \left\langle \tfrac{1}{2} | \boldsymbol{a} \cdot \boldsymbol{n} | [\![w^h]\!], [\![\phi']\!] \right\rangle_{\Gamma_0} + \left\langle \boldsymbol{a} \cdot \boldsymbol{n}\, w^h, \phi^h \right\rangle_{\partial\Omega^+}$$
$$+ \left(\nabla w^h, \kappa \nabla \phi^h \right)_{\tilde{\Omega}} + \left\langle [\![w^h]\!], \{\!\{ \kappa \nabla \phi' \}\!\} \right\rangle_{\Gamma_D} + \left\langle \{\!\{ \kappa \nabla w^h \}\!\}, [\![\phi']\!] \right\rangle_{\Gamma_D} - \left\langle \kappa \eta [\![w^h]\!], [\![\phi']\!] \right\rangle_{\Gamma_D}$$
$$= \left(w^h, f \right)_{\tilde{\Omega}} + \left\langle w^h, g_N \right\rangle_{\partial\Omega_N} - \left\langle \boldsymbol{a} \cdot \boldsymbol{n}\, w^h, \phi_D \right\rangle_{\partial\Omega_D^-} \,. \tag{9.26}$$

We proceed by using a priori knowledge of the a posteriori solutions ϕ^h, ϕ' and λ. From Lemma 9.1 we know that $\lambda = -\kappa \partial_n \phi$ on $\partial\Omega_D$ and $\lambda = \{\!\{ \boldsymbol{a}\, \phi - \kappa \nabla \phi \}\!\} \cdot \tilde{\boldsymbol{n}}$ on Γ_0. Additionally, the enforced transmission condition Eq. (9.9c) will mean that $[\![\phi']\!] = -[\![\phi^h]\!]$, and, finally, on the Dirichlet boundary the fine-scale solution satisfies $\phi' = \phi_D - \phi^h$. Use of those identities results in a discontinuous Galerkin formulation of the advection-diffusion problem with Interior Penalty treatment of the diffusive flux (from Section 8.1) and upwind

treatment of the advective flux (from Section 8.2):

Find $\phi^h \in \mathcal{W}^h$ s.t. $\forall\, w^h \in \mathcal{W}^h$:

$$- (\boldsymbol{a} \cdot \nabla w^h, \phi^h + \phi')_{\tilde{\Omega}} + \langle [\![w^h]\!], \{\!\!\{\boldsymbol{a}\phi'\}\!\!\}\rangle_{\Gamma_0} + \langle [\![w^h]\!], (\boldsymbol{a}\phi^h)^+\rangle_{\Gamma_0} + \langle \boldsymbol{a} \cdot \boldsymbol{n}\, w^h, \phi^h\rangle_{\partial\Omega^+}$$

$$+ (\nabla w^h, \kappa\nabla\phi^h)_{\tilde{\Omega}} - \langle [\![w^h]\!], \{\!\!\{\kappa\nabla\phi^h\}\!\!\}\rangle_{\Gamma_D} - \langle \{\!\!\{\kappa\nabla w^h\}\!\!\}, [\![\phi^h]\!]\rangle_{\Gamma_D} + \langle \eta[\![w^h]\!], [\![\phi^h]\!]\rangle_{\Gamma_D}$$

$$= (w^h, f)_{\tilde{\Omega}} - \langle \boldsymbol{a} \cdot \boldsymbol{n}\, w^h, \phi_D\rangle_{\partial\Omega_D^-} + \langle w^h, g_N\rangle_{\partial\Omega_N}$$

$$- \langle \kappa\partial_n w^h, \phi_D\rangle_{\partial\Omega_D} + \langle w^h, \kappa\eta\phi_D\rangle_{\partial\Omega_D}. \tag{9.27}$$

This formulation still includes two fine-scale terms: a volumetric term and an interface term. We will develop suitable models for these terms in Chapter 11.

9.3   A short study on the Interior Penalty projector

To determine appropriate modeling assumptions, we require more a priori understanding of the fine-scale solution. The Interior Penalty optimality condition, Eq. (9.25), represents the projection operator and defines the fine-scale space $\mathcal{W}'$. It dictates a finite number of functional constraints satisfied by every $w' \in \mathcal{W}'$. Recall from Section 3.2 that this set of functionals forms the basis for the image of the adjoint of $\mathscr{P}_{IP}$:

$$\mathscr{P}_{IP}\phi = \phi^h \text{ such that:}$$
$$\langle \mu_i, \phi - \phi^h\rangle = 0 \qquad i = 1, ..., \dim(\mathcal{W}^h). \tag{9.28}$$

Following the approach outlined in Section 3.2.2, the functionals μ_i can be determined from the fine-scale constraint equation, Eq. (9.25), by substituting in place of v^h the different bases that span $\mathcal{W}^h$. In the one-dimensional case we can derive a particularly insightful set of functionals after slight modification of the constraint equation. We integrate by parts on each element such that the equation may be written as:

$$(\nabla \cdot \kappa\nabla v^h, \phi')_{\tilde{\Omega}} + \langle [\![\kappa\nabla v^h]\!], \{\!\!\{\phi'\}\!\!\}\rangle_{\Gamma_N} + \langle [\![v^h]\!], \{\!\!\{\kappa\nabla\phi'\}\!\!\}\rangle_{\Gamma_D} - \langle [\![v^h]\!], \eta^*[\![\phi']\!]\rangle_{\Gamma} = 0, \tag{9.29}$$

which is required to hold for all $v^h \in \mathcal{W}^h$.

In the one-dimensional case, Γ is the set of nodes $\{x_0, x_2, .., x_n\}$ and $\tilde{\Omega}$ is the union of the set of open element domains $\mathcal{T} = \{(x_0, x_1), .., (x_{n-1}, x_n)\}$. We assume κ to be constant. If we then substitute the function of Figure 9.1a as a test function, Eq. (9.29) requires

$\phi' = 0$ on $\partial\Omega_N$. Similarly, we may substitute the piecewise linear test function illustrated in Figure 9.1b to conclude that the fine scales average vanish on all of Γ_N. Next, we substitute the piecewise continuous functions for which $[\![v^h]\!]$ is nonzero at a node in Γ_D. Examples are shown in Figures 9.1c and 9.1d. This requires $\{\!\{\kappa\nabla\phi'\}\!\} - \eta^*[\![\phi']\!] = 0$ on each node in Γ_D. Finally, by substituting the bubble functions of Figures 9.1e and 9.1f the first term in Eq. (9.29) requires that ϕ' has vanishing moments up to order $p - 2$, p being the polynomial order of the approximation space.

These statements combine into the set of μ_i functionals that represent the Interior Penalty projector:

$\mathscr{P}_{IP}\phi = \phi^h$ such that:

$$\langle\mu, \phi - \phi^h\rangle = \{\!\{\phi - \phi^h\}\!\}\big|_{x_i} = 0 \qquad\qquad \text{for } x_i \in \Gamma_N, \qquad (9.30\text{a})$$

$$\langle\mu, \phi - \phi^h\rangle = [\{\!\{\kappa\nabla(\phi - \phi^h)\}\!\} - (\eta + \tfrac{1}{2}|\boldsymbol{a}\cdot\boldsymbol{n}|)[\![\phi - \phi^h]\!]]\big|_{x_i} = 0 \quad \text{for } x_i \in \Gamma, \quad (9.30\text{b})$$

$$\langle\mu, \phi - \phi^h\rangle = \kappa\partial_n(\phi - \phi^h)\big|_{x_i} - \kappa\eta(\phi - \phi^h)\big|_{x_i} = 0 \qquad\qquad \text{for } x_i \in \partial\Omega_D^-, \qquad (9.30\text{c})$$

$$\langle\mu, \phi - \phi^h\rangle = \kappa\partial_n(\phi - \phi^h)\big|_{x_i} - (\kappa\eta + \boldsymbol{a}\cdot\boldsymbol{n})(\phi - \phi^h)\big|_{x_i} = 0 \quad \text{for } x_i \in \partial\Omega_D^+, \qquad (9.30\text{d})$$

$$\langle\mu, \phi - \phi^h\rangle = \int_K(\phi - \phi^h)x^q = 0 \qquad\qquad \text{for } K \in \mathcal{T} \text{ and } 0 \le q \le p-2. \qquad (9.30\text{e})$$

The first statement is especially important: it says that the average of the fine scales is zero on interior nodes. This will be used for modeling the remaining fine scales in the coarse-scale formulation. The functionals of Equations (9.30c) to (9.30e) were also obtained for the Nitsche projector (Eq. (5.19)), and those of Eq. (9.30e) also correspond to the H_0^1-projector.

(a) Vanishing averages, example 1.

(b) Vanishing averages, example 2.

(c) Balancing jumps and gradients, example 1.

(d) Balancing jumps and gradients, example 2.

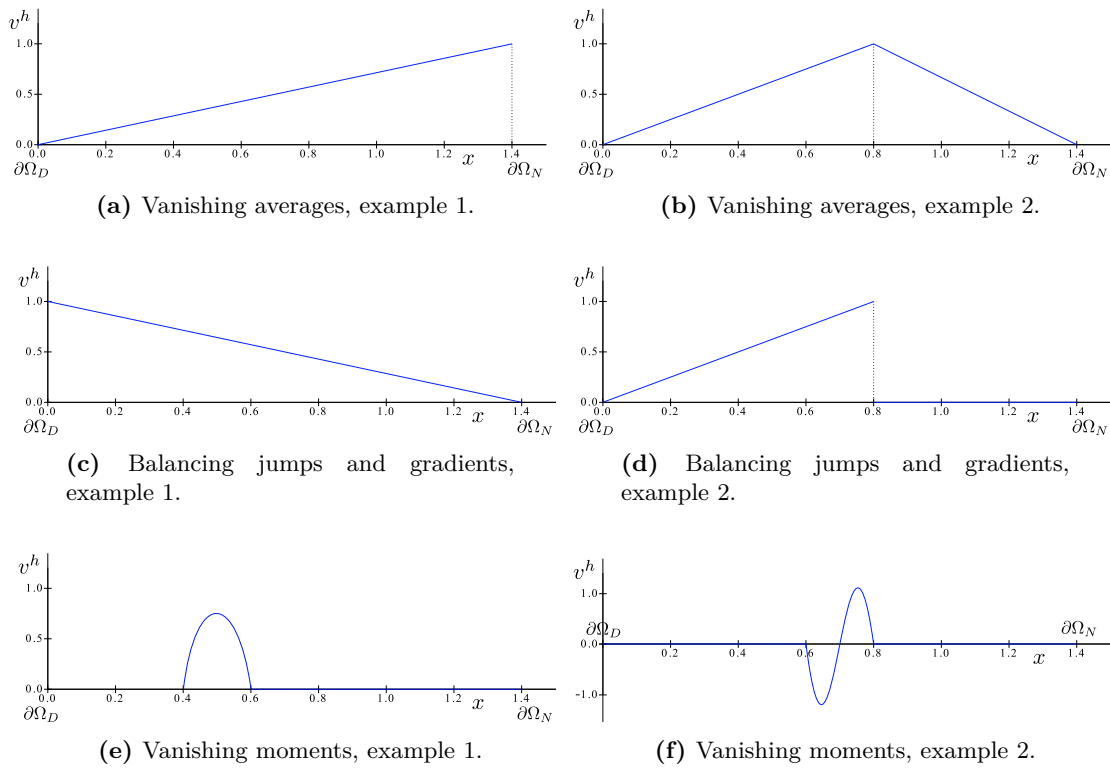**(e)** Vanishing moments, example 1.

(f) Vanishing moments, example 2.

Fig. 9.1: *Example test functions to arrive at the functional constraints imposed by the Interior Penalty projector, collected in Eq. (9.30).*

The variational multiscale decomposition for mixed and broken spaces

10.1 The mixed multiscale weak formulation with broken spaces

The advection-diffusion equation may be written in a reduced order form in one of two ways. Either we choose the auxiliary variable to be the diffusive flux $-\kappa\nabla\phi$ or we include the advective term in the flux-definition as $\boldsymbol{a}\phi - \kappa\nabla\phi$. We choose to proceed with the former, which reflects a separation of hyperbolic and elliptic terms. In Chapter 8, we showed that the same separation is employed for devising stable discontinuous Galerkin methods. The mixed form of the scalar conservation law reads:

$$-\nabla\phi = \kappa^{-1}\boldsymbol{\sigma} \qquad \text{in} \quad \Omega\,, \tag{10.1a}$$

$$\boldsymbol{a}\cdot\nabla\phi + \nabla\cdot\boldsymbol{\sigma} = f \qquad \text{in} \quad \Omega\,, \tag{10.1b}$$

$$\phi = \phi_D \qquad \text{on} \quad \partial\Omega_D\,, \tag{10.1c}$$

$$-\boldsymbol{\sigma}\cdot\boldsymbol{n} = g_N \qquad \text{on} \quad \partial\Omega_N^+\,, \tag{10.1d}$$

$$-\boldsymbol{\sigma}\cdot\boldsymbol{n} - \boldsymbol{a}\cdot\boldsymbol{n}\phi = g_N \qquad \text{on} \quad \partial\Omega_N^-\,, \tag{10.1e}$$

where we again consider data of the form $f \in L^2(\Omega)$, $g_N \in H^{1/2}(\partial\Omega_N)$ and

$\phi_D \in H^{1/2}(\partial\Omega_D)$. The advection field $\boldsymbol{a} = \boldsymbol{a}(\boldsymbol{x})$ is assumed solenoidal $(\nabla \cdot \boldsymbol{a} = 0)$ and the diffusion coefficient κ is assumed to be a positive constant.

We denote by $\mathcal{T} = \{\Omega_i\}_{i=1}^n$ some tessellation of the domain Ω. Recall from the previous two chapters the following domain definitions based on this tessellation:

$$\tilde{\Omega} := \bigcup_{i=1}^n K_i \,, \tag{10.2a}$$

$$\Gamma := \bigcup_{i=1}^n \partial K_i \,, \tag{10.2b}$$

$$\Gamma_0 := \Gamma \setminus \partial\Omega \,, \tag{10.2c}$$

$$\Gamma_D := \Gamma \setminus \partial\Omega_N = \Gamma_0 \cup \partial\Omega_D \,, \tag{10.2d}$$

$$\Gamma_N := \Gamma \setminus \partial\Omega_D = \Gamma_0 \cup \partial\Omega_N \,. \tag{10.2e}$$

Our objective is to state Eqs. (10.1a) and (10.1b) on the individual subdomains while enforcing transmission conditions on Γ_0 that ensure equivalence with Eq. (10.1). These conditions are based on the required regularity of the solutions $\boldsymbol{\sigma}$ and ϕ. In the mixed form of the advection-diffusion equation these must be members of $H(\mathrm{div};\Omega)$ and $H^1(\Omega)$, respectively. Note that this is different from the mixed formulation of the Poisson problem, discussed in Chapter 4, which permitted $\phi \in L^2(\Omega)$. In the current case we remain restricted to $H^1(\Omega)$ due to the advective term. The transmission conditions follow directly from the required continuity implied by these spaces::

$$[\![\phi]\!] = \boldsymbol{0} \qquad \text{on } \Gamma_0 \,, \tag{10.3a}$$

$$[\![\boldsymbol{\sigma}]\!] = 0 \qquad \text{on } \Gamma_0 \,. \tag{10.3b}$$

Equations (10.1a) and (10.1b) are posed on each subdomain, while we add the constraints from Eq. (10.3) to couple the solution across domain interfaces:

$$-\nabla\phi = \kappa^{-1}\boldsymbol{\sigma} \qquad \text{in} \quad \tilde{\Omega} \,, \tag{10.4a}$$

$$\boldsymbol{a} \cdot \nabla\phi + \nabla \cdot \boldsymbol{\sigma} = f \qquad \text{in} \quad \tilde{\Omega} \,, \tag{10.4b}$$

$$[\![\boldsymbol{\sigma}]\!] = 0 \qquad \text{on} \quad \Gamma_0 \,, \tag{10.4c}$$

$$[\![\phi]\!] = \boldsymbol{0} \qquad \text{on} \quad \Gamma_0 \,, \tag{10.4d}$$

$$\phi = \phi_D \qquad \text{on} \quad \partial\Omega_D \,, \tag{10.4e}$$

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = g_N \qquad \text{on} \quad \partial\Omega_N^+ \,, \tag{10.4f}$$

$$-\boldsymbol{\sigma} \cdot \boldsymbol{n} - \boldsymbol{a} \cdot \boldsymbol{n}\phi = g_N \qquad \text{on} \quad \partial\Omega_N^- \,. \tag{10.4g}$$

If we use this representation of the PDE as a point of departure, then we are permitted to work with broken spaces in our weak formulation. The continuity requirements of Eqs. (10.4c) and (10.4d) may be enforced with Lagrange multipliers [11, 12], as we did in Chapter 9.

Consider the following weak form:

Find $\phi, \boldsymbol{\sigma}, \lambda_1, \lambda_2 \in \mathcal{W} \times \boldsymbol{\Sigma} \times \mathcal{Q}_D \times \mathcal{Q}_{N+}$ s.t. $\forall w, \boldsymbol{\tau}, q_1, q_2 \in \mathcal{W} \times \boldsymbol{\Sigma} \times \mathcal{Q}_D \times \mathcal{Q}_{N+}$:

$$- (\kappa^{-1}\boldsymbol{\tau}, \boldsymbol{\sigma})_{\tilde{\Omega}} + (\nabla \cdot \boldsymbol{\tau}, \phi)_{\tilde{\Omega}} - \langle \boldsymbol{\tau} \cdot \boldsymbol{n}, \phi \rangle_{\partial\Omega_N^-} + \langle \lambda_2, [\![\boldsymbol{\tau}]\!] \rangle_{\Gamma_N^+} = \langle \boldsymbol{\tau} \cdot \boldsymbol{n}, \phi_D \rangle_{\partial\Omega_D}, \quad (10.5\text{a})$$

$$- (\nabla w, \boldsymbol{a}\phi)_{\tilde{\Omega}} + \langle \boldsymbol{a} \cdot \boldsymbol{n} w, \phi \rangle_{\partial\Omega^+} + (w, \nabla \cdot \boldsymbol{\sigma})_{\tilde{\Omega}} - \langle w, \boldsymbol{\sigma} \cdot \boldsymbol{n} \rangle_{\partial\Omega_N^-}$$
$$+ \langle \lambda_1, [\![w]\!] \cdot \tilde{\boldsymbol{n}} \rangle_{\Gamma_D} = (w, f)_{\tilde{\Omega}} - \langle \boldsymbol{a} \cdot \boldsymbol{n} w, \phi_D \rangle_{\partial\Omega_D^-} - \langle w, g_N \rangle_{\partial\Omega_N^-}, \quad (10.5\text{b})$$

$$\langle q_1, [\![\phi]\!] \cdot \tilde{\boldsymbol{n}} \rangle_{\Gamma_D} = \langle q_1, \phi_D \rangle_{\partial\Omega_D}, \quad (10.5\text{c})$$

$$\langle q_2, [\![\boldsymbol{\sigma}]\!] \rangle_{\Gamma_N^+} = \langle q_2, g_N \rangle_{\partial\Omega_N^+}, \quad (10.5\text{d})$$

where, again, $\tilde{\boldsymbol{n}}$ arbitrarily refers to $\boldsymbol{n}^+$ or $\boldsymbol{n}^-$ on interior facets, and is the outward facing normal on $\partial\Omega$. The suitable function(al) spaces may be identified as:

$$\mathcal{W} = H^1(\tilde{\Omega}) := \left\{ w \in L^2(\Omega) : w|_K \in H^1(K) \ \forall K \in \mathcal{T} \right\}, \quad (10.6\text{a})$$

$$\boldsymbol{\Sigma} = H(\text{div}, \tilde{\Omega}) := \left\{ \boldsymbol{\sigma} \in [L^2(\Omega)]^d : \boldsymbol{\sigma}|_K \in H(\text{div}; K) \ \forall K \in \mathcal{T} \right\}, \quad (10.6\text{b})$$

$$\mathcal{Q}_D = H^{-1/2}(\Gamma_D), \quad (10.6\text{c})$$

$$\mathcal{Q}_{N+} = H^{-1/2}(\Gamma_N^+). \quad (10.6\text{d})$$

The proof of the well-posedness of this weak formulation is outlined in the following lemma.

Lemma 10.1. *The weak formulation of Equation* (10.5) *is well-posed, it is equivalent to the strong form of Eq.* (10.1)*, and this equivalence dictates that* $\lambda_1 = \boldsymbol{a} \cdot \tilde{\boldsymbol{n}} \{\!\{\phi\}\!\}$ *and* $\lambda_2 = -\{\!\{\phi\}\!\}$.

Proof. The start of the proof is similar to that of Lemma 9.1: the Lagrange multiplier parts constitute coercive (and thus inf-sup stable) bilinear forms, meaning that they will permit unique solutions λ_1 and λ_2 once the set $(\phi, \boldsymbol{\sigma})$ is known. Their expressions follow from reverse integration by parts while making use of the imposed continuity conditions.

This allows us to recast Eqs. (10.5a) and (10.5b) in a weak formulation where $\boldsymbol{\Sigma} = H_{g_N}(\mathrm{div}; \Omega, \partial\Omega_N^+)$ (where g_N is enforced strongly on $\partial\Omega_N^+$) and $\mathcal{W} = H_{\phi_D}^1(\Omega, \partial\Omega_D)$ (ϕ_D being enforced strongly on $\partial\Omega_D$).

As usual we proof existence and uniqueness for homogeneous strong boundary data such that we are operating in vector spaces. The alternative case follows from decomposing the solutions into homogeneous and non-homogeneous parts.

With the new spaces, the mixed formulation may be written in operator form as:

Find $\phi, \boldsymbol{\sigma} \in H_0^1(\Omega, \partial\Omega_D) \times H_0(\mathrm{div}; \Omega, \partial\Omega_N)$:

$$- \mathcal{A}\boldsymbol{\sigma} + \mathcal{B}\phi = 0 \,, \tag{10.7a}$$

$$\mathcal{B}^T \boldsymbol{\sigma} + \mathcal{C}\phi = \mathcal{F} \,. \tag{10.7b}$$

The operator $\mathcal{A}$ is bijective as its corresponding bilinear form is coercive. It thus has an inverse. A unique solution $\boldsymbol{\sigma}$ may be obtained for any solution ϕ as:

$$\boldsymbol{\sigma} = \mathcal{A}^{-1}\mathcal{B}\phi \,. \tag{10.8}$$

Substituting into Eq. (10.7b) while writing the operators in duality pairing forms gives:

Find $\phi \in H_0^1(\Omega, \partial\Omega_D)$ s.t. $\forall w \in H_0^1(\Omega, \partial\Omega_D)$:

$$\langle \mathcal{B}^T(\mathcal{A}^{-1}\mathcal{B}\phi), w \rangle + \langle \mathcal{C}\phi, w \rangle = \langle \mathcal{F}, w \rangle \,. \tag{10.9}$$

With the enforced regularity and homogeneous boundary conditions, the duality pairing $\langle \mathcal{B}^T \boldsymbol{\tau}, w \rangle$ represents the L^2-inner product $(\boldsymbol{\tau}, \nabla w)_\Omega$. The $\langle \mathcal{C}\phi, w \rangle$ pairing represents the advective bilinear form.

Equation (10.9) is thus completely analogous to the standard advection-diffusion weak formulation in $H_0^1(\Omega, \partial\Omega_D)$ with zero Neumann conditions on $\partial\Omega_N^+$. Coercivity of this weak form, and hence unique solvability of ϕ, has been proven in earlier chapters. $\qquad\square$

Next, we can introduce finite-dimensional (coarse-scale) subspaces of $\boldsymbol{\Sigma}$ and $\mathcal{W}$. These will be the finite element approximation spaces for the eventual finite element methods. The following spaces correspond to discontinuous Galerkin methods:

$$\mathcal{W}^h = \{w \in L^2(\Omega) : w|_K \in \mathbb{P}^p(K) \; \forall \, K \in \mathcal{T}\} \subset \mathcal{W}, \tag{10.10a}$$

$$\boldsymbol{\Sigma}^h = \{\boldsymbol{\sigma} \in [L^2(\Omega)]^d : \boldsymbol{\sigma}|_K \in [\mathbb{P}^q(K)]^d \; \forall \, K \in \mathcal{T}\} \subset \boldsymbol{\Sigma}, \tag{10.10b}$$

with p and q the polynomial order of the approximation spaces.

A projector must be defined in order to establish the split of the exact solutions into coarse-scale and fine-scale components. In this mixed case, the projector should map an element $(\phi, \boldsymbol{\sigma}) \in \mathcal{W} \times \boldsymbol{\Sigma}$ to the mixed coarse-scale space $\mathcal{W}^h \times \boldsymbol{\Sigma}^h$:

$$\begin{aligned} \mathscr{P} : \mathcal{W} \times \boldsymbol{\Sigma} &\to \mathcal{W}^h \times \boldsymbol{\Sigma}^h \\ \phi, \boldsymbol{\sigma} &\mapsto \mathscr{P}(\phi, \boldsymbol{\sigma}). \end{aligned} \tag{10.11}$$

The coarse-scale and fine-scale components of ϕ and $\boldsymbol{\sigma}$ are then defined as:

$$(\phi^h, \boldsymbol{\sigma}^h) := \mathscr{P}(\phi, \boldsymbol{\sigma}), \tag{10.12a}$$

$$(\phi', \boldsymbol{\sigma}') := (\mathscr{I} - \mathscr{P})(\phi, \boldsymbol{\sigma}), \tag{10.12b}$$

with $\mathscr{I}$ the identity projector $\mathscr{I}(\phi, \boldsymbol{\sigma}) = (\phi, \boldsymbol{\sigma})$.

From Eq. (10.12b) it follows that the fine scales live in the range of $\mathscr{I} - \mathscr{P}$, or the kernel of $\mathscr{P}$, which we define as the mixed fine-scale space:

$$(\phi', \boldsymbol{\sigma}') \in \ker \mathscr{P} =: (\mathcal{W} \times \boldsymbol{\Sigma})'. \tag{10.13}$$

Recall from Lemma 4.1 that $(\mathcal{W} \times \boldsymbol{\Sigma})'$ does generally *not* decouple into fine-scale spaces $\mathcal{W}'$ and $\boldsymbol{\Sigma}'$. With this consideration in mind, the scale decomposition of the weak formulation

reads:

Find $\phi^h, \boldsymbol{\sigma}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h$ and $\phi', \boldsymbol{\sigma}' \in (\mathcal{W} \times \boldsymbol{\Sigma})'$ and $\lambda_1, \lambda_2 \in \mathcal{Q}_D \times \mathcal{Q}_{N+}$

$\quad$ s.t. $\forall\, w^h, \boldsymbol{\tau}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h$ and $\forall\, w', \boldsymbol{\tau}' \in (\mathcal{W} \times \boldsymbol{\Sigma})'$ and $\forall\, q_1, q_2 \in \mathcal{Q}_D \times \mathcal{Q}_{N+}$:

$$- (\kappa^{-1}\boldsymbol{\tau}^h, \boldsymbol{\sigma}^h + \boldsymbol{\sigma}')_{\tilde{\Omega}} + (\nabla \cdot \boldsymbol{\tau}^h, \phi^h + \phi')_{\tilde{\Omega}} - \langle \boldsymbol{\tau}^h \cdot \boldsymbol{n}, \phi^h + \phi' \rangle_{\partial\Omega_N^-}$$
$$+ \langle \lambda_2, [\![\boldsymbol{\tau}^h]\!] \rangle_{\Gamma_N^+} = \langle \boldsymbol{\tau}^h \cdot \boldsymbol{n}, \phi_D \rangle_{\partial\Omega_D}, \tag{10.14a}$$

$$- (\kappa^{-1}\boldsymbol{\tau}', \boldsymbol{\sigma}^h + \boldsymbol{\sigma}')_{\tilde{\Omega}} + (\nabla \cdot \boldsymbol{\tau}', \phi^h + \phi')_{\tilde{\Omega}} - \langle \boldsymbol{\tau}' \cdot \boldsymbol{n}, \phi^h + \phi' \rangle_{\partial\Omega_N^-}$$
$$+ \langle \lambda_2, [\![\boldsymbol{\tau}']\!] \rangle_{\Gamma_N^+} = \langle \boldsymbol{\tau}' \cdot \boldsymbol{n}, \phi_D \rangle_{\partial\Omega_D}, \tag{10.14b}$$

$$- (\boldsymbol{a} \cdot \nabla w^h, \phi^h + \phi')_{\tilde{\Omega}} + \langle \boldsymbol{a} \cdot \boldsymbol{n} w^h, \phi^h + \phi' \rangle_{\partial\Omega^+} + \langle \lambda_1, [\![w^h]\!] \cdot \tilde{\boldsymbol{n}} \rangle_{\Gamma_D}$$
$$+ (w^h, \nabla \cdot \boldsymbol{\sigma}^h + \nabla \cdot \boldsymbol{\sigma}')_{\tilde{\Omega}} - \langle w^h, \boldsymbol{\sigma}^h \cdot \boldsymbol{n} + \boldsymbol{\sigma}' \cdot \boldsymbol{n} \rangle_{\partial\Omega_N^-} \tag{10.14c}$$
$$= (w^h, f)_{\tilde{\Omega}} - \langle \boldsymbol{a} \cdot \boldsymbol{n} w^h, \phi_D \rangle_{\partial\Omega_D^-} - \langle w^h, g_N \rangle_{\partial\Omega_N^-},$$

$$- (\boldsymbol{a} \cdot \nabla w', \phi^h + \phi')_{\tilde{\Omega}} + \langle \boldsymbol{a} \cdot \boldsymbol{n} w', \phi^h + \phi' \rangle_{\partial\Omega^+} + \langle \lambda_1, [\![w']\!] \cdot \tilde{\boldsymbol{n}} \rangle_{\Gamma_D}$$
$$+ (w', \nabla \cdot \boldsymbol{\sigma}^h + \nabla \cdot \boldsymbol{\sigma}')_{\tilde{\Omega}} - \langle w', \boldsymbol{\sigma}^h \cdot \boldsymbol{n} + \boldsymbol{\sigma}' \cdot \boldsymbol{n} \rangle_{\partial\Omega_N^-} \tag{10.14d}$$
$$= (w', f)_{\tilde{\Omega}} - \langle \boldsymbol{a} \cdot \boldsymbol{n} w', \phi_D \rangle_{\partial\Omega_D^-} - \langle w', g_N \rangle_{\partial\Omega_N^-},$$

$$\langle q_1, [\![\phi^h + \phi']\!] \cdot \tilde{\boldsymbol{n}} \rangle_{\Gamma_D} = \langle q_1, \phi_D \rangle_{\partial\Omega_D}, \tag{10.14e}$$

$$\langle q_2, [\![\boldsymbol{\sigma}^h + \boldsymbol{\sigma}']\!] \rangle_{\Gamma_N^+} = \langle q_2, g_N \rangle_{\partial\Omega_N^+}. \tag{10.14f}$$

If the Lagrange multipliers and fine scales are treated as data, then Eqs. (10.14a) and (10.14c) represent the mixed coarse-scale problem. In the same way, Eqs. (10.14b) and (10.14d) form the mixed fine-scale problem for given Lagrange multipliers and coarse scales.

10.2 Discontinuous Galerkin methods as partial fine-scale closures

Some fine-scale closure is required to obtain a stand-alone finite element method. What we choose to substitute in place of the fine-scale terms in Eqs. (10.14a) and (10.14c) directly affects the solution behavior, and should be determined based on the projector operator used to decompose $\mathcal{W}$ and $\boldsymbol{\Sigma}$. In this section, we derive a number of well known discontinuous Galerkin finite element formulations from this variational multiscale perspective. We pose a 'sensible' projection operator, determine the associated fine-scale constraints,

and out comes the discontinuous Galerkin method. We base these 'sensible' projectors off of those imposed by conformal finite element formulations of the Poisson problem, i.e, those of Section 4.2.

Bassi-Rebay's 1st method

Consider the following projection:

$$
\mathscr{P}_{BR} : \mathcal{W} \times \boldsymbol{\Sigma} \to \mathcal{W}^h \times \boldsymbol{\Sigma}^h
$$

$$
\phi, \boldsymbol{\sigma} \;\mapsto\; \underset{\phi^h \in \mathcal{W}^h}{\arg\inf} \, \underset{\boldsymbol{\sigma}^h \in \boldsymbol{\Sigma}^h}{\arg\sup} \int_{\tilde{\Omega}} \frac{1}{2} \| -\kappa\nabla\phi + \kappa\nabla\phi^h \|^2
$$

$$
- \int_{\tilde{\Omega}} \frac{1}{2} \| \boldsymbol{\sigma} - \boldsymbol{\sigma}^h + \kappa\nabla\phi - \kappa\nabla\phi^h \|^2 \tag{10.15}
$$

$$
+ \int_{\Gamma_D} \{\!\!\{ \boldsymbol{\sigma} - \boldsymbol{\sigma}^h \}\!\!\} \cdot [\![\kappa\phi - \kappa\phi^h]\!] \,,
$$

where the norms are Euclidean norms of the vectors. Note the similarity between this projector and the one implied by the mixed formulation of the Poisson problem in Eqs. (4.39) and (4.40).

Solutions $(\phi^h, \boldsymbol{\sigma}^h)$ of $\mathscr{P}(\phi, \boldsymbol{\sigma})$ are equilibrium points of the functional in Eq. (10.15). They may be determined by simultaneously equating to zero the first variation with respect to $\boldsymbol{\sigma}^h$ and the first variation with respect to ϕ^h. This results in the following two optimality conditions, which have to be satisfied for all $\boldsymbol{\nu} \in \boldsymbol{\Sigma}^h$ and all $v^h \in \mathcal{W}^h$:

$$
- (\boldsymbol{\nu}^h, -\boldsymbol{\sigma} + \boldsymbol{\sigma}^h - \kappa\nabla\phi + \kappa\nabla\phi^h)_{\tilde{\Omega}} - \langle \{\!\!\{ \boldsymbol{\nu} \}\!\!\}^h, [\![\kappa\phi - \kappa\phi^h]\!] \rangle_{\Gamma_D} = 0 \,, \tag{10.16a}
$$

$$
\big(-\kappa\nabla v^h, -\boldsymbol{\sigma} + \boldsymbol{\sigma}^h \big)_{\tilde{\Omega}} - \langle [\![\kappa v^h]\!], \{\!\!\{ \boldsymbol{\sigma} - \boldsymbol{\sigma}^h \}\!\!\} \rangle_{\Gamma_D} = 0 \,, \tag{10.16b}
$$

or, by replacing all occurrences of $\boldsymbol{\sigma} - \boldsymbol{\sigma}^h$ and $\phi - \phi^h$ by $\boldsymbol{\sigma}'$ and ϕ', respectively:

$$
(\boldsymbol{\nu}^h, \boldsymbol{\sigma}' + \kappa\nabla\phi')_{\tilde{\Omega}} - \langle \{\!\!\{ \boldsymbol{\nu}^h \}\!\!\}, [\![\kappa\phi']\!] \rangle_{\Gamma_D} = 0 \qquad \forall\, \boldsymbol{\nu}^h \in \boldsymbol{\Sigma}^h \,, \tag{10.17a}
$$

$$
\big(-\kappa\nabla v^h, -\boldsymbol{\sigma}' \big)_{\tilde{\Omega}} - \langle [\![\kappa v^h]\!], \{\!\!\{ \boldsymbol{\sigma}' \}\!\!\} \rangle_{\Gamma_D} = 0 \qquad \forall\, v^h \in \mathcal{W}^h \,, \tag{10.17b}
$$

which represents the fine-scale constraint that defines all the functions that live in $\ker(\mathscr{P}_{BR}) = (\mathcal{W} \times \boldsymbol{\Sigma})'$.

Remark 10.1. *Note that the second term of the error potential in Eq. (10.15) adds both $\boldsymbol{\sigma}$ and $\kappa\nabla\phi$. For the true solution these would cancel. So, if they were omitted*

from the potential, then Eq. (10.17) would still hold for the particular pair $(\phi, \boldsymbol{\sigma})$ that solves the PDE. However, the projector may act on any pair in $\mathcal{W} \times \Sigma$. These terms are thus required to arrive at the fine-scale constraints of Eq. (10.17), which now hold for all $\phi', \boldsymbol{\sigma}' \in (\mathcal{W} \times \Sigma)'$.

We divide both equations by κ (which we assume constant), and perform integration by parts on both the gradient terms:

$$\left(\kappa^{-1}\boldsymbol{\nu}^h, \boldsymbol{\sigma}'\right)_{\tilde{\Omega}} - \left(\nabla \cdot \boldsymbol{\nu}^h, \phi'\right)_{\tilde{\Omega}} + \left\langle [\![\boldsymbol{\nu}^h]\!], \{\!\{\phi'\}\!\} \right\rangle_{\Gamma_N} = 0 \qquad \forall \boldsymbol{\nu}^h \in \boldsymbol{\Sigma}^h\,, \tag{10.18a}$$

$$- \left(v^h, \nabla \cdot \boldsymbol{\sigma}'\right)_{\tilde{\Omega}} + \left\langle \{\!\{v^h\}\!\}, [\![\boldsymbol{\sigma}']\!] \right\rangle_{\Gamma_N} = 0 \qquad \forall v^h \in \mathcal{W}^h\,. \tag{10.18b}$$

This representation of the fine-scale constraints may directly be used as a partial closure in the coarse-scale problem. By choosing $v^h = w^h$ and $\boldsymbol{\nu}^h = \boldsymbol{\tau}^h$, and by adding the results to Eqs. (10.14a) and (10.14c), we obtain:

Find $\phi^h, \boldsymbol{\sigma}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h$ s.t. $\forall w^h, \boldsymbol{\tau}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h$:

$$\begin{aligned} &- \left(\kappa^{-1}\boldsymbol{\tau}^h, \boldsymbol{\sigma}^h\right)_{\tilde{\Omega}} + \left(\nabla \cdot \boldsymbol{\tau}^h, \phi^h\right)_{\tilde{\Omega}} + \left\langle [\![\boldsymbol{\tau}^h]\!], \{\!\{\phi'\}\!\} \right\rangle_{\Gamma_N^+} + \left\langle \lambda_2, [\![\boldsymbol{\tau}^h]\!] \right\rangle_{\Gamma_N^+} \\ &\qquad - \left\langle \boldsymbol{\tau}^h \cdot \boldsymbol{n}, \phi^h \right\rangle_{\partial\Omega_N^-} = \left\langle \boldsymbol{\tau}^h \cdot \boldsymbol{n}, \phi_D \right\rangle_{\partial\Omega_D}\,, \end{aligned} \tag{10.19a}$$

$$\begin{aligned} &- \left(\boldsymbol{a} \cdot \nabla w^h, \phi^h + \phi'\right)_{\tilde{\Omega}} + \left\langle \boldsymbol{a} \cdot \boldsymbol{n} w^h, \phi^h + \phi' \right\rangle_{\partial\Omega_N^+} + \left\langle \lambda_1, [\![w^h]\!] \cdot \tilde{\boldsymbol{n}} \right\rangle_{\Gamma_D^+} \\ &\qquad + \left(w^h, \nabla \cdot \boldsymbol{\sigma}^h\right)_{\tilde{\Omega}} + \left\langle \{\!\{w^h\}\!\}, [\![\boldsymbol{\sigma}']\!] \right\rangle_{\Gamma_N^+} - \left\langle w^h, \boldsymbol{\sigma}^h \cdot \boldsymbol{n} \right\rangle_{\partial\Omega_N^-} \\ &\qquad\qquad = \left(w^h, f\right)_{\tilde{\Omega}} - \left\langle \boldsymbol{a} \cdot \boldsymbol{n} w^h, \phi_D \right\rangle_{\partial\Omega_D^-} - \left\langle w^h, g_N \right\rangle_{\partial\Omega_N^-}\,. \end{aligned} \tag{10.19b}$$

According to Eq. (10.4c) the coarse and fine scales corresponding to the true solution will satisfy the transmission condition $[\![\boldsymbol{\sigma}']\!] = -[\![\boldsymbol{\sigma}^h]\!]$ on the internal edges. Additionally, on the Dirichlet boundary we know that $\phi' = \phi_D - \phi^h$ and on the Neumann boundary $\boldsymbol{\sigma}' \cdot \boldsymbol{n} = g_N - \boldsymbol{\sigma}^h \cdot \boldsymbol{n}$. Finally, the Lagrange multipliers are $\lambda_1 = \boldsymbol{a} \cdot \tilde{\boldsymbol{n}} \left(\{\!\{\phi^h\}\!\} + \{\!\{\phi'\}\!\}\right)$ and $\lambda_2 = -\{\!\{\phi^h\}\!\} - \{\!\{\phi'\}\!\}$ as discussed in Lemma 10.1. By using this knowledge of the solution, we arrive at the following finite element formulation:

Find $\phi^h, \boldsymbol{\sigma}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h$ s.t. $\forall w^h, \boldsymbol{\tau}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h$:

$$- \left(\kappa^{-1}\boldsymbol{\tau}^h, \boldsymbol{\sigma}^h\right)_{\tilde{\Omega}} + \left(\nabla \cdot \boldsymbol{\tau}^h, \phi^h\right)_{\tilde{\Omega}} - \left\langle [\![\boldsymbol{\tau}^h]\!], \{\!\{\phi^h\}\!\} \right\rangle_{\Gamma_N} = \left\langle \boldsymbol{\tau}^h \cdot \boldsymbol{n}, \phi_D \right\rangle_{\partial\Omega_D}\,, \tag{10.20a}$$

$$\begin{aligned} &- \left(\boldsymbol{a} \cdot \nabla w^h, \phi^h + \phi'\right)_{\tilde{\Omega}} + \left\langle [\![w^h]\!], \boldsymbol{a}\{\!\{\phi^h\}\!\} + \boldsymbol{a}\{\!\{\phi'\}\!\} \right\rangle_{\Gamma \cup \partial\Omega^+} + \left(w^h, \nabla \cdot \boldsymbol{\sigma}^h\right)_{\tilde{\Omega}} \\ &\qquad - \left\langle \{\!\{w^h\}\!\}, [\![\boldsymbol{\sigma}^h]\!] \right\rangle_{\Gamma_N} = \left(w^h, f\right)_{\tilde{\Omega}} - \left\langle \boldsymbol{a} \cdot \boldsymbol{n} w^h, \phi_D \right\rangle_{\partial\Omega_D^-} - \left\langle w^h, g_N \right\rangle_{\partial\Omega_N}\,. \end{aligned} \tag{10.20b}$$

This represents a discontinuous Galerkin formulation of the advection-diffusion equation where the diffusive term is treated in the way proposed by Bassi and Rebay in [15] and the advective term is treated with average fluxes.

In the limit of $\|\boldsymbol{a}\| \to 0$ Eq. (10.20) becomes Bassi-Rebay's 1$^{\text{st}}$ formulation for the diffusion problem. We have thus shown that by using this finite element formulation for the Poisson problem, one formally obtains the coarse-scale solution that best approximates the true solution according to the saddle point problem of Eq. (10.15). Similarly, if the remaining fine-scale term in Eq. (10.20) is handled correctly, we would also obtain this 'optimal' solution for the advection-diffusion problem.

Remark 10.2. *It is well known that this formulation is unstable. This is a result of both the Bassi-Rebay treatment of the numerical fluxes, and the treatment of the advective flux with average values. However, the mixed multiscale formulation of Eq. (10.14) was well-posed (see Lemma 10.1), and all subsequent manipulation was consistent with the true solution. The instability of the finite element formulation can be traced back to the optimality condition of Eq. (10.15). In its current form, this inf-sup optimality is ill-posed and hence so is the resulting discontinuous Galerkin method.*

Remark 10.3. *Despite the mixed formulation, only the fine scales of the primal solution show up in this formulation. These will have to be modeled. Again, they occur as weighted terms in element volumes, and on element interfaces. In the element volume, the weighting differential operator is the SUPG operator $(-\boldsymbol{a} \cdot \nabla)$.*

Bassi-Rebay's 2$^{\text{nd}}$ method

As addressed in Remark 10.2, Bassi-Rebay's 1$^{\text{st}}$ formulation is not well-posed, which effectively means that the projection operator of Eq. (10.15) is unstable. To remedy this situation, Bassi, Rebay and coworkers themselves suggested the addition of the term $\sum_{e \in \Gamma} \left(\eta \, r_e(\llbracket \phi^h \rrbracket), r_e(\llbracket w^h \rrbracket) \right)_{\Omega}$, where η is a tunable stabilization parameter [16]. In this stabilization term, $r_e(\boldsymbol{\varphi})$ is the so-called lifting operator that lifts some function $\boldsymbol{\varphi}$ defined on an edge e to a function in $\boldsymbol{\Sigma}^h$, $r_e : [L^1(e)]^d \to \boldsymbol{\Sigma}^h$, according to:

$$\langle \boldsymbol{\varphi}, \boldsymbol{\tau} \rangle_e = \left(r_e(\boldsymbol{\varphi}), \boldsymbol{\tau} \right)_{\Omega} \qquad \forall \, \boldsymbol{\tau} \in \boldsymbol{\Sigma}^h. \tag{10.21}$$

Their new formulation can be fitted to our variational multiscale framework via addition of $\sum_{e\in\Gamma}\int_\Omega \frac{1}{2}\eta\, r_e([\![\phi-\phi^h]\!])^2$ to the projector of Eq. (10.15), and by adding $-\sum_{e\in\Gamma}\left(\eta\, r_e([\![\phi']\!]), r_e([\![v^h]\!])\right)_\Omega$ to the fine-scale constraint of Eq. (10.17).

Local Discontinuous Galerkin method

Another mechanism to stabilize the optimality condition of the projection operator is the introduction of penalties on the inter-element solution jumps. In the primal form, we have shown that such an approach relates to Nitsche's method and to the Interior Penalty method. In a mixed formulation, we investigate the Local Discontinuous Galerkin method and the Hybridizable Discontinuous Galerkin method.

First, we add a weighting to the jump of $\phi - \phi^h$ and on the mixed jumps of $\phi - \phi^h$ and $\boldsymbol{\sigma} - \boldsymbol{\sigma}^h$. To remain dimensionally consistent with previous penalty terms, the first term must be multiplied by κ^2 and the second term by κ. The projection operator becomes:

$$
\begin{aligned}
\mathscr{P}_{LDG} : &\, \mathcal{W} \times \boldsymbol{\Sigma} \to \mathcal{W}^h \times \boldsymbol{\Sigma}^h \\
\phi, \boldsymbol{\sigma} \;\mapsto\; &\operatorname*{arg\,inf}_{\phi^h\in\mathcal{W}^h}\operatorname*{arg\,sup}_{\boldsymbol{\sigma}^h\in\boldsymbol{\Sigma}^h} \int_{\tilde{\Omega}} \frac{1}{2}\| -\kappa\nabla\phi + \kappa\nabla\phi^h\|^2 \\
&\qquad - \int_{\tilde{\Omega}} \frac{1}{2}\|\boldsymbol{\sigma} - \boldsymbol{\sigma}^h + \kappa\nabla\phi - \kappa\nabla\phi^h\|^2 \\
&\qquad + \int_{\Gamma_D}\left\{ \{\!\{\boldsymbol{\sigma}-\boldsymbol{\sigma}^h\}\!\}\cdot[\![\kappa\phi - \kappa\phi^h]\!] + \frac{1}{2}\eta^*[\![\kappa\phi - \kappa\phi^h]\!]^2\right\} \\
&\qquad - \int_{\Gamma_0}\boldsymbol{\beta}\cdot[\![\kappa\phi - \kappa\phi^h]\!][\![\boldsymbol{\sigma} - \boldsymbol{\sigma}^h]\!]\,,
\end{aligned}
\tag{10.22}
$$

where we make use of the following penalty parameter definitions:

$$
\eta^* = \begin{cases}
0 & \text{on } \partial\Omega_N^-\,, \\[4pt]
\frac{\boldsymbol{a}\cdot\boldsymbol{n}}{\kappa} & \text{on } \partial\Omega_N^+\,, \\[4pt]
\eta & \text{on } \partial\Omega_D^-\,, \\[4pt]
\eta + \frac{\boldsymbol{a}\cdot\boldsymbol{n}}{\kappa} & \text{on } \partial\Omega_D^+\,, \\[4pt]
\eta + \frac{1}{2}\frac{|\boldsymbol{a}\cdot\boldsymbol{n}|}{\kappa} & \text{on } \Gamma_0\,,
\end{cases}
\tag{10.23}
$$

and:

$$\boldsymbol{\beta} = S^+ \boldsymbol{n}^+ + S^- \boldsymbol{n}^- \,. \tag{10.24}$$

Typically, the double valued 'switch' parameter $S^\pm$ satisfies $S^+ + S^- = \frac{1}{2}$, and often also either $S^+ = 0$ or $S^- = 0$ [153].

To determine the fine-scale constraints from Eq. (10.22), we write the saddle point problem in terms of the fine-scale solutions. Of course, these are equivalent to the discretization errors. The inf-sup statement thus says something about how the error is weighted:

$$\mathscr{P}_{LDG} : \mathcal{W} \times \boldsymbol{\Sigma} \to \mathcal{W}^h \times \boldsymbol{\Sigma}^h$$

$$\phi, \boldsymbol{\sigma} \ \mapsto \ \underset{\phi^h \in \mathcal{W}^h}{\arg\inf} \ \underset{\boldsymbol{\sigma}^h \in \boldsymbol{\Sigma}^h}{\arg\sup} \int_{\tilde{\Omega}} \left\{ \frac{1}{2} \| -\kappa \nabla \phi' \|^2 - \frac{1}{2} \| \boldsymbol{\sigma}' + \kappa \nabla \phi' \|^2 \right\}$$

$$+ \int_{\Gamma_D} \left\{ \{\!\!\{\boldsymbol{\sigma}'\}\!\!\} \cdot [\![\kappa \phi']\!] + \frac{1}{2} \eta^* [\![\kappa \phi']\!]^2 \right\} - \int_{\Gamma_0} \boldsymbol{\beta} \cdot [\![\kappa \phi']\!] [\![\boldsymbol{\sigma}']\!] \,. \tag{10.25}$$

In Eq. (10.25), ϕ' and $\boldsymbol{\sigma}'$ are defined as functions of ϕ^h and $\boldsymbol{\sigma}^h$ via $\phi' = \phi - \phi^h$ and $\boldsymbol{\sigma}' = \boldsymbol{\sigma} - \boldsymbol{\sigma}^h$. In these relations, ϕ and $\boldsymbol{\sigma}$ are treated as data. We obtain the fine-scale constraints by equating the first variations of ϕ^h and $\boldsymbol{\sigma}^h$ to zero. Then, the following relations hold $\forall \boldsymbol{\nu}^h \in \boldsymbol{\Sigma}^h$ and $\forall v^h \in \mathcal{W}^h$. They define the kernel of the LDG projector:

$$(\boldsymbol{\nu}^h, \boldsymbol{\sigma}')_{\tilde{\Omega}} + (\boldsymbol{\nu}^h, \kappa \nabla \phi')_{\tilde{\Omega}} - \langle \{\!\!\{\boldsymbol{\nu}^h\}\!\!\}, [\![\kappa \phi']\!] \rangle_{\Gamma_D} + \langle [\![\boldsymbol{\nu}^h]\!], \boldsymbol{\beta} \cdot [\![\kappa \phi']\!] \rangle_{\Gamma_0} = 0 \,, \tag{10.26a}$$

$$(\kappa \nabla v^h, \boldsymbol{\sigma}')_{\tilde{\Omega}} - \langle [\![\kappa v^h]\!], \{\!\!\{\boldsymbol{\sigma}'\}\!\!\} \rangle_{\Gamma_D} + \langle \boldsymbol{\beta} \cdot [\![\kappa v^h]\!], [\![\boldsymbol{\sigma}']\!] \rangle_{\Gamma_0} - \langle \eta^* [\![\kappa v^h]\!], [\![\kappa \phi']\!] \rangle_{\Gamma_D} = 0 \,. \tag{10.26b}$$

We divide both equations by κ and integrate the gradient terms by parts such that for all $(v^h, \boldsymbol{\nu}^h) \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h$:

$$(\kappa^{-1} \boldsymbol{\nu}^h, \boldsymbol{\sigma}')_{\tilde{\Omega}} - (\nabla \cdot \boldsymbol{\nu}^h, \phi')_{\tilde{\Omega}} + \langle [\![\boldsymbol{\nu}^h]\!], \{\!\!\{\phi'\}\!\!\} \rangle_{\Gamma_N} + \langle [\![\boldsymbol{\nu}^h]\!], \boldsymbol{\beta} \cdot [\![\phi']\!] \rangle_{\Gamma_0} = 0 \,, \tag{10.27a}$$

$$- (v^h, \nabla \cdot \boldsymbol{\sigma}')_{\tilde{\Omega}} + \langle \{\!\!\{v^h\}\!\!\}, [\![\boldsymbol{\sigma}']\!] \rangle_{\Gamma_N} + \langle \boldsymbol{\beta} \cdot [\![v^h]\!], [\![\boldsymbol{\sigma}']\!] \rangle_{\Gamma_0} - \langle \kappa \eta^* [\![v^h]\!], [\![\phi']\!] \rangle_{\Gamma_D} = 0 \,. \tag{10.27b}$$

Just like those in Eq. (10.26), the relations from Eq. (10.27) hold for all pairs $(\phi', \boldsymbol{\sigma}') \in \ker(\mathscr{P}_{LDG})$. After choosing $v^h = w^h$ and $\boldsymbol{\nu}^h = \boldsymbol{\tau}^h$ we may add the expressions to the

multiscale coarse-scale formulations of Eqs. (10.14a) and (10.14c). We obtain:

Find $\phi^h, \boldsymbol{\sigma}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h$ s.t. $\forall\, w^h, \boldsymbol{\tau}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h$:

$$- (\kappa^{-1}\boldsymbol{\tau}^h, \boldsymbol{\sigma}^h)_{\tilde{\Omega}} + (\nabla \cdot \boldsymbol{\tau}^h, \phi^h)_{\tilde{\Omega}} + \langle [\![\boldsymbol{\tau}^h]\!], \{\!\{\phi'\}\!\}\rangle_{\Gamma_N^+} + \langle \lambda_2, [\![\boldsymbol{\tau}^h]\!]\rangle_{\Gamma_N^+}$$
$$- \langle \boldsymbol{\tau}^h \cdot \boldsymbol{n}, \phi^h\rangle_{\partial\Omega_N^-} + \langle [\![\boldsymbol{\tau}^h]\!], \boldsymbol{\beta} \cdot [\![\phi']\!]\rangle_{\Gamma_0} = \langle \boldsymbol{\tau}^h \cdot \boldsymbol{n}, \phi_D\rangle_{\partial\Omega_D}, \tag{10.28a}$$

$$- (\boldsymbol{a} \cdot \nabla w^h, \phi^h + \phi')_{\tilde{\Omega}} + \langle \boldsymbol{a} \cdot \boldsymbol{n} w^h, \phi^h + \phi'\rangle_{\partial\Omega_N^+} + \langle \lambda_1, [\![w^h]\!] \cdot \tilde{\boldsymbol{n}}\rangle_{\Gamma_D^+}$$
$$+ (w^h, \nabla \cdot \boldsymbol{\sigma}^h)_{\tilde{\Omega}} + \langle \{\!\{w^h\}\!\}, [\![\boldsymbol{\sigma}']\!]\rangle_{\Gamma_N^+} - \langle w^h, \boldsymbol{\sigma}^h \cdot \boldsymbol{n}\rangle_{\partial\Omega_N^-}$$
$$+ \langle \boldsymbol{\beta} \cdot [\![w^h]\!], [\![\boldsymbol{\sigma}']\!]\rangle_{\Gamma_0} - \langle \kappa\eta^*[\![w^h]\!], [\![\phi']\!]\rangle_{\Gamma_D} \tag{10.28b}$$
$$= (w^h, f)_{\tilde{\Omega}} - \langle \boldsymbol{a} \cdot \boldsymbol{n} w^h, \phi_D\rangle_{\partial\Omega_D^-} - \langle w^h, g_N\rangle_{\partial\Omega_N^-}.$$

Many of the remaining fine scales may still be related to coarse scales: from the transmission conditions of Eqs. (10.4c) and (10.4d) we know that the solutions will satisfy $[\![\phi']\!] = -[\![\phi^h]\!]$ and $[\![\boldsymbol{\sigma}']\!] = -[\![\boldsymbol{\sigma}^h]\!]$. On the Dirichlet boundary $\phi' = \phi_D - \phi^h$ and on the Neumann boundary $\boldsymbol{\sigma}' \cdot \boldsymbol{n} = g_N - \boldsymbol{\sigma}^h \cdot \boldsymbol{n}$, and the Lagrange multipliers are $\lambda_1 = \boldsymbol{a} \cdot \tilde{\boldsymbol{n}}(\{\!\{\phi^h\}\!\} + \{\!\{\phi'\}\!\})$ and $\lambda_2 = -\{\!\{\phi^h\}\!\} - \{\!\{\phi'\}\!\}$. By making use of this known solution behavior, while also substituting the penalty parameter definition of Eq. (10.23), we arrive at the following weak formulation for the coarse-scale solution:

Find $\phi^h, \boldsymbol{\sigma}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h$ s.t. $\forall\, w^h, \boldsymbol{\tau}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h$:

$$- (\kappa^{-1}\boldsymbol{\tau}^h, \boldsymbol{\sigma}^h)_{\tilde{\Omega}} + (\nabla \cdot \boldsymbol{\tau}^h, \phi^h)_{\tilde{\Omega}} - \langle [\![\boldsymbol{\tau}^h]\!], \{\!\{\phi^h\}\!\}\rangle_{\Gamma_N} - \langle [\![\boldsymbol{\tau}^h]\!], \boldsymbol{\beta} \cdot [\![\phi^h]\!]\rangle_{\Gamma_0}$$
$$= \langle \boldsymbol{\tau}^h \cdot \boldsymbol{n}, \phi_D\rangle_{\partial\Omega_D}, \tag{10.29a}$$

$$- (\boldsymbol{a} \cdot \nabla w^h, \phi^h + \phi')_{\tilde{\Omega}} + \langle [\![w^h]\!], \boldsymbol{a}\{\!\{\phi'\}\!\}\rangle_{\Gamma_0} + \langle [\![w^h]\!], (\boldsymbol{a}\phi^h)^+\rangle_{\Gamma\cup\partial\Omega^+}$$
$$+ (w^h, \nabla \cdot \boldsymbol{\sigma}^h)_{\tilde{\Omega}} - \langle \{\!\{w^h\}\!\}, [\![\boldsymbol{\sigma}^h]\!]\rangle_{\Gamma_N} - \langle \boldsymbol{\beta} \cdot [\![w^h]\!], [\![\boldsymbol{\sigma}^h]\!]\rangle_{\Gamma_0}$$
$$+ \langle \kappa\eta\,[\![w^h]\!], [\![\phi^h]\!]\rangle_{\Gamma_D} = (w^h, f)_{\tilde{\Omega}} - \langle \boldsymbol{a} \cdot \boldsymbol{n} w^h, \phi_D\rangle_{\partial\Omega_D^-} \tag{10.29b}$$
$$- \langle w^h, g_N\rangle_{\partial\Omega_N} + \langle \kappa\eta\,w^h, \phi_D\rangle_{\partial\Omega_D}.$$

This represents a discontinuous Galerkin formulation where the diffusive term is treated with a Local Discontinuous Galerkin method as proposed by Cockburn et al. in [58], and the advective term is treated with an upwind flux.

Hybridizable Discontinuous Galerkin method

A natural extension of the projection operator from the Local Discontinuous Galerkin method is to add a weighting to the jump of $\boldsymbol{\sigma} - \boldsymbol{\sigma}^h$. Since we know that $\boldsymbol{\sigma}$ is continuous this will add a penalty to the jump of the coarse-scale solution $\boldsymbol{\sigma}^h$. For dimensional consistency, we multiply the weighting term by the element diameter h.

$$
\begin{aligned}
\mathscr{P}_{HDG} : \mathcal{W} \times \boldsymbol{\Sigma} &\rightarrow \mathcal{W}^h \times \boldsymbol{\Sigma}^h \\
\phi, \boldsymbol{\sigma} \quad &\mapsto \underset{\phi^h \in \mathcal{W}^h}{\arg\inf} \underset{\boldsymbol{\sigma}^h \in \boldsymbol{\Sigma}^h}{\arg\sup} \int_{\tilde{\Omega}} \frac{1}{2} \| - \kappa\nabla\phi + \kappa\nabla\phi^h \|^2 \\
&\quad - \int_{\tilde{\Omega}} \frac{1}{2} \| \boldsymbol{\sigma} - \boldsymbol{\sigma}^h + \kappa\nabla\phi - \kappa\nabla\phi^h \|^2 \\
&\quad + \int_{\Gamma_D} \left\{ \{\!\{ \boldsymbol{\sigma} - \boldsymbol{\sigma}^h \}\!\} \cdot [\![\kappa\phi - \kappa\phi^h]\!] + \frac{1}{2}\eta^* [\![\kappa\phi - \kappa\phi^h]\!]^2 \right\} \\
&\quad - \int_{\Gamma_0} \left\{ \boldsymbol{\beta} \cdot [\![\kappa\phi - \kappa\phi^h]\!] [\![\boldsymbol{\sigma} - \boldsymbol{\sigma}^h]\!] + \frac{1}{2}Ch[\![\boldsymbol{\sigma} - \boldsymbol{\sigma}^h]\!]^2 \right\},
\end{aligned}
\tag{10.30}
$$

where η^* is defined in Eq. (10.23).

By replacing $\phi - \phi^h$ by ϕ' and $\boldsymbol{\sigma} - \boldsymbol{\sigma}^h$ by $\boldsymbol{\sigma}'$, the projector may equivalently be written as a saddle point problem for the fine-scale solutions:

$$
\begin{aligned}
\mathscr{P}_{HDG} : \mathcal{W} \times \boldsymbol{\Sigma} &\rightarrow \mathcal{W}^h \times \boldsymbol{\Sigma}^h \\
\phi, \boldsymbol{\sigma} \quad &\mapsto \underset{\phi^h \in \mathcal{W}^h}{\arg\inf} \underset{\boldsymbol{\sigma}^h \in \boldsymbol{\Sigma}^h}{\arg\sup} \int_{\tilde{\Omega}} \left\{ \frac{1}{2} \| - \kappa\nabla\phi' \|^2 - \frac{1}{2} \| \boldsymbol{\sigma}' + \kappa\nabla\phi' \|^2 \right\} \\
&\quad + \int_{\Gamma_D} \left\{ \{\!\{ \boldsymbol{\sigma}' \}\!\} \cdot [\![\kappa\phi']\!] + \frac{1}{2}\eta^* [\![\kappa\phi']\!]^2 \right\} \\
&\quad - \int_{\Gamma_0} \left\{ \boldsymbol{\beta} \cdot [\![\kappa\phi']\!] [\![\boldsymbol{\sigma}']\!] + Ch[\![\boldsymbol{\sigma}']\!]^2 \right\}.
\end{aligned}
\tag{10.31}
$$

The fine-scale constraints are obtained by equating to zero the first variations of the coarse-scale fields:

$$
\begin{aligned}
(\boldsymbol{\nu}^h, \boldsymbol{\sigma}')_{\tilde{\Omega}} + (\boldsymbol{\nu}^h, \kappa\nabla\phi')_{\tilde{\Omega}} &- \langle \{\!\{ \boldsymbol{\nu}^h \}\!\}, [\![\kappa\phi']\!] \rangle_{\Gamma_D} + \langle [\![\boldsymbol{\nu}^h]\!], \boldsymbol{\beta} \cdot [\![\kappa\phi']\!] \rangle_{\Gamma_0} \\
&+ \langle Ch\,[\![\boldsymbol{\nu}^h]\!], [\![\boldsymbol{\sigma}']\!] \rangle_{\Gamma_0} = 0,
\end{aligned}
\tag{10.32a}
$$

$$
\begin{aligned}
(\kappa\nabla v^h, \boldsymbol{\sigma}')_{\tilde{\Omega}} &- \langle [\![\kappa v^h]\!], \{\!\{ \boldsymbol{\sigma}' \}\!\} \rangle_{\Gamma_D} + \langle \boldsymbol{\beta} \cdot [\![\kappa v^h]\!], [\![\boldsymbol{\sigma}']\!] \rangle_{\Gamma_0} \\
&- \langle \eta^* [\![\kappa v^h]\!], [\![\kappa\phi']\!] \rangle_{\Gamma_D} = 0.
\end{aligned}
\tag{10.32b}
$$

When we proceed in the same fashion as we did for Bassi-Rebay's 1st method and the Local Discontinuous Galerkin method, then we obtain the following discontinuous Galerkin formulation, which (in the case that the penalty parameter C relates to $\boldsymbol{\beta}$ and η as $4C\kappa\eta = 1 - \boldsymbol{\beta} \cdot \boldsymbol{\beta}$) corresponds to Hybridizable Discontinuous Galerkin treatment of the diffusive flux [48]:

$$
\begin{aligned}
&\text{Find } \phi^h, \boldsymbol{\sigma}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h \text{ s.t. } \forall w^h, \boldsymbol{\tau}^h \in \mathcal{W}^h \times \boldsymbol{\Sigma}^h : \\
&\quad -\left(\kappa^{-1}\boldsymbol{\tau}^h, \boldsymbol{\sigma}^h\right)_{\tilde{\Omega}} + \left(\nabla \cdot \boldsymbol{\tau}^h, \phi^h\right)_{\tilde{\Omega}} - \left\langle [\![\boldsymbol{\tau}^h]\!], \{\!\{\phi^h\}\!\}\right\rangle_{\Gamma_N} - \left\langle [\![\boldsymbol{\tau}^h]\!], \boldsymbol{\beta} \cdot [\![\phi^h]\!]\right\rangle_{\Gamma_0} \\
&\qquad - \left\langle Ch[\![\boldsymbol{\tau}^h]\!], [\![\boldsymbol{\sigma}^h]\!]\right\rangle_{\Gamma_0} = \left\langle \boldsymbol{\tau}^h \cdot \boldsymbol{n}, \phi_D\right\rangle_{\partial\Omega_D},
\end{aligned}
\tag{10.33a}
$$

$$
\begin{aligned}
&\quad -\left(\boldsymbol{a} \cdot \nabla w^h, \phi^h + \phi'\right)_{\tilde{\Omega}} + \left\langle [\![w^h]\!], \boldsymbol{a}\{\!\{\phi'\}\!\}\right\rangle_{\Gamma_0} + \left\langle [\![w^h]\!], (\boldsymbol{a}\phi^h)^+\right\rangle_{\Gamma \cup \partial\Omega^+} \\
&\qquad + \left(w^h, \nabla \cdot \boldsymbol{\sigma}^h\right)_{\tilde{\Omega}} - \left\langle \{\!\{w^h\}\!\}, [\![\boldsymbol{\sigma}^h]\!]\right\rangle_{\Gamma_N} - \left\langle \boldsymbol{\beta} \cdot [\![w^h]\!], [\![\boldsymbol{\sigma}^h]\!]\right\rangle_{\Gamma_0} \\
&\qquad + \left\langle \kappa\eta\,[\![w^h]\!], [\![\phi^h]\!]\right\rangle_{\Gamma_D} = \left(w^h, f\right)_{\tilde{\Omega}} - \left\langle \boldsymbol{a} \cdot \boldsymbol{n}w^h, \phi_D\right\rangle_{\partial\Omega_D^-} \\
&\qquad - \left\langle w^h, g_N\right\rangle_{\partial\Omega_N} + \left\langle \kappa\eta\,w^h, \phi_D\right\rangle_{\partial\Omega_D}.
\end{aligned}
\tag{10.33b}
$$

Remark 10.4. *It should be noted that all aforementioned projectors can be obtained as minimizers of error potentials due to the symmetric nature of the resulting discontinuous Galerkin methods (for the diffusive term). Such a potential minimization interpretation of the projector is not possible for nonsymmetric schemes such as the Baumann-Oden method or the Nonsymmetric Interior Penalty method. Coincidentally, these schemes are non-conservative and their stability characteristics are suboptimal, as addressed in Section 8.2.*

10.3 A short study on the Bassi-Rebay, LDG and HDG fine-scale constraints

Next, we investigate the fine-scale constraints imposed by the Bassi-Rebay, LDG, and HDG projectors of Eqs. (10.15), (10.22) and (10.30) more closely. Knowledge of the behavior of the fine scales will guide the fine-scale modeling. At the same time, since the fine-scale solution may equivalently be interpreted as the approximation error, the fine-scale constraints provide predictive information regarding the error behavior. Each of the fine-scale constraint equations (Eqs. (10.17), (10.27) and (10.32)) may be written

in the following general form:

$$(\kappa^{-1}\boldsymbol{\nu}^h, \boldsymbol{\sigma}')_{\tilde{\Omega}} - (\nabla \cdot \boldsymbol{\nu}^h, \phi')_{\tilde{\Omega}} + \langle [\![\boldsymbol{\nu}^h]\!], \mathcal{A}' \rangle_{\Gamma_0} + \langle \boldsymbol{\nu}^h \cdot \boldsymbol{n}, \mathcal{N}' \rangle_{\partial\Omega_N} = 0\,, \qquad (10.34\text{a})$$

$$(\nabla v^h, \boldsymbol{\sigma}')_{\tilde{\Omega}} - \langle [\![v^h]\!], \boldsymbol{\mathcal{B}}' \rangle_{\Gamma_0} - \langle v^h \boldsymbol{n}, \boldsymbol{\mathcal{D}}' \rangle_{\partial\Omega_D} = 0\,, \qquad (10.34\text{b})$$

where $\mathcal{A}'$ and $\boldsymbol{\mathcal{B}}'$ are functions of ϕ' and $\boldsymbol{\sigma}'$ on Γ, and $\mathcal{N}'$ and $\boldsymbol{\mathcal{D}}'$ are functions of ϕ' and $\boldsymbol{\sigma}'$ on $\partial\Omega_N$ and $\partial\Omega_D$, respectively. For the Hybridizable Discontinuous Galerkin projector the expressions are:

$$\mathcal{A}' = \{\!\{\phi'\}\!\} + \boldsymbol{\beta} \cdot [\![\phi']\!] + C[\![\boldsymbol{\sigma}']\!]\,, \qquad (10.35\text{a})$$

$$\mathcal{N}' = \phi'\,, \qquad (10.35\text{b})$$

$$\boldsymbol{\mathcal{B}}' = \{\!\{\boldsymbol{\sigma}'\}\!\} - \boldsymbol{\beta}[\![\boldsymbol{\sigma}']\!] + \kappa\,\eta^*[\![\phi']\!]\,, \qquad (10.35\text{c})$$

$$\boldsymbol{\mathcal{D}}' = \boldsymbol{\sigma}' + \kappa\,\eta^*\phi'\boldsymbol{n}\,. \qquad (10.35\text{d})$$

Those for the Local Discontinuous Galerkin method follow from setting C to zero, and those for Bassi-Rebay's 1ˢᵗ method from setting η^* and $\boldsymbol{\beta}$ to zero as well.

Error expressions can be obtained by choosing specific functions for $\boldsymbol{\nu}^h$ and v^h and substituting these into the fine-scale constraint equations. For example, substituting a test function v^h that is one inside elements K and zero outside of the element results in:

$$\int\limits_{\partial K\backslash\partial\Omega} \boldsymbol{\mathcal{B}}' \cdot \boldsymbol{n} + \int\limits_{\partial K\cap\partial\Omega_D} \boldsymbol{\mathcal{D}}' \cdot \boldsymbol{n} = 0\,. \qquad (10.36)$$

Similarly, when we choose some constant normalized vector $\boldsymbol{\nu}^h = \mathbf{1}$ inside element K and the zero vector outside of K, it follows:

$$-\int\limits_{K} \kappa^{-1}\boldsymbol{\sigma}' \cdot \mathbf{1} = \int\limits_{\partial K\backslash\partial\Omega} \mathcal{A}'\,\mathbf{1} \cdot \boldsymbol{n} + \int\limits_{\partial K\cap\partial\Omega_N} \mathcal{N}'\,\mathbf{1} \cdot \boldsymbol{n}\,. \qquad (10.37)$$

These relations predict how the error of the finite element solution behaves, under the assumption that all the remaining fine scales are accurately modeled. Equation (10.36) shows that the average quantity of some error measure on the element boundary equals zero, while Eq. (10.37) relates the total integrated error of $\boldsymbol{\sigma}^h$ in an element to error expressions on the element boundary.

We can obtain more detailed expressions by examining a simpler case. If we consider a one-dimensional problem with a Dirichlet condition on one end, then, by taking K to be

the complete domain Ω, equation Eq. (10.36) dictates:

$$\boldsymbol{\mathcal{D}}' = 0 \qquad \text{at } \partial\Omega_D \,. \tag{10.38}$$

By substituting this result back into Eq. (10.36) and by replacing K with a domain that runs from the Dirichlet boundary until any node in the internal mesh skeleton Γ_0 we obtain:

$$\boldsymbol{\mathcal{B}}'\big|_{x_i} = 0 \qquad \forall\, x_i \in \Gamma_0 \,. \tag{10.39}$$

Substitution of Eqs. (10.38) and (10.39) in Eq. (10.34b) results in the following vanishing integrals of $\boldsymbol{\sigma}'$:

$$(\boldsymbol{\sigma}', \nabla v^h)_{\tilde{\Omega}} = 0 \quad \forall\, v^h \in \mathcal{W}^h \,. \tag{10.40}$$

Under the assumption that the polynomial order of the finite element construction of $\mathcal{W}^h$ is at least $p \geq 1$, the first term in Eq. (10.37) cancels for piecewise constant $\boldsymbol{\nu}^h$. Repeating the same analysis as above, but using the test function $\boldsymbol{\nu}^h$ instead of v^h, yields:

$$\mathcal{N}' \quad = 0 \qquad \text{on } \partial\Omega_N \,, \tag{10.41a}$$

$$\mathcal{A}'\big|_{x_i} = 0 \qquad \forall\, x_i \in \Gamma_0 \,. \tag{10.41b}$$

Substituting these identities into the fine-scale constraint of Eq. (10.34a) and choosing any remaining function $\boldsymbol{\nu}^h$, we find:

$$(\boldsymbol{\sigma}', \boldsymbol{\nu}^h)_{\tilde{\Omega}} + (\phi', \nabla \cdot \boldsymbol{\nu}^h)_{\tilde{\Omega}} = 0 \quad \forall\, \boldsymbol{\nu}^h \in \boldsymbol{\Sigma}^h \,. \tag{10.42}$$

Table 10.1 summarizes all expressions where $\mathcal{A}'$, $\boldsymbol{\mathcal{B}}'$, $\boldsymbol{\mathcal{D}}'$ and $\mathcal{N}'$ have been filled in for the Bassi-Rebay, Local Discontinuous Galerkin, and Hybridizable Discontinuous Galerkin methods. These expressions represent the functional constraints imposed by the projectors if the domain is one-dimensional. They are the tabulated equivalents of Eq. (9.30) for the Interior Penalty method, and Eq. (5.19) for Nitsche's method.

Remark 10.5. *These expressions bear resemblance to the well known exactness of the numerical traces of conservative discontinuous Galerkin methods, and the ensuing super-convergence properties, see, e.g., [42] and the references therein.*

Table 10.1: *Functional constraints (i.e., fine-scale identities) imposed by the projection operators for different discontinuous Galerkin methods in the one-dimensional case.*

	Bassi-Rebay's 1st	**LDG**	**HDG**
On Γ:	$\{\phi'\} = 0$	$\{\phi'\} + \boldsymbol{\beta} \cdot [\![\phi']\!] = 0$	$\{\phi'\} + \boldsymbol{\beta} \cdot [\![\phi']\!] + C[\![\boldsymbol{\sigma'}]\!] = 0$
On Γ:	$\{\boldsymbol{\sigma'}\} = 0$	$\{\boldsymbol{\sigma'}\} - \boldsymbol{\beta}[\![\boldsymbol{\sigma'}]\!] + \kappa\eta^*[\![\phi']\!] = 0$	$\{\boldsymbol{\sigma'}\} - \boldsymbol{\beta}[\![\boldsymbol{\sigma'}]\!] + \kappa\eta^*[\![\phi']\!] = 0$
On $\partial\Omega_N$:	$\phi' = 0$	$\phi' = 0$	$\phi' = 0$
On $\partial\Omega_D$:	$\boldsymbol{\sigma'} \cdot \boldsymbol{n} = 0$	$\boldsymbol{\sigma'} \cdot \boldsymbol{n} - \frac{\kappa\eta^*}{h}\phi' = 0$	$\boldsymbol{\sigma'} \cdot \boldsymbol{n} - \frac{\kappa\eta^*}{h}\phi' = 0$
$\forall K \in \mathcal{T}$:		$\int_K \boldsymbol{\sigma'} x^{\hat{p}-1} = 0 \qquad \forall \hat{p} \in \mathbb{N} : 1 \le \hat{p} \le p$	
$\forall K \in \mathcal{T}$:		$\int_K \boldsymbol{\sigma'} x^{\hat{q}} + \hat{q}\,\phi' x^{\hat{q}-1} = 0 \qquad \forall \hat{q} \in \mathbb{N} : 1 \le \hat{q} \le q$	

We illustrate that the identities stated in Table 10.1 do in fact hold pointwise in the one-dimensional case. To avoid having to model any remaining fine scales in the discontinuous Galerkin methods of Section 10.2, we focus on the purely diffusive case (i.e., $\boldsymbol{a} = \boldsymbol{0}$). For simplicity, we choose $\kappa = 1$ and we discretize the domain $\Omega = [-1, 4]$ with only 3 elements. We employ the Dirichlet condition $\phi_D = 6.5$ and the Neumann condition $\boldsymbol{\sigma}_N = 1.8$ at the right and the left boundary, respectively, and the second order polynomial $f = -6\,x + 2.4\,x^2$ as the source function. We use linear basis functions for the construction of $\mathcal{W}^h$ and quadratics for the construction of $\boldsymbol{\Sigma}^h$.

Figure 10.1 shows the resulting coarse-scale solution for Bassi-Rebay's 1st method. We observe that the average coarse-scale solutions coincide with the exact solution at element interfaces, rendering the average fine-scale solutions zero at element interfaces. One can also see that the solution of the coarse-scale primal variable is equal to the exact solution at the Neumann boundary, and that the solution of the coarse-scale auxiliary variable is exact at the Dirichlet boundary. This is all in line with the predictions in Table 10.1. Figures 10.2 and 10.3 show the results for the Local Discontinuous Galerkin and Hybridizable Discontinuous Galerkin methods. The parameter choices are $\eta = 3$ and $\boldsymbol{\beta} = -0.25$ (which leads to $C = \frac{5}{64}$ in Figure 10.3a). As a proof of concept, η is chosen as 0 on the Dirichlet boundary for the HDG computation. The computations illustrate that identities in Table 10.1 hold.
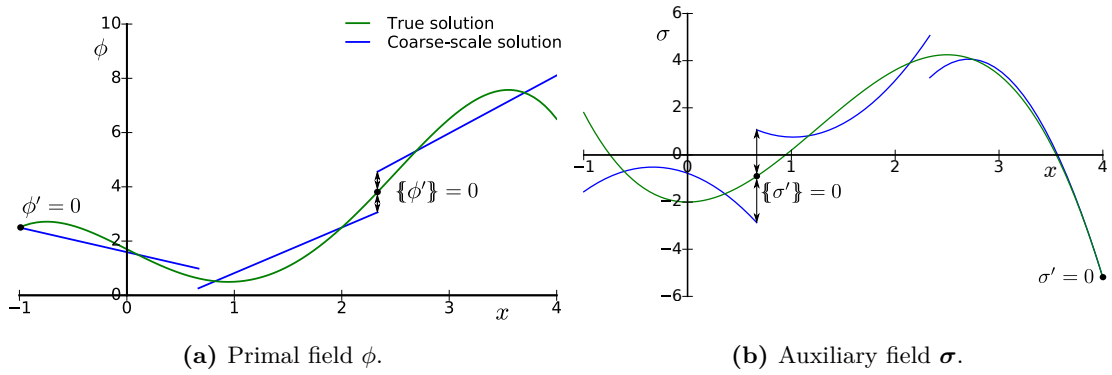
(a) Primal field ϕ.
    (b) Auxiliary field $\boldsymbol{\sigma}$.

Fig. 10.1: *One-dimensional numerical test results that illustrate the pointwise fine-scale identities for Bassi-Rebay's 1st method.*

(a) Primal field ϕ.
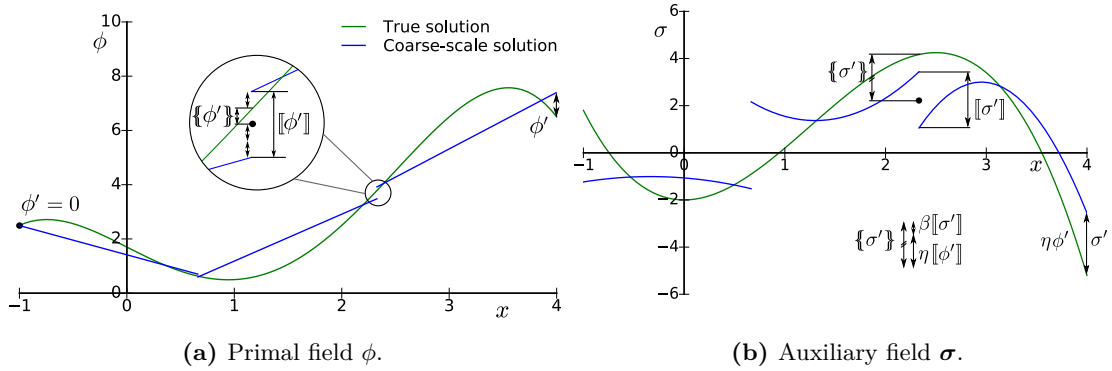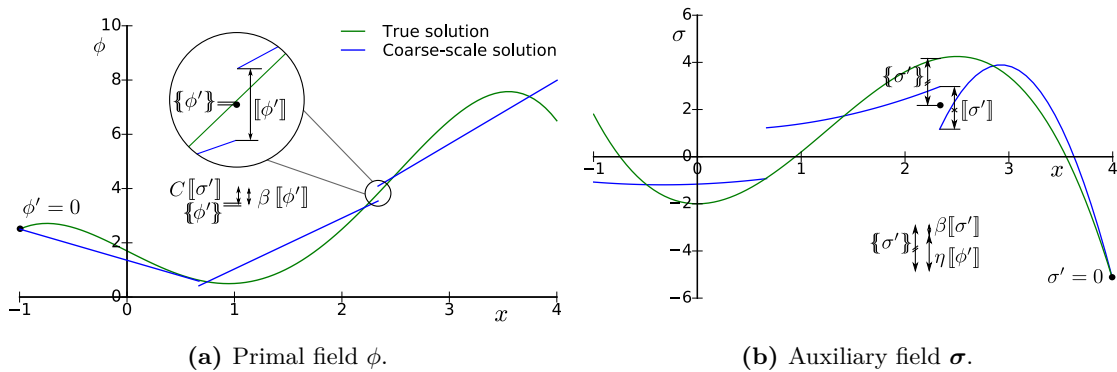   **(b)** Auxiliary field $\boldsymbol{\sigma}$.

Fig. 10.2: *One-dimensional numerical test results that illustrate the pointwise fine-scale identities for the Local Discontinuous Galerkin method ($\eta = 3$, $\boldsymbol{\beta} = -0.25$).*

(a) Primal field ϕ.
    (b) Auxiliary field $\boldsymbol{\sigma}$.

Fig. 10.3: *One-dimensional numerical test results that illustrate the pointwise fine-scale identities for the Hybridizable Discontinuous Galerkin method ($\eta = 3$ on Γ_0 and $\eta = 0$ on $\partial\Omega_D$, $\boldsymbol{\beta} = -0.25$).*

143

CHAPTER 11

Extending the residual-based model for inflow boundaries

Each of the coarse-scale discontinuous Galerkin formulations derived in the previous two chapters still include two terms that involve the fine-scale solution:

On interior boundaries: $\qquad \left\langle [\![w^h]\!], \boldsymbol{a} \{\!\!\{ \phi' \}\!\!\} \right\rangle_{\Gamma_0},$ (11.1a)

In element volumes: $\qquad - \left(\boldsymbol{a} \cdot \nabla w^h, \phi' \right)_{\tilde{\Omega}}.$ (11.1b)

In this chapter we develop models for these terms suitable for discontinuous Galerkin finite element formulations. The model for the term in Eq. (11.1a) relies on the analysis of the implied projection operators performed in Sections 9.3 and 10.3, and the model for the term in Eq. (11.1b) is an extension of the model developed in Chapter 6. This chapter heavily relies on the theory developed in Chapter 6, which is a necessary precursor for understanding Sections 11.2 and 11.3.

11.1 Fine scales on interior element boundaries

We start with the fine scales in Eq. (11.1a). A similar term did not occur in Part I nor in Part II of this dissertation, as those concern H^1-conforming coarse-scale spaces where $[\![w^h]\!] = 0$.

To find an expression for the fine scales on element boundaries, we make use of the conditions imposed by the different projection operators. For each of the methods, we performed

a short analysis in the one-dimensional case. This revealed a set of identities that the fine scales in $\mathcal{W}'$ must adhere to. For the Interior Penalty method these are stated in Eq. (9.30), and for the methods with a mixed origin they are collected in Table 10.1. These may be written as relations for the fine-scale averages. For the four methods we obtain:

$$\{\phi'\} = 0 \quad \text{on } \Gamma_0 \qquad\qquad\qquad \text{for the BR method,} \qquad (11.2a)$$

$$\{\phi'\} = 0 \quad \text{on } \Gamma_0 \qquad\qquad\qquad \text{for the IP method,} \qquad (11.2b)$$

$$\{\phi'\} = -\boldsymbol{\beta} \cdot [\![\phi']\!] = \boldsymbol{\beta} \cdot [\![\phi^h]\!] \quad \text{on } \Gamma_0 \qquad \text{for the LDG method,} \qquad (11.2c)$$

$$\{\phi'\} = -\boldsymbol{\beta} \cdot [\![\phi']\!] - C[\![\boldsymbol{\sigma}']\!] = \boldsymbol{\beta} \cdot [\![\phi^h]\!] + C[\![\boldsymbol{\sigma}^h]\!] \text{ on } \Gamma_0 \text{ for the HDG method.} \qquad (11.2d)$$

In Eqs. (11.2c) and (11.2d) we use the known solution behavior $[\![\phi']\!] = -[\![\phi^h]\!]$ and $[\![\boldsymbol{\sigma}']\!] = -[\![\boldsymbol{\sigma}^h]\!]$ from the transmission conditions. We thus use the substitution of the fine-scale averages of Eq. (11.2) into Eq. (11.1a) as a closure model. This is a modeling approximation that becomes an identity in the one-dimensional case, which has been the general aim for all modeling approximations throughout this dissertation.

Remark 11.1. *The above suggestion is relatively straightforward and easy to implement. Still, it is a simplification that necessarily leads to approximation in the multidimensional case. We could take a step back and obtain a model for Eq.* (11.1a) *directly from the fine-scale constraints imposed by the projector. A preliminary step in this direction was taken in Eq.* (10.37). *As the equation illustrates, this would shift the modeling effort to $\boldsymbol{\sigma}'$ in the element interior.*

11.2   Fine scales in element volumes

We proceed with modeling the fine scales in Eq. (11.1b). Modeling this term was the focus of Chapter 6, and much of the same theory applies. In summary, we go through the following steps:

↪ We phrase the fine-scale problem in terms of the coarse-scale residual and the Lagrange multiplier.

↪ We define the fine-scale Green's function as the test function that induces a point sampling of ϕ'.

↪ Substitution back into the fine-scale problem solves the fine scales in terms of the coarse-scale residual and the Lagrange multiplier.

↪ To remove the Lagrange multiplier trial function, we decompose the fine-scale Green's function into the one associated to the H_0^1-projector and a remainder.

↪ The H_0^1 fine-scale Green's function corresponds to a very similar fine-scale problem. This problem formulation requires the enforcement of conditions on $\mathcal{W}$ such that it becomes a subset of H_0^1, for which we again make use of a Lagrange multiplier formulation. The expression of the new Lagrange multiplier as a function of the H_0^1 fine-scale Green's function is known.

↪ Substituting the decomposition into the expression for the fine scales allows us to cancel the unknown Lagrange multiplier term while adding the new (known) Lagrange multiplier term. This new term takes into account the fine-scale boundary values, which are otherwise assumed to vanish.

In Sections 6.1 and 6.2 we go through these steps in detail. If we adopt the same procedure while making use of discontinuous approximation spaces, then we obtain:

$$\phi'(x) = \langle \mathcal{R}_{\phi^h}, g'_{H_0^1} \rangle + \langle \lambda_0, \phi' \rangle_{\Gamma_D} \,, \tag{11.3}$$

where $\mathcal{R}_{\phi^h}$ is the coarse-scale residual, $g'_{H_0^1}$ is fine-scale Green's function associated to the H_0^1-projector and λ_0 is the Lagrange multiplier whose expression depends on $g'_{H_0^1}$.

We substitute this expression into Eq. (11.1b), the term that has to be modeled. In the one-dimensional case, the H_0^1 fine-scale Green's function decouples from element to element. By also writing the test function and residual on each element as the polynomial expansions $w^h = \sum_{i=1}^p \hat{w}_i \, x^i$ and $\mathcal{R}_{\phi^h} = \sum_{i=1}^p \hat{R}_i \, x^i$, the term simplifies to:

$$-\int_K \boldsymbol{a} \cdot \nabla w^h \, \phi' \, \mathrm{d}x = -\int_K a \, \hat{w}_p p \, h^{2p-2} \tau \, \hat{R}_p \, \mathrm{d}\hat{x} - \int_{\partial K} a \, \hat{w}_p p \, h^{p-1} \gamma \, \phi' \, \mathrm{d}\hat{x} \,. \tag{11.4}$$

The steps from Eq. (11.3) to Eq. (11.4) require vanishing integral expressions of the H_0^1 fine-scale Green's function. The reader is referred to Section 6.3, where we perform the derivation in detail. The model parameters γ and τ are based on the fine-scale Green's function as:

$$\tau = \frac{1}{|K|} \int_K \int_K \frac{x^{p-1}}{h^{p-1}} g'_{H_0^1}(x,y) \frac{y^{p-1}}{h^{p-1}} \, \mathrm{d}y \, \mathrm{d}x \,, \tag{11.5a}$$

$$\gamma = \frac{1}{|F|} \int_K \int_F \frac{x^{p-1}}{h^{p-1}} \mathbb{H} g'_{H_0^1}(x,y) \, \mathrm{d}y \, \mathrm{d}x \,, \tag{11.5b}$$

where the operator $\mathbb{H}$ depends on the PDE at hand. It relates to λ_0, the Lagrange multi-

plier for the problem that defines the H_0^1 fine-scale Green's function. For the advection-diffusion equation we obtain:

$$\mathbb{H} = -\kappa \, \boldsymbol{n} \cdot \nabla_y \,, \tag{11.6}$$

Equation (11.4) involves an integral over the element boundary, with the related parameter γ from Eq. (11.5b). It represents the homogenized effect of the fine-scale boundary value onto the element interior. In Part II of this dissertation, we made use of the same term on the Dirichlet boundary. There, the exact fine-scale solution is known in terms of coarse scales as $\phi_D - \phi^h$. Now that we are using discontinuous coarse scales, Eq. (11.4) requires us to know the fine-scale boundary values on all element boundaries. From the definition of the average and jump operators we may write the fine-scale field on either side of an interface as:

$$\phi'^{\pm} = \{\!\!\{\phi'\}\!\!\} + \tfrac{1}{2}[\![\phi']\!] \cdot \boldsymbol{n}^{\pm} \,. \tag{11.7}$$

The continuity condition requires that the solution satisfies $[\![\phi']\!] = -[\![\phi^h]\!]$. We thus only need a closure for the fine-scale averages, for which we make use of the results from Section 11.1. We obtain:

$$\phi'^{\pm} = -\tfrac{1}{2}[\![\phi^h]\!] \cdot \boldsymbol{n}^{\pm} \quad \text{on } \Gamma_0 \qquad\qquad \text{for the BR method,} \tag{11.8a}$$

$$\phi'^{\pm} = -\tfrac{1}{2}[\![\phi^h]\!] \cdot \boldsymbol{n}^{\pm} \quad \text{on } \Gamma_0 \qquad\qquad \text{for the IP method,} \tag{11.8b}$$

$$\phi'^{\pm} = (\boldsymbol{\beta} - \tfrac{1}{2}\boldsymbol{n}^{\pm}) \cdot [\![\phi^h]\!] \quad \text{on } \Gamma_0 \qquad\qquad \text{for the LDG method,} \tag{11.8c}$$

$$\phi'^{\pm} = (\boldsymbol{\beta} - \tfrac{1}{2}\boldsymbol{n}^{\pm}) \cdot [\![\phi^h]\!] + C[\![\boldsymbol{\sigma}^h]\!] \quad \text{on } \Gamma_0 \qquad \text{for the HDG method.} \tag{11.8d}$$

which are exact in the one-dimensional case. Notice that with the choice $\boldsymbol{\beta} = \tfrac{1}{2}\boldsymbol{n}^-$ or $\boldsymbol{\beta} = \tfrac{1}{2}\boldsymbol{n}^+$ the one-dimensional Local Discontinuous Galerkin method would yield exact coarse scales at inflow or outflow boundaries of an element.

The above sets of equations represent an exact closure of the fine scales in the coarse-scale equations. Some simplifications (and hence approximations) have to be introduced to make them applicable for multidimensional problems. We propose:

$$-\int_K a\,\hat{w}_p p\,h^{p-1}\tau \hat{R}_p h^{p-1}\,\mathrm{d}\hat{x} \approx -\left(\boldsymbol{a}\cdot\nabla w^h, \tau_{\text{eff}}\mathcal{R}_{\phi^h}\right)_K, \tag{11.9a}$$

$$-\int_{\partial K} a\,\hat{w}_p p\,h^{p-1}\gamma\,\phi'\,\mathrm{d}\hat{x} \approx -\left\langle\boldsymbol{a}\cdot\nabla w^h, \gamma_{\text{eff}}^-\phi'^-\right\rangle_{\partial K^-} - \left\langle\boldsymbol{a}\cdot\nabla w^h, \gamma_{\text{eff}}^+\phi'^+\right\rangle_{\partial K^+}. \tag{11.9b}$$

The validity of these approximations depends on the quality of the estimates of the parameters τ_{eff}, γ_{eff}^+ and γ_{eff}^-. For one, they must be accurate approximations of the multidimensional fine-scale Green's function per Eq. (11.5). For higher order polynomials, another factor of importance is the change in differential operator. This important constituent of the model parameters is discussed in Section 6.5.1. The multiplicative factors derived in that section carry over to the parameters for discontinuous Galerkin methods:

$$\tau_{\text{eff}}, \gamma_{\text{eff}}^+, \gamma_{\text{eff}}^- \approx \begin{cases} \tau, & \gamma^+, & \gamma^- & \text{for } p = 1\,, \\ 12\,\tau, & \sqrt{3}\,\gamma^+, & \sqrt{3}\,\gamma^- & \text{for } p = 2\,, \\ 180\,\tau, & 2\sqrt{5}\,\gamma^+, & 2\sqrt{5}\,\gamma^- & \text{for } p = 3\,. \end{cases} \tag{11.10}$$

In Part II, we focused exclusively on the outflow boundary. That is where the boundary layer leads to large fine-scale boundary values. Ignoring the inflow part of Eq. (11.9b) is no longer appropriate when discontinuous Galerkin methods are considered. As a direct consequence of Eq. (11.8), discontinuities in the domain interior lead to non-vanishing fine scales on both the inflow and the outflow side of any element boundary. Moreover, the inflow fine-scale boundary value is the one that is advected into the element, and will have the most significant impact on the element interior. We thus require a new estimate for the parameter γ^-, the inflow γ.

11.3 γ^--parameter approximation for $p \in \{1, 2, 3\}$

Just like the γ approximations for the outflow boundary, the inflow γ follows from the exact expressions. Equation (6.39) stated:

$$\gamma_1 = \frac{h}{2}\left(\frac{2 + 2Pe_n - 2\exp(Pe_n)}{Pe_n - Pe_n\exp(Pe_n)}\right) =: \frac{h}{2}\,\zeta_1(Pe_n)\,, \tag{11.11a}$$

$$\gamma_2 = \frac{h}{12}\left(\frac{12 + 8Pe_n + 2Pe_n^2 - (12 - 4Pe_n)\exp(Pe_n)}{2Pe_n + Pe_n^2 - (2Pe_n - Pe_n^2)\exp(Pe_n)}\right) =: \frac{h}{2}\,\zeta_2(Pe_n)\,, \tag{11.11b}$$

$$\gamma_3 = \frac{h}{60}\left(\frac{120 + 72Pe_n + 18Pe_n^2 + 2Pe_n^3 - (120 - 48Pe_n + 6Pe_n^2)\exp(Pe_n)}{12Pe_n + 6Pe_n^2 + Pe_n^3 - (12Pe_n - 6Pe_n^2 + Pe_n^3)\exp(Pe_n)}\right)$$
$$=: \frac{h}{2}\,\zeta_3(Pe_n), \tag{11.11c}$$

148

for $p = 1, 2$ and 3, respectively. On the inflow boundary we obtain the advective and diffusive limits by taking $Pe_n = \frac{a \cdot n\, h}{\kappa} \leq 0$. The limits are:

$$\gamma_{1,a}^- := \lim_{Pe_n \to -\infty} \gamma_1 = h \,, \qquad \gamma_{1,d}^- := \lim_{Pe_n \to 0^-} \gamma_1 = \frac{h}{2} \,, \tag{11.12a}$$

$$\gamma_{2,a}^- = \frac{h}{6} \,, \qquad\qquad \gamma_{2,d}^- = \frac{h}{12} \,, \tag{11.12b}$$

$$\gamma_{3,a}^- = \frac{h}{30} \,, \qquad\qquad \gamma_{3,d}^- = \frac{h}{60} \,. \tag{11.12c}$$

The diffusive limits are the same as those for the outflow boundary. That is to be expected: if diffusion dominates, then there is no distinction between inflow and outflow. The advective limits are quite different. For the outflow boundary these limits were zero, as the boundary data no longer affects the element interior. Of course, the opposite is true for the inflow boundary: with increasing advective dominance, the inflow boundary data becomes more and more important in terms of interior solution behavior.

Recall from Section 6.5.3 that the approximation for γ followed from its definition in Eq. (11.5b) as:

$$\gamma^- \propto |\mathbb{H}|\, \hat{\tau}^- \,, \tag{11.13}$$

where we took $|\mathbb{H}| \propto \sqrt{\kappa / \tau_d}$ for the scaling of $\mathbb{H}$. The expression for $\hat{\tau}^-$ takes into account the change of integration domain of γ compared to that of τ. For the inflow boundary we find that the following approximation is suitable:

$$\hat{\tau}^- = \sqrt{\frac{1 + 4\frac{c_2}{c_1}\left(\frac{\tau_{1,d}}{\tau_{1,a}}\right)^2}{c_1 \tau_{1,a}^{-2} + c_2 \tau_{1,d}^{-2}}} \,, \tag{11.14}$$

where we have chosen to relate $\hat{\tau}^-$ to the stabilization parameters for the linear case. Suitable choices of c_1 and c_2 will result in the correct scaling of the approximate γ^-. Substitution of Eq. (11.14) into Eq. (11.13) results in the following approximation of γ^-:

$$\gamma^- \approx c_s \sqrt{\frac{\kappa + 4\kappa\frac{c_1}{c_2}\left(\frac{\tau_{1,d}}{\tau_{1,a}}\right)^2}{c_1 \tau_{1,d}\tau_{1,a}^{-2} + c_2 \tau_{1,d}^{-1}}} \,, \tag{11.15}$$

where c_s takes into account the effect of the shape of the element.

In the one-dimensional case, the expressions for $\tau_{1,d}$ and $\tau_{1,a}$ are known to be $h^2/(12\kappa)$ and

149

$h/(2\|\boldsymbol{a}\|)$, respectively (see Section 6.5.2). Substitution thereof into Eq. (11.15) results in an approximation that depends on the element Péclet number and scales linearly with h:

$$\gamma^- \approx \frac{h}{2} \sqrt{\frac{\frac{1}{9}\frac{c_1}{c_2}Pe_n^2 + 1}{\frac{1}{12}c_1 Pe_n^2 + 3\,c_2}} =: \frac{h}{2}\,\tilde{\zeta}^-(Pe_n). \tag{11.16}$$

By taking the limit $Pe_n \to 0^-$ we obtain $h/\sqrt{12c_2}$. Matching the limiting values to those of Eq. (11.12) requires $c_2 = 1/3$, 12 and 300 for $p = 1$, 2 and 3, respectively. With these choices, the limits $Pe_n \to -\infty$ comply with those of Eq. (11.12) (which is the result of the ansatz of Eq. (11.14)). The remaining parameter c_1 will be used to shift the graph such that the approximation exhibits symmetry through the correct midpoint in the log-log plot. With empirical fitting we obtain $c_1 = 8/10$, 10 and 120 for $p = 1$, 2 and 3.

Figure 11.1 shows the exact and approximate $\tilde{\zeta}^-$ functions. For reference, the scaling of the outflow function $\tilde{\zeta}^+$ from Section 6.5.3 is shown in Figure 11.1b. The figure confirms that the limits of both functions overlap for all polynomial orders. The approximate function is rotationally symmetric around the correct midpoint. The general shape of the approximation function is also correct, although we observe undershoots before the midpoint, and overshoots after the midpoint. This is a result of the exponents of two in the ansatz function of Eqs. (11.14) and (11.16). A better fit could be achieved for a different, smaller, exponent. However, to avoid overfitting for this specific case (one-dimensional linear advection-diffusion), we will proceed with the functions from Figure 11.1a. An exponent of two was also used in the construction of the τ approximation and for the γ^+ approximation in Sections 6.5.2 and 6.5.3 and is widely used in literature as the 'switching' exponent in the harmonic mean of typical stabilized methods [19, 186, 187].
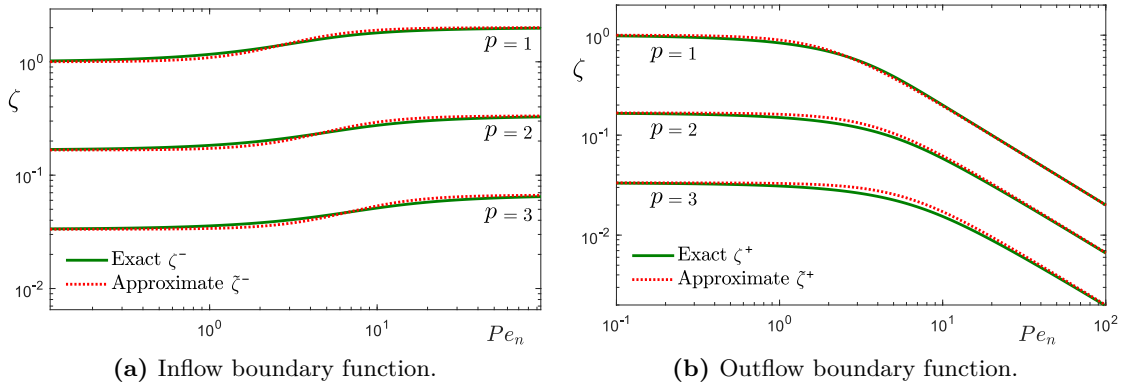
(a) Inflow boundary function.

(b) Outflow boundary function.

Fig. 11.1: *Exact ζ^- and ζ^+ functions and their approximations for the one-dimensional case. Showing the correct scaling of $\gamma^{\pm}$.*

11.4 Summary of fine-scale model for discontinuous Galerkin methods

The discontinuous Galerkin formulations for the IP, Bassi-Rebay, LDG and HDG methods, Eqs. (9.27), (10.20), (10.29) and (10.33), respectively, were derived from a variational multiscale perspective and require the addition of a fine-scale model. The additional terms that we propose to add to the 'plain' formulations are:

$$
\begin{aligned}
B_{\mathrm{VMS,DG}} = &-(\boldsymbol{a}\cdot\nabla w^h, \phi')_{\tilde{\Omega}} + \langle \llbracket w^h \rrbracket, \boldsymbol{a}\{\!\{\phi'\}\!\}\rangle_{\Gamma_0} \\
\approx &-\sum_{K\in\mathcal{T}} \Big\{ \ (\boldsymbol{a}\cdot\nabla w^h, \tau_{\mathrm{eff}}\mathcal{R}_{\phi^h})_K \\
&+ \langle (\boldsymbol{a}\cdot\nabla w^h)^-, \gamma_{\mathrm{eff}}^-(\{\!\{\phi'\}\!\} - \tfrac{1}{2}\llbracket\phi^h\rrbracket\cdot\boldsymbol{n}^-)\rangle_{\partial K^-\backslash\partial\Omega} \\
&+ \langle (\boldsymbol{a}\cdot\nabla w^h)^+, \gamma_{\mathrm{eff}}^+(\{\!\{\phi'\}\!\} - \tfrac{1}{2}\llbracket\phi^h\rrbracket\cdot\boldsymbol{n}^+)\rangle_{\partial K^+\backslash\partial\Omega} \\
&+ \langle \boldsymbol{a}\cdot\nabla w^h, \gamma_{\mathrm{eff}}^+(\phi_D - \phi^h)\rangle_{\partial K\cap\partial\Omega_D^+} \ \Big\} + \langle \llbracket w^h \rrbracket, \boldsymbol{a}\{\!\{\phi'\}\!\}\rangle_{\Gamma_0},
\end{aligned}
\tag{11.17}
$$

where the closure for the fine-scale averages depends on the method, and followed from Section 11.1 as:

$$
\{\!\{\phi'\}\!\} = 0 \quad \text{on } \Gamma_0 \qquad\qquad\qquad \text{for the BR method,} \tag{11.18a}
$$

$$
\{\!\{\phi'\}\!\} = 0 \quad \text{on } \Gamma_0 \qquad\qquad\qquad \text{for the IP method,} \tag{11.18b}
$$

$$
\{\!\{\phi'\}\!\} = \boldsymbol{\beta}\cdot\llbracket\phi^h\rrbracket \quad \text{on } \Gamma_0 \qquad\quad \text{for the LDG method,} \tag{11.18c}
$$

$$
\{\!\{\phi'\}\!\} = \boldsymbol{\beta}\cdot\llbracket\phi^h\rrbracket + C\llbracket\boldsymbol{\sigma}^h\rrbracket \quad \text{on } \Gamma_0 \qquad \text{for the HDG method.} \tag{11.18d}
$$
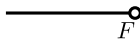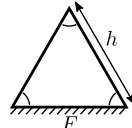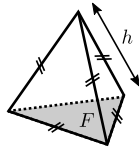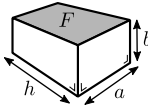
The model parameters, τ_{eff}, γ_{eff}^+ and γ_{eff}^- are collected in Table 11.1 for different polynomial orders and different element types. We relate all model parameters to $\tau_{1,a}$ and $\tau_{1,d}$, the advective and diffusive limits of the classical stabilization parameter for linear elements. The preferred method for estimating these limits is up to the analyst (e.g., with the use of the Jacobian of the element mapping as proposed in [19, 23], or with the definitions of the local length-scales and Péclet numbers of [187]). In the numerical experiments of the following chapter we use the analytical expressions from the one-dimensional case:

$$
\tau_{1,a} = \frac{h}{2\|\boldsymbol{a}\|}, \tag{11.19a}
$$

$$
\tau_{1,d} = \frac{h^2}{12\kappa}, \tag{11.19b}
$$

while using the element diameter h and an elementwise constant diffusivity and velocity magnitude.

Table 11.1: *Overview of τ_{eff}, γ_{eff}^+ and γ_{eff}^- expressions for different elements and polynomial degrees.*

	$p = 1$	$p = 2$	$p = 3$
τ_{eff} :	$\sqrt{\dfrac{1}{\tau_{1,a}^{-2} + \tau_{1,d}^{-2}}}$	$\sqrt{\dfrac{1}{9\,\tau_{1,a}^{-2} + 25\,\tau_{1,d}^{-2}}}$	$\sqrt{\dfrac{1}{25\,\tau_{1,a}^{-2} + \frac{1225}{9}\,\tau_{1,d}^{-2}}}$
γ_{eff}^+ :	$c_s\sqrt{\dfrac{\kappa}{3\tau_{1,d}\tau_{1,a}^{-2} + \frac{1}{3}\tau_{1,d}^{-1}}}$	$c_s\sqrt{\dfrac{\kappa}{9\tau_{1,d}\tau_{1,a}^{-2} + 4\tau_{1,d}^{-1}}}$	$c_s\sqrt{\dfrac{\kappa}{15\tau_{1,d}\tau_{1,a}^{-2} + 15\tau_{1,d}^{-1}}}$
γ_{eff}^- :	$c_s\sqrt{\dfrac{\kappa + 9.6\kappa\left(\frac{\tau_{1,d}}{\tau_{1,a}}\right)^2}{0.8\tau_{1,d}\tau_{1,a}^{-2} + \frac{1}{3}\tau_{1,d}^{-1}}}$	$c_s\sqrt{\dfrac{\kappa + \frac{10}{3}\kappa\left(\frac{\tau_{1,d}}{\tau_{1,a}}\right)^2}{10\tau_{1,d}\tau_{1,a}^{-2} + 12\tau_{1,d}^{-1}}}$	$c_s\sqrt{\dfrac{\kappa + 1.6\kappa\left(\frac{\tau_{1,d}}{\tau_{1,a}}\right)^2}{120\tau_{1,d}\tau_{1,a}^{-2} + 300\tau_{1,d}^{-1}}}$
Element:			
c_s (Eq. (6.43)):	1	$\frac{4}{\sqrt{3}}$	$2\sqrt{\frac{2}{3}}$ h/b

CHAPTER 12

Numerical experiments

In Part II of this dissertation we used a continuous Galerkin treatment of the PDE, while incorporating the Dirichlet conditions weakly. Without a suitable fine-scale model (that is, stabilization technique), the continuous Galerkin method yields a highly oscillatory solution. The discontinuous Galerkin method is more naturally stable, which begs the question, "do we still need a fine-scale model?" In this chapter, we answer this question with numerical experiments. We also compare the relative importance of the different components of the residual-based model: the volumetric, the outflow and the inflow term.

12.1 Numerical verification for a one-dimensional model problem

We begin with a number of one-dimensional simulations. All approximations become identities, allowing us to verify the derivations of the previous sections. We also see how the different projection operators affect the solution quality.

12.1.1 Outflow boundary layer

One of the promises of discontinuous Galerkin methods is that they are better suited for approximating solutions with sharp layers. The lack of continuity means that the interpolation error can be minimized in each element separately. A pure L^2-projection

of a Heaviside function would only spoil the result in a single element, whereas it would produce strong oscillations throughout the domain when (higher-order) continuous bases are used. In practice, we do not obtain L^2-projections of exact solutions, nor do we aim to. The earlier analysis revealed exactly the projections that we do aim to obtain. In this section, we perform numerical experiments concerning an outflow boundary layer. We make observations regarding the suitability of the different projectors when sharp layers are present, and we determine the impact of the fine-scale model.

We consider an advection dominated case with $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.01$. For each of the four discontinuous Galerkin methods, we show results for three pairs of polynomial order bases. These are $p, q = (1, 2)$, $p, q = (2, 3)$ and $p, q = (3, 4)$, where p and q are the polynomial orders corresponding to $\mathcal{W}^h$ and Σ^h, respectively. We show the primal solution variable on the left, and the auxiliary variable on the right. In the background of each graph, the exact solution is shown transparently in green and the plain method in gray. On the foreground, the formulations include the residual-based model: blue for only the classical (volumetric) variational multiscale model, purple for the augmented model when only the outflow is included, and red for the augmented model when both the inflow and the outflow terms are included.

Figures 12.1 to 12.3 show the results for Bassi-Rebay's 1$^{\text{st}}$ method. Inf-sup stability of the Bassi-Rebay projector of Eq. (10.15) cannot be proven in general, but unique solutions do exists for these particular model set-ups. As Bassi-Rebay's 1$^{\text{st}}$ method and the Bassi-Rebay projector represent the simplest case, this provides a baseline for interpreting all the following numerical results. In the figures, we observe a significant difference between each of the variational multiscale models, most significantly so for the simulations of Figure 12.1 where the lowest order basis functions are used. Recall from Table 10.1 that the Bassi-Rebay projector in one dimension yields coarse scales with exact averages on element boundaries. As the jumps are not penalized, the solution exhibits large fine-scale boundary values on both sides of each node. Hence, also the inflow part of the VMS model has a meaningful impact. Indeed, only the simulations that make use of the complete augmented model (i.e., the volumetric term, the outflow term and the inflow term) satisfy the nodal identities of Table 10.1 and are thus the coarse-scale solutions defined by the Bassi-Rebay projector of Eq. (10.15). This verifies the correctness of the residual-based model developed in Chapter 11.
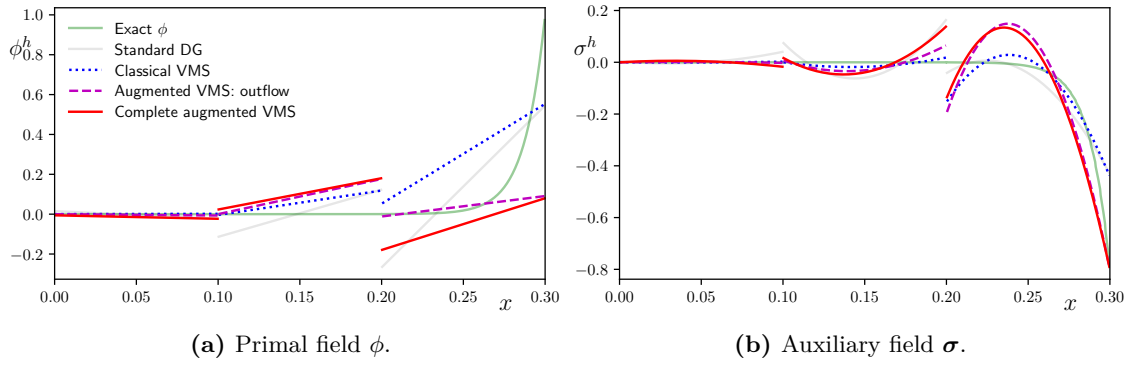
(a) Primal field ϕ.

(b) Auxiliary field $\boldsymbol{\sigma}$.

Fig. 12.1: *One-dimensional results of Bassi-Rebay's 1st method using three $p, q = (1, 2)$ elements.*

(a) Primal field ϕ.

(b) Auxiliary field $\boldsymbol{\sigma}$.

Fig. 12.2: *One-dimensional results of Bassi-Rebay's 1st method using three $p, q = (2, 3)$ elements.*

(a) Primal field ϕ.

(b) Auxiliary field $\boldsymbol{\sigma}$.
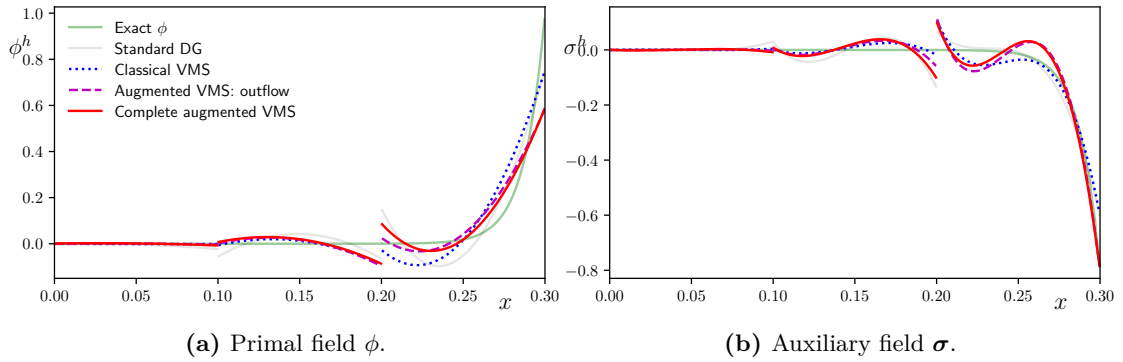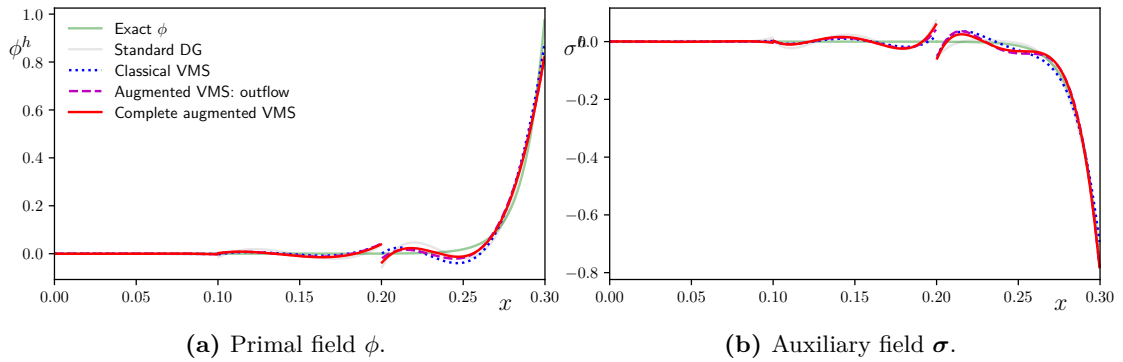
Fig. 12.3: *One-dimensional results of Bassi-Rebay's 1st method using three $p, q = (3, 4)$ elements.*

155

In the subsequent three sets of figures, Figures 12.4 to 12.6, we use the Interior Penalty method with a penalty parameter η of $2p/h$. Again, there is a notable distinction between the various degrees of VMS model. Interestingly, and different from the simulations that involved Bassi-Rebay's 1st method, this distinction becomes more pronounced with polynomial order. For the simulations of Figure 12.4, which make use of the lowest order polynomials, the obtained solution without any VMS model is already very satisfactory. However, with increasing polynomial order the solution starts to exhibit oscillations within the element. To some extend this may be countered by adding the volumetric part of the VMS model, as illustrated by the blue dotted line in Figure 12.6. However, the figure also show that in order to arrive at a good solution we also need to include the augmented VMS terms. Those are the solutions that satisfy the constraints dictated by the Interior Penalty projector (Eqs. (9.22) and (9.30)) and are thus the exact coarse-scale solutions. The identities of Eq. (9.30) have been verified numerically.

In all these figures, the purple dashed line and the red line are nearly overlapping. That means that the impact of the inflow part of the VMS model is almost negligible. As the Interior Penalty projector dictates vanishing fine-scale average of the primal field while penalizing the solution jump, the fine-scale solution is necessarily small at the interior nodes. Naturally, the inflow and outflow parts of the model have little impact in elements when the fine-scale boundary values are small. Only at the domain outflow node does the fine-scale solution still have a large nonzero value, such that it is the outflow part of the VMS model that has a significant impact.

(a) Primal field ϕ. **(b)** Auxiliary field $\boldsymbol{\sigma}$.
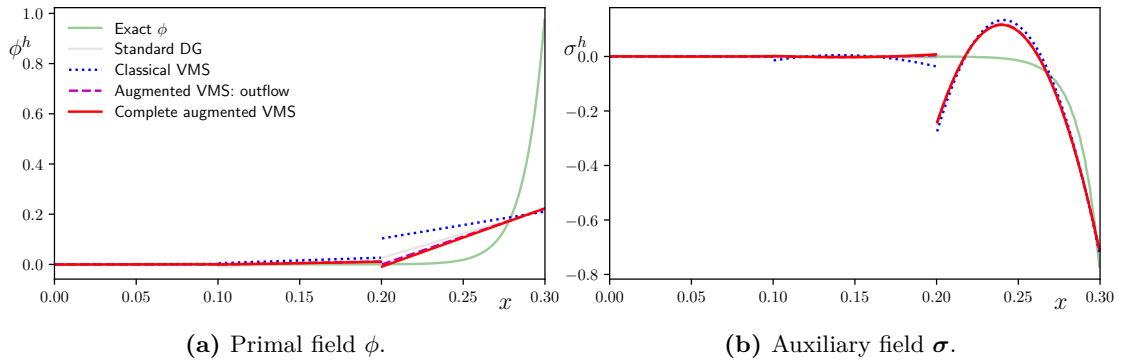
Fig. 12.4: *One-dimensional results of the Interior Penalty method using three $p, q = (1, 2)$ elements.*

(a) Primal field ϕ.

(b) Auxiliary field $\boldsymbol{\sigma}$.
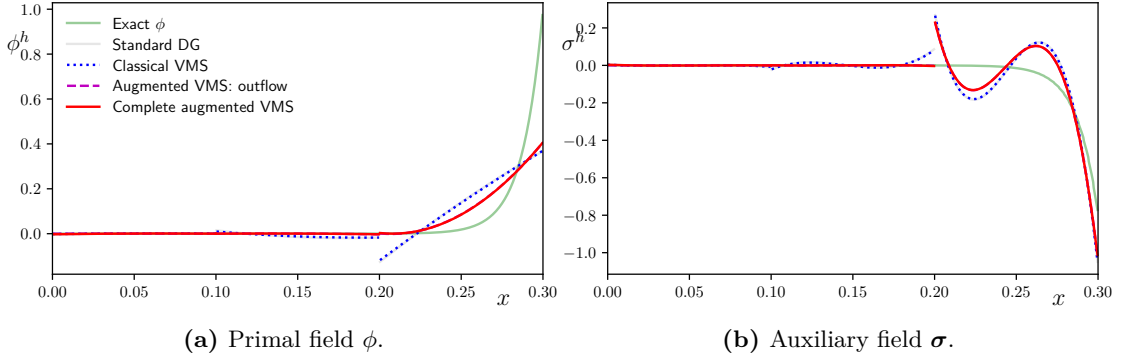
Fig. 12.5: *One-dimensional results of the Interior Penalty method using three $p, q = (2, 3)$ elements.*

(a) Primal field ϕ.

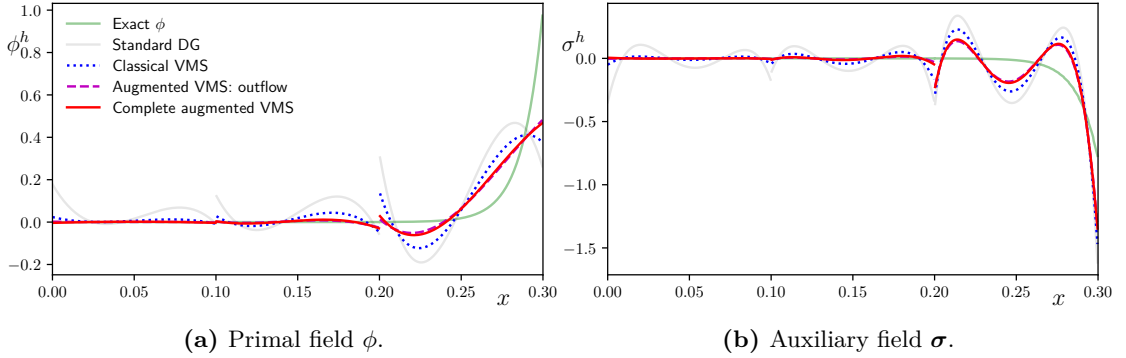**(b)** Auxiliary field $\boldsymbol{\sigma}$.

Fig. 12.6: *One-dimensional results of the Interior Penalty method using three $p, q = (3, 4)$ elements.*

Next, we use the Local Discontinuous Galerkin method and obtain the results shown in Figs. 12.7 to 12.9. We make use of $\eta = 2p/h$ and $\boldsymbol{\beta} = \frac{1}{2}$. The qualitative behavior of these solution is a mixture of that of the earlier computations. Just like for the Bassi-Rebay simulations, the VMS modeling becomes less impactful for the higher-order polynomial bases. Still, each of the solution fields remains distinguishable, which means that each part of the VMS model has a contribution. Only when all the terms are added to the VMS model do we obtain the coarse-scale solution defined by the LDG projector of Eq. (10.22). For this one-dimensional case the projector dictates the constraints on the fine-scale field from Table 10.1. Different from Bassi-Rebay's 1st method, and more like the Interior Penalty method, these fine-scale constraints lead to high quality coarse-scale solutions. Specifically, the complete interpolation error is constrained to a single element. Notice that, due to the choice of $\boldsymbol{\beta} = \frac{1}{2}$, the first fine-scale constraint in Table 10.1 says $\{\!\{\phi'\}\!\} +$

$\frac{1}{2}\boldsymbol{n}^+[\![\phi']\!] = \phi'^+ = 0$ on interior nodes. That means that the coarse-scale solution is exact at the outflow node of the element. As a result, the poor solution behavior that is a necessary evil in the boundary layer does not spoil the elements in the bulk of the domain interior.

Remark 12.1. *Typically, $\boldsymbol{\beta}$ is chosen relatively arbitrarily as a 'switch vector' according to the definition in Eq. (10.24). However, the previous observation provides a logic behind the choice of the switch orientation: for the Local Discontinuous Galerkin method it should point in the direction of the steep layers. Typically, this means that it should act in the direction of the flow, meaning $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{n}^+$.*

(a) Primal field ϕ. **(b)** Auxiliary field $\boldsymbol{\sigma}$.
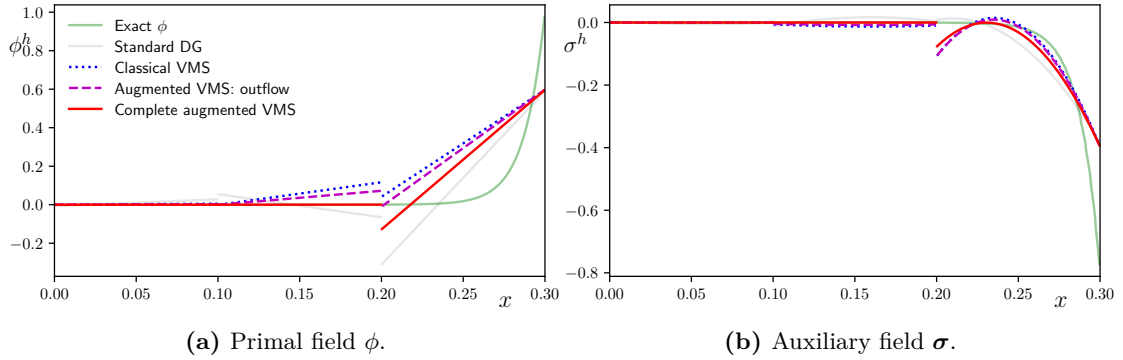
Fig. 12.7: *One-dimensional results of the Local Discontinuous Galerkin method using three $p, q = (1, 2)$ elements.*

(a) Primal field ϕ. **(b)** Auxiliary field $\boldsymbol{\sigma}$.
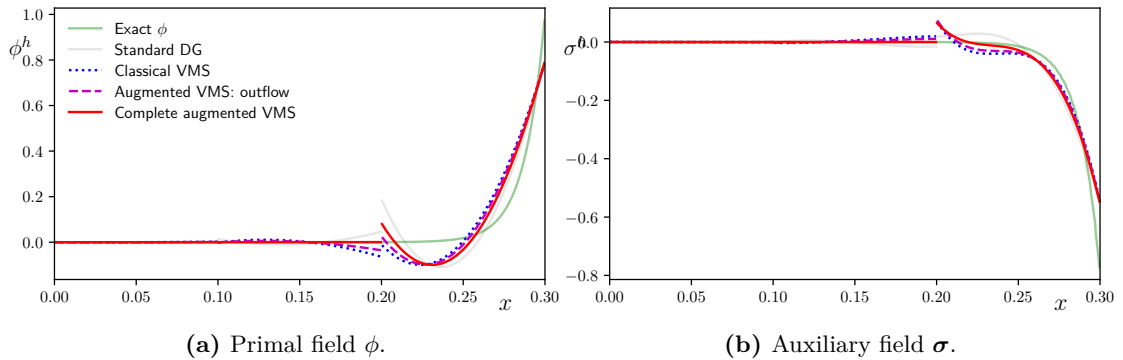
Fig. 12.8: *One-dimensional results of the Local Discontinuous Galerkin method using three $p, q = (2, 3)$ elements.*

(a) Primal field ϕ.

(b) Auxiliary field $\boldsymbol{\sigma}$.
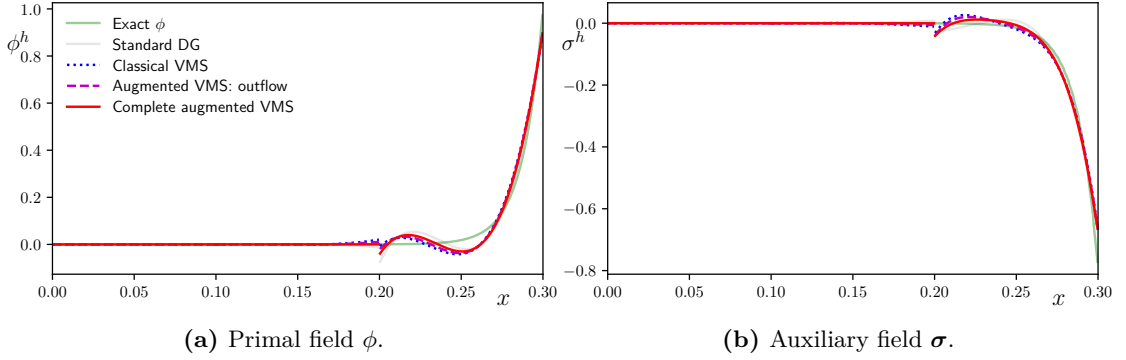
Fig. 12.9: *One-dimensional results of the Local Discontinuous Galerkin method using three $p, q = (3, 4)$ elements.*

Finally, the simulations in Figures 12.10 to 12.12 involve the Hybridizable Discontinuous Galerkin method. We again use $\eta = 2p/h$ and $\boldsymbol{\beta} = \frac{1}{2}$, and define the penalty parameter C as $(1 - \boldsymbol{\beta} \cdot \boldsymbol{\beta})/(4\kappa\eta)$. The conclusions regarding solution quality and model importance are roughly the same as those of for the Local Discontinuous Galerkin method just discussed. However, the added penalty term in the definition of the HDG-projection operator of Eq. (10.30) no longer leads to the favorable fine-scale constraint that caused $\phi'^+ = 0$ for the LDG method. The result is that the solution in the domain interior is still (more or less so) impacted by the under resolved boundary layer. This is the price that is paid for making the discontinuous Galerkin method hybridizable.

In summary, for these boundary layer experiments, the impact of the residual-based fine-scale model has focused on outflow boundary. This is where the residual is necessarily large. Additionally, since the boundary conditions are enforced weakly, this is also where the fine-scale solution shows large boundary values. Hence, the outflow part of residual-based model is crucial for obtaining the exact coarse-scale solution. These conclusions were also drawn in Part II and they are no different for the discontinuous Galerkin method just investigated (more or less so depending on the method under consideration).

(a) Primal field ϕ.

(b) Auxiliary field $\boldsymbol{\sigma}$.
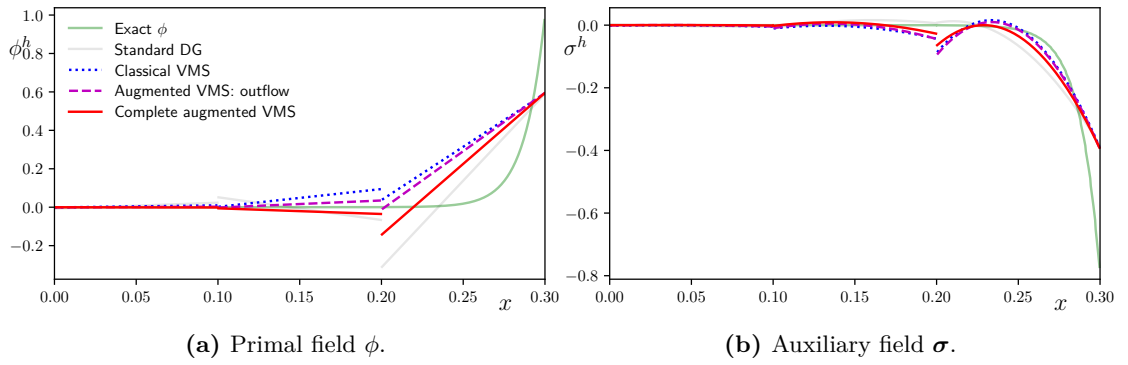
Fig. 12.10: *One-dimensional results of the Hybridizable Discontinuous Galerkin method using three $p, q = (1, 2)$ elements.*
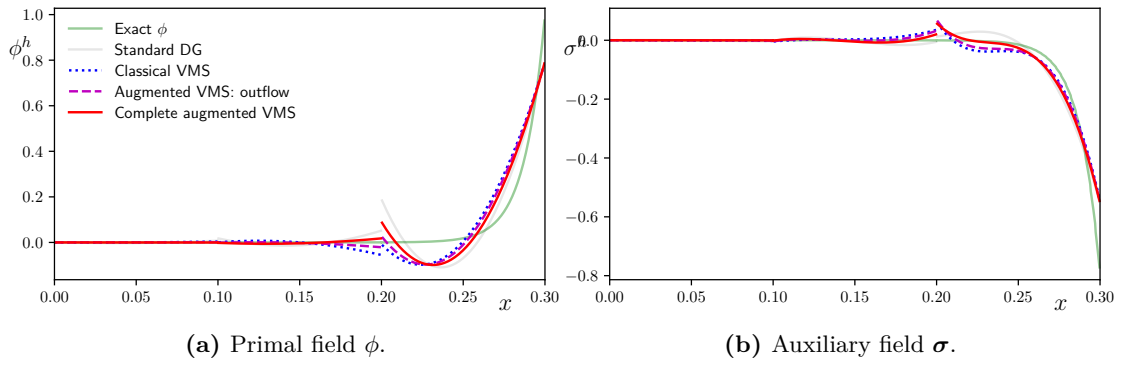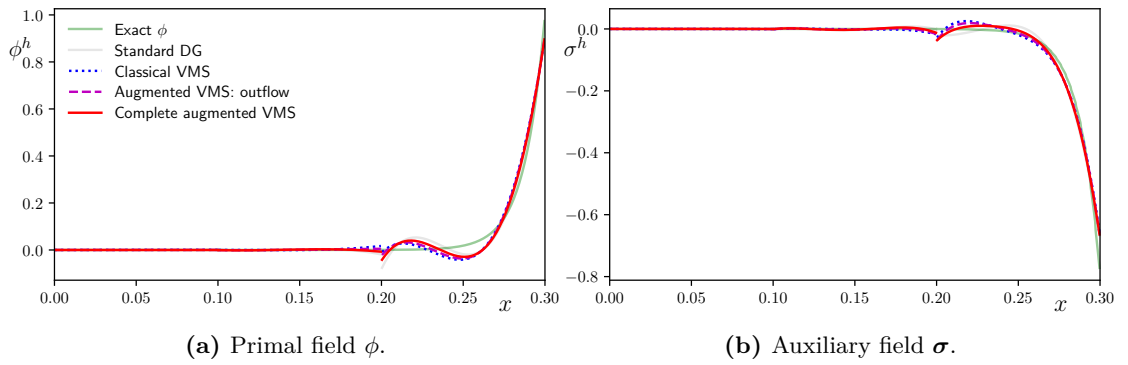

(a) Primal field ϕ.

(b) Auxiliary field $\boldsymbol{\sigma}$.

Fig. 12.11: *One-dimensional results of the Hybridizable Discontinuous Galerkin method using three $p, q = (2, 3)$ elements.*

(a) Primal field ϕ.

(b) Auxiliary field $\boldsymbol{\sigma}$.

Fig. 12.12: *One-dimensional results of the Hybridizable Discontinuous Galerkin method using three $p, q = (3, 4)$ elements.*

12.1.2 Bubble force

To get a sense of the importance of the inflow part of the model, we construct a study case where the coarse-scale solution exhibits a non-negligible jump around the inflow section of the domain. We do so by adding a large non-zero source in the first element of the domain. To ensure that the residual remains a polynomial function, we choose the source function f to be the quadratic function $2300x^2$ in the first one-third of the domain, and zero in the remaining two-third. This function is illustrated in Figure 12.13.
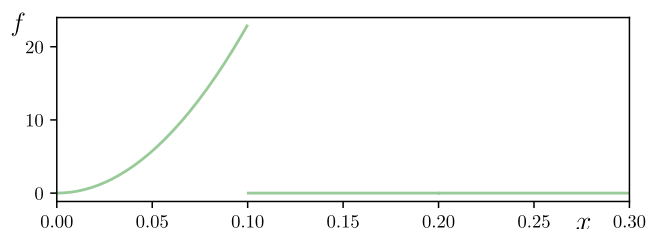
Fig. 12.13: *Bubble force at the inflow section of the domain.*

Due to the quadratic forcing, the residual in the first element will be at least quadratic. The residual-based model would no longer be exact if the polynomial approximation for ϕ would be of order 1 or 2. For each of the four methods, we thus show results only for the pair $p, q = (3, 4)$. Again, the primal solution is shown in the figure on the left, and the auxiliary variable in the figure on the right. In the background of each graph, the exact solution is plotted in transparent green and the solution of the plain discontinuous Galerkin method in transparent gray. In the foreground, the classical volumetric VMS model is graphed as a dotted blue line, the model that includes the volumetric term and the outflow term is graphed as a dashed purple line, and the solid red line represents the complete augmented VMS model.
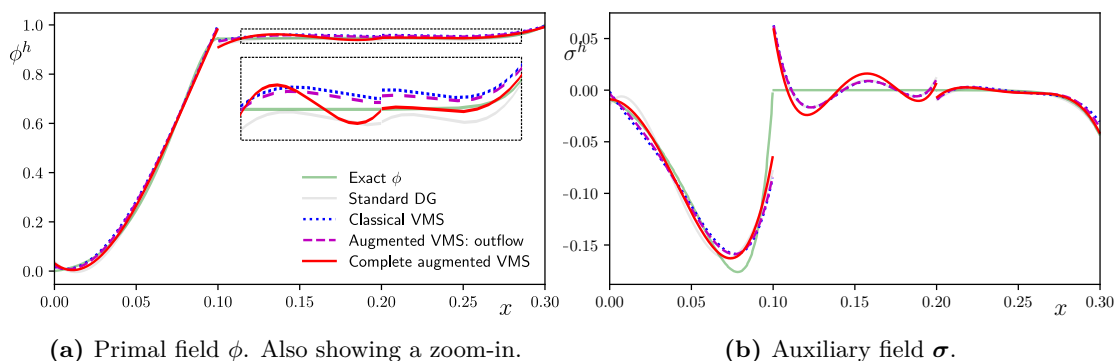
(a) Primal field ϕ. Also showing a zoom-in.   **(b)** Auxiliary field $\boldsymbol{\sigma}$.

Fig. 12.14: *Results of Bassi-Rebay's 1st method for the bubble force problem.*

(a) Primal field ϕ. Also showing a zoom-in.

(b) Auxiliary field $\boldsymbol{\sigma}$.
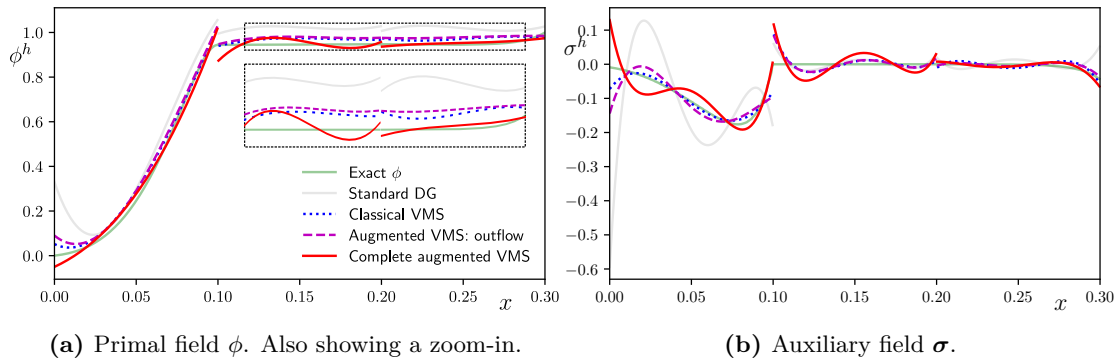
Fig. 12.15: *Results of the Interior Penalty method for the bubble force problem.*

(a) Primal field ϕ. Also showing a zoom-in.

(b) Auxiliary field $\boldsymbol{\sigma}$.
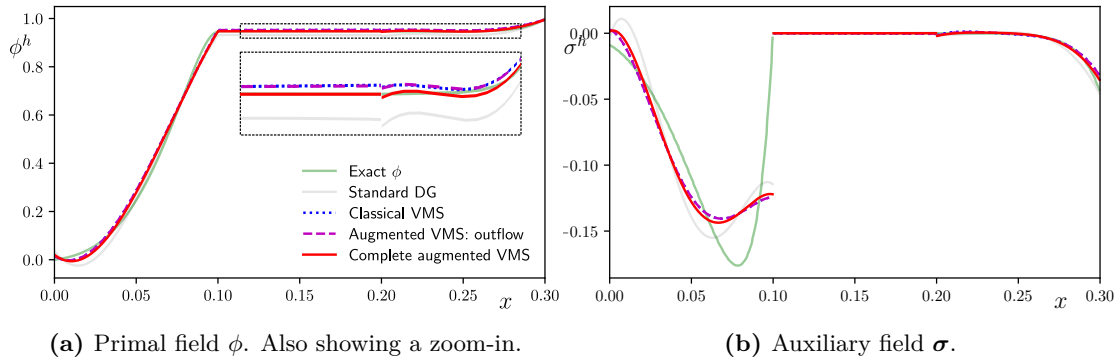
Fig. 12.16: *Results of the Local Discontinuous Galerkin method for the bubble force problem.*

(a) Primal field ϕ. Also showing a zoom-in.

(b) Auxiliary field $\boldsymbol{\sigma}$.
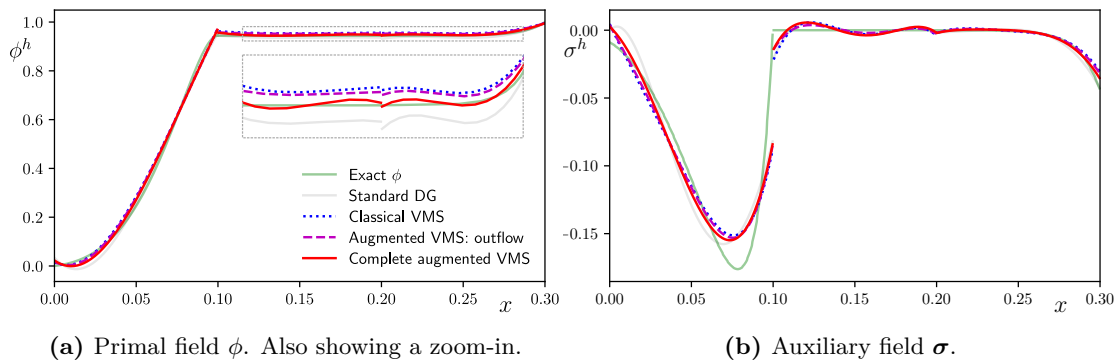
Fig. 12.17: *Results of the Hybridizable Discontinuous Galerkin method for the bubble force problem.*

The close-ups in these graphs show that the overshoot produced by the sharp layer in the first element causes a lasting overprediction of the primal variable. This effect is particularly strong for the Interior Penalty method, where the gray line overpredicts the true value of ϕ by almost 10% in the second and third element. While the classical VMS model partly remedies the situation, the effect is only completely mitigated by adding the inflow term of the VMS model. Interestingly, and somewhat in line with the results from Section 12.1.1, the plain Local Discontinuous Galerkin and Hybridizable Discontinuous Galerkin methods are quite robust for this high polynomial order, and benefit relatively little from any of the VMS models. The improvement only becomes clear after significant zoom-in.

12.2 Numerical experiments for a two-dimensional model problem

For the one-dimensional computations, the fine-scale equation could be inverted exactly and the exact effect of the fine scales on the coarse scales was known. This is no longer the case when the spatial domain is two-dimensional. All the modeling assumptions come into play as approximations. In this section, we use the same benchmark problem as we did in Section 7.2.2: a square domain with a polygonal cut-out. The problem is described in Figure 12.18a, and the true solution is shown in Figure 12.18b.

First, to get a sense of the qualitative behavior of the different discontinuous Galerkin methods, we solve the problem on a coarse mesh with the IP, LDG and HDG formulations. Unless stated otherwise, we use $\eta = 2p^2/h$ in all the following, and $\boldsymbol{\beta}$ as either $\frac{1}{2}\boldsymbol{n}^+$ or $\frac{1}{2}\boldsymbol{n}^-$ which is randomly selected. Figure 12.19 shows the primary solution fields for each of these
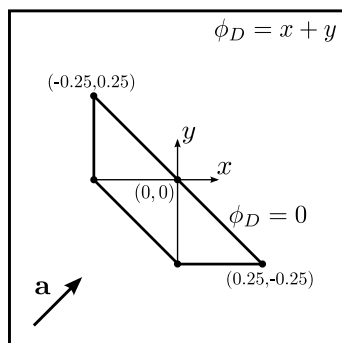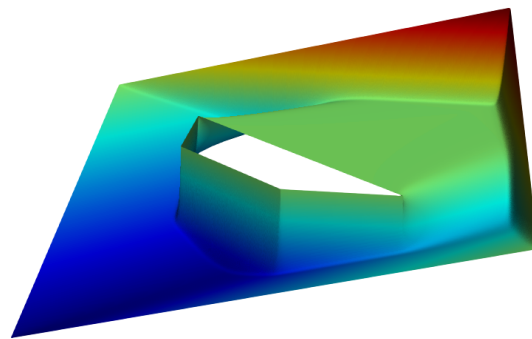
(a) Problem specification.   (b) Solution for $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.003$.

Fig. 12.18: *Two-dimensional model problem.*

(a) Interior Penalty method.   **(b)** Local Discontinuous Galerkin method.   **(c)** Hybridizable Discontinuous Galerkin method.
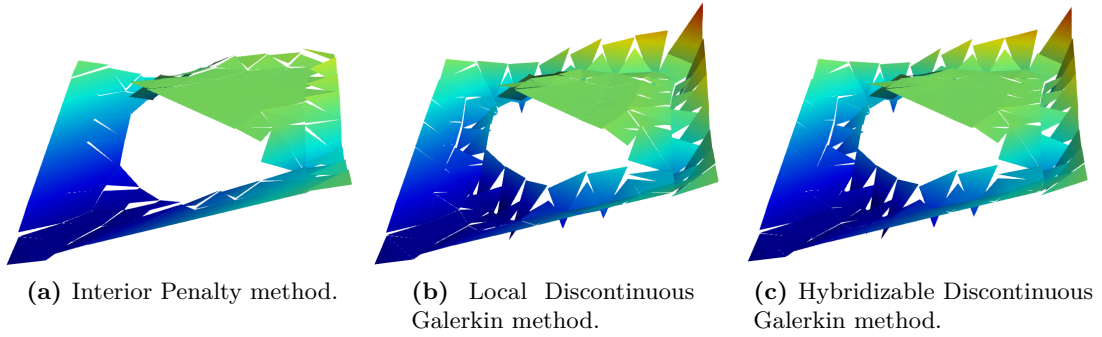
Fig. 12.19: *Discontinuous Galerkin solutions of the two-dimensional model problem.*

methods with the pair of polynomial orders $p, q = (1, 2)$. Even though no residual-based stabilization is used, these methods yield stable solutions. We observe that the IP solution manages to ignore the effect of the boundary layer quite effectively, while the LDG and HDG solutions appear to place more emphasis on the outflow boundary condition. As a result, the latter two methods show some local undershoots within the elements adjacent to the outflow boundary. We also note that the LDG and HDG solutions appear nearly identical for this particular model set-up.

Next, we investigate the impact of the fine-scale models. For each discontinuous Galerkin method we compute the solutions after addition of the volumetric part of the residual-based model, and once more when we add the complete augmented residual-based model. The one-dimensional experiments have shown that both the inflow and the outflow part of the augmented model have contributions, and their relative importance depends on the problem at hand. For these two-dimensional simulations we thus no longer make the distinction between inflow and outflow modeling terms. We show results for the fine-scale solution fields of the primary variable, for the polynomial orders $p, q = (1, 2)$ and $p, q = (3, 4)$.

Figures 12.20 to 12.22 show the obtained fine-scale solutions for the Interior Penalty method. In Figures 12.20 and 12.21, the fine-scale solutions are shown in a two-dimensional plot, and in Figure 12.22 they are shown along a cut-plane. According to these results, the plain Interior Penalty method shows high quality solutions for both polynomial orders. For the low order computation, adding the classical fine-scale model has virtually no impact. Addition of the augmented model pushes the solution error a bit further towards the boundary. In all cases, the error is constrained to a single row of elements. This is also the case in Figure 12.21, where higher-order basis functions are used. Addition of the classical VMS model results in a thickening of the error, which is countered by the additional terms in the augmented model.
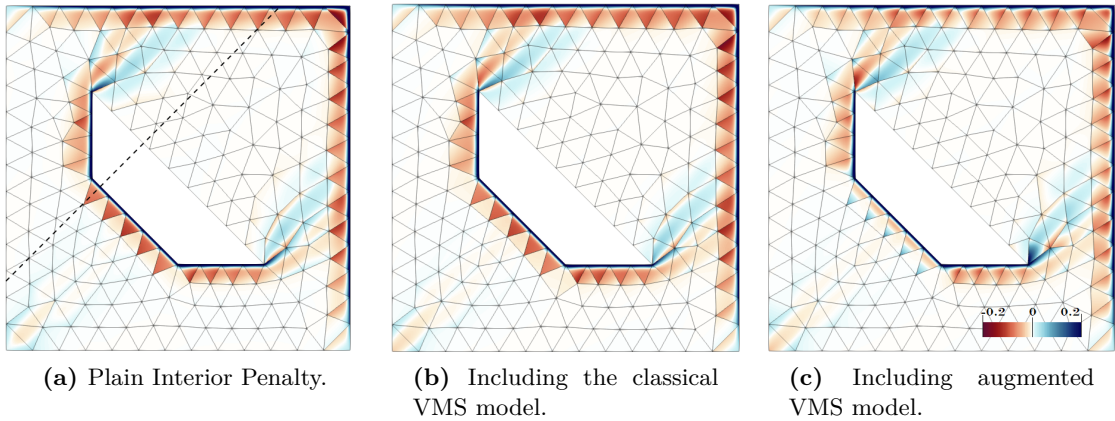
(a) Plain Interior Penalty.

(b) Including the classical VMS model.

(c) Including augmented VMS model.

Fig. 12.20: *Fine-scale solutions $\phi - \phi^h$ (errors) for linear basis functions with the Interior Penalty method and $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.003$.*
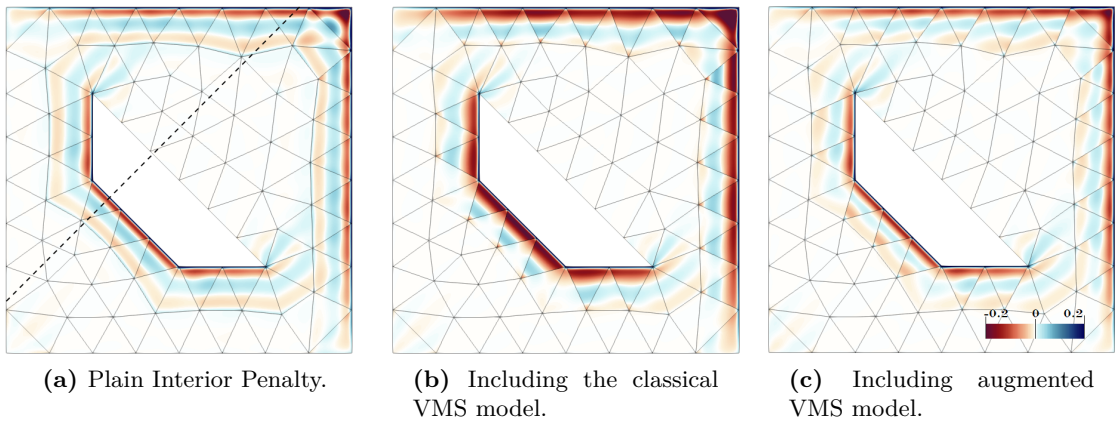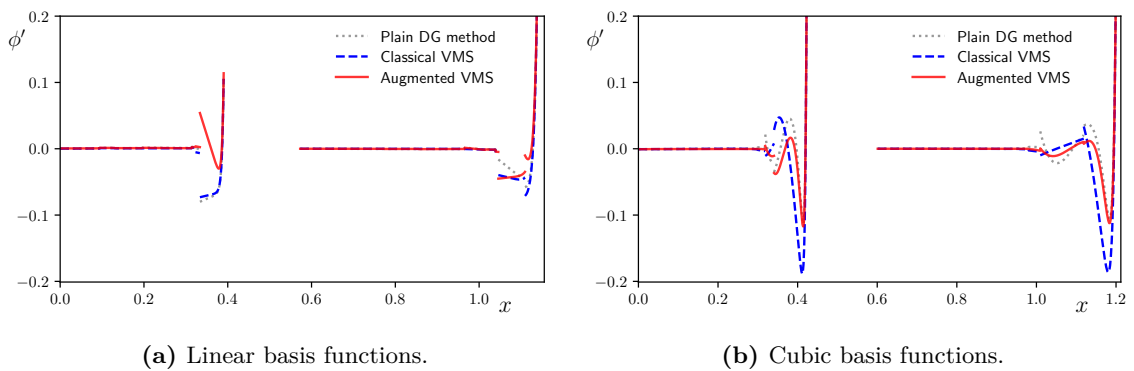

(a) Plain Interior Penalty.

(b) Including the classical VMS model.

(c) Including augmented VMS model.

Fig. 12.21: *Fine-scale solutions $\phi - \phi^h$ (errors) for cubic basis functions with the Interior Penalty method and $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.003$.*

(a) Linear basis functions.

(b) Cubic basis functions.

Fig. 12.22: *Error solutions along the cut-planes illustrated in Figures 12.20a and 12.21a.*

165

(a) Plain Interior Penalty.

(b) Including the classical VMS model.
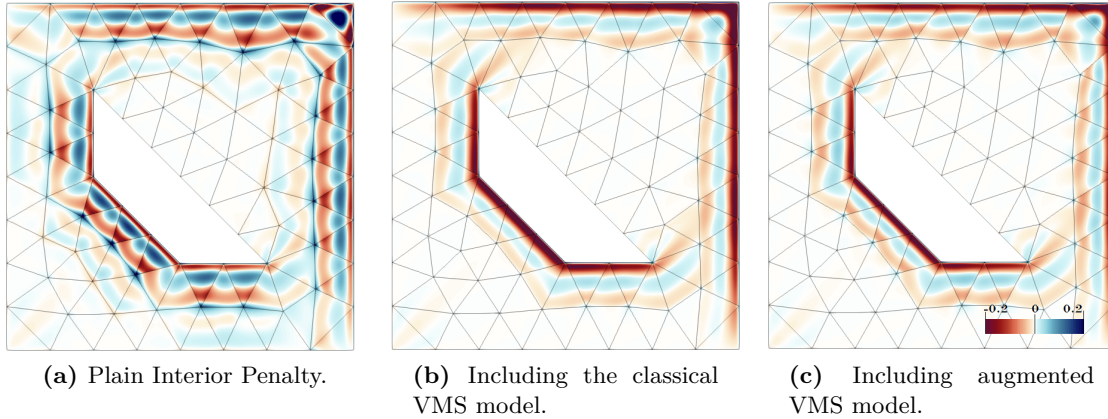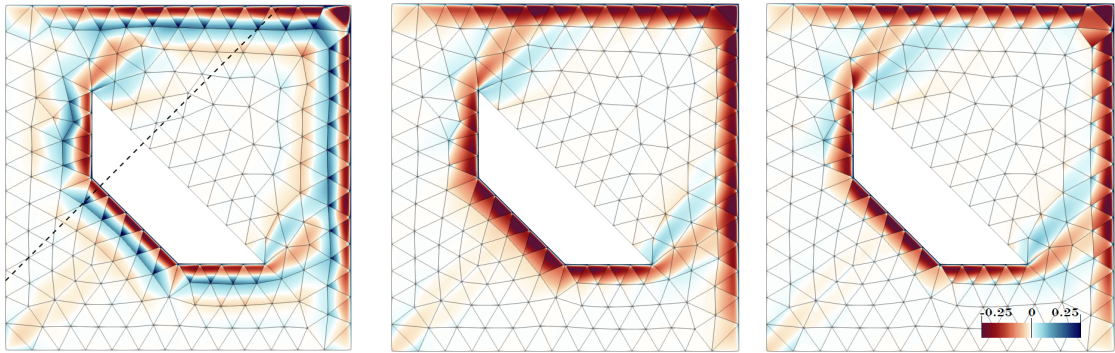
(c) Including augmented VMS model.

Fig. 12.23: *Fine-scale solutions $\phi - \phi^h$ (errors) for cubic basis functions with the Interior Penalty method with a three times higher penalty parameter of $\eta = 6p^2/h$.*

We wish to emphasize that the good solution quality in Figure 12.21a of the plain Interior Penalty method is a result of the choice of penalty parameter. If we increase η by a factor of three to $6p^2/h$, then we obtain the results shown in Figure 12.23. The solution fields shows strong oscillations that propagate through multiple rows of elements. This is largely suppressed by adding the volumetric part of the fine-scale model. The error is once again almost exclusively contained in the elements that include the boundary layer. Adding the remaining augmented terms of the model appears to have little impact. A closer look reveals that the oscillations in the fine-scale solution are compressed towards the boundary. The same thing was observed in Section 7.2.2, where we showed that this finite element solution, in fact, bears a closer resemblance to the 'true' coarse-scale solution.

Figures 12.24 to 12.26 show the results for the Local Discontinuous Galerkin method. Different from the Interior Penalty method, the obtained solution for the low order discretization shows significant defects. We observe strong overshoots and strong under-shoots. The errors seeps into the domain a few element rows. By adding the residual based model, the undershoots are almost completely canceled, at the cost of an excess error. This reminds strongly of the 'overly diffused boundary layer' interpretation proposed in Chapter 7. Indeed, this excess diffusion is countered by adding the augmented terms in the residual-based model, as shown in Figure 12.24c. The results with the cubic basis functions are very similar to those obtained for the Interior Penalty method of Figure 12.23, and the conclusions are the same.

Finally, Figs. 12.24 to 12.26 show the results for the Hybridizable Discontinuous Galerkin method. These results are almost completely identical to the ones of the Local Discontinuous Galerkin method; apparently the added penalty term has little impact on the solution. The same comment was made in relation to Figure 12.19.
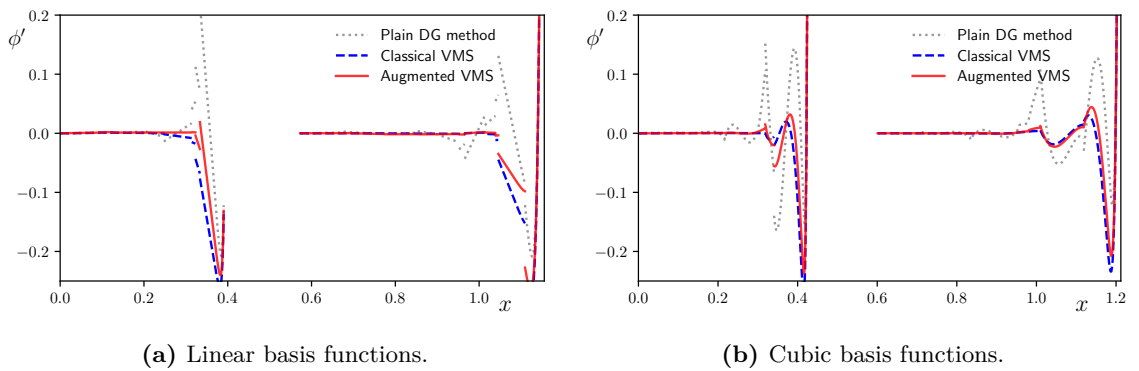
(a) Plain Local Discontinuous Galerkin.

(b) Including the classical VMS model.

(c) Including augmented VMS model.

Fig. 12.24: *Fine-scale solutions $\phi - \phi^h$ (errors) for linear basis functions with the Local Discontinuous Galerkin method and $\|\mathbf{a}\| = 0.8$ and $\kappa = 0.003$.*
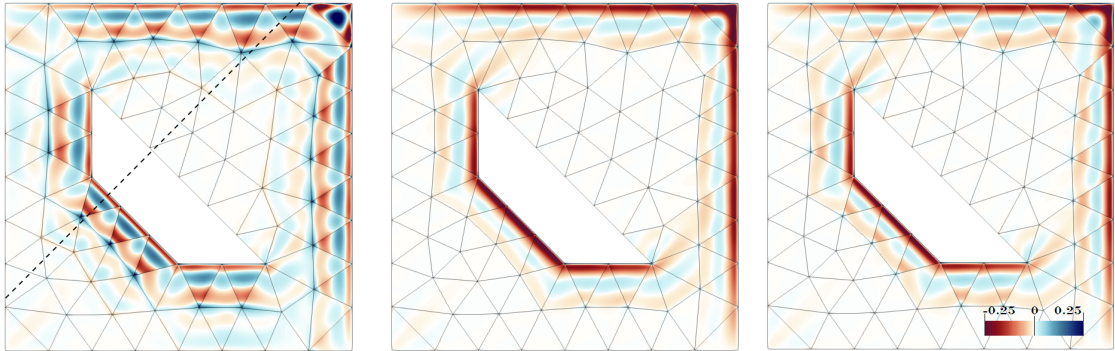
(a) Plain Local Discontinuous Galerkin.

(b) Including the classical VMS model.

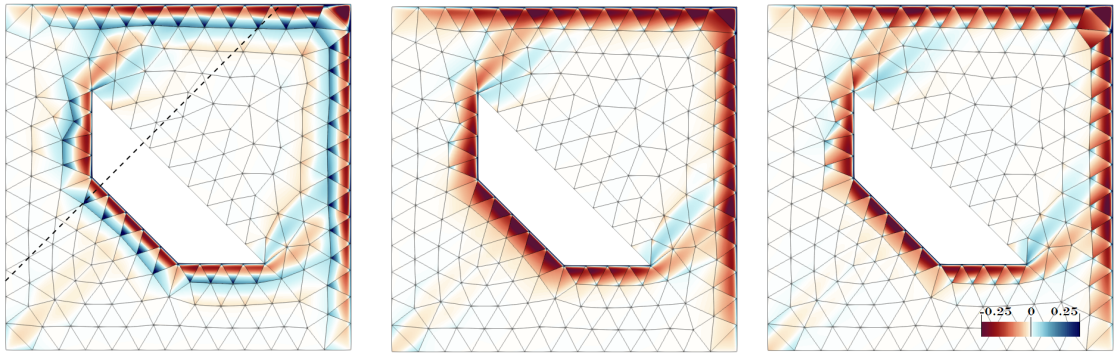**(c)** Including augmented VMS model.

Fig. 12.25: *Fine-scale solutions $\phi - \phi^h$ (errors) for cubic basis functions with the Local Discontinuous Galerkin method and $\|\mathbf{a}\| = 0.8$ and $\kappa = 0.003$.*
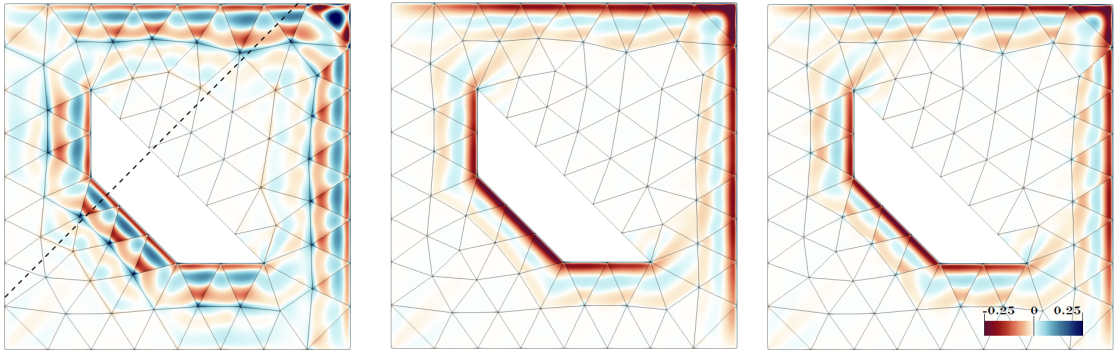

(a) Linear basis functions.

(b) Cubic basis functions.

Fig. 12.26: *Error solutions along the cut-planes illustrated in Figures 12.24a and 12.25a.*

(a) Plain Hybridizable Discontinuous Galerkin.

(b) Including the classical VMS model.

(c) Including augmented VMS model.

Fig. 12.27: *Fine-scale solutions $\phi - \phi^h$ (errors) for linear basis functions with the Hybridizable Discontinuous Galerkin method and $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.003$.*

(a) Plain Hybridizable Discontinuous Galerkin.

(b) Including the classical VMS model.

(c) Including augmented VMS model.

Fig. 12.28: *Fine-scale solutions $\phi - \phi^h$ (errors) for cubic basis functions with the Hybridizable Discontinuous Galerkin method and $\|\boldsymbol{a}\| = 0.8$ and $\kappa = 0.003$.*

(a) Linear basis functions.

(b) Cubic basis functions.

Fig. 12.29: *Error solutions along the cut-planes illustrated in Figures 12.27a and 12.28a.*
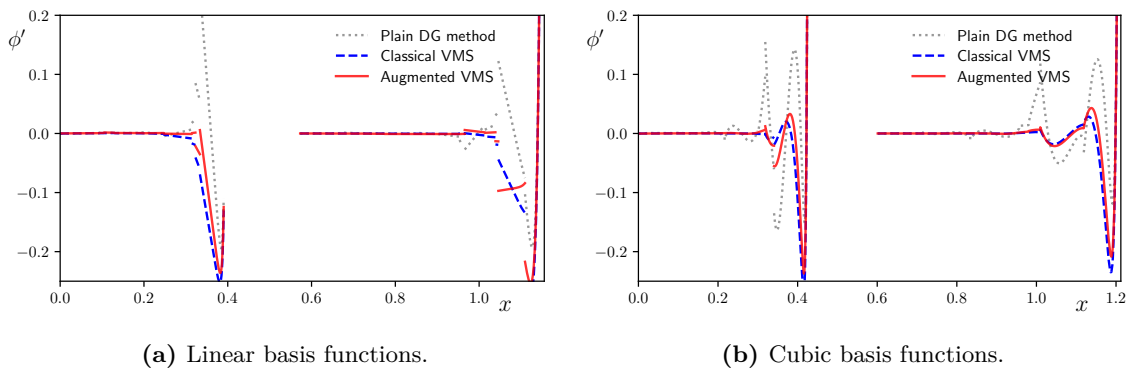
Conclusion

The overarching theme of this part of the dissertation has been the extension of the variational multiscale concept to formulations with discontinuous basis functions. As illustrated in Chapter 8, discontinuous Galerkin formulations are normally derived by choosing numerical fluxes. These numerical fluxes are interpreted as "approximations to the exact solution on the boundary of the element", as paraphrased from [9]. Stable methods rely on the introduction of stabilization (i.e., coercive) terms in these fluxes. From the VMS perspective, different discontinuous Galerkin methods originate from different projection operators. We have shown that these projectors represent optimality conditions of error potentials. The stability of the discontinuous Galerkin method then depends on additional elliptic terms in the error potential, which increases its convexity/concavity. In the case of a pure Poisson problem, these two perspectives are equivalent. For non-elliptic PDEs, however, this perspective provides a means for tracking the effects of the fine scales on the coarse-scale approximation. We then know exactly what fine-scale effects have to be included to arrive at the optimal solution in the discontinuous Galerkin approximation space.

In relevant cases, the exact expressions for these fine-scale effects are not known and approximations need to be introduced. We have extended the classical residual-based fine-scale model described in Section 3.3 to take into account the non-vanishing inflow and outflow fine-scale boundary values. In Part II, we developed this model for elements that neighbor Dirichlet boundaries where the relation between coarse scales and fine scales is known explicitly. In the current context we derived these relations on interior element boundaries from the optimality conditions implied by the the projection operators. We also developed new parameter estimates associated to the inflow of the fine-scale solution.

We explored the impact of the fine-scale model and its various components with numerical experiments. The discontinuous Galerkin methods that we considered are Bassi-Rebay's 1st method, the Interior Penalty method, the Local Discontinuous Galerkin method and the Hybridizable Discontinuous Galerkin method. We first performed a number of one-dimensional experiments. By introducing the new residual-based variational multiscale

model, we retrieved the *exact* coarse-scale solutions associated to the projection operator implied by each of these methods. We observed that the outflow part of the model is needed to ensure that the boundary layer does not cause unstable results in the domain interior, while the inflow part is important for ensuring that solution fluctuations near the domain inflow do not cause a lasting over/under prediction in the remaining domain.

The two-dimensional experiments further corroborated the importance of fine-scale modeling in discontinuous Galerkin frameworks. When we used linear basis functions, the plain discontinuous Galerkin methods produced oscillations towards the outflow boundary. The classical variational multiscale model mitigated this issue but produced an 'overly diffused' thick layer of error around that boundary. The augmented modeling terms balanced that effect. For higher-order polynomials, the instabilities became more significant. The classical model suppressed these, such that the error was contained in a single row of elements. The augmented terms changed the profile of the coarse-scale solution in those elements. We made similar observations in the numerical experiments of Part II, where we saw that this altered profile more accurately resembles the 'true' coarse-scale solution.

While we believe that this part of the dissertation has made a big step towards unification of the variational multiscale paradigm and discontinuous Galerkin methods, we do not claim to be exhaustive. Many important aspects of the discontinuous Galerkin framework are still left to be addressed. Examples are shock capturing mechanisms and the numerical flux functions for more complex hyperbolic PDEs. We provide some initial thoughts regarding these points in the conclusion of this thesis.

CHAPTER 13

Thesis conclusion and outlook

In this dissertation, we have developed a formalism by which we can use the variational multiscale method for finite element approximation spaces with low-order regularity. We have shown that this allows us to retrieve various existing finite element formulations as particular choices of scale decomposition. These include Nitsche's method for weak enforcement of essential boundary conditions, Bassi-Rebay's 1$^{\text{st}}$ method, the Interior Penalty method, the Local Discontinuous Galerkin method and the Hybridizable Discontinuous Galerkin method. Despite the occurrences of penalties and their non-conformal origin, these methods thus assume a natural place in the variational multiscale framework.

13.1 Synopsis

The first step in our approach is to reformulate the continuous variational formulation such that the discontinuous Galerkin approximations spaces *are* conformal. We release the continuity requirements on the relevant spaces, and add Lagrange multipliers to re-enforce them. This results in a well-posed formulation where the variational multiscale decomposition may lead to coarse-scale equations that resemble discontinuous Galerkin formulations. Of course, these still involve occurrences of the fine scales as well as the Lagrange multiplier. Many of the remaining fine scales may be eliminated by a suitable choice of projector. Our choices of projectors are inspired by the optimality conditions implied by the corresponding purely elliptic problem. Stable formulations require the addition of weighting on the solution jump in the error potential. Different choices of

weightings yield different finite element formulations. Finally, the Lagrange multiplier is eliminated by substituting its known relation with the coarse-scale and fine-scale solutions.

This methodology leaves us with a coarse-scale set of equations where the *exact* impact of the fine-scale solution is known. To obtain finite element solutions that closely resemble the 'true' coarse-scale solution, these remaining fine scales need to be taken into account. In this direction we have shown that much of the existing literature applies despite the change of projection operator. We solve the fine-scale solution in terms of the coarse-scale solution and the Lagrange multiplier by means of a fine-scale Green's function associated to the relevant projection operator. Then, we decompose that fine-scale Green's function into the classical H_0^1-projector related fine-scale Green's function and a remainder. Manipulation of the resulting expression eliminates the Lagrange multiplier, while adding element boundary terms that involve the fine-scale solution. Based on these findings, we devise a new residual-based fine-scale model that includes the effect of these fine-scale boundary values. We propose parameter estimations based on the fine-scale Green's function and develop an approximation strategy whereby the fine-scale model is also effective for higher-order discretization.

13.2 Recommendations

While the formalism developed in this dissertation is a step towards unification of the variational multiscale paradigm and discontinuous Galerkin type methods, there are still many concepts native to the discontinuous Galerkin family of formulations that we have not addressed. Below, we discuss a few points and provide some initial thoughts.

↪ In Chapter 8, we referenced a number of flux formulations for the hyperbolic problem and the elliptic problem. For the advection-diffusion model problem, all hyperbolic flux functions simplify to the upwind flux, as pointed out in Remark 8.1. In the variational multiscale framework, this upwind flux can be incorporated as an additional weighting on interface jumps in the error potential. A similar approach may be possible for these fluxes for other types of problems. Possibly, this introduces additional or different fine-scale terms on element boundaries. The model of Section 11.1 may be tailored to treat those terms.

↪ The fluxes for the elliptic term that we did not consider are either inconsistent or non-symmetric. We analyzed some non-symmetric formulations in earlier work [176]. For those methods, the projector does not correspond to an optimality condition of a potential, but we can still find associated fine-scale constraints.

↪ The mixed nature of the discontinuous Galerkin formulations for the elliptic problem introduces many excess degrees of freedom. These may be eliminated by solving a localized problem that relates the auxiliary variable to the primal variable. Since our proposed fine-scale model only impacts the equations corresponding to the primal test function, we anticipate that this strategy is still viable in the current framework.

↪ If the auxiliary field *is* computed, then the additional information captured by that approximation may in some cases be used to post-process the primal variable to achieve a superconvergent solution field [47, 49, 51]. This approach relies on superconvergent properties of the coarse scales: the auxiliary field better approximates the true flux than the flux of the discrete primal solution, and the average of the primal variable converges faster than its L^2-norm. This superconverging average relates closely to the optimality conditions implied by the projectors in Table 10.1. As the objective of the variational multiscale method is to find precisely that optimal solution, the superconvergence post-processing procedure might again become valid, or become much more effective, for non-elliptic problems after incorporating the appropriate fine-scale model.

↪ Discontinuous Galerkin methods become competitive in terms of computational time only when we make use of their parallel implementation capabilities and if efficient implementation strategies are employed [108, 133]. One approach is the hybridization of the scheme, and subsequent static condensation of the majority of the degrees of freedom [44, 48, 92]. As the name suggests, the Hybridizable Discontinuous Galerkin method may be manipulated as such [48]. Suitable formulations have been derived for various PDEs, such as the advection-diffusion equation [149], the Stokes equations [150], the Navier-Stokes equations [151], the equations of linear elasticity [173] and of nonlinear elasticity [130]. Of importance in these formulations is that the basis functions on the element interior only communicate with the single-valued "hybrid variable" that is defined on the element boundary. Our residual-based model would have to be reworked to satisfy this requirement. Conceivably, the difference between the element interior value and the hybrid variable may be leveraged as a fine-scale boundary value indicator, such that the model again becomes purely element local.

↪ As emphasized in Chapter 8, solving hyperbolic problems requires some slope limiting mechanism or shock capturing terms to ensure stability of the numerical scheme. In one of their earlier works, the authors that pioneered the variational multiscale framework suggested that it could motivate and guide the design of shock capturing terms [20], and recently ten Eikelder et al. have proposed a framework for doing so [185]. This framework might be merged with the formalism proposed in this dissertation.

These points, to a large extend, represent our recommendations for future research. As a final suggestion, we speculate that a highly localized fine-scale modeling procedure might be viable in discontinuous Galerkin frameworks. We have shown that the knowledge of the projector is incorporated in the discontinuous Galerkin formulation through substitution of the fine-scale constraints (this procedure is summarized in the second paragraph of this concluding chapter, and refers to Sections 3.2.3, 5.2, 9.2 and 10.2). For the remaining fine-scale occurrences we thus merely need to substitute *some* expression that *does not conflict* with the projector. It does not need to represent the unique operation corresponding to the fine-scale Green's operator. Our approach of using the H_0^1 fine-scale Green's function instead of the fine-scale Green's function corresponding to the relevant projection operator illustrates this point. This notion was addressed in Remark 6.1, and might have profound impact. Due to the potential of choosing element local test functions, we can rephrase the fine-scale equations as some element local problem. This locality necessarily makes the problem ill-posed, which means that it would permit multiple (infinitely many) solutions. However, we do not wish to compute the fine scales. Rather, we need to find a *relation* between coarse scales and fine scales that *is not wrong.* The element local (ill-posed) formulation implies infinitely many relations that *are not wrong.* This suggests an element-local procedure that allows us to retrieve a finite element solution that is arbitrarily close to the true coarse-scale solution defined by the projection operator.

List of journal publications

Below follows an overview of the first-authored and coauthored journal publications submitted and/or published by the author of this dissertation during his time at the University of Minnesota.

Under review: **S.K.F. Stoter**, M.F.P. ten Eikelder, F. de Prenter, I. Akkerman, E.H. van Brummelen, C.V. Verhoosel and D. Schillinger. A unification of variational multiscale analysis and Nitsche's method, and a resulting boundary layer fine-scale model. *Submitted to Computer Methods in Applied Mechanics and Engineering.*

Under review: Q. Zhu, **S.K.F. Stoter**, M. Heisel, C.E. French, M. Guala, L.E. Linderman and D. Schillinger: "Reducing wind-induced vibrations of road sign structures through aerodynamic modifications: a computational pilot study for a practical example. *Submitted to Journal of Wind Engineering and Industrial Aerodynamics.*

Z. Han, **S.K.F. Stoter**, C.T. Wu, C. Cheng, A. Mantzaflaris, S. Mogilevskaya and D. Schillinger. Consistent discretization of higher-order interface models for thin layers and elastic material surfaces, enabled by isogeometric cut-cell methods. *Computer Methods in Applied Mechanics and Engineering*, 350:245-267, 2019.

S.K.F. Stoter, S.R. Turteltaub, S.J. Hulshoff, and D. Schillinger. Residual-based variational multiscale modeling in a discontinuous Galerkin framework. *Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal*, 88(5):217-238, 2018.

S.K.F. Stoter, S.R. Turteltaub, S.J. Hulshoff, and D. Schillinger. A discontinuous Galerkin residual-based variational multiscale method for modeling subgrid-scale behavior of the viscous Burgers equation. *International Journal for Numerical Methods in Fluids*, 16(3):1333-1364, 2018.

L.H. Nguyen, **S.K.F. Stoter**, M. Ruess, M.A. Sanchez Uribe, D. Schillinger. The diffuse Nitsche method: Dirichlet constraints on phase-field boundaries. *International Journal for Numerical Methods in Engineering*, 113(4):601-633, 2018.

L.H. Nguyen, **S.K.F. Stoter**, T. Baum, J.S. Kirschke, M. Ruess, Z. Yosibash, D. Schillinger. Phase-field boundary conditions for the voxel finite cell method: surface-free stress analysis of CT-based bone structures. *International Journal for Numerical Methods in Biomedical Engineering*, doi:10.1002/cnm.2880, 2017.

S.K.F. Stoter, P. Müller, L. Cicalese, M. Tuveri, D. Schillinger, T.J.R. Hughes. A diffuse interface method for the Navier-Stokes/Darcy equations: Perfusion profile for a patient-specific human liver based on MRI scans. *Computer Methods in Applied Mechanics and Engineering*, 321:70-102, 2017.

D. Schillinger, I. Harari, M.-C. Hsu, D. Kamensky, **S.K.F. Stoter**, Y. Yue, Y. Zhao. The non-symmetric Nitsche method for the parameter-free imposition of weak boundary and coupling conditions in immersed finite elements. *Computer Methods in Applied Mechanics and Engineering*, 309:625-652, 2016.

References

[1] Adams, R. A. and Fournier, J. J. F. (1975). *Sobolev Spaces*. Academic Press, New York.

[2] Akkerman, I., Bazilevs, Y., Calo, V. M., Hughes, T. J. R., and Hulshoff, S. J. (2008). The role of continuity in residual–based variational multiscale modeling of turbulence. *Computational Mechanics*, 41(3):371–378.

[3] Alt, H. W. (1992). *Linear functional analysis - An application oriented introduction*. Springer, Berlin.

[4] Arnold, D., Falk, R., and Winther, R. (2010). Finite element exterior calculus: from hodge theory to numerical stability. *Bulletin of the American mathematical society*, 47(2):281–354.

[5] Arnold, D. N. (1982). An interior penalty finite element method with discontinuous elements. *SIAM journal on numerical analysis*, 19(4):742–760.

[6] Arnold, D. N., Bochev, P. B., Lehoucq, R. B., Nicolaides, R. A., and Shashkov, M. (2007). *Compatible spatial discretizations*, volume 142. Springer Science & Business Media.

[7] Arnold, D. N. and Brezzi, F. (1985). Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. *ESAIM: Mathematical Modelling and Numerical Analysis*, 19(1):7–32.

[8] Arnold, D. N., Brezzi, F., Cockburn, B., and Marini, D. (2000). Discontinuous Galerkin methods for elliptic problems. In *Discontinuous Galerkin Methods*, pages 89–101. Springer.

[9] Arnold, D. N., Brezzi, F., Cockburn, B., and Marini, L. D. (2002). Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM Journal on Numerical Analysis*, 39(5):1749–1779.

[10] Arnold, D. N., Falk, R. S., and Winther, R. (2006). Finite element exterior calculus, homological techniques, and applications. *Acta numerica*, 15:1–155.

[11] Babuška, I. (1973). The finite element method with Lagrangian multipliers. *Numerische Mathematik*, 20(3):179–192.

[12] Babuška, I. and Gatica, G. N. (2003). On the mixed finite element method with Lagrange multipliers. *Numerical Methods for Partial Differential Equations*, 19(2):192–210.

[13] Baker, G. A. (1977). Finite element methods for elliptic equations using nonconforming elements. *Mathematics of Computation*, 31(137):45–59.

[14] Bassi, F., Crivellini, A., Rebay, S., and Savini, M. (2005). Discontinuous Galerkin solution of the Reynolds–averaged Navier–Stokes and k–ω turbulence model equations. *Computers & Fluids*, 34(4-5):507–540.

[15] Bassi, F. and Rebay, S. (1997). A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier–Stokes equations. *Journal of Computational Physics*, 131(2):267 – 279.

[16] Bassi, F., Rebay, S., Mariotti, G., Pedinotti, S., and Savini, M. (1997). A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows. In Decuypere, R. and Dibelius, G., editors, *Proceedings of 2nd European Conference on Turbomachinery, Fluid Dynamics and Thermodynamicst*, pages 99–108, Antwerpen, Belgium.

[17] Baumann, C. E. and Oden, J. T. (1999). A discontinuous *hp*-finite element method for convection-diffusion problems. *Computer Methods in Applied Mechanics and Engineering*, 175(3–4):311–341.

[18] Bazilevs, Y. (2006). *Isogeometric analysis of turbulence and fluid-structure interaction*. PhD thesis, The University of Texas at Austin.

[19] Bazilevs, Y., Calo, V. M., Cottrell, J. A., Hughes, T. J. R., Reali, A., and Scovazzi, G. (2007a). Variational multiscale residual-based turbulence modeling for large eddy simulation of incompressible flows. *Computer Methods in Applied Mechanics and Engineering*, 197(1-4):173–201.

[20] Bazilevs, Y., Calo, V. M., Tezduyar, T. E., and Hughes, T. J. R. (2007b). YZβ discontinuity capturing for advection-dominated processes with application to arterial drug delivery. *International Journal for Numerical Methods in Fluids*, 54:593–608.

[21] Bazilevs, Y., Hsu, M.-C., and Scott, M. A. (2012). Isogeometric fluid–structure interaction analysis with emphasis on non-matching discretizations, and with application to wind turbines. *Computer Methods in Applied Mechanics and Engineering*, 249:28–41.

[22] Bazilevs, Y. and Hughes, T. J. R. (2007). Weak imposition of Dirichlet boundary conditions in fluid mechanics. *Computers & Fluids*, 36(1):12–26.

[23] Bazilevs, Y., Michler, C., Calo, V. M., and Hughes, T. J. R. (2007c). Weak Dirichlet boundary conditions for wall-bounded turbulent flows. *Computer Methods in Applied Mechanics and Engineering*, 196(49-52):4853–4862.

[24] Bazilevs, Y., Michler, C., Calo, V. M., and Hughes, T. J. R. (2010). Isogeometric variational multiscale modeling of wall-bounded turbulent flows with weakly enforced boundary conditions on unstretched meshes. *Computer Methods in Applied Mechanics and Engineering*, 199(13-16):780–790.

[25] Belytschko, T., Stolarski, H., and Carpenter, N. (1984). A $\mathcal{C}^0$ triangular plate element with one-point quadrature. *International Journal for Numerical Methods in Engineering*, 20(5):787–802.

[26] Boffi, D. (2006). Compatible discretizations for eigenvalue problems. In Arnold, D. N., Bochev, P. B., Lehoucq, R. B., Nicolaides, R. A., and Shashkov, M., editors, *Compatible Spatial Discretizations*, pages 121–142, New York, NY. Springer New York.

[27] Boffi, D., Brezzi, F., and Fortin, M. (2013). *Mixed finite element methods and applications*, volume 44. Springer.

[28] Bramble, J. H. (2003). A proof of the inf–sup condition for the Stokes equations on Lipschitz domains. *Mathematical Models and Methods in Applied Sciences*, 13(03):361–371.

[29] Brezzi, F. (1974). On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *Publications mathématiques et informatique de Rennes*, S4:1–26.

[30] Brezzi, F., Cockburn, B., Marini, L. D., and Süli, E. (2006). Stabilization mechanisms in discontinuous Galerkin finite element methods. *Computer Methods in Applied Mechanics and Engineering*, 195(25-28):3293–3310.

[31] Brezzi, F., Douglas, J., Durán, R., and Fortin, M. (1987). Mixed finite elements for second order elliptic problems in three variables. *Numerische Mathematik*, 51(2):237–250.

[32] Brezzi, F., Douglas, J., and Marini, L. D. (1985). Two families of mixed finite elements for second order elliptic problems. *Numerische Mathematik*, 47(2):217–235.

[33] Brezzi, F. and Falk, R. S. (1991). Stability of higher-order Hood–Taylor methods. *SIAM Journal on Numerical Analysis*, 28(3):581–590.

[34] Brezzi, F. and Fortin, M. (2012). *Mixed and hybrid finite element methods*, volume 15. Springer Science & Business Media.

[35] Brezzi, F., Franca, L. P., Hughes, T. J. R., and Russo, A. (1997). $b = \int g$. *Computer Methods in Applied Mechanics and Engineering*, 145(3-4):329–339.

[36] Brezzi, F., Lipnikov, K., and Simoncini, V. (2005). A family of mimetic finite difference methods on polygonal and polyhedral meshes. *Mathematical Models and Methods in Applied Sciences*, 15(10):1533–1551.

[37] Brezzi, F., Manzini, G., Marini, D., Pietra, P., and Russo, A. (2000). Discontinuous Galerkin approximations for elliptic problems. *Numerical Methods for Partial Differential Equations*, 16(4):365–378.

[38] Brezzi, F. and Russo, A. (1994). Choosing bubbles for advection-diffusion problems. *Mathematical Models and Methods in Applied Sciences*, 4(04):571–587.

[39] Brooks, A. N. and Hughes, T. J. R. (1982). Streamline upwind/Petrov–Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier–Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 32(1-3):199–259.

[40] Burman, E. (2012). A penalty-free nonsymmetric Nitsche-type method for the weak imposition of boundary conditions. *SIAM Journal on Numerical Analysis*, 50(4):1959–1981.

[41] Calo, V. M. (2004). *Residual–based multiscale turbulence modeling: Finite volume simulations of bypass transition*. PhD thesis, Stanford University.

[42] Celiker, F. and Cockburn, B. (2007). Superconvergence of the numerical traces of discontinuous Galerkin and hybridized methods for convection–diffusion problems in one space dimension. *Mathematics of Computation*, 76(257):67–96.

[43] Ciarlet, P., Dunkl, C. F., and Sauter, S. A. (2018). A family of Crouzeix–Raviart finite elements in 3D. *Analysis and Applications*, 16(05):649–691.

[44] Ciarlet, P. G. (1978). A mixed finite element method. In Lions, J. L., Papanicolaou, G., and Rockafellar, R. T., editors, *The finite element method for elliptic problems*, chapter 7, pages 381–424. North Holland publishing company, Amsterdam.

[45] Clough, R. W. (1960). The finite element method in plane stress analysis. In *Proceedings of 2nd ASCE Conference on Electronic Computation*, Pittsburgh PA, USA.

[46] Cockburn, B. (1998). An introduction to the discontinuous Galerkin method for convection-dominated problems. In *Advanced numerical approximation of nonlinear hyperbolic equations*, pages 150–268. Springer.

[47] Cockburn, B. (2018). Discontinuous Galerkin methods for computational fluid dynamics. *Encyclopedia of Computational Mechanics Second Edition*, pages 1–63.

[48] Cockburn, B., Gopalakrishnan, J., and Lazarov, R. (2009a). Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM Journal on Numerical Analysis*, 47(2):1319–1365.

[49] Cockburn, B., Guzmán, J., and Wang, H. (2009b). Superconvergent discontinuous Galerkin methods for second-order elliptic problems. *Mathematics of Computation*, 78(265):1–24.

[50] Cockburn, B., Hou, S., and Shu, C.-W. (1990). The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws IV: The multidimensional case. *Mathematics of Computation*, 54(190):545–581.

[51] Cockburn, B., Kanschat, G., Perugia, I., and Schötzau, D. (2001). Superconvergence of the local discontinuous Galerkin method for elliptic problems on Cartesian grids. *SIAM Journal on Numerical Analysis*, 39(1):264–285.

[52] Cockburn, B., Kanschat, G., and Schötzau, D. (2005). A locally conservative LDG method for the incompressible Navier-Stokes equations. *Mathematics of Computation*, 74(251):1067–1095.

[53] Cockburn, B., Karniadakis, G. E., and Shu, C.-W. (2012). *Discontinuous Galerkin methods: theory, computation and applications*, volume 11. Springer Science & Business Media.

[54] Cockburn, B., Li, F., and Shu, C.-W. (2004). Locally divergence-free discontinuous Galerkin methods for the Maxwell equations. *Journal of Computational Physics*, 194(2):588–610.

[55] Cockburn, B., Lin, S.-Y., and Shu, C.-W. (1989). TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws III: One-dimensional systems. *Journal of Computational Physics*, 84(1):90 – 113.

[56] Cockburn, B. and Shu, C.-W. (1989). TVB Runge-Kutta Local Projection Discontinuous Galerkin Finite Element Method for Conservation Laws II: General Framework. *Mathematics of Computation*, 52(186):411–435.

[57] Cockburn, B. and Shu, C.-W. (1991). The Runge-Kutta local projection P^1-discontinuous-Galerkin finite element method for scalar conservation laws. *Modélisation mathématique et analyse numérique*, 25(3):337–361.

[58] Cockburn, B. and Shu, C.-W. (1998a). The local discontinuous Galerkin method for time–dependent convection–diffusion systems. *SIAM Journal on Numerical Analysis*, 35(6):2440–2463.

[59] Cockburn, B. and Shu, C.-W. (1998b). The Runge–Kutta discontinuous Galerkin method for conservation laws V: multidimensional systems. *Journal of Computational Physics*, 141(2):199–224.

[60] Cockburn, B. and Shu, C.-W. (2001). Runge–Kutta discontinuous Galerkin methods for convection–dominated problems. *Journal of Scientific Computing*, 16(3):173–261.

[61] Codina, R., Principe, J., and Baiges, J. (2009). Subscales on the element boundaries in the variational two–scale finite element method. *Computer Methods in Applied Mechanics and Engineering*, 198:838–852.

[62] Codina, R., Principe, J., Guasch, O., and Badia, S. (2007). Time dependent subscales in the stabilized finite element approximation of incompressible flow problems. *Computer Methods in Applied Mechanics and Engineering*, 196(21):2413–2430.

[63] Coley, C. and Evans, J. A. (2018). Variational multiscale modeling with discontinuous subscales: analysis and application to scalar transport. *Meccanica*, 53(6):1241–1269.

[64] Collis, S. S. (2001). Monitoring unresolved scales in multiscale turbulence modeling. *Physics of Fluids*, 13(6):1800–1806.

[65] Collis, S. S. (2002a). Discontinuous Galerkin methods for turbulence simulation. In *Studying Turbulence Using Numerical Simulation Databases – IX: Proceedings of the 2002 Summer Program*, Stanford, California.

[66] Collis, S. S. (2002b). The DG/VMS method for unified turbulence simulation. In *32nd AIAA Fluid Dynamics Conference and Exhibit*, Reston, Virigina.

[67] Collis, S. S. and Ramakrishnan, S. (2005). The local variational multiscale method. *Third MIT Conference on Computational Fluid and Solid Dynamics*.

[68] Courant, R. (1943). Variational methods for the solution of problems of equilibrium and vibrations. *Bulletin of the American Mathematical Society*, 49(1):1–23.

[69] Crouzeix, M. and Raviart, P.-A. (1973). Conforming and nonconforming finite element methods for solving the stationary Stokes equations. *Revue française d'automatique informatique recherche opérationnelle. Mathématique*, 7(R3):33–75.

[70] Darcy, H. (1856). *Les fontaines publiques de la ville de Dijon: exposition et application des Principes a Suivre et des Formulesa Employer dans les Questions de Distribution d'Eau.* Dalmont, E., Paris.

[71] Desbrun, M., Hirani, A. N., Leok, M., and Marsden, J. E. (2005). Discrete exterior calculus. *arXiv preprint math/0508341*.

[72] Donéa, J. and Huerta, A. (2003). *Finite element methods for flow problems.* John Wiley & Sons, Ltd, Hoboken, New Jersey.

[73] Douglas, J. and Dupont, T. (1976). Interior penalty procedures for elliptic and parabolic Galerkin methods. In *Computing methods in applied sciences*, pages 207–216. Springer.

[74] Embar, A., Dolbow, J., and Harari, I. (2010). Imposing Dirichlet boundary conditions with Nitsche's method and spline based finite elements. *International Journal for Numerical Methods in Engineering*, 83(7):877–898.

[75] Engquist, B. and Osher, S. (1981). One-sided difference approximations for nonlinear conservation laws. *Mathematics of Computation*, 36(154):321–351.

[76] Ern, A. and Guermond, J.-L. (2004). *Theory and practice of finite elements*, volume 159. Springer, Berlin.

[77] Evans, J. A. and Hughes, T. J. R. (2013a). Isogeometric divergence-conforming B-splines for the Darcy–Stokes–Brinkman equations. *Mathematical Models and Methods in Applied Sciences*, 23(04):671–741.

[78] Evans, J. A. and Hughes, T. J. R. (2013b). Isogeometric divergence-conforming B-splines for the steady Navier–Stokes equations. *Mathematical Models and Methods in Applied Sciences*, 23(08):1421–1478.

183

[79] Evans, J. A. and Hughes, T. J. R. (2013c). Isogeometric divergence-conforming B-splines for the unsteady Navier–Stokes equations. *Journal of Computational Physics*, 241:141–167.

[80] Evans, J. A., Hughes, T. J. R., and Sangalli, G. (2009). Enforcement of constraints and maximum principles in the variational multiscale method. *Computer Methods in Applied Mechanics and Engineering*, 199(1-4):61–76.

[81] Evans, L. C. (1998). *Partial differential equations*, volume 19. Providence: American Mathematical Society.

[82] Farhat, C., Chapman, T., and Avery, P. (2015). Structure-preserving, stability, and accuracy properties of the energy-conserving sampling and weighting method for the hyper reduction of nonlinear finite element dynamic models. *International Journal for Numerical Methods in Engineering*, 102(5):1077–1110.

[83] Fortin, M. (1977). An analysis of the convergence of mixed finite element methods. *RAIRO. Analyse numérique*, 11(4):341–354.

[84] Fraeijs de Veubeke, B. M. (1965). Displacement and equilibrium models in the finite element method. In Zienkiewicz, O. C. and Holister, G. S., editors, *Stress Analysis*, chapter 9, pages 145–197. John Wiley & Sons, Ltd.

[85] Fraeijs De Veubeke, B. M. (1973). *Diffusive equilibrium models, Lecture notes.* University of Calgary, Calgary.

[86] Fraeijs De Veubeke, B. M. and Hogge, M. (1972). Dual analysis for heat conduction problems by finite elements. *International Journal for Numerical Methods in Engineering*, 5:65–82.

[87] Gamnitzer, P., Gravemeier, V., and Wall, W. A. (2010). Time-dependent subgrid scales in residual-based large eddy simulation of turbulent channel flow. *Computer Methods in Applied Mechanics and Engineering*, 199(13):819 – 827.

[88] Germano, M., Piomelli, U., Moin, P., and Cabot, W. H. (1991). A dynamic subgrid-scale eddy viscosity model. *Physics of Fluids A: Fluid Dynamics*, 3(7):1760.

[89] Godunov, S. K. (1959). A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik*, 89(3):271–306.

[90] Gottlieb, S. and Shu, C.-W. (1998). Total variation diminishing Runge-Kutta schemes. *Mathematics of computation of the American Mathematical Society*, 67(221):73–85.

[91] Gravemeier, V. and Wall, W. A. (2011). Residual-based variational multiscale methods for laminar, transitional and turbulent variable-density flow at low Mach number. *International Journal for Numerical Methods in Fluids*, 65(10):1260–1278.

[92] Guyan, R. J. (1965). Reduction of stiffness and mass matrices. *AIAA journal*, 3(2):380–380.

[93] Hansbo, P., Lovadina, C., Perugia, I., and Sangalli, G. (2005). A Lagrange multiplier method for the finite element solution of elliptic interface problems using non-matching meshes. *Numerische Mathematik*, 100(1):91–115.

[94] Harari, I. and Albocher, U. (2018). Spectral investigations of Nitsche's method. *Finite Elements in Analysis and Design*, 145:20–31.

[95] Harari, I. and Hughes, T. J. R. (1992). What are C and h?: Inequalities for the analysis and design of finite element methods. *Computer Methods in Applied Mechanics and Engineering*, 97(2):157–192.

[96] Harten, A., Lax, P. D., and van Leer, B. (1983). On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM review*, 25(1):35–61.

[97] Hauke, G., Fuster, D., and Doweidar, M. H. (2008). Variational multiscale a-posteriori error estimation for multi-dimensional transport problems. *Computer Methods in Applied Mechanics and Engineering*, 197(33-40):2701–2718.

[98] Hauke, G. and Garcá-Olivares, A. (2001). Variational subgrid scale formulations for the advection–diffusion-reaction equation. *Computer Methods in Applied Mechanics and Engineering*, 190(51-52):6847–6865.

[99] Heisenberg, W. (1946). Zur statistischen Theorie der Turbulenz. *Zeitschrift für Physik*, 124(7–12):628–657.

[100] Hesthaven, J. S. and Warburton, T. (2007). *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*. Springer Science & Business Media.

[101] Hiptmair, R., Li, L., Mao, S., and Zheng, W. (2018). A fully divergence-free finite element method for magnetohydrodynamic equations. *Mathematical Models and Methods in Applied Sciences*, 28(04):659–695.

[102] Hirani, A. N. (2003). *Discrete exterior calculus*. PhD thesis, California Institute of Technology.

185

[103] Hoang, T., Verhoosel, C. V., Qin, C.-Z., Auricchio, F., Reali, A., and van Brummelen, E. H. (2019). Skeleton-stabilized immersogeometric analysis for incompressible viscous flow problems. *Computer Methods in Applied Mechanics and Engineering*, 344:421–450.

[104] Holmen, J., Hughes, T. J. R., Oberai, A. A., and Wells, G. N. (2004). Sensitivity of the scale partition for variational multiscale large-eddy simulation of channel flow. *Physics of Fluids*, 16(3):824–827.

[105] Hrennikoff, A. (1941). Solution of problems of elasticity by the framework method. *Journal of Applied Mechanics*, 8:A169–175.

[106] Hsu, M.-C., Kamensky, D., Bazilevs, Y., Sacks, M. S., and Hughes, T. J. R. (2014). Fluid–structure interaction analysis of bioprosthetic heart valves: significance of arterial wall deformation. *Computational mechanics*, 54(4):1055–1071.

[107] Hsu, M.-C., Wang, C., Xu, F., Herrema, A. J., and Krishnamurthy, A. (2016). Direct immersogeometric fluid flow analysis using B-rep CAD models. *Computer Aided Geometric Design*, 43:143–158.

[108] Huerta, A., Angeloski, A., Roca, X., and Peraire, J. (2013). Efficiency of high-order elements for continuous and discontinuous Galerkin methods. *International Journal for Numerical Methods in Engineering*, 96(9):529–560.

[109] Hughes, T. J. R. (1995). Multiscale phenomena: Green's functions, the Dirichlet–to–Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Computer Methods in Applied Mechanics and Engineering*, 127(1-4):387–401.

[110] Hughes, T. J. R. (2000). *The finite element method: linear static and dynamic finite element analysis.* Courier Corporation.

[111] Hughes, T. J. R. and Allik, H. (1969). Finite elements for compressible and incompressible continua. In *Proceedings of the Symposium on Civil Engineering*, pages 27–62. Vanderbilt University Nashville, TN.

[112] Hughes, T. J. R., Calo, V. M., and Scovazzi, G. (2005). Variational and multiscale methods in turbulence. *Mechanics of the 21st Century*, pages 153–163.

[113] Hughes, T. J. R., Engel, G., Mazzei, L., and Larson, M. G. (2000a). A comparison of discontinuous and continuous Galerkin methods based on error estimates, conservation, robustness and efficiency. In *Discontinuous Galerkin Methods*, pages 135–146. Springer.

[114] Hughes, T. J. R., Engel, G., Mazzei, L., and Larson, M. G. (2000b). The continuous Galerkin method is locally conservative. *Journal of Computational Physics*, 163(2):467–488.

[115] Hughes, T. J. R., Feijóo, G. R., Mazzei, L., and Quincy, J.-B. (1998). The variational multiscale method – a paradigm for computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 166(1-2):3–24.

[116] Hughes, T. J. R., Franca, L. P., and Balestra, M. (1986). A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuška-Brezzi condition: A stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolations. *Computer Methods in Applied Mechanics and Engineering*, 59(1):85–99.

[117] Hughes, T. J. R., Franca, L. P., and Hulbert, G. M. (1989). A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/least–squares method for advective–diffusive equations. *Computer Methods in Applied Mechanics and Engineering*, 73(2):173–189.

[118] Hughes, T. J. R. and Mallet, M. (1986). A new finite element formulation for computational fluid dynamics: III. The generalized streamline operator for multidimensional advective-diffusive systems. *Computer Methods in Applied Mechanics and Engineering*, 58(3):305–328.

[119] Hughes, T. J. R., Mazzei, L., and Jansen, K. E. (2000c). Large eddy simulation and the variational multiscale method. *Computing and Visualization in Science*, 3(1-2):47–59.

[120] Hughes, T. J. R., Mazzei, L., Oberai, A. A., and Wray, A. A. (2001a). The multiscale formulation of large eddy simulation: Decay of homogeneous isotropic turbulence. *Physics of Fluids*, 13(2):505–512.

[121] Hughes, T. J. R., Oberai, A. A., and Mazzei, L. (2001b). Large eddy simulation of turbulent channel flows by the variational multiscale method. *Physics of Fluids*, 13(6):1784–1799.

[122] Hughes, T. J. R. and Sangalli, G. (2007). Variational multiscale analysis: The fine-scale Green's function, projection, optimization, localization, and stabilized methods. *Society for Industrial and Applied Mathematics*, 45(2):539–557.

[123] Hughes, T. J. R., Scovazzi, G., Bochev, P. B., and Buffa, A. (2006). A multiscale discontinuous Galerkin method with the computational structure of a continuous Galerkin method. *Computer Methods in Applied Mechanics and Engineering*, 195(19):2761–2787.

[124] Hughes, T. J. R., Scovazzi, G., and Franca, L. P. (2004a). Multiscale and stabilized methods. In Stein, E., De Borst, R., and Hughes, T. J. R., editors, *Encyclopedia of computational mechanics*, chapter 4. John Wiley & Sons, Ltd.

[125] Hughes, T. J. R. and Stewart, J. R. (1996). A space–time formulation for multiscale phenomena. *Journal of Computational and Applied Mathematics*, 74(1-2):217–229.

[126] Hughes, T. J. R., Wells, G. N., and Wray, A. A. (2004b). Energy transfers and spectral eddy viscosity in large-eddy simulations of homogeneous isotropic turbulence: Comparison of dynamic Smagorinsky and multiscale models over a range of discretizations. *Citation: Physics of Fluids*, 16(11):4044–4052.

[127] Hulshoff, S. J. (2004). Implicit subgrid-scale models in space-time variational-multiscale discretizations. *International Journal for Numerical Methods in Fluids*, 47(10-11):1093–1099.

[128] Jagalur Mohan, J. J., Sahni, O., Doostan, A., and Oberai, A. A. (2014). Variational multiscale analysis: the fine-scale Green's function for stochastic partial differential equations. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):397–422.

[129] Jansen, K. E., Collis, S. S., Whiting, C., and Shaki, F. (1999). A better consistency for low-order stabilized finite element methods. *Computer Methods in Applied Mechanics and Engineering*, 174(1-2):153–170.

[130] Kabaria, H., Lew, A. J., and Cockburn, B. (2015). A hybridizable discontinuous Galerkin formulation for non-linear elasticity. *Computer Methods in Applied Mechanics and Engineering*, 283:303–329.

[131] Kamensky, D., Hsu, M.-C., Schillinger, D., Evans, J. A., Aggarwal, A., Bazilevs, Y., Sacks, M. S., and Hughes, T. J. R. (2015). An immersogeometric variational framework for fluid–structure interaction: Application to bioprosthetic heart valves. *Computer methods in applied mechanics and engineering*, 284:1005–1053.

[132] Kanschat, G. and Riviere, B. (2010). A strongly conservative finite element method for the coupling of Stokes and Darcy flow. *Journal of Computational Physics*, 229(17):5933–5943.

[133] Kirby, R. M., Sherwin, S. J., and Cockburn, B. (2012). To CG or to HDG: A comparative study. *Journal of Scientific Computing*, 51(1):183–212.

[134] Kolmogorov, A. N. (1941a). Dissipation of energy in the locally isotropic turbulence. *Proceedings: Mathematical and Physical Sciences*, 434(1890):15–17.

[135] Kolmogorov, A. N. (1941b). Local structure of turbulence in an incompressible fluid at very high Reynolds number. *Doklady Akademiia Nauk SSSR*, 30:301–305.

[136] Landmann, B., Kessler, M., Wagner, S., and Krämer, E. (2008). A parallel, high-order discontinuous Galerkin code for laminar and turbulent flows. *Computers & Fluids*, 37(4):427–438.

[137] Lilly, D. K. (1967). Representation of small scale turbulence in numerical simulation experiments. In *IBM Scientific Computing Symposium on Environmental Sciences*, White Plains, New York.

[138] Lilly, D. K. (1992). A proposed modification of the Germano subgrid-scale closure method. *Physics of Fluids A: Fluid Dynamics*, 4(3):633–635.

[139] Loula, A. F. D., Rochinha, F. A., and Murad, M. A. (1995). Higher-order gradient post-processings for second-order elliptic problems. *Computer Methods in Applied Mechanics and Engineering*, 128(3-4):361–381.

[140] MacNeal, R. H. (1970). *The NASTRAN theoretical manual*, volume 221. Scientific and Technical Information Office, National Aeronautics and Space Administration.

[141] Masud, A. and Hughes, T. J. R. (2002). A stabilized mixed finite element method for Darcy flow. *Computer Methods in Applied Mechanics and Engineering*, 191(39-40):4341–4370.

[142] Masud, A. and Khurram, R. A. (2006). A multiscale finite element method for the incompressible Navier–Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 195(13-16):1750–1777.

[143] Moro, D., Nguyen, N. C., and Peraire, J. (2011). Navier-Stokes solution using hybridizable discontinuous Galerkin methods. In *20th AIAA computational fluid dynamics conference*, page 3407.

[144] Munts, E. A., Hulshoff, S. J., and De Borst, R. (2004). A space-time variational multiscale discretization for LES. In *34th AIAA Aerospace Sciences Meeting and Exhibit*, Portland, Oregon.

[145] Nguyen, L. H. and Schillinger, D. (2018). A multiscale predictor/corrector scheme for efficient elastoplastic voxel finite element analysis, with application to CT-based bone strength prediction. *Computer Methods in Applied Mechanics and Engineering*, 330:598–628.

[146] Nguyen, L. H. and Schillinger, D. (2019a). The multiscale finite element method for nonlinear continuum localization problems at full fine-scale fidelity, illustrated through phase-field fracture and plasticity. *Journal of Computational Physics*, 396:129–160.

[147] Nguyen, L. H. and Schillinger, D. (2019b). A residual-driven local iterative corrector scheme for the multiscale finite element method. *Journal of Computational Physics*, 377:60–88.

[148] Nguyen, N., Persson, P.-O., and Peraire, J. (2007). RANS solutions using high order discontinuous Galerkin methods. In *45th AIAA Aerospace Sciences Meeting and Exhibit*, page 914.

[149] Nguyen, N. C., Peraire, J., and Cockburn, B. (2009). An implicit high-order hybridizable discontinuous Galerkin method for linear convection–diffusion equations. *Journal of Computational Physics*, 228(9):3232–3254.

[150] Nguyen, N. C., Peraire, J., and Cockburn, B. (2010). A hybridizable discontinuous Galerkin method for Stokes flow. *Computer Methods in Applied Mechanics and Engineering*, 199(9):582–597.

[151] Nguyen, N. C., Peraire, J., and Cockburn, B. (2011). An implicit high-order hybridizable discontinuous Galerkin method for the incompressible Navier-Stokes equations. *Journal of Computational Physics*, 230(4):1147–1170.

[152] Nitsche, J. (1971). Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. In *Abhandlungen aus dem mathematischen Seminar der Universität Hamburg*, volume 36, pages 9–15. Springer.

[153] Peraire, J. and Persson, P.-O. (2007). The compact discontinuous Galerkin (CDG) method for elliptic problems. *SIAM Journal on Scientific Computing*, 30(4):1806–1824.

[154] Peraire, J. and Persson, P.-O. (2011). High-Order Discontinuous Galerkin Methods for CFD. In Wang, Z. J., editor, *Adaptive high-order methods in computational fluid dynamics*, volume 2, chapter 5.

190

[155] Pope, S. B. (2004). Ten questions concerning the large-eddy simulation of turbulent flows. *New Journal of Physics*, 6(6).

[156] Principe, J., Codina, R., and Henke, F. (2010). The dissipative structure of variational multiscale methods for incompressible flows. *Computer Methods in Applied Mechanics and Engineering*, 199(13):791 – 801.

[157] Ramakrishnan, S. and Collis, S. S. (2004). Multiscale modeling for turbulence simulation in complex geometries. In *42nd AIAA Aerospace Sciences Meeting and Exhibit*, page 241.

[158] Raviart, P. A. and Thomas, J. M. (1977). A mixed finite element method for second order elliptic problems. In *Proceedings of Symposia in Mathematical Aspects of the Finite Element Method (Rome, 1975)*, volume 606 of *Lecture Notes in Math*, pages 292–315. Springer-Verlag, Berlin, Germany.

[159] Reddy, B. D. (2013). *Introductory functional analysis: with applications to boundary value problems and finite elements*, volume 27. Springer.

[160] Reed, W. H. and Hill, T. R. (1973). Triangular mesh methods for the neutron transport equation. *Proceedings of the American Nuclear Society*, 836:1–23.

[161] Rivière, B., Wheeler, M. F., and Girault, V. (1999). Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. Part I. *Computational Geosciences*, 3(3):337–360.

[162] Roe, P. L. (1981). Approximate Riemann solvers, parameter vectors, and difference schemes. *Journal of computational physics*, 43(2):357–372.

[163] Ruess, M., Schillinger, D., Oezcan, A. I., and Rank, E. (2014). Weak coupling for isogeometric analysis of non-matching and trimmed multi-patch geometries. *Computer Methods in Applied Mechanics and Engineering*, 269:46–71.

[164] Sagaut, P. (2006). *Large-eddy simulation for incompressible flows: an introduction.* Scientific Computation. Springer, Berlin, Heidelberg.

[165] Schillinger, D., Dede, L., Scott, M. A., Evans, J. A., Borden, M. J., Rank, E., and Hughes, T. J. R. (2012). An isogeometric design-through-analysis methodology based on adaptive hierarchical refinement of NURBS, immersed boundary methods, and T-spline CAD surfaces. *Computer Methods in Applied Mechanics and Engineering*, 249:116–150.

[166] Schillinger, D., Gangwar, T., Gilmanov, A., Heuschele, J. D., and Stolarski, H. K. (2018). Embedded shell finite elements: Solid–shell interaction, surface locking, and application to image-based bio-structures. *Computer Methods in Applied Mechanics and Engineering*, 335:298–326.

[167] Schillinger, D., Harari, I., Hsu, M.-C., Kamensky, D., Stoter, S. K. F., Yu, Y., and Zhao, Y. (2016). The non-symmetric Nitsche method for the parameter-free imposition of weak boundary and coupling conditions in immersed finite elements. *Computer Methods in Applied Mechanics and Engineering*, 309:625–652.

[168] Schillinger, D. and Ruess, M. (2015). The Finite Cell Method: A review in the context of higher-order structural analysis of CAD and image-based geometric models. *Archives of Computational Methods in Engineering*, 22(3):391–455.

[169] Shahbazi, K. (2005). An explicit expression for the penalty parameter of the interior penalty method. *Journal of Computational Physics*, 205(2):401–407.

[170] Shakib, F., Hughes, T. J. R., and Johan, Z. (1991). A new finite element formulation for computational fluid dynamics: X. The compressible Euler and Navier-Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 89(1-3):141–219.

[171] Shu, C.-W. (1988). Total-variation-diminishing time discretizations. *SIAM Journal on Scientific and Statistical Computing*, 9(6):1073–1084.

[172] Smagorinsky, J. (1963). General circulation experiments with the primitive equations. *Monthly Weather Review*, 91(3):99–164.

[173] Soon, S.-C., Cockburn, B., and Stolarski, H. K. (2009). A hybridizable discontinuous Galerkin method for linear elasticity. *International journal for numerical methods in engineering*, 80(8):1058–1092.

[174] Stolarski, H. and Belytschko, T. (1982). Membrane locking and reduced integration for curved elements. *Journal of Applied Mechanics*, 49(1):172–176.

[175] Stolarski, H., Belytschko, T., and Lee, S. H. (1995). A review of shell finite elements and corotational theories. *Computational mechanics advances*, 2(2):125–212.

[176] Stoter, S. K. F. (2017). *Residual-based variational multiscale modeling in a discontinuous Galerkin framework.* Master's thesis, Delft University of Technology.

[177] Stoter, S. K. F., ten Eikelder, M. F. P., de Prenter, F., Akkerman, I., van Brummelen, E. H., Verhoosel, C. V., and Schillinger, D. A unification of variational multiscale analysis and Nitsche's method, and a resulting boundary layer fine-scale model. *Submitted to: Computer Methods in Applied Mechanics and Engineering.*

[178] Stoter, S. K. F., Turteltaub, S. R., Hulshoff, S. J., and Schillinger, D. (2018a). A discontinuous Galerkin residual-based variational multiscale method for modeling subgrid-scale behavior of the viscous Burgers equation. *International Journal for Numerical Methods in Fluids*, 88(5):217–238.

[179] Stoter, S. K. F., Turteltaub, S. R., Hulshoff, S. J., and Schillinger, D. (2018b). Residual-based variational multiscale modeling in a discontinuous Galerkin framework. *Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal*, 16(3):1333–1364.

[180] Strang, G. and Fix, G. J. (1973). *An analysis of the finite element method*, volume 212. Prentice-hall, Englewood Cliffs, NJ.

[181] Takizawa, K., Tezduyar, T. E., McIntyre, S., Kostov, N., Kolesar, R., and Habluetzel, C. (2014). Space–time VMS computation of wind-turbine rotor and tower aerodynamics. *Computational Mechanics*, 53(1):1–15.

[182] Taylor, C. and Hood, P. (1973). A numerical solution of the Navier-Stokes equations using the finite element technique. *Computers & Fluids*, 1(1):73–100.

[183] ten Eikelder, M. F. P. and Akkerman, I. (2018a). Correct energy evolution of stabilized formulations: The relation between VMS, SUPG and GLS via dynamic orthogonal small-scales and isogeometric analysis. I: The convective–diffusive context. *Computer Methods in Applied Mechanics and Engineering*, 331:259–280.

[184] ten Eikelder, M. F. P. and Akkerman, I. (2018b). Correct energy evolution of stabilized formulations: The relation between VMS, SUPG and GLS via dynamic orthogonal small-scales and isogeometric analysis. II: The incompressible Navier-Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 340:1135–1154.

[185] ten Eikelder, M. F. P., Bazilevs, Y., and Akkerman, I. (2019). A theoretical framework for discontinuity capturing: Joining variational multiscale analysis and variation entropy theory. *Computer Methods in Applied Mechanics and Engineering*, page 112664.

[186] Tezduyar, T. E. (1991). Stabilized finite element formulations for incompressible flow computations. *Advances in Applied Mechanics*, 28(C):1–44.

[187] Tezduyar, T. E. and Osawa, Y. (1999). Finite element stabilization parameters computed from element matrices and vectors. *Computer Methods in Applied Mechanics and Engineering*, 190(3-4):411–430.

[188] Tsinober, A. (2009). *An informal introduction to turbulence*, volume 92 of *Fluid Mechanics and Its Applications*. Springer, Berlin, Heidelberg.

[189] Turner, M., Clough, R. W., Martin, H. C., and Topp, L. J. (1956). Stiffness and deflection analysis of complex structures. *Journal of the Aeronautical Sciences*, 23(9):805–823.

[190] Uranga, A., Persson, P.-O., Drela, M., and Peraire, J. (2011). Implicit large eddy simulation of transition to turbulence at low Reynolds numbers using a discontinuous Galerkin method. *International Journal for Numerical Methods in Engineering*, 87(1-5):232–261.

[191] van Leer, B. (1974). Towards the ultimate conservative difference scheme. II. Monotonicity and conservation combined in a second-order scheme. *Journal of computational physics*, 14(4):361–370.

[192] van Leer, B. (1977a). Towards the ultimate conservative difference scheme III. Upstream-centered finite-difference schemes for ideal compressible flow. *Journal of Computational Physics*, 23(3):263–275.

[193] van Leer, B. (1977b). Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection. *Journal of computational physics*, 23(3):276–299.

[194] van Leer, B. (1979). Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method. *Journal of computational Physics*, 32(1):101–136.

[195] Wang, Z. J., Fidkowski, K., Abgrall, R., Bassi, F., Caraeni, D., Cary, A., Deconinck, H., Hartmann, R., Hillewaert, K., Huynh, H. T., Kroll, N., May, G., Persson, P.-O., van Leer, B., and Visbal, M. (2013). High-Order CFD Methods: Current Status and Perspective. *International Journal for Numerical Methods in Fluids*, 72(8):811–845.

[196] Wheeler, M. F. (1978). An elliptic collocation-finite element method with interior penalties. *SIAM Journal on Numerical Analysis*, 15(1):152–161.

[197] Wu, M. C. H., Kamensky, D., Wang, C., Herrema, A. J., Xu, F., Pigazzini, M. S., Verma, A., Marsden, A. L., Bazilevs, Y., and Hsu, M.-C. (2017). Optimizing fluid–structure interaction systems with immersogeometric analysis and surrogate modeling: Application to a hydraulic arresting gear. *Computer Methods in Applied Mechanics and Engineering*, 316:668–693.

[198] Xu, F., Schillinger, D., Kamensky, D., Varduhn, V., Wang, C., and Hsu, M.-C. (2016). The tetrahedral finite cell method for fluids: Immersogeometric analysis of turbulent flow around complex geometries. *Computers & Fluids*, 141:135–154.

[199] Zhao, Y., Schillinger, D., and Xu, B.-X. (2017). Variational boundary conditions based on the Nitsche method for fitted and unfitted isogeometric discretizations of the mechanically coupled Cahn–Hilliard equation. *Journal of Computational Physics*, 340:177–199.

[200] Zienkiewicz, O. C., Taylor, R. L., Nithiarasu, P., and Zhu, J. Z. (1977). *The finite element method*, volume 3. McGraw-hill, London.